
Ontology of Designing

ISSN 2223-9537 (P)

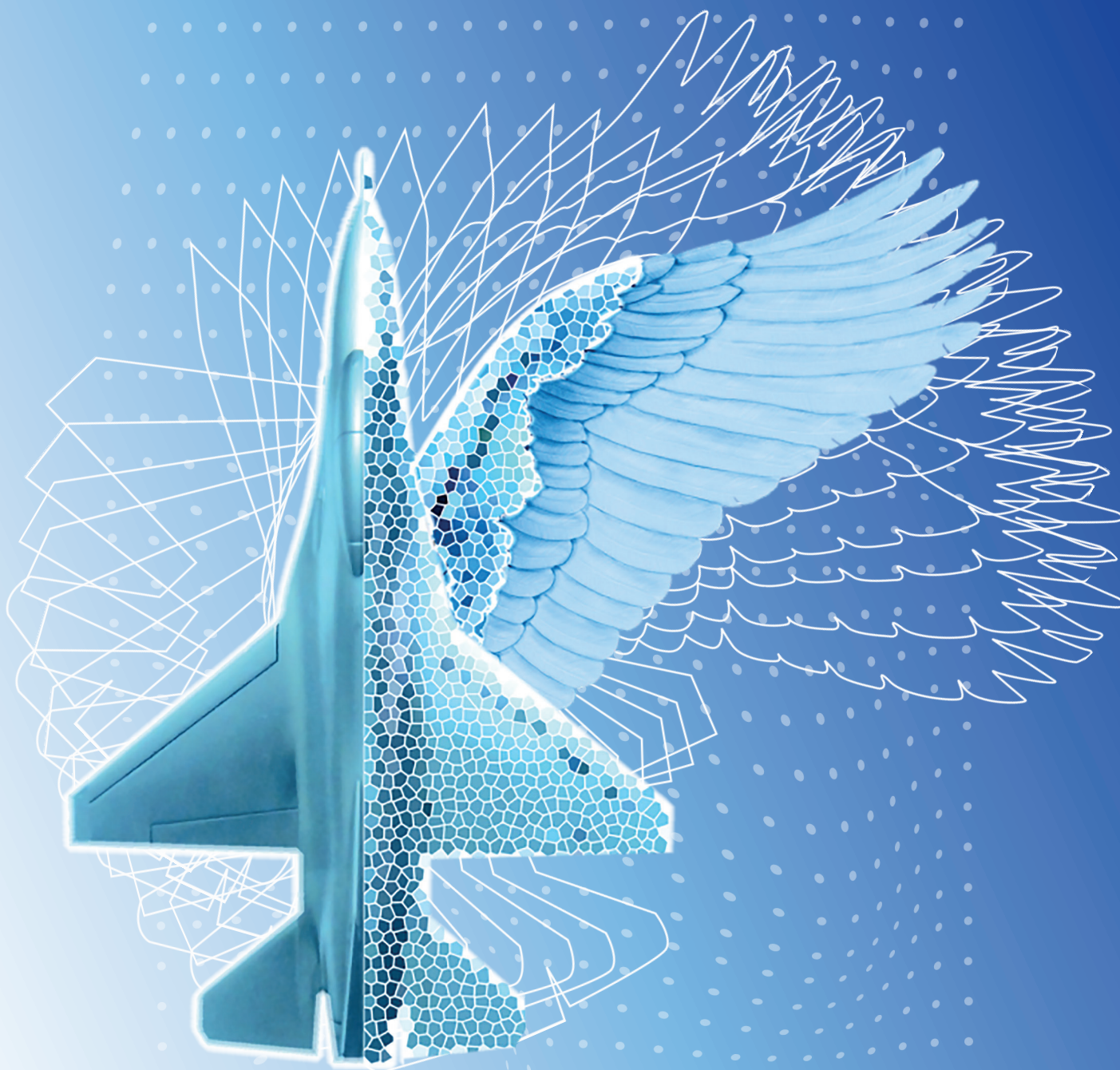
ISSN 2313-1039 (E)

ОНТОЛОГИЯ

Vol 13
2023
№1

Научный журнал -
Scientific journal

ПРОЕКТИРОВАНИЯ



Передовые
инженерные
школы

Scientific journal

Volume 13

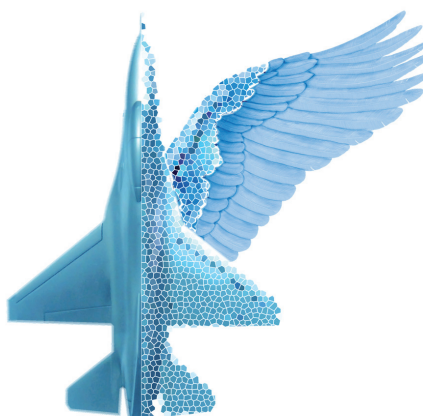
N 1

ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ

Научный журнал

Том 13

№ 1



Editorial Board - Редакционная коллегия

Nikolay M. **Borgest***, Ph.D., Associate Professor, Samara University, Member of IAOA, AAAl. Samara, Russia
Stanislav N. **Vasiliev***, Doctor of Phys. and Math. Sciences, Professor, Academician of RAS, ICS RAS, Moscow, Russia
Tatiana A. **Gavrilova***, Doctor of Technical Sciences, Professor, GSOM SPbU, St.-Petersburg, Russia
Vladimir G. **Gainutdinov**, Doctor of Technical Sciences, Professor, KNITU-KAI, Kazan, Russia
Vladimir V. **Golenkov***, Doctor of Technical Sciences, Professor, BSUIR, Minsk, Belarus
Vladimir I. **Gorodetsky***, Doctor of Technical Sciences, Professor, JSC «EVRIKA», St. Petersburg, Russia
Valeriya V. **Gribova***, Doctor of Technical Sciences, Corresponding Member of RAS, Senior Researcher, IAPU of the Far Eastern Branch of RAS, Vladivostok, Russia
Yury A. **Zagorulkov***, Ph.D., Senior Researcher, ISI of the Siberian Branch of RAS, Novosibirsk, Russia
Valery A. **Komarov**, Doctor of Technical Sciences, Professor, Samara University, Samara, Russia
Vladik **Kreinovich**, Ph.D., Professor, University of Texas at El Paso, El Paso, USA
Venedikt S. **Kuzmichev**, Doctor of Technical Sciences, Professor, Samara University, Samara, Russia
Victor M. **Kureichik***, Doctor of Technical Sciences, Professor, Southern Federal University, Taganrog, Russia
Dmitry V. **Lande***, Doctor of Technical Sciences, Senior Researcher, IPRI NASU, Kiev, Ukraine
Paulo **Leitao**, Professor at Polytechnic Institute of Bragança, Bragança, Portugal
Vladimir **Marik**, Professor, Scientific Director of the CIIRC of the Czech Technical University in Prague, Praha, Czech Republic
Lyudmila V. **Massel***, Doctor of Technical Sciences, Professor, ISEM of the Siberian Branch of RAS, Irkutsk, Russia
Aleksandr Yu. **Nesterov**, Doctor of Philosophy, Professor, Samara University, Samara, Russia
Dmitry A. **Novikov**, Doctor of Technical Sciences, Professor, Academician of RAS, ICS RAS, Moscow, Russia
Alexander V. **Palagin**, Doctor of Technical Sciences, Professor, Academician of the NASU, Ins. of Cybernetics, Kiev, Ukraine
Semyon A. **Piyavsky**, Doctor of Technical Sciences, Professor, Moscow City Pedagogical University, Samara, Russia
Yury M. **Reznik**, Doctor of Philosophy, Professor, Institute of Philosophy of RAS, Moscow
George **Rzevski**, Professor, Open University, London, UK
Peter O. **Skobelev***, Doctor of Technical Sciences, «Smart solutions» Scientific Production C., Samara, Russia
Sergey V. **Smirnov***, Doctor of Technical Sciences, ICCS RAS, member of IAOA, Samara, Russia
Dzhavdet S. **Suleymanov***, Doctor of Technical Sciences, Professor, Academician of the AS of the RT, Kazan, Russia
Boris E. **Fedunov***, Doctor of Technical Sciences, Professor, State Research Institute of Aviation Systems, Moscow, Russia
Altynbek **Sharipbay***, Doctor of Technical Sciences, Professor, Institute of Artificial Intelligence, Astana, Kazakhstan

Боргест Николай Михайлович*, к.т.н., доцент, Самарский университет, член IAOA, AAAl. Самара, Россия
Васильев Станислав Николаевич*, д.ф.-м.н., профессор, академик РАН, ИПУ РАН, Москва, Россия
Гаврилова Татьяна Альбертовна*, д.т.н., профессор, ВШМ СПбУ, Санкт-Петербург, Россия
Гайнутдинов Владимир Григорьевич, д.т.н., профессор, КНИТУ-КАИ, Казань, Россия
Голенков Владимир Васильевич*, д.т.н., профессор, БГУИР, Минск, Беларусь
Городецкий Владимир Иванович*, д.т.н., профессор, АО «Эврика», Санкт-Петербург, Россия
Грибова Валерия Викторовна*, д.т.н., член-корреспондент РАН, г.н.с., ИАПУ ДВО РАН, Владивосток, Россия
Загорулков Юрий Алексеевич*, к.т.н., с.н.с., ИСИ СО РАН, Новосибирск, Россия
Комаров Валерий Андреевич, д.т.н., профессор, Самарский университет, Самара, Россия
Крейнвич Владик, профессор, Техасский университет Эль Пасо, Эль Пасо, США
Кузьмичев Венедикт Степанович, д.т.н., профессор, Самарский университет, Самара, Россия
Курейчик Виктор Михайлович*, д.т.н., профессор, Южный федеральный университет, Таганрог, Россия
Ландэ Дмитрий Владимирович*, д.т.н., с.н.с., ИПРИ НАН Украины, Киев, Украина
Лейтао Пауло, профессор, Политехнический институт Браганса, Браганса, Португалия
Марик Владимир, профессор, научный директор ЧИИРК Чешского технического университета, Прага, Республика Чехия
Массель Людмила Васильевна*, д.т.н., профессор, ИСЭМ СО РАН, Иркутск, Россия
Нестеров Александр Юрьевич, д.филос.н., профессор, Самарский университет, Самара, Россия
Новиков Дмитрий Александрович, д.т.н., профессор, академик РАН, ИПУ РАН, Москва, Россия
Палагин Александр Васильевич, д.т.н., профессор, академик НАН Украины, Ин-т кибернетики, Киев, Украина
Пиявский Семён Авраамович, д.т.н., профессор, Московский город. педагог. университет, Самара, Россия
Резник Юрий Михайлович, д.филос.н., профессор, Институт философии РАН, Москва, Россия
Ржевский Георгий, профессор, Открытый университет, Лондон, Великобритания
Скобелев Петр Олегович*, д.т.н., ФИЦ СамНЦ РАН, Самара, Россия
Смирнов Сергей Викторович*, д.т.н., ИПУСС РАН – СамНЦ РАН, член IAOA, Самара, Россия
Сулейманов Джавдет Шевкетович*, д.т.н., профессор, академик АН РТ, Казань, Россия
Федунов Борис Евгеньевич*, д.т.н., профессор, ГосНИИ авиационных систем, Москва, Россия
Шарипбай Алтынбек*, д.т.н., профессор, Ин-т искусственного интеллекта, Астана, Казахстан

* - members of the Russian Association of Artificial Intelligence - члены Российской ассоциации искусственного интеллекта - http://www.raai.org/about/about.shtml?raai_list

Executive Editorial Board - Исполнительная редакция

Chief Editor	P.O. Skobelev	Samara, Russia	Главный редактор	Скобелев П.О. ФИЦ СамНЦ РАН, Самара, Россия
Deputy Chief Editor	S.V. Smirnov	Samara, Russia	Зам. главного редактора	Смирнов С.В. ИПУСС РАН – СамНЦ РАН, Самара, Россия
Executive Editor	N.M. Borgest	Samara, Russia	Выпускающий редактор	Боргест Н.М. Самарский университет, Самара, Россия
Editor	D.M. Kozlov	Samara, Russia	Редактор	Козлов Д.М. Самарский университет, Самара, Россия
Technical Editor	D.N. Borgest	Samara, Russia	Технический редактор	Боргест Д.Н. Самарский университет, Самара, Россия
Executive Secretary	S.A. Vlasov	Samara, Russia	Ответственный секретарь	Власов С.А. Самарский университет, Самара, Россия

The journal has entered into an electronic licensing relationship with EBSCO Publishing, the world's leading aggregator of full text journals, magazines and eBooks. The full text of JOURNAL can be found in the **EBSCOhost™** databases. The journal has been successfully evaluated in the evaluation procedure for the **ICI Journals Master List** 2014-2019 and journal received the ICV (Index Copernicus Value).

Журнал размещен в коллекции «Издания по естественным наукам» на платформе **EastView**.

Журнал включён в перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёных степеней кандидата и доктора наук (Письмо Департамента аттестации научных и научно-педагогических работников Минобрнауки РФ от 01.12.2015 № 13-6518) по научным специальностям 05.13.17 (с 26.03.19); 1.2.2, 2.3.1, 2.5.13, 2.5.15 (с 1.02.2022).

Журнал включен в список журналов, входящих в базу данных **Russian Science Citation Index (RSCI)** на платформе **Web of Science**. Пятилетний импакт-фактор РИНЦ **1,00** (2013), **0,92** (2014), **1,30** (2015), **1,08** (2016), **1,00** (2017), **1,18** (2018), **0,85** (2019), **1,08** (2020), **1,00** (2021).

Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство ПИ № ФС 77-70157 от 16.06.2017 г. (ранее выданное свидетельство ПИ № ФС 77-46447 от 07.09.2011 г.)

СОДЕРЖАНИЕ

ОТ РЕДАКЦИИ

Онтологии в компьютерных науках

5-9

ОБЩИЕ ВОПРОСЫ ФОРМАЛИЗАЦИИ ПРОЕКТИРОВАНИЯ: ОНТОЛОГИЧЕСКИЕ АСПЕКТЫ

С.Н. Васильев

10-28

Интерактивное порождение новых знаний на основе автоматических средств логического вывода

С.В. Микони

29-43

Подход к оцениванию уровня интеллектуальности информационной системы

ПРИКЛАДНЫЕ ОНТОЛОГИИ ПРОЕКТИРОВАНИЯ

Л.В. Массель, А.Г. Массель

44-54

Семантическое моделирование при построении цифровых двойников энергетических объектов и систем

А.Е. Колоденкова

55-74

Онтология идентификации человека по движениям тела и лицу в видеонаблюдениях

И.С. Ткаченко, С.Л. Сафронов, М.Д. Коровин, М.А. Иванушкин, А.В. Крестина

75-89

Информационная система поддержки предварительного проектирования малых космических аппаратов

Хамза Мазин Абдулаали Хамза

90-98

Проектирование системы сбора полётной информации для расчёта аэродинамических характеристик БПЛА

А.В. Соловов, А.А. Меньшикова

99-112

Проектирование онтологии содержания электронного учебного курса

ИНЖИНИРИНГ ОНТОЛОГИЙ

А.С. Серый

113-124

Доверие к данным при пополнении онтологий и графов знаний

МЕТОДЫ И ТЕХНОЛОГИИ ПРИНЯТИЯ РЕШЕНИЙ

В.Е. Гвоздев, О.Я. Бежаева, Г.Р. Сафина

125-138

Многоаспектное моделирование ситуаций в задачах обеспечения функциональной безопасности аппаратно-программных комплексов

И.Т. Кимяев

139-149

Онтологический нечётко-возможностный подход к созданию модели управления печью пиролиза

Рекомендуемые издания

150-152

Журнал ориентирован на учёных и специалистов, работающих по научным направлениям: онтологические аспекты общих вопросов формализации проектирования, прикладные онтологии проектирования, инжиниринг онтологий, методы и технологии принятия решений.

Правила подготовки рукописей статей размещены на сайте журнала «Онтология проектирования»:

http://agora.guru.ru/scientific_journal/.

Контент журнала распространяется по лицензии CC-BY 4.0 (Creative Commons Attribution 4.0 International License).



Контакты учредителей

ИПУСС РАН – СамНЦ РАН: 443020, Самара, ул. Садовая, 61, тел./факс.: +7 (846) 333 27 70, Смирнов С.В., smirnov@iccs.ru.

Самарский университет: 443086, Самара, Московское шоссе 34, корп. 10, тел.: +7 (846) 267 46 47, Боргест Н.М., borgest@yandex.ru.

ООО «Новая техника» (издательство): 443010, Самара, ул. Фрунзе, 145, тел.: +7 (846) 332 67 84, факс: +7 (846) 332 67 81.

Отпечатано в ООО «Новая техника», г. Самара, пр. К. Маркса, 24-76. Дата выхода 07.03.2023. Тираж 200 экз. Свободная цена. (6+).

CONTENTS

EDITORIAL

Ontologies in computer science	5-9
--------------------------------	-----

GENERAL ISSUES OF FORMALIZATION IN THE DESIGNING: ONTOLOGICAL ASPECTS

S.N. Vassilyev	10-28
Interactive generation of new knowledge based on automatic means of logical inference	
S.V. Mikoni	29-43
Approach to assessing the level of intelligence of an information system	

APPLIED ONTOLOGIES OF DESIGNING

L.V. Massel, A.G. Massel	44-54
Semantic modeling in the construction of digital twins of energy objects and systems	
A.E. Kolodenkova	55-74
Ontology of human identification by face and body motions in video surveillance systems	
I.S. Tkachenko, S.L. Safronov, M.D. Korovin, M.A. Ivanushkin, A.V. Krestina	75-89
An information system to support preliminary design of small spacecraft	
Hamzah Mazin Abdulaali Hamzah	90-98
Designing a flight information collection system for calculating the aerodynamic characteristics of a UAV	
A.V. Solovov, A.A. Menshikova	99-112
Designing an ontology of the e-learning course content	

ONTOLOGY ENGINEERING

A.S. Sery	113-124
Data credibility when populating ontologies and knowledge graphs	

METHODS AND TECHNOLOGIES OF DECISION MAKING

V.E. Gvozdev, O.Ya. Bezhaeva, G.R. Safina	125-138
Multi-aspect modeling of situation in the functional safety control tasks of hardware and software complexes	
I.T. Kimyaev	139-149
Ontological fuzzy-possibility approach to creating a pyrolysis furnace control model	

Recommended Books	150-152
--------------------------	---------

The journal is aimed at scientists and specialists working in the following research areas: ontological aspects of general issues of design formalization, applied design ontologies, ontology engineering, methods and technologies of decision making.

The current version of the Rules for the preparation of manuscripts of articles for the journal «Ontology of Designing» is on the journal website:

http://agora.guru.ru/scientific_journal/

The content of the scientific journal is distributed under a license **CC-BY 4.0**
(Creative Commons Attribution 4.0 International License)



Contacts of the Founders

Samara Scientific Center of the RAS: 61, Sadovaya st., Samara, 443020, Russia. Tel.: +7 (846) 333 27 70, S.V. Smirnov, smirnov@iccs.ru
Samara University: 34, Moskovskoye shosse, bldg. 10, Samara, 443086, Russia. Tel.: +7 (846) 267 46 47, N.M. Borgest, borgest@yandex.ru
New Engineering LLC (publishing house): 145, Frunze st., Samara, 443010, Russia. Tel.: +7 (846) 332 67 84, fax: +7 (846) 332 67 81



ОТ РЕДАКЦИИ

Онтологии в компьютерных науках

Ontologies in computer science

«To be, or not to be, that is the question...»

*William Shakespeare,
«Hamlet». 1600*

«... to be is to be the value of a variable»

*Willard Van Orman Quine
«On What There Is». 1948*

Дорогой наш читатель, уважаемые авторы и члены редакционной коллегии!

Многозначность и многозначительность всего и вся, скрытый и неявный смысл посланий, льющихся из разных источников, всё это позволило обратить свой взор к классикам, вопрошающим «быть или не быть» и постулирующим «быть, значит быть значением переменной». Быть значимым и иметь значение! Именно бытие и будущее, вытекающее из него, есть предмет изучения в онтологии проектирования, в лабиринтах которой авторы публикаций стремятся познать и раскрыть неизвестные ранее формализмы процессов, протекающих в различных предметных областях (ПрО), а добытые ими знания, став всеобщим достоянием, откроют возможность развития этих идей в работах других исследователей.

Перед Вами, дорогой наш читатель, первый номер нового 13-ого тома. Число 13 в некоторых культурных традициях не столь почитаемо, но в науке различного рода суеверия не находят своего подтверждения. Поэтому наши надежды с выпуском нового тома - радужные. Ведь число 13 в эзотерике и нумерологии означает скорее число трансформации и символизирует не конец жизни (на чём мы твёрдо настаиваем), а наоборот - переход из одного состояния в другое, в начало новой фазы жизни¹. Именно так мы трактуем и хотим видеть будущее нашего журнала, эволюционирующего, как и всё живое.

Мысли и особенно слова – материальны. Поэтому печатным словом посылаем позитив всем сущностям в онтологии проектирования и далеко за её пределами. Пусть живые сущности, акторы выстраивают лишь позитивные отношения, демонстрируют лишь свои лучшие качества и атрибуты, пусть процессы в ПрО соответствуют колониальным принципам, в которых есть место и идеальному государству Платона, и городу Солнца Томмазо Кампанеллы, и коммунизму Маркса, и эвергетике Владимира Виттиха. Пусть критериями процессов в разных ПрО и выбираемых значений атрибутов создаваемых артефактов всегда будут векторы жизненно важных показателей, дающие ростки и смыслы всем сущностям.

¹ В отличие от ожидаемого числа 13, которое сразу следует после 12, цивилизационные сдвиги, свидетелями которых мы являемся, наступили неожиданно и были трудно предсказуемы. Причём это не природные землетрясения в Турции и Сирии, унёсшие десятки тысяч жизней в феврале этого года, а человеческий, чисто субъективный фактор, который привёл к разрушению сложившихся отношений практически во всех сферах жизни. Жертв и разрушений от сформированных элитами идей, взглядов, устоев и принципов, от их неспособности найти консенсус, договориться в нашем высокотехнологичном мире, может быть значительно больше, чем от природных катаклизмов. Мы научились бороться со многими стихийными бедствиями, предсказывать их, лечить самые страшные болезни, но не смогли выстроить систему устойчивого развития, мирного сосуществования. Капитализм (общество, основанное на прибыли и потреблении) жадно истощает земные ресурсы, не заботясь об образовании человека. А ведь именно человек, как субъект в онтологии проектирования любой ПрО, определяет ход всех процессов в ПрО. Именно он - главный актор, который формирует принципы и критерии развития, он является лицом, принимающим решение, он должен нести и несёт ответственность за происходящее в ПрО.

Краткая история онтологии в информатике

Продолжая тему чисел, наступивший **2023** год можно было бы считать юбилейным для онтологии в информатике, если за отсчёт взять широко известную и часто цитируемую статью «К принципам проектирования онтологий, используемых для обмена знаниями» американского учёного Т. Грубера². Работа была представлена на международном семинаре по формальным онтологиям в **1993** году³ и переиздана в 1995 году в журнале «Человеко-машинные исследования». Т. Грубер использовал онтологию как технический термин в информатике, тесно связанный с более ранней идеей семантических сетей и таксономий, и определил этот термин как спецификацию концептуализации. Онтология - это описание понятий и отношений, которые могут формально существовать для агента или сообщества агентов. Онтологии часто отождествляются с таксономическими иерархиями классов, определениями классов и отношениями включения, но онтологии не обязательно должны ограничиваться этими формами. Онтологии также не ограничиваются только терминологией, определениями в традиционном логическом смысле. Для определения концептуализации необходимо указать аксиомы, которые ограничивают возможные интерпретации для определённых терминов.

Но эта была не первая работа по онтологиям в информатике. Т. Грубер ещё в 1992 году выпустил технический отчёт «Подход к переводу спецификаций переносимых онтологий»⁴, в котором есть ссылки на его более ранние работы по онтологиям⁵ и работы других авторов, использующих понятие онтологии⁶.



Нельзя не упомянуть, что уже в **1993** году для русскоязычной аудитории стал доступен перевод на русский язык книги С. Шлеер и С. Меллора «Объектно-ориентированный анализ: моделирование мира в состояниях»⁷, в которой изложены первые этапы процесса разработки сложных программных, технических или других систем. Методология объектно-ориентированного анализа разработана в книге как метод отождествления важных сущностей в задачах реального мира, объяснения и понимания того, как они взаимодействуют. Фактически в книге представлен метод построения онтологии ПрО, который изложен в виде примеров и руководящих принципов; при этом авторы, применяя онтологический анализ ПрО,

не употребляют этот термин.

Позднее К. Фейлмайр и В. Вёсс, уточняя трактовку Т. Грубера, определили онтологию как «формальную, явную спецификацию общей концептуализации, которая характеризуется высокой семантической выразительностью, необходимой для повышенной сложности»⁸.

² Tom Gruber. <https://tomgruber.org/>.

³ Gruber T. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies*. 43 (5–6): 907–928. doi:10.1006/ijhc.1995.1081.

⁴ Thomas R. Gruber. A Translation Approach to Portable Ontology Specifications. Knowledge Systems Laboratory September 1992 Technical Report KSL 92-71 Revised April 1993. 27 p.

⁵ Gruber T.R. (1991). The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases. In J. A. Allen, R. Fikes, & E. Sandewall (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, Cambridge, MA, pages 601-602, Morgan Kaufmann.

Gruber T.R. (1992). Ontolingua: A mechanism to support portable ontologies. Technical Report KSL 91-66, Stanford University, Knowledge Systems Laboratory. Revision.

⁶ Guarino N. (1992). Concepts, Attributes, and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases. *Data and Knowledge Engineering*, 8.

Bradshaw J.M., Holm P.D., Boose J.H. (1992). Sharable ontologies as a basis for communication and collaboration in conceptual modeling. *Proceedings of the 7th Knowledge Acquisition for Knowledge-based Systems Workshop*, Banff, Canada.

⁷ Sally Shlaer and Stephen J. Mellor. *Object Lifecycles: Modeling the World in States*, Prentice-Hall, 1992. 251 p.

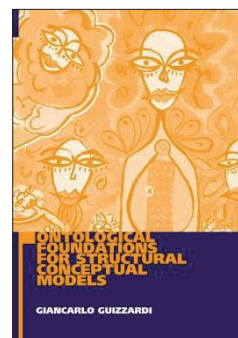
⁸ Feilmayr Christina, Wöß Wolfram (2016). An analysis of ontologies and their success factors for application to business. *Data & Knowledge Engineering*. 101: 1–23. doi:10.1016/j.datak.2015.11.003.

Эти детализирующие и конкретизирующие определения, позволяющие в конечном итоге построить модель бытия, отличаются от философских трактовок, от философской онтологии, где рассматривается «бесконечный и живой» поток бытия. «Бесконечный» в своих элементах, аспектах, контекстах, значениях, связях, комбинациях и «живой» – в силу связности, постоянных изменений, трансформаций, логики развития, включая переход в свои диалектические противоположности⁹. Это явный отличительный признак философской онтологии от онтологии в информатике, которая предполагает модельное упрощение.

Поэтому поиск первоисточника онтологии в информатике привёл всё-таки к философам. Именно они сформировали те концепции, которые впоследствии развили специалисты, разрабатывающие информационные технологии и в конечном итоге добравшиеся до создания систем с искусственным интеллектом (ИИ). Можно предположить, что таким философом мог быть Уиллард Ван Орман Куайн (1908-2000)¹⁰. Во многих отношениях он продолжил подход Бертрانا Рассела (1872-1970) с его акцентом на формальную логику и онтологию. Здесь же следует отметить статью Куайна 1948 года «О том, что есть»¹¹, впервые опубликованную в журнале «Обзор метафизики» и впоследствии включённую в его книгу «С логической точки зрения»¹². Именно на работу Куайна ссылался компьютерный консультант из Массачусетса Джордж Х. Мили в своей работе «Другой взгляд на данные»¹³, используя понятие *онтология* при описании своей модели. Предложенные им теоретическая модель данных и модель обработки данных представляют собой систему наборов сущностей, значений, карт данных и карт процедур. Сущности соответствуют объектам реального мира, данные о которых записываются или вычисляются. Карты данных присваивают значения атрибутам сущностей. Структурные данные, как особый тип карты данных, представляют собой набор самих сущностей; структурные карты состоят из указателей. Процедуры — это операции над картами данных, создающие новые карты данных. Обработка данных происходит в памяти вычислительной системы, которая в идеале является представлением моделируемой реальной или абстрактной системы. Описание данных — это спецификация систем и представлений машинных данных; тип данных — это фрагмент описания данных, описывающий объект и его применимые карты. В конце статьи автор утверждает, что стандартизация методов описания данных может оказаться гораздо более важной, чем стандартизация методов представления данных и спецификации процедур. Тем самым указывая на важность построения онтологии ПрО.

Исторический анализ развития онтологии от философских начал до онтологии в компьютерных и информационных науках представлен в докторской диссертации Джанкарло Гиззарди «Онтологические основы структурных концептуальных моделей»¹⁴, где рассмотрены онтологии в информационных системах, различных ПрО, ИИ, семантической сети. Там же даны терминологические уточнения понятия онтологии и её формальные характеристики, применительно к информатике.

Согласно английской Википедии¹⁵, с середины 1970-х годов исследователи в области ИИ признали, что инженерия знаний является ключом к созданию систем ИИ. Исследователи ИИ полагали, что создание новых онтологий в качестве



⁹ Александр Клейн. Проблема онтологии и человека. 8 января 2023. (Письмо в редакцию журнала). Редакция надеется познакомить читателей журнала с обстоятельной позицией по онтологии А. Клейна в ближайшем номере.

¹⁰ Willard Van Orman Quine. https://en.wikisource.org/wiki/Author:Willard_Van_Orman_Quine.

¹¹ Willard Van Orman Quine (1948). https://en.wikisource.org/wiki/On_What_There_Is.

¹² Willard Van Orman Quine (1953). From a logical point of view; 9 logico-philosophical essays. Harvard Univ. Press.

¹³ Mealy G.H. Another Look at Data, Proceedings of the Fall Joint Computer Conference, November 14-16, Anaheim, California (AFIPS Conference Proceedings, Volume 31), Washington, DC: Thompson Books, London: Academic Press, 525-534, 1967.

¹⁴ Giancarlo Guizzardi. Ontological Foundations for Structural Conceptual Models. Enschede, The Netherlands, 2005. 441 p.

¹⁵ Ontology (computer science). [https://en.wikipedia.org/wiki/Ontology_\(computer_science\)#cite_ref-17](https://en.wikipedia.org/wiki/Ontology_(computer_science)#cite_ref-17).

вычислительных моделей позволит выполнять определённые виды автоматизированных рассуждений. Но там же¹⁵ первенство в использовании термина «онтология» для обозначения теории моделируемого мира и компонент систем, основанных на знаниях, ошибочно отдаёт австралийскому учёному Дэвиду Пауэрсу, который ввёл его в ИИ для обозначения реального мира или «роботизированного заземления»¹⁶. Действительно, в работе «Интеллект роботов», опубликованной в 1983 году, особое внимание уделяется построению всех значимых атрибутов онтологии ПрО для роботов. Правда, сам термин онтология в этой работе не используется, хотя в последующих статьях автор уже не смог обойтись без него.

В России история онтологии в компьютерных науках развивалась со своим временным лагом, поспешая за коллегами из-за рубежа. Впервые онтологии в ПрО и онтологии как форму представления знаний о ПрО в информационных системах можно было найти в работах Т.А. Гавриловой, В.А. Виттиха, С.В. Смирнова¹⁷, В.Ф. Хорошевского¹⁸ и др. Первая докторская диссертация по онтологиям в России – это работа С.В. Смирнова «Онтологический анализ в системах компьютерного моделирования»¹⁹. Первые учебные курсы по дисциплинам «Онтологии производственной сферы» (2006), по «Онтологии проектирования» (2010) разработаны в Самарском университете.

Подводя итог краткому экскурсу по истории онтологии в информатике и учитывая нашу любовь к поиску повода для праздника, можно смело говорить о юбилее онтологии в компьютерных науках. Поэтому 2023 год – это юбилейный год! Это 30 лет по Груберу, 40 лет по Пауэрсу и 75 лет по Куайну! С праздником Вас, дорогие товарищи онтологи!

Прогноз, который не ждали

«Кому - война, а кому - мать родна». Кто празднует юбилей и ищет повод отвлечься, а кто создаёт апокалиптические прогнозы ближайшего будущего.

В опубликованном в начале года отчёте «Генеративные языковые модели и автоматизированные операции влияния: возникающие угрозы и потенциальные меры по их смягчению»²⁰ содержится стратегическая аналитика катастрофического сценария, описанного в обращении Кай-Фу Ли к мировому сообществу разработчиков ИИ (см. также о его книге «ИИ-2041. Десять образов нашего будущего» в нашем журнале №3, 2022 г.). Кай-Фу Ли рассказал, что в 2023 году мир столкнётся с риском более серьёзным, чем риски ядерной войны, голода или пандемии. Это риск того, что технологии генерации контента с помощью ИИ (*Artificial Intelligence Content Generation, AICG*) сделают неотличимыми правду и ложь, как для индивидов, так и для общества. Новый отчёт расписывает в деталях, как, скорее всего, будет происходить процесс падения человечества в пропасть *AICG*. При этом вышедшая статья в медийном журнале о трендах ИИ, также поставившая на первое место генеративный ИИ, завершается словами «кто знает, может быть, даже эта статья была написана какой-нибудь генеративной нейронной сетью»²¹.

¹⁶ Powers David MW (1983) Robot Intelligence. Electronics Today International.

Powers David MW (1984) Natural Language the Natural Way. Computer Compacts. 2 (3–4): 100–109.

Powers David MW (1991) Goals, Issues and Directions in Machine Learning of Natural Language and Ontology. AAAI Spring Symposium on Machine Learning of Natural Language and Ontology. Stanford CA (March 1991).

¹⁷ Смирнов С.В. Онтологическая относительность и технология компьютерного моделирования сложных систем. *Известия СамНЦ РАН*. 2000. Т.2(1). С.66–71.

Смирнов С.В. Онтологический анализ предметных областей моделирования. *Известия СамНЦ РАН*. 2001. Т.3(1). С.62–70.

¹⁸ Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. СПб.: Питер. 2000. 384 с.

¹⁹ Смирнов С.В. Онтологический анализ в системах компьютерного моделирования. Д.т.н. Специальность ВАК РФ 05.13.01. 2002. 348 с. <https://www.dissercat.com/content/ontologicheskii-analiz-v-sistemakh-kompyuternogo-modelirovaniya>.

²⁰ Josh A. Goldstein, Girish Sastry, Micah Musser, Renée DiResta, Matthew Gentzel, Katerina Sedova. Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. January 2023. 82 p. arXiv:2301.04246v1 [cs.CY] 10 Jan 2023.

²¹ Alexander Glukhov. Top 5 trends in AI. Dec 29, 2022. <https://medium.com/letavc/top-5-trends-in-ai-47f919543800>.

Кратко о языке. Важные замечания для авторов

Государственная Дума РФ приняла закон²² о защите русского языка от чрезмерных иностранных заимствований. Закон вводит положения о нормативных словарях, справочниках и грамматиках, где фиксируются нормы современного русского литературного языка, которые необходимо соблюдать при использовании русского языка, в т.ч. в СМИ, т.е. напрямую касается и научных журналов. При использовании русского языка не допускается употребление слов и выражений, не соответствующих нормам современного русского литературного языка, за исключением иностранных слов, которые не имеют общепотребительных аналогов в русском языке и перечень которых содержится в указанных нормативных словарях. *Коллеги, строго следим за языком!*

Обложка: краткие итоги опроса читателей

Новый том и новая радужная обложка мотивируют нас на новое качество публикуемых материалов. В конце прошлого года, в 4-ом номере журнала нашим читателям было предложено²³ оценить варианты разрабатываемой новой обложки нашего журнала. На суд читателей было вынесено три варианта концепций (см. рисунок). На каждый вариант предложенной обложки нашлись свои почитатели. Но большинство голосов отдали предпочтение 3-ему варианту обложки. Что мы и можем видеть на лицевой части нашего нового тома.



1

2

3

В номере

Открывает новый том статья академика РАН С.Н. Васильева (Москва), члена редколлегии нашего журнала с момента его образования. Статья посвящена актуальной теме порождения новых знаний на основе автоматических средств логического вывода. Это дальнейшее развитие работ, выполняемых автором в области предложения им исчисления позитивно-образованных стандартизованных формул (пос-формул).

Наибольшую дискуссию в редколлегии вызвала статья профессора С.В. Микони (Санкт-Петербург) о предложенном им подходе к оцениванию уровня интеллектуальности информационной системы (ИС). В эпоху повсеместного внедрения ИС и ИИ, эта тема достойна развития. Поэтому редакция надеется, что публикация этой статьи даст импульс критической оценке подхода и расширит спектр предложений по оценке уровня интеллектуальности ИС.

Раздел «Прикладные онтологии проектирования» богат на предлагаемые решения в различных ПрО. Это: семантическое моделирование энергетических объектов (проф. Л.В. Массель и доц. А.Г. Массель, Иркутск), онтология идентификации в видеонаблюдениях (проф. А.Е. Колоденкова, Самара), ИС поддержки предварительного проектирования малых космических аппаратов (коллектив авторов под руководством доц. И.С. Ткаченко, Самара), проектирование системы сбора полётной информации БПЛА (Хамза Мазин, иракский аспирант из Казани), на английском языке представлен подход к проектированию онтологии содержания электронного учебного курса (проф. А.В. Соловов и доц. А.А. Меньшикова, Самара).

В разделе «Инжиниринг онтологий» статья о доверии к данным при пополнении онтологий и графов знаний (м.н.с. А.С. Серый, Новосибирск). Автор полагает, что предлагаемая модель оценки доверия к информации способна демонстрировать эффективные показатели на задачах оценки численных данных.

В разделе «Методы и технологии принятия решений» рассмотрено моделирование ситуаций в задачах обеспечения функциональной безопасности аппаратно-программных комплексов (проф. В.Е. Гвоздев, к.т.н. О.Я. Бежаева и Г.Р. Сафина, Уфа), а также онтологический нечётко-возможностный подход к созданию модели управления печью пиролиза (к.т.н. И.Т. Кимяев, «Норникель Спутник», Москва).

Завершает номер рекомендованное издание, в переводе которого на русский язык принял участие член редколлегии нашего журнала проф. В.И. Городецкий. Там же приведён фрагмент разработанного им толкового англо-русского словаря терминов науки о данных.

Наш журнал - место для научных дискуссий. Мы ждём новых интересных результатов исследований, критического анализа и развития уже опубликованных работ.

Dum spiro, spero! Ontologists and designers of all countries and subject areas, join us!

²² Проект Федерального закона №221977-8 «О внесении изменений в ФЗ «О государственном языке РФ».

²³ От редакции. Онтология Армагеддона: поиск выхода. Онтология проектирования, том 12, №4, 2022. С.425-429.

ОБЩИЕ ВОПРОСЫ ФОРМАЛИЗАЦИИ ПРОЕКТИРОВАНИЯ: ОНТОЛОГИЧЕСКИЕ АСПЕКТЫ

УДК 004.83

Научная статья

DOI: 10.18287/2223-9537-2023-13-1-10-28



Интерактивное порождение новых знаний на основе автоматических средств логического вывода

© 2023, С.Н. Васильев

Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия

Аннотация

Работа относится к области представления знаний и алгоритмизации их обработки. Для описания предмета рассмотрения совокупностью утверждений и их обработки используются логические средства первого порядка. В отличие от известных достижений в области автоматического доказательства (АД) в формальных аксиоматических теориях, автоматический синтез текстов утверждений затруднён плохой формализуемостью понятия ценности знания, содержащегося в тексте. В работе на основе развития исчисления позитивно-образованных стандартизованных формул (пос-формул) исследуются возможности автоматизации анализа описания предмета рассмотрения и интерактивного порождения содержательных утверждений для совершенствования этого описания. Предлагаемый метод интерактивного порождения знаний объединяет возможности АД с алгоритмизированным человеко-машинным синтезом текстов новых знаний. АД в исчислении пос-формул помогает человеку не только в анализе свойств, но и в выборе требуемых свойств описания предмета рассмотрения. При выявлении посредством АД отсутствия в анализируемых знаниях свойства, желательного для человека, осуществляется формирование гипотезы, логически гарантирующей наличие требуемого свойства. Автоматически синтезируются в предварительном виде варианты гипотез, дорабатываемые затем интерактивно для повышения содержательности порождаемого текста нового знания. Обосновываются правила синтеза. Их применение продолжает вывод в ситуациях неприменимости базового правила вывода в исчислении пос-формул. Расширение исчисления правилами синтеза преобразует его как первоепорядковое исчисление дедуктивного типа в средство дедуктивно-абдуктивного вывода, которое отличается от известных работ, использующих абдукцию, отсутствием ряда ограничений. На иллюстративных примерах показаны представление, анализ и порождение знаний. На этих примерах при интерактивной обработке знаний видна важность наглядности обрабатываемого текста, обеспечиваемой крупноблочной структурой пос-формул, построенных из типовых кванторов. В заключение сформулированы итоги работы и возможные направления развития полученных результатов.

Ключевые слова: представление знаний, порождение знаний, логический синтез гипотез, абдуктивный вывод, автоматическое доказательство теорем.

Цитирование: Васильев С.Н. Интерактивное порождение новых знаний на основе автоматических средств логического вывода // Онтология проектирования. 2023. Т.13, №1(47). С.10-28. DOI: 10.18287/2223-9537-2023-13-1-10-28.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

Представление и обработка знаний обычно сопровождаются оценкой свойств сформированной с использованием некоторого языка совокупности знаний, как описания некоторой предметной области (ПрО). В числе анализируемых свойств могут быть взаимная независимость или, наоборот, следование одних знаний из других, непротиворечивость и другие

свойства. Обработка знаний, как элементов описания ПрО, включает компоновку состава с удалением одних и порождением новых. Данная работа посвящена методу порождения знаний с применением логических средств формализации и обоснования утверждений в языке, полном относительно выразительности языка исчисления предикатов (ИП) первого порядка.

Известны разные формальные аксиоматические теории [1] и средства логического вывода для автоматизации доказательства (АД) теорем, например, [2-8] и др. В данной работе для представления и обработки знаний используются логический язык L и исчисление J позитивно-образованных стандартизованных формул (пос-формул) [9, 10], хотя в метаязыке удобно использовать и традиционную терминологию ИП; далее ИП – первого порядка [1] (собственные аксиомы теории первого порядка, если они есть, предполагаются в составе описания ПрО).

В отличие от АД, задача синтеза текстов новых знаний трудно разрешима с точки зрения автоматизации. В данной работе рассматривается порождение утверждений, которые не являются просто логическими следствиями известного знания, т.е. допускающими своё получение механической дедукцией следствий. Причина трудностей автоматизации синтеза текста нового знания – в некорректности общей постановки этой задачи, и в первую очередь из-за трудности формализации интуитивных предпочтений специалиста в отношении ценности порождаемого знания. Не гарантируют этой ценности и такие, представимые синтаксически, критерии ценности теоремы, как высокая сложность (большая длина) доказательства или, наоборот, компактность текста теоремы (один из критериев элегантности), вхождение в тексты теоремы и её доказательства ключевых слов актуальной области исследований. С другой стороны, при использовании АД в интерактивном обучении человека уделяется больше внимания не только и не столько доказательству, как свидетельству корректности теоремы, сколько её понятности. Это может означать, например, что в задачах синтеза теорем для повышения ценности синтезируемого оно, возможно, должно быть структурировано с разбивкой на несколько утверждений (лемм и теорем) по критерию совокупной понятности. И это – при том, что автоматически синтезировать тексты теорем, формально верных, не трудно (тем более, что в число таких теорем входят и простые логические следствия), что было продемонстрировано, например, в [11].

Сказанное подводит к целесообразности использования мер обеспечения содержательности синтезируемого уже в самой постановке задачи. Так, в контексте анализа знаний и конкретно в ситуации выяснения наличия некоторого свойства F анализируемой совокупности знаний, как некоторого описания ПрО, при отсутствии этого свойства может возникнуть задача формирования некоторого непротиворечивого условия H , гарантирующего наличие F , т.е. задача синтеза теоремы $H \rightarrow F$ (здесь F – свойство, отличное от непротиворечивости, проверяемой просто средствами АД без обращения к синтезу H). Например, такая задача может возникнуть, когда выявлены одновременно непротиворечивость и недоказуемость F или когда выясняемое свойство F состоит в следовании некоторых знаний из других, а H – условие, логически достаточное для наличия F . При этом важно то, что в отличие от общей постановки задачи порождения новых знаний, трудности которой были отмечены выше, теперь текст утверждения $H \rightarrow F$, как нового знания, наследует содержательность исходного текста F как части нового текста.

В данной работе на основе интерактивного подхода к проблеме порождения новых знаний предлагается метод синтеза теорем $H \rightarrow F$. В этом методе человек выбирает версию условия H в ответ на предлагаемые ему опции со стороны средств АД. В настоящее время, несмотря на достижения в области АД, интерактивность в доказательстве теорем тоже широко используется; причины – не только в возможной недостаточности ресурсов (времени,

памяти и др.), но и в полурешимости ИП и, возможно, в широте собственной аксиоматики, вызывающей нерешимость теории, объемлющей ИП [1, 12-14].

В интерактивном режиме синтеза знаний важна наглядность текста в языке формализации. Используемый язык пос-формул, в силу их крупноблочности, как и крупноблочности средств обработки, облегчает восприятие текста человеком, по сравнению, например, с неестественной клаузуальной формой представления и обработки знаний резолюционными решателями [5, 15]. Важно также то, что эвристическая структура знания, выраженная в естественном языке, сохраняется при переводе на язык пос-формул. В языке пос-формул формулы построены из *типовых кванторов* (ТК) при том, что в естественном языке можно видеть аналоги не обычных кванторов логического языка, а аналоги именно ТК. Действительно, в естественном языке квантификация предметных переменных (т.е. использование выражений, подобных фразам «для всех», «существует») сопровождается указанием некоторого условия, ограничивающего множество допустимых значений предметной переменной. Это условие в типовых кванторах именуется *типовым условием*. Такое частичное подобие пос-формул естественному языку и крупноблочность формульных преобразований в исчислении пос-формул повышают наглядность текста для восприятия и семантической оценки человеком [10, 16-18].

Понятие «ТК» введено в [19]. Частными случаями ТК являются *ограниченные (специализированные)* кванторы в формулировках свойств многоосновных алгебраических систем и моделей [20, 21]. Типовыми условиями здесь являются условия принадлежности значений кванторных переменных тому или иному множеству из некоторого семейства *основных множеств*. В языках многосортных логик эти множества, как сорта переменных, используются для повышения эффективности АД, проверки проектных ограничений и др. [22, 23]. Прямым обобщением ограниченных кванторов являются кванторы с типовыми условиями в форме принадлежности значений кванторных переменных так называемым *ступеням* [19], построенным из основных множеств с помощью операций декартового произведения и образования множества всех подмножеств. Дальнейшее обобщение кванторов со ступенями использовано при решении логико-алгебраических уравнений [24, 25].

Типово-кванторный язык L пос-формул в так называемом *стандартизованном* виде (см. раздел 1) и соответствующее исчисление J были предложены автором совместно с А.К. Жерловым [9, 10]. Типово-кванторное представление знаний автор использовал в предложенном им языке позитивно-образованных формул, но *не стандартизованных*. Автор ввёл этот язык, именуемый здесь как язык *пон-формул* для развития алгоритмов метода сравнения В.М. Матросова и синтеза теорем о сохранении свойств математических моделей [26-29]. Этот язык выше первого порядка и этим обеспечивалась представимость в нём, например, свойства устойчивости по Ляпунову. Язык пон-формул может быть и частично-формализованным – до некоторой глубины структуры своих формул, а именно без формализации типовых условий.

Развиваемое здесь в языке L пос-формул исчисление J_+ для синтеза теорем $H \rightarrow F$ выходит за рамки чисто дедуктивного подхода, реализованного в исчислении J [9, 10]. В отличие от двухместного правила резолюции [5], правило дедукции в J является одноместным, что снижает комбинаторность поиска выводов. Что касается систем АД не резолюционного типа, то типичными их недостатками являются многоместность и неединственность их правил вывода (см., например, [8]).

С логической точки зрения исчисление J_+ является дедуктивно-абдуктивным, и вывод в нём с логической точки зрения достоверен, хотя синтезируемые гипотезы при обычной ограниченности описания ПрО могут быть лишь правдоподобными, почему и нужна их оценка человеком. Известны системы логического вывода, похожим образом ориентированные на

задачи синтеза гипотез, не противоречащих теории и «объясняющих наблюдаемое», например, в стиле задач диагностики «причина-эффект» в первопорядковой логике [30]. В них в роли свойства F выступает логическое следование наблюдаемого факта из базовой совокупности знаний (теории), а H – условие этого следования как объяснение наблюдаемого. При этом в разработках по абдуктивному выводу, помимо или вместо привлечения эксперта, часто используется некоторое множество потенциально возможных вариантов эмпирических гипотез, предполагаемое априори заданным на основе имеющегося прошлого опыта. Из альтернативно синтезированных тем или иным способом гипотез выбирается гипотеза, принадлежащая этому множеству.

В известных работах по выводу с абдукцией формульное представление знаний часто бывает ограничено требованием хорновости [31], типичным для логического программирования [32, 33]. Ограничением другого типа часто выступает требование, чтобы знания теории представлялись эквивалентностями в словаре переменных-причин и переменных-эффектов (с возможной частичной упорядоченностью эквивалентностей между собой импликациями). Таковыми в пропозициональном случае являются эквивалентности переменных-причин дизъюнкциям элементарных конъюнкций, составленных из переменных-эффектов и их отрицаний [30, 34].

Программные разработки на базе предложенного в работе метода интерактивного синтеза новых знаний с участием эксперта могут нуждаться в совершенствовании с учётом опыта использования, например, при реализации метода в составе некоторой экспертной или оперативно-советующей системы в медицине, авиации, учебном процессе и других приложениях [35-37]. Это, например, может быть связано с возможными ситуациями сомнения или несогласия конечного пользователя (оператора, врача, пилота, учителя и т.п.) с синтезированным объяснением или рекомендацией системы даже при её способности объяснять и наблюдаемое, и свои решения. Аккумуляирование таких случаев и объяснений пользователя, обладающего системным, холистическим взглядом на ПрО, с верификацией оснований для сомнений пользователя может быть полезным для совершенствования системы.

В разделе 1 данной работы вводятся основные понятия и формулируется постановка рассматриваемой задачи. В разделе 2 вводится исчисление дедуктивно-абдуктивного типа J_+ . Излагаются и обосновываются свойства его правил синтеза с точки зрения необходимости и достаточности синтезируемого. В разделе 3 приводятся примеры анализа свойств F описания ПрО и синтеза гипотез H как условий доказуемости F . В заключении формулируются итоги работы и задачи на будущее.

1 Основные понятия и постановка задачи

Язык пос-формул [10] построен на базе *позитивных* ТК, а именно: ТК всеобщности (ТКВ) $\forall X : W$ и ТК существования (ТКС) $\exists X : W$. Здесь $X = \{x_1, \dots, x_k\}$ – множество предметных переменных ($k \geq 0$), а W – типовое условие, накладываемое как на эти переменные, так и, может быть, на переменные, связанные другими ТК, в область действия которых в пос-формуле входит рассматриваемый ТК. ТК именуется позитивным, если типовое условие W – конъюнкт, т.е. конъюнкция атомов или пропозициональная константа t либо f (*true* и *false*, соответственно). Далее рассматриваемые ТК считаются позитивными.

Определение 1 (пос-формулы). Пусть $Q \in \{\forall, \exists\}$, а через $\bar{Q}$ обозначен символ $\exists$, если $Q = \forall$, и символ $\forall$, если $Q = \exists$. Тогда:

1) $QX : W$ – Q -формула;

2) если $G_1, \dots, G_n$ – Q -формулы, то $\bar{Q}X : W \{G_1, \dots, G_n\}$ – $\bar{Q}$ -формула;

3) все Q -формулы являются пос-формулами; других пос-формул нет.

Отсутствие в синтаксисе пос-формул символа отрицания объясняет название формул. Язык L состоит из пос-формул. Подмножество языка L , состоящее из $\exists$ -формул, обозначается $L^\exists$. Существенное отличие языка L от других языков представления знаний состоит в сборке пос-формул только из ТК.

Семантика пос-формул F , $F \in L$, определяется [10] через семантику соответствующих им формул языка ИП. Представление пос-формулы F в языке ИП обозначается F' . Соответствие задаётся следующим отображением:

1) $(\forall X : W \{G_1, \dots, G_n\})' = \forall x_1 \dots \forall x_k (W \rightarrow G'_1 \vee \dots \vee G'_n)$ (при этом $G_1, \dots, G_n$ – $\exists$ -формулы);

2) $(\exists X : W \{G_1, \dots, G_n\})' = \exists x_1 \dots \exists x_k (W \wedge G'_1 \wedge \dots \wedge G'_n)$ (здесь $G_1, \dots, G_n$ – $\forall$ -формулы);

3) $(\forall X : W)' = (\forall X : W \Phi)'$, где Φ – пусто, поэтому $(\forall X : W)' = (\forall X : W \Phi)' \leftrightarrow$

$\leftrightarrow \forall x_1 \dots \forall x_k (W \rightarrow f) \leftrightarrow \forall x_1 \dots \forall x_k \neg W$ (дизъюнкция набора Φ формул $G_1, \dots, G_n$, пустого при $n = 0$, ложна);

4) $(\exists X : W)' = (\exists X : W \Phi)' \leftrightarrow \exists x_1 \dots \exists x_k (W \wedge t) \leftrightarrow \exists x_1 \dots \exists x_k W$ (конъюнкция пустого набора Φ пос-формул истинна).

Таким образом, структура пос-формулы может ветвиться дизъюнктивно после ТКВ и конъюнктивно после ТКС, а ТКВ и ТКС могут входить в ветвь пос-формулы только поочередно. Далее считается, что обозначения кванторных переменных двух ТК, один из которых входит в область действия другого, не пересекаются. Метод трансляции формул языка ИП в формулы языка L описан, например, в [18]. Язык L полон [10] относительно выразительной силы языка ИП.

Обычные понятия логической эквивалентности, следования и другие будут пониматься в языке L как в языке ИП, так как L можно считать собственным подмножеством множества формул языка ИП. Представление формулы F языка ИП в L обозначается $(F)^L$. Если в метаязыке выражение построено из пос-формул как подформулы с помощью некоторых из логических символов $\rightarrow, \leftrightarrow, \neg, \wedge, \vee$, то вместо пос-формул следует понимать их образы в языке ИП. Так, если из контекста ясно, что $F_1, F_2 \in L$, то запись $F_1 \rightarrow F_2$ понимается как $(F_1)' \rightarrow (F_2)'$, т.е. как общезначимая, или истинная в любой интерпретации, формула, а в силу полноты ИП относительно выводимости это означает и выводимость в ИП.

Замечание 1. Язык *позитивных* формул [21, 38], как собственное подмножество языка ИП, не использует ТК и существенно уже языка пос-формул и, тем более, пон-формул. Позитивные формулы в смысле [21, 38] используют только логические связки $\wedge, \vee$ и обычные кванторы, не используя отрицания и импликации. Позитивность пос-формул проявляется в двух смыслах: 1) структура допускает только $\wedge$ - и $\vee$ -ветвления; 2) структура собрана из ТК, типовые условия которых суть конъюнкты без отрицаний атомов [9]. При переводе пос- или пон-формул в язык ИП проявляются импликации ТКВ. Это обеспечивает полноту языка пос-формул относительно выразительности языка ИП и, в том числе, подмножества его формул с отрицаниями. Язык пон-формул: а) позитивен в первом смысле, б) может быть более высокого порядка, чем первый порядок, в) может быть частично формализованным и г) предложен для решения логических уравнений, имеющих структуру теорем о сохранении свойств связанных математических моделей.

Если при символе Q пос-формулы кванторные переменные отсутствуют ($k=0$), то символ Q не удаляется для конкретизации смысла связки $\rightarrow$ или $\wedge$, ассоциируемой с посылочным

или утвердительным смыслом типового условия W в ТКВ и ТКС соответственно. Если в пос-формуле кванторные переменные отсутствуют (при всех символах Q и во всех типовых условиях), то множество таких пос-формул может рассматриваться как подмножество пропозиционального языка, полное относительно выразительности пропозиционального языка.

Определение 2. Каноническим видом пос-формулы V , $V \in L$, считается её представление с корневым ТКВ $\forall: \mathbf{t}$, т.е. в виде $V = \forall: \mathbf{t} \Omega$, где $\Omega \subset L^\exists$ и Ω – непустое множество $\exists$ -формул, именуемых *базовыми подформулами* (БП).

В множестве Ω канонического представления формулы V выделена одна подформула G как образец преобразований всех $\exists$ -формул из $\Omega = \{G, \Sigma\}$ по описываемому ниже правилу вывода ω . Здесь Σ – подмножество прочих БП (возможно, пустое), строение которых аналогично следующему строению формулы G :

$$G = \exists X: A \Phi, \quad \Phi = \{\forall Y_1: B_1 \Psi_1, \dots, \forall Y_n: B_n \Psi_n\}, \quad n \geq 0. \quad (1)$$

Каждый ТКС, как узел структуры пос-формулы, смежный корню $\forall: \mathbf{t}$, является корневым для БП так же, как ТКС $\exists X: A$ для Φ . Он именуется *базой (фактов)*, а его конъюнкт (в ТКС $\exists X: A$ это конъюнкт A) – *базовым конъюнктом*. Узлы, дочерние базе, т.е. ТКВ (для ТКС $\exists X: A$ это ТКВ $\forall Y_i: B_i$), называются *вопросами к базе*, а подформулы, находящиеся в области действия базы (для ТКС $\exists X: A$ это подформулы $\forall Y_i: B_i \Psi_i$), – *закономерностями базы* (они у некоторых баз могут отсутствовать, когда $n = 0$). Пос-формулы, все БП которых не содержат ТКВ с ветвлением (т.е. ТКВ с дизъюнктивными ветвлениями), называются *хорновскими*, по аналогии с [31].

Отображение $\theta: (Y_i) \rightarrow X$ (частичное из множества Y_i в множество X) именуется *подстановкой* (контексты употребления символа $\rightarrow$ одновременно в теоретико-множественном и логическом смыслах исключают коллизии). Через $B_i\theta$ обозначается результат замены в B_i переменных из множества Y_i на некоторые переменные из X . Если $B_i\theta \subseteq A$, то отображение θ называется *ответной подстановкой*, а вопрос $\forall Y_i: B_i$ – *уместным вопросом*.

Определение 3 (правило вывода $\omega: L \rightarrow L$). Пусть пос-формула V отлична от $\perp$ и представлена в каноническом виде $V = \forall: \mathbf{t} \Omega$, где $\Omega = \{G, \Sigma\}$, G имеет вид (1), но $n \neq 0$, т.е. в G множество закономерностей Φ базы не пусто, и для некоторого $i \in \overline{1, n}$ $B_i\theta \subseteq A$, а

$\Psi_i = \{\exists Z_{i1}: C_{i1} \Psi_{i1}, \dots, \exists Z_{im}: C_{im} \Psi_{im}\}$. Тогда

$$\omega V = \forall: \mathbf{t} \{\Xi_1, \dots, \Xi_m, \Sigma\}, \quad \forall s \in \overline{1, m} \quad \Xi_s = \exists X \cup Z_{is}^*: A \cup C_{is}\theta(Z_{is}/Z_{is}^*) \{\Psi_{is}\theta(Z_{is}/Z_{is}^*), \Phi\},$$

где Z_{is}/Z_{is}^* – разыменованье кванторных переменных, т.е. замена элементов множества Z_{is} на новые переменные из Z_{is}^* .

Правило $\omega: L \rightarrow L$ является логически эквивалентным преобразованием [10]. В языке L пос-формул с применением этого правила введено исчисление $J = (\perp, \omega)$ [9, 10], где формула $\perp = \forall: \mathbf{t} \exists: \mathbf{f}$ – признак завершения вывода. Она является противоречием, поскольку в исчислении J доказательство всякой теоремы F ведётся от противного, как, например, и в [5]: опровергается отрицание $V = (\neg F)^L$ доказываемого. Формула V из L называется *выводимой в J* (это записывается $\vdash_J V$ или $V \vdash_J \perp$), если $\exists s: s \geq 1 \quad \omega^s V = \perp$, т.е. на некотором шаге

применения правила ω получается $\perp$. Исчисление J полно относительно выводимости [10]: отрицание всякой теоремы в ИП, представленное пос-формулой в языке L , выводимо в ис-

числении J . Наоборот, если пос-формула выводима в исчислении J , то отрицание её образа в языке ИП – теорема ИП, т.е. общезначимо. Простейший пример невыводимой формулы возникает в случае, когда какая-то БП не содержит закономерностей ($n = 0$). Необходимым условием опровержимости пос-формулы является наличие листового узла $\exists: f$ как потенциального «источника» проникновения в базу фактов атома f .

В случае наличия прикорневого ветвления структуры пос-формулы V , задача опровержения V ввиду дизъюнктивности ветвления сводится к независимым и аналогичным друг другу подзадачам опровержения своей БП. Проникновение в базовый конъюнкт константы f означает логическую эквивалентность базового конъюнкта и всей БП константе f ; поэтому в случае единственности этой БП процесс опровержения завершён, а при неединственности она удаляется (это является логически эквивалентным преобразованием) с переходом к уместным вопросам других баз, и вывод по замыслу должен продолжаться до получения аналогичного эффекта во всех базах, если F была противоречивой.

Пусть F – формула в языке ИП. Её дедуктивная проверка на доказуемость в исчислении $J = (\perp, \omega)$ состоит в опровержении пос-формулы $(\neg F)^L$, т.е. в выводе $(\neg F)^L \vdash_J \perp$ или $\vdash F$ в ИП.

2 Синтез гипотез

Пусть $V \in L$ и в канонической форме $V = \forall: t \{G, \Sigma\}$, где G – выделенная БП вида (1), а Σ – множество прочих БП.

В исчислении $J_+ = (\perp, \omega, \delta)$ вводится первый тип правил δ . Это определяемые ниже правила α *первичного преобразования* пос-формул. Им будут соответствовать *гипотезы H первичного типа*. Правила α расширяют множество закономерностей некоторой базы обрабатываемой пос-формулы гипотезой H , корневой узел которой является заведомо уместным вопросом. Поэтому в ситуации непродолжимости вывода в исчислении $J = (\perp, \omega)$, т.е. неприменимости правила ω , применение правила α заведомо продолжает вывод.

Определение 4 (правила α первичного преобразования пос-формул). Правила α – это преобразования $\alpha: L \rightarrow L$, $\alpha(V) = V(\Phi / \{\Phi, H\})$, где H – пос-формула, предполагаемая замкнутой (не содержащей свободных вхождений кванторных переменных), а преобразование α в зависимости от H имеет вид:

$$\text{а) } \alpha = \alpha_0, \text{ когда } H = H_0 = \forall X^*: A \quad \exists: f \text{ (при } n = 0 \text{);} \quad (2)$$

$$\text{б) } \alpha = \alpha_{\vee}, \text{ когда } H = H_{\vee} = \forall X^*: A \quad \{ \exists Y_1^*: B_1, \dots, \exists Y_n^*: B_n \} \text{ (при } n > 1 \text{);} \quad (3)$$

$$\text{в) } \forall i \in \overline{1, n} \quad \alpha = \alpha_i, \text{ когда } H = H_i = \forall X^*: A \quad \exists Y_i^*: B_i \text{ (при } n \geq 1 \text{).} \quad (4)$$

Здесь верхний индекс $*$ означает разыменовывание в H кванторных переменных (для исключения их совпадения с переменными, входящими в корень $\exists X: A$ и его дочерние узлы $\forall Y_i: B_i$ ($i \in \overline{1, n}$) из G). При позитивности конъюнктов A, A_i, B_i , вытекающей из определения пос-формул, гипотезы H непротиворечивы. Логические связи свойства F и гипотез H представлены в следующей теореме.

Теорема 1 (о связи свойства F и гипотез H первичного типа). Пусть F – изучаемое свойство описания ПрО, а H – гипотеза первичного типа, $F = (\neg F)^L$, $S = \omega^{s_2} \alpha \omega^{s_1} F$, где $s_i \geq 0$ ($i \in \overline{1, 2}$). Тогда $\vdash_J T$, где $T = \neg(\neg S \wedge H \rightarrow F)$, т.е. $\neg S \wedge H \rightarrow F$. В частности:

1) если $S = \perp$, то $H \rightarrow F$ (условие H достаточно для F);

2) если $\alpha \in \{\alpha_0, \alpha_\vee\}$, то $F \rightarrow H$ (условия H_0 и $H_\vee$ необходимы для F).

Доказательство. Необходимо показать достаточность условия $\neg S \wedge H$ для F при любом $\alpha \in \{\alpha_0, \alpha_\vee, \alpha_i\}$. Пусть синтез гипотезы H выполнен применением правила α к пос-формуле V , полученной из F s_1 -кратным применением правила ω , т.е. к формуле $V = \omega^{s_1} F = \forall: \mathbf{t} \{G, \Sigma\}$, где $G = \exists X: A \ \Phi$ и $\Phi = \{\forall Y_1: B_1 \ \Xi_1, \dots, \forall Y_n: B_n \ \Xi_n\}$. Тогда

$$S = \omega^{s_2} \alpha(V) \leftrightarrow \alpha(V) = V(G / (G(\Phi / \{\Phi, H\}))) = \forall: \mathbf{t} \{G(\Phi / \{\Phi, H\}), \Sigma\},$$

а так как $H = H(X^*)$ и не содержит вхождений переменной X , то

$$G(\Phi / \{\Phi, H\}) \leftrightarrow \exists X(A' \wedge \Phi' \wedge H') \leftrightarrow G' \wedge H'$$

и поэтому $S \leftrightarrow (G \wedge H) \vee \Sigma$. Это значит, что $\neg S \rightarrow \neg((G \wedge H) \vee \Sigma) \leftrightarrow (H \rightarrow \neg G) \wedge \neg \Sigma$. Отсюда следует $\neg S \wedge H \rightarrow \neg G \wedge \neg \Sigma \leftrightarrow \neg V \leftrightarrow \neg F = \mathbf{F}$. С учётом этого и того, что при $S = \perp$ справедливо $\neg S \wedge H \leftrightarrow H$, получается утверждение 1) доказываемой теоремы.

Пусть теперь $\alpha \in \{\alpha_0, \alpha_\vee\}$. Требуется показать необходимость соответствующих гипотез первичного типа H_0 и $H_\vee$ (см. (2) и (3)). Если $\alpha = \alpha_0$, то $\neg F \leftrightarrow \neg V \leftrightarrow \neg G \wedge \neg \Sigma = \forall X: A \ \exists: \mathbf{f} \wedge \neg \Sigma$, откуда очевидным образом следует H_0 .

Если же $\alpha = \alpha_\vee$, то $\neg F \leftrightarrow \neg V \leftrightarrow \neg G \wedge \neg \Sigma = \forall X: A \ \{\exists Y_1: B_1 \neg \Xi_1, \dots, \exists Y_n: B_n \neg \Xi_n\} \wedge \neg \Sigma$. Из сравнения этого с условием $H_\vee$ видно, что $\neg F \rightarrow H_\vee$. Теорема доказана.

Замечание 2. По определению (2) правила α_0 при $n = 0$, т.е. при пустом множестве Φ , формировалась гипотеза $H = H_0 = \forall X^*: A \ \exists: \mathbf{f}$ как условие противоречивости базового конъюнкта A . Подразумевается, что эта гипотеза без дальнейших преобразований включается в конъюнктивный набор гипотез, синтезированных по всем БП из $\Omega = \{G, \Sigma\}$, с отслеживанием неоявления гипотезы, обуславливающей $\exists X: A$ и потому противоречащей H_0 . Далее в БП $n \geq 1$.

Замечание 3. При $n > 1$ согласно (3) структура гипотезы $H_\vee$ нехорновская, т.е. содержит $\vee$ -ветвление: $H_\vee = \forall X^*: A \ \{\exists Y_1^*: B_1, \dots, \exists Y_n^*: B_n\}$. Эксперту предлагается возможность корректировки и выбора $n+1$ вариантов гипотезы из множества $\{H_\vee\} \cup \{H_i\}_{i \in \overline{1, n}}$, где $\forall i \in \overline{1, n} \ H_i = \forall X^*: A \ \exists Y_i^*: B_i$. В общем случае эксперт может рассматривать и варианты, получаемые из $H_\vee$ удалением меньшего количества дизъюнктивно связанных ветвей, чем $n-1$. Ниже рассматриваются гипотезы H_i вида (4), как более «естественные» для теорем $H \rightarrow F$, чем нехорновская гипотеза $H_\vee$. В частности, при $Y_i^* = \emptyset$ гипотезы H_i имеют вид $(H_i)' = \forall X(A(X) \rightarrow B_i(X))$, т.е. простых условий типа «если ..., то ...», не только без $\vee$ -ветвления, но и без квантора существования. Использование гипотез (4) предпочтительнее, чем вида (3), тем более при использовании правил в задачах с конструктивной семантикой [10, 36]. Финальный вид гипотез H_i достигается применением в исчислении $J_+ = (\perp, \omega, \delta)$ правил δ второго типа. Это правила β преобразования $\forall$ -формулы H_i , именуемые правилами *вторичного преобразования* пос-формулы, как частичные отображения $\beta: (L^\forall) \rightarrow L$, $\beta(H_i) = H_i^*$. Пос-формулы H_i^* называются *гипотезами вторичного типа*.

Определение 5 (операции вторичного преобразования β пос-формул). Для всякого H_i вида (4) преобразование $\beta(H_i)$ состоит в некоторой суперпозиции определяемых ниже операций $\{O_i\}_{i \in \overline{1,5}}$, $O_i : (L^\forall) \rightarrow L$. Если $V = \forall Z^1 : C \exists Z^2 : D$ – формула H_i или общее обозначение результата применения предыдущих операций внутри преобразования β , то операциями преобразования β являются:

- 1) операция O_1 удаления некоторых атомов корневого конъюнкта C ;
- 2) операция O_2 удаления из листового конъюнкта D повторов, т.е. атомов, возможно, вошедших в корневой конъюнкт C на предыдущих операциях;
- 3) операция O_3 подстановки $\theta : (Z^2) \rightarrow Z^1$;
- 4) операция O_4 удаления из множества Z^1 кванторных переменных корневого узла $\forall Z^1 : C$ всех *независимых* переменных, т.е. не входящих $\forall : C \exists Z^2 : D$;
- 5) если конъюнкты C, D формулы $V = \forall Z^1 : C \exists Z^2 : D$ имеют вид $C = C_1(Z_1^1) \cup C_2(Z_2^1)$ и $D(Z_1^1)$, где $Z^1 = Z_1^1 \cup Z_2^1$, $Z_1^1 \cap Z_2^1 = \emptyset$, т.е. атомы C_1 и D не содержат переменных из Z_2^1 , а C_2 не содержит переменных из Z_1^1 , то $O_5(V) = V_1 \wedge V_2$, $V_1 = \forall Z_1^1 : C_1 \exists Z^2 : D$, $V_2 = \exists Z_2^1 : C_2$.

Используемый далее для краткости термин «*преобразование ПДУ*» понимается как преобразование Перехода к Достаточному Условию. Т.е, если формула U_1 получена из U_2 преобразованием, обеспечивающим логическую импликацию $U_1 \rightarrow U_2$, то это преобразование именуется как преобразование «перехода от U_2 к достаточному условию U_1 ». При этом U_1 не слабее, чем U_2 , так как всякая логическая импликация включает и случай эквивалентности. В частности, для всякой формулы V , $V \in L$, $\alpha_i(V)$ логически не слабее, чем $\alpha_\forall(V)$.

Теорема 2 (о связи свойства F и гипотез H_i^* вторичного типа). Если к гипотезе H_i вида (4) применено преобразование $\beta : (L^\forall) \rightarrow L$, $\beta(H_i) = H_i^*$, в виде той или иной суперпозиции некоторых операций O_i из множества $\{O_i\}_{i \in \overline{1,5}}$, $O_i : (L^\forall) \rightarrow L$, то $(\neg(\beta(H_i) \rightarrow F))^L \vdash_J \perp$, т.е. синтезированный текст $H_i^* \rightarrow F$ – теорема.

Доказательство. Требуется проверить, что каждая из операций $O_i \in \{O_i\}_{i \in \overline{1,5}}$ преобразования β сопровождается ПДУ. Пусть $V = \forall Z^1 : C^1 \exists Z^2 : D$ – гипотеза H_i или результат предыдущих операций внутри преобразования β .

Применение операции O_1 удаления из C^1 некоторых атомов не сужает заданной типовым условием C^1 области значений кванторных переменных множества Z^1 , и $O_1(V)$ не слабее, чем V , что означает ПДУ. Операция O_2 удаления повторов из конъюнкта D атомов, входящих в посылку C является логически эквивалентным преобразованием. Операция O_3 подстановки в D вместо вхождений переменных множества Z^2 соответствующих переменных из Z^1 – это снова ПДУ для V .

В связи с операциями O_4 и O_5 преобразования β следует отметить, что удаления независимых кванторов – типовых и обычных – различаются по результату. Например, пусть имеются две формулы $U_1 = \forall x A[x]$, $U_2 = \forall x (W(x) \rightarrow A[x])$, где $A[x]$ обозначает отсут-

ствие в формуле A свободных вхождений переменной x , отчего обычный квантор $\forall x$ и ТКВ $\forall x:W(x)$ именуются *независимыми*. После их удаления из U_1 и U_2 результат A соотносится с формулами U_1 и U_2 по разному: $A \rightarrow U_1 \wedge U_2$, $U_1 \rightarrow A$, но U_2 влечёт A лишь при дополнительном условии $\exists xW(x)$, т.е. $\exists xW(x) \rightarrow (U_2 \rightarrow A)$.

В силу сказанного, операция O_5 , в отличие от операции O_4 , не является логически эквивалентным преобразованием. Так как $O_5(V) = \exists Z_2^1 C_2 \wedge \forall Z_1^1 (C_1 \rightarrow \exists Z^2 D)$ и $V = \forall Z^1 : C \exists Z^2 : D = \forall Z_1^1 \cup Z_2^1 : C_1(Z_1^1) \cup C_2(Z_2^1) \exists Z^2 : D \leftrightarrow \forall Z_2^1 (C_2 \rightarrow \forall Z_1^1 (C_1 \rightarrow \exists Z^2 D))$, то $O_5(V) \rightarrow V$ и поэтому обеспечивается ПДУ. При этом удаляется независимый ТКВ $\forall Z_2^1 : C_2$, что и потребовало конъюнктивного включения в гипотезу $O_5(V)$ условия $\exists Z_2^1 C_2$.

Таким образом, все операции $\{O_i\}_{i \in \overline{1,5}}$ преобразования β сопровождаются переходами к достаточным условиям и поэтому $(H_i^*)' \rightarrow F$, где $H_i^* = \beta(H_i)$, т.е. $(\neg((H_i^*)' \rightarrow F))^L \vdash_J \perp$, что и требовалось доказать.

На простых иллюстративных примерах можно рассмотреть представление знаний в языке L и их обработку с интеграцией ролевых функций правил ω, δ исчисления $J_+ = (\perp, \omega, \delta)$ и эксперта в интерактивном анализе и синтезе знаний.

3 Примеры

Пусть предметом интереса человека является доступность самолётам высот тропосферы, как нижнего слоя атмосферы Земли, и стратосферы, как соседнего, более удалённого от планеты слоя. Предположим, что у человека-эксперта имеется предварительное описание предмета в форме некоторой совокупности V утверждений-знаний, возможно из разных источников этой ПрО. Описание является заготовкой к началу работы эксперта по интерактивному анализу этого описания с точки зрения свойств непротиворечивости, избыточности и других. Предполагается, что работа эксперта ведётся с применением исчисления $J_+ = (\perp, \omega, \delta)$ для АД свойств и порождения новых знаний.

Пример 1 (проверка V на непротиворечивость). Пусть описание предмета в естественном языке представлено совокупностью следующих утверждений:

«Отдельным самолётам доступны некоторые высоты стратосферы. Для всякой пары (самолёт, высота), связанной отношением доступности, следует, что и все более низкие высоты доступны этому самолёту. Существуют самолёты, которым доступны все тропосферные высоты, а также некоторые высоты стратосферы».

Для логической формализации этого описания задействуются следующие предикаты:

$$\begin{aligned} P(x) &\Leftrightarrow \text{«} x \text{ — самолет »}, \\ T(y) &\Leftrightarrow \text{«} y \text{ — высота в тропосфере »}, \\ S(y) &\Leftrightarrow \text{«} y \text{ — высота в стратосфере »}, \\ R(x, y) &\Leftrightarrow \text{«} \text{для } x \text{ доступно } y \text{ »}, \\ L(y, z) &\Leftrightarrow \text{«} z \text{ — ниже, чем } y \text{ »} \end{aligned}$$

(здесь $\Leftrightarrow$ означает «...тогда и только тогда, когда...», а обозначения предикатов — от английских слов «Plane», «Troposphere», «Stratosphere», «Reachable», «Lower than»). В языке ИП (первого порядка) данное описание предмета рассмотрения примет вид

$$V = \exists x \exists y (P(x) \wedge S(y) \wedge R(x, y)) \wedge \\ \wedge \forall x \forall y (P(x) \wedge R(x, y) \rightarrow (\forall z (L(y, z) \rightarrow R(x, z))) \wedge \\ \wedge \exists x (P(x) ((\forall y (T(y) \rightarrow R(x, y)) \wedge \exists y (S(y) \wedge R(x, y)))),$$

что в переводе на язык L пос-формул может быть записано в следующей канонической форме (с двумя $\wedge$ -ветвлениями и тремя БП как базовыми закономерностями):

$$V^L = \forall: \mathbf{t} \quad \exists: \mathbf{t} \quad \left\{ \begin{array}{l} \forall: \mathbf{t} \quad \exists x, y: P(x), S(y), R(x, y) \\ \forall x, y: P(x), R(x, y) \quad \exists: \mathbf{t} \quad \forall z: L(y, z) \quad \exists: R(x, z) \\ \forall: \mathbf{t} \quad \exists x: P(x) \quad \left\{ \begin{array}{l} \forall y: T(y) \quad \exists: R(x, y) \\ \forall: \mathbf{t} \quad \exists y: S(y), R(x, y). \end{array} \right. \end{array} \right.$$

Для уменьшения потребности в разыменованиях кванторных переменных при применении правила ω (см. в разделе 1) необходимо исключать совпадение кванторных переменных не только в тех случаях, когда один ТК находится в области действия другого, но и в случаях, когда кванторные переменные входят в разные ветви. Поэтому далее используется представление с разыменованными кванторными переменными:

$$V^L = \forall: \mathbf{t} \quad \exists: \mathbf{t} \quad \left\{ \begin{array}{l} \forall: \mathbf{t} \quad \exists x, y: P(x), S(y), R(x, y) \\ \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ \forall: \mathbf{t} \quad \exists x_2: P(x_2) \quad \left\{ \begin{array}{l} \forall y_3: T(y_3) \quad \exists: R(x_2, y_3) \\ \forall: \mathbf{t} \quad \exists y_4: S(y_4), R(x_2, y_4). \end{array} \right. \end{array} \right. \quad (5)$$

Пусть $V^L = \forall: \mathbf{t} \quad \exists: \mathbf{t} \quad \{V_1, V_2, V_3\}$ при следующих обозначениях для закономерностей:

$$\begin{aligned} V_1 &= \forall: \mathbf{t} \quad \exists x, y: P(x), S(y), R(x, y), \\ V_2 &= \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2), \\ V_3 &= \forall: \mathbf{t} \quad \exists x_2: P(x_2) \quad \left\{ \begin{array}{l} \forall y_3: T(y_3) \quad \exists: R(x_2, y_3) \\ \forall: \mathbf{t} \quad \exists y_4: S(y_4), R(x_2, y_4). \end{array} \right. \end{aligned} \quad (6)$$

Здесь и далее БП нумеруются сверху вниз (как далее и при нумерации корневых узлов из БП, т.е. вопросов).

Утверждение (5) непротиворечиво, и это свойство в общем случае устанавливается автоматически в пос-исчислении J путём проверки того, что вывод $(V)^L \vdash_J \perp$ не осуществим, т.е.

формула $(V)^L$ не опровержима. Действительно, через три шага применения правила ω с пустыми ответными подстановками получается следующее:

$$\omega^3(V^L) = \forall: \mathbf{t} \quad \exists x, y, x_2, y_4: P(x), \quad \left\{ \begin{array}{l} \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ S(y), R(x, y), \\ P(x_2), S(y_4), R(x_2, y_4) \quad \left\{ \begin{array}{l} \forall y_3: T(y_3) \quad \exists: R(x_2, y_3). \end{array} \right. \end{array} \right.$$

Теперь только первый вопрос $\forall x_1, y_1: P(x_1), R(x_1, y_1)$ - уместный. Ответная подстановка $\theta_1: x_1 \rightarrow x, y_1 \rightarrow y$ приводит к формуле

$$\omega^4(V^L) = \forall: \mathbf{t} \quad \exists x, y, x_2, y_4: P(x), \quad \left\{ \begin{array}{l} \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ S(y), R(x, y), \\ P(x_2), S(y_4), R(x_2, y_4) \quad \left\{ \begin{array}{l} \forall y_3: T(y_3) \quad \exists: R(x_2, y_3) \\ \forall y_5: L(y, y_5) \quad \exists: R(x, y_5). \end{array} \right. \end{array} \right.$$

После второго ответа на тот же вопрос с подстановкой $\theta_2: x_1 \rightarrow x_2, y_1 \rightarrow y_4$ получается

$$\omega^5(V^L) = \forall: \mathbf{t} \quad \begin{matrix} \exists x, y, x_2, y_4: P(x), \\ S(y), R(x, y), \\ P(x_2), S(y_4), R(x_2, y_4) \end{matrix} \quad \begin{cases} \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ \forall y_3: T(y_3) \quad \exists: R(x_2, y_3) \\ \forall y_5: L(y, y_5) \quad \exists: R(x, y_5) \\ \forall y_6: L(y_4, y_6) \quad \exists: R(x_2, y_6). \end{cases}$$

Теперь первый вопрос больше неуместен, так как новых ответных подстановок в его конъюнкт нет. Остальные три вопроса тоже неуместны, поэтому вывод оказался в тупике непродолжимости, означающей, что в исчислении J не выводимо $\perp$, т.е. формула V непротиворечива.

Следует заметить, что приведённое рассуждение – лишь иллюстрация применения правила ω в исчислении J ; в действительности в этом частном случае исходной формулы V^L вида (5) непротиворечивость усматривается сразу, а именно по факту отсутствия в V^L ни одного листового узла в виде ТКС $\exists: \mathbf{f}$. Стоит заметить, что этот признак не является достаточным: пос-формула может содержать ТКС $\exists: \mathbf{f}$, но не быть противоречивой и поэтому в таких случаях проверка невыводимости $\perp$ является обязательной как проверка попадания в ситуацию непродолжимости вывода, что и сделано выше.

Пример 2 (проверка следования одних утверждений из других, синтез условия следования). Например, для каждого утверждения V_i из $(V)^L$ (см. (6)) может делаться проверка его логического следования из остальных. Если такое следование имеет место, то может выявляться минимальный набор утверждений, из которых следует рассматриваемое утверждение. Таких наборов может быть несколько. Результаты такой проверки на избыточность сообщаются эксперту.

Описание $(V)^L$ предмета рассмотрения избыточно, так как V_1 следует из минимального одноэлементного набора $\{V_3\}$. Эксперт может удалить V_1 или исключить из V_3 подформулу $\forall: \mathbf{t} \quad \exists y_4: S(y_4), R(x_2, y_4)$. Эксперт, видя, что V_3 не следует из $V_1 \wedge V_2$, может заинтересоваться, например, синтезом условия H , достаточного для $V_1 \wedge V_2 \rightarrow V_3$, с тем, чтобы, если условие H будет лаконичнее, чем V_3 , то заменить V_3 на H .

Проверка свойства $F = V_1 \wedge V_2 \rightarrow V_3$ осуществляется автоматически как проверка выводимости $\perp$ из $F = \neg F = (V_1 \wedge V_2 \wedge \neg V_3)^L$ в исчислении $J = (\omega, \perp)$. Формула F в языке L представима с одной БП Ω :

$$F = \forall: \mathbf{t} \quad \exists: \mathbf{t} \quad \Omega = \forall: \mathbf{t} \quad \exists: \mathbf{t} \quad \begin{cases} \forall x, y: P(x), S(y), R(x, y) \\ \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ \forall x_2: P(x_2) \quad \begin{cases} \exists y_3: T(y_3) \quad \forall: R(x_2, y_3) \quad \exists: \mathbf{f} \\ \exists: \mathbf{t} \quad \forall y_4: S(y_4), R(x_2, y_4) \quad \exists: \mathbf{f}. \end{cases} \end{cases} \quad (7)$$

Так как третья закономерность в (7) имеет $\vee$ -ветвление, то она нехорновского типа, и поэтому в дальнейшем БП расщепится на две БП. Пока же после ответа на тривиальный вопрос из V_1 (с пустой подстановкой θ) получится

$$\omega(F) = \forall: \mathbf{t} \quad \begin{matrix} \exists x, y: P(x), S(y), \\ R(x, y) \end{matrix} \quad \begin{cases} \forall x_1, y_1: P(x_1), R(x_1, y_1) \quad \exists: \mathbf{t} \quad \forall y_2: L(y_1, y_2) \quad \exists: R(x_1, y_2) \\ \forall x_2: P(x_2) \quad \begin{cases} \exists y_3: T(y_3) \quad \forall: R(x_2, y_3) \quad \exists: \mathbf{f} \\ \exists: \mathbf{t} \quad \forall y_4: S(y_4), R(x_2, y_4) \quad \exists: \mathbf{f}. \end{cases} \end{cases} \quad (8)$$

Теперь возможны ответы на оба вопроса.

При ответе на первый с ответной подстановкой $\theta_3 : x_1 \rightarrow x, y_1 \rightarrow y$ получается

$$\omega^2(F) = \forall : t \exists x, y : P(x), S(y), R(x, y) \left\{ \begin{array}{l} \forall x_1, y_1 : P(x_1), R(x_1, y_1) \exists : t \forall y_2 : L(y_1, y_2) \exists : R(x_1, y_2) \\ \forall x_2 : P(x_2) \left\{ \begin{array}{l} \exists y_3 : T(y_3) \forall : R(x_2, y_3) \exists : f \\ \exists : t \forall y_4 : S(y_4), R(x_2, y_4) \exists : f \end{array} \right. \\ \forall y_2 : L(y, y_2) \exists : R(x, y_2). \end{array} \right.$$

При ответе на вопрос $\forall x_2 : P(x_2)$ с ответной подстановкой $\theta_4 : x_2 \rightarrow x$ БП из $\omega^2(F)$ расщепится: $\omega^3(F) = \forall : t \{ \Omega_1, \Omega_2 \}$, где

$$\Omega_1 = \exists x, y, y_5 : P(x), S(y), R(x, y), T(y_5) \left\{ \begin{array}{l} \Sigma \\ \forall : R(x, y_5) \exists : f, \end{array} \right. \quad (9)$$

$$\Omega_2 = \exists x, y : P(x), S(y), R(x, y) \left\{ \begin{array}{l} \Sigma \\ \forall y_5 : S(y_5), R(x, y_5) \exists : f, \end{array} \right.$$

$$\Sigma = \left\{ \begin{array}{l} \forall x_1, y_1 : P(x_1), R(x_1, y_1) \exists : t \forall y_2 : L(y_1, y_2) \exists : R(x_1, y_2) \\ \forall x_2 : P(x_2) \left\{ \begin{array}{l} \exists y_3 : T(y_3) \forall : R(x_2, y_3) \exists : f \\ \exists : t \forall y_4 : S(y_4), R(x_2, y_4) \exists : f \end{array} \right. \\ \forall y_2 : L(y, y_2) \exists : R(x, y_2). \end{array} \right. \quad (10)$$

В отличие от первой подзадачи, вторая (опровержения Ω_2) разрешается положительно за один шаг применения ω с ответной подстановкой $\theta_5 : y_5 \rightarrow y$, приводя к $\perp$, что не достигается в первой подзадаче (опровержения Ω_1). Действительно, вопрос $\forall x_2 : P(x_2)$ перестал быть уместным, так как, несмотря на пополнение базы атомом $T(y_5)$, ответной подстановки, отличной от уже использованной $\theta_4 : x_2 \rightarrow x$, не появилось. Поэтому в исчислении $J = (\perp, \omega)$ БП Ω_1 , а значит и в целом F неопровержимы, так как $F \leftrightarrow \omega^3(F)^L = \forall : t \{ \Omega_1, \Omega_2 \}$. Таким образом, F , как отрицание импликации $V_1 \wedge V_2 \rightarrow V_3$, неопровержимо и V_3 не следует из $V_1 \wedge V_2$. При другом порядке ответа на вопросы из (8) результат обязан быть тем же.

Эксперта может интересовать, какого условия H не хватает для $V_1 \wedge V_2 \rightarrow V_3$ с тем, чтобы при лаконичности этого H заменить им «громоздкое» V_3 . В продолжение примера 1 для синтеза этого H применяется исчисление $J_+ = (\perp, \omega, \delta)$, $\delta \in \{ \alpha, \beta \}$, $\alpha \in \{ \alpha_0, \alpha_v, \alpha_i \}$, $\beta \in \{ \beta_j \}_{j \in \overline{1,6}}$.

При выборе $\delta = \alpha_v$

$$H_v = \forall x^*, y^*, y_5^* : P(x^*), S(y^*), R(x^*, y^*), T(y_5^*) \left\{ \begin{array}{l} \exists x_1^*, y_1^* : P(x_1^*), R(x_1^*, y_1^*) \\ \exists x_2^* : P(x_2^*) \\ \exists y_2^* : L(y^*, y_2^*) \\ \exists : R(x^*, y_5^*) \end{array} \right. \quad (11)$$

$$\alpha_v \Omega_1 = \exists x, y, y_5 : P(x), S(y), R(x, y), T(y_5) \{ \Sigma, (\forall : R(x, y_5) \exists : f), H_v \}.$$

В $\exists$ -формуле Ω_1 вида (9-10) содержатся четыре закономерности, из которых две последние кажутся перспективными для опровержения Ω_1 . Это

$$\forall y_2 : L(y, y_2) \exists : R(x, y_2), \quad \forall : R(x, y_5) \exists : f, \quad (12)$$

активизируемые на основе начальных гипотез H_3 и H_4 по правилам α_i .

Согласно гипотезе первичного типа H_3 :

$$H_3 = \forall x^*, y^*, y_5^* : P(x^*), S(y^*), R(x^*, y^*), T(y_5^*) \exists y_2^* : L(y^*, y_2^*) \quad (13)$$

Действительно, в применении к (9) гипотеза H_3 может позволить нарастить базу в Ω_1 атомом $L(\cdot, \cdot)$, задействуя вначале третью закономерность из (10), а далее, после попадания в базу атома $R(x, y_5)$, и четвёртую закономерность из (9) с опровержением Ω_1 . Это будет означать опровержимость и в целом сделанного предположения $F = (V_1 \wedge V_2 \wedge \neg V_3)^L$, которое логически эквивалентно $\omega^2(F)^L = \forall : t \{ \Omega_1, \Omega_2 \}$. Можно проверить реализуемость синтеза требуемого условия для замещения V_3 . По правилу α_3 формируется расширение закономерностей в БП Ω_1 гипотезой (13):

$$\alpha_3 \Omega_1 = \exists x, y, y_5 : P(x), S(y), R(x, y), T(y_5) \{ \Sigma, (\forall : R(x, y_5) \exists : f), H_3 \}.$$

После подстановки $\theta_6 : y_2^* \rightarrow y_5^*$ в (13) получается утверждение

$$O_3(H_3) = \forall x^*, y^*, y_5^* : P(x^*), S(y^*), R(x^*, y^*), T(y_5^*) \exists : L(y^*, y_5^*),$$

означающее: «Для любой стратосферной высоты и любого самолёта, которому эта высота доступна, любая тропосферная высота – ниже». С позиции упрощения гипотезы эксперт заметит нецелесообразную ограничительность посылки и применит операцию O_1 удаления атомов $P(x^*), R(x^*, y^*)$. После этого квантор $\forall x^*$ становится независимым и удаляется согласно O_4 . Получится гипотеза

$$O_4 O_1 O_3(H_3) = \forall y^*, y_5^* : S(y^*), T(y_5^*) \exists : L(y^*, y_5^*), \quad (14)$$

означающая, что «Тропосферные высоты ниже стратосферных».

После подстановки $\theta_7 : y^* \rightarrow y, y_5^* \rightarrow y_5$ база в (9) дополнится атомом $L(y, y_5)$, приведя к уместности третьего вопроса в (9). Ответная подстановка $\theta_8 : y_2 \rightarrow y_5$ обеспечит попадание в базу из (9) атома $R(x, y_5)$ и уместность вопроса четвёртой закономерности из (9). Ответ на него с пустой подстановкой приведёт к проникновению в базу из (9) константы f , т.е. к опровержению БП Ω_1 .

По теоремам 1 и 2 гипотеза (вторичного типа) H_3^* вида (14) является достаточным условием доказуемости свойства $F = V_1 \wedge V_2 \rightarrow V_3$. Поэтому в описании предмета рассмотрения (см. $(V)^L$ вида (6)) утверждение V_3 можно заменить на (14).

Пример 3 (синтез гипотезы, альтернативной гипотезе (14)). Из ранее выделенных в (12) двух закономерностей для синтеза гипотезы, достаточной для $V_1 \wedge V_2 \rightarrow V_3$, можно выбрать четвёртую закономерность, для задействования которой в множестве усиления (точнее логических неослаблений) условия (11), получаемых разделением ветвей, по правилу α_4 выбрать гипотезу

$$H_4 = \forall x^*, y^*, y_5^* : P(x^*), S(y^*), R(x^*, y^*), T(y_5^*) \exists : R(x^*, y_5^*). \quad (15)$$

Получится

$$\alpha_4 \Omega_1 = (\alpha_\vee \Omega_1)(H_\vee / H_4) = \exists x, y, y_5 : P(x), S(y), R(x, y), T(y_5) \{ \Sigma, (\forall : R(x, y_5) \exists : f), H_4 \}.$$

Ответная подстановка $\theta_9 : x^* \rightarrow x, y_5^* \rightarrow y_5$ с применением правила ω дополнит базу из Ω_1 атомом $R(x, y_5)$ и с учётом закономерности $\forall : R(x, y_5) \exists : f$ из Ω_1 (см. (9)) приводит к жела-

емому опровержению БП Ω_1 , причём более коротким путём, чем в примере 2. Гипотеза первичного типа H_4 в форме (15) означает: «Для любого самолёта и любой доступной ему стратосферной высоты любая тропосферная высота тоже доступна», или, что эквивалентно: «Для любой стратосферной высоты и любого самолёта, которому эта высота доступна, любая тропосферная высота тоже доступна».

Из операций преобразования β применима только операция O_1 . Если эксперт операцией O_1 удалит атомы $S(y^*), R(x^*, y^*)$, то потом операцией O_4 удалится и квантор $\forall y^*$, становящийся независимым. Это приведёт к семантически более простой гипотезе вторичного типа

$$O_4 O_1(H_4) = \forall x^*, y_5^* : P(x^*), T(y_5^*) \exists : R(x^*, y_5^*), \quad (16)$$

означающей: «Самолётам все тропосферные высоты доступны».

Гипотеза (16) является примером гипотезы логически верной с точки зрения её достаточности для доказуемости свойства $V_1 \wedge V_2 \rightarrow V_3$, но неверной при учёте знаний, не охваченных описанием предмета $(V)^L$. Это так, поскольку разнообразие самолётов в описании $(V)^L$ не ограничено и, в частности, включает модели низковысотные, которым верхние высоты тропосферы недоступны. Эксперту придётся задуматься над содержанием используемого им понятия «Самолёт» и при определённой широте этого понятия отказаться от замещения утверждения V_3 в описании $(V)^L$ ПрО (см. (6)) условием (16) и вместо этого использовать ранее синтезированную гипотезу (14)

$$H_3^* = O_4 O_1 O_3(H_3) = \forall y^*, y_5^* : S(y^*), T(y_5^*) \exists : L(y^*, y_5^*),$$

согласно которой «Тропосферные высоты ниже стратосферных».

Заключение

В работе исследованы возможности интерактивного синтеза гипотез как новых знаний на основе расширения исчисления позитивно-образованных стандартизованных формул (пос-формул), разработанного ранее для задач автоматизации логического вывода.

Предложенный метод синтеза новых знаний объединяет возможности анализа знаний на основе АД с алгоритмизированным человеко-машинным формированием текстов утверждений. При автоматическом выявлении в описании ПрО отсутствия свойства, желательного для человека, задействуется расширение исчисления дополнительными правилами, которые обеспечивают синтез условия-гипотезы, логически гарантирующей наличие требуемого свойства. На их основе варианты гипотетических условий синтезируются автоматически, но в предварительном виде. Для оценки их приемлемости с учётом знаний ПрО, не охваченных её описанием, и повышения содержательности дальнейший синтез выполняется интерактивно.

Обоснованы свойства правил синтеза. Их применение дополняет множество закономерностей, представленных в обрабатываемой пос-формуле, и гарантированно продолжает вывод в ситуациях неприменимости базового правила вывода в исчислении пос-формул.

Расширение первопорядкового исчисления пос-формул правилами синтеза преобразует его в исчисление дедуктивно-абдуктивного типа, свободное от ряда ограничений.

Рассмотрены примеры анализа знаний и синтеза условий наличия требуемых свойств в описании ПрО. Использование крупноблочного типово-кванторного представления знаний, свойственного языку пос-формул, обеспечивает наглядность обрабатываемого текста.

К перспективным направлениям дальнейшей работы относится анализ возможностей использования и развития полученных результатов:

- в области онтологий проектирования с порождением знаний шире, чем только в классе логических следствий известного, и с понижением, может быть поначалу, выразительности языка формализации для совместимости с другими разработками этой области;
- в задачах диагностики наблюдаемого с привлечением экспериментальных данных для поддержки экспертной оценки синтезируемых объяснений;
- применительно к неклассическим логикам с семантикой конструктивности и немонотонности в приложениях компьютерных систем диагностики и поддержки принятия решений.

Список источников

- [1] *Мендельсон Э.* Введение в математическую логику. М.: Наука, 1971. 320 с.
- [2] *Gentzen G.* Untersuchungen uber das logische schliessen // *Mathematische Zeitschrift*. 1935. Vol.39. P.176-210; DOI:10.1007/BF01201353.
- [3] *Маслов С.Ю.* Обратный метод установления выводимости в классическом исчислении предикатов // *ДАН СССР*. 1964. Т.159, №1. С.17-20.
- [4] *Шанин Н.А.* Алгоритм машинного поиска естественного логического вывода в исчислении высказываний / Н.А. Шанин, Г.В. Давыдов, С.Ю. Маслов, Г.Е. Минц и др. // М.-Л.: Наука, 1965. 39 с.
- [5] *Робинсон Д.* Машинно-ориентированная логика, основанная на методе резолюции // Ляпунов А.А., Лупанов О.Б. (ред.). *Киберн. сб., Нов. сер. Вып.7.* М.: Мир, 1970. С.180-218.
- [6] *D'Agostino M.* Handbook of Tableau Methods / M. D'Agostino, D.M. Gabbay, R. Hahnle, J. Posegga, eds. // Dordrecht: Kluwer Academic Publishers, 1999. 680 p.
- [7] *Kovács L., Voronkov A.* First-order theorem proving and Vampire // In: Sharygina, N., Veith, H. (eds.). *Proc. of the 25th Conf. on Computer Aided Verification, LNCS.* Springer, Heidelberg. 2013. Vol.8044. P.1–35. DOI:10.1007/978-3-642-39799-8_1.
- [8] *Otten J.* NanoCoP: a Non-clausal Connection Prover // *Proc. Int. Joint Conference on Automated Reasoning*. 2016. P.302-312.
- [9] *Васильев С.Н., Жерлов А.К.* Об исчислении типово-кванторных формул // *Докл. РАН*. 1995. Т.343(5). С.583-585.
- [10] *Васильев С.Н., Жерлов А.К., Федосов Е.А., Федунев Б.Е.* Интеллектуальное управление динамическими системами. М.: ФИЗМАТЛИТ, 2000. 352 с.
- [11] *Wang H.* Toward mechanical mathematics // *IBM J. Res. Develop.* 1960. Vol. 4(1). P.2-22.
- [12] *Gödel K.* Die Vollständigkeit der Axiome des Logischen Funktionenkalküls // *Monatsh. Math. Phys.* 1930. 37. P.349-360.
- [13] *Church A.* A Note on the Entscheidungsproblem // *J. Symbolic Logic*. 1936. 1. P.40-41.
- [14] *Turing A.M.* On Computable Numbers, with an Application to the Entscheidungsproblem // *Proc. Lond. Math. Soc., Ser. 2*. 1936. 42. P.230-265; 1935. 43. P.544-546.
- [15] *Чень Ч., Ли Р.* Математическая логика и автоматическое доказательство теорем. М.: Наука, 1983. 360 с.
- [16] *Черкашин Е.А.* Программная система «КВАНТ/1» для автоматического доказательства теорем. Иркутск: ИДСТУ СО РАН, 1999. 16 с.
- [17] *Larionov A.A., Davydov A.V., Cherkashin E.A.* The calculus of positively constructed formulas, its features, strategies and implementation // In: *Proc. of the 36th Intern. Convention on Information and Communication Technology, Electronics and Microelectronics (2013, Opatija, Croatia)*. N.Y.: IEEE, 2013. P.1023-1028.
- [18] *Давыдов А.В., Ларионов А.А., Черкашин Е.А.* Метод трансляции первопорядковых логических формул в позитивно-образованные формулы // *Программные продукты и системы*. 2019. Т.32(4). С.556–564. DOI:10.15827/0236-235X.128.556-564.
- [19] *Бурбаки Н.* Теория множеств. М.: Мир, 1965. 465 с.
- [20] *Мальцев А.И.* Модельные соответствия // *Изв. АН СССР. Математика*. 1959. 23(3). С.313-336.
- [21] *Мальцев А.И.* Алгебраические системы. М.: Наука, 1970. 329 с.
- [22] *Walther C.* Many-sorted unification // *J. of the ACM*. 1988. 35(1). P.1–17.
- [23] *Palacz W., Grabska E., Slusarczyk G.* Ontological Approach to Design Reasoning with the Use of Many-Sorted First Order Logic // In: *Proc. of the 15th Int. Conf. on Artificial Intelligence and Soft Computing (Zakopane, Poland, 2016), Part II*, 2016. P.364–374. DOI:10.1007/978-3-319-39384-1_31.
- [24] *Нагул Н.В.* Метод логико-алгебраических уравнений в динамике систем // *Алгебра и анализ*. 2012. 24(4). С.156-181. DOI:10.1090/S1061-0022-2013-01258-1.

- [25] **Нагул Н.В.** Генерация условий для сохранения свойств управляемых дискретных систем событий // Автоматика и телемеханика. 2016. 77 (4). С.153–172. DOI:10.1134/S0005117916040111.
- [26] **Матросов В.М., Анапольский Л.Ю., Васильев С.Н.** Метод сравнения в математической теории систем. Новосибирск: Наука, 1980. 481 с.
- [27] **Васильев С.Н.** Метод сравнения в анализе систем. I-IV // Дифференциальные уравнения. 1981. 17(9). С.1562-1573; 1981.17(11). С.1945-1954; 1982.18(2). С.197-205; 1982. 18(6). С.938-947.
- [28] **Vassilyev S.N.** Machine synthesis of mathematical theorems // J. Logic Program. 1990. 9(2–3). P.235–266.
- [29] **Васильев С.Н.** Об импликации свойств связанных систем: метод получения условий импликации и примеры применения // Изв. РАН. Теория и системы управления. 2020. 4. С.3-17. DOI:10.31857/S0002338820040149.
- [30] **Kowalski R.** History of Logic Programming. In: Computational Logic, D. Gabbay, J. Woods (eds.). Elsevier. 2014. P.523-569.
- [31] **Horn A.** On sentences which are true on direct unions of algebras // J. Symbolic Logic. 1951. 16. P.14-21.
- [32] **Calmerauer A., Kanoui H., Pasero R., Roussel P.** Un Systeme de Communication Homme-Machine en Francais // Rapport, Groupe d'Intelligence Artificielle, Universite d'Aix-Marseille II, 1973.
- [33] **Kowalski R.** Predicate Logic as a Programming Language // In: Proc. of the IFIP-74 Congress. 1974. P.569-574.
- [34] **Poole D.** Representing Diagnosis Knowledge // J. Annals of Mathematics and Artificial Intelligence. 1994. 11. P.33-50.
- [35] **Kobriniskii B.A.** Images in Logical-and-Linguistic Artificial Intelligence // J. Biomedical Engineering and Medical Imaging. 2019. 6(1). P.1-8. DOI: 10.14738/jbemi.61.61.61.
- [36] **Buzikov M., Galyaev A., Guryev Yu., Titov K., Yakushenko E., Vassilyev S.** Intelligent Control of Autonomous and Anthropocentric On-board Systems // Procedia Computer Science. 2019. 150. P.10-18. DOI:10.1016/j.procs.2019.02.004.
- [37] **Kowalski R., Sadri F.** Reactive computing as model generation // New Generat. Comput. 2015. 33: 33-67. DOI:10.1007/s00354-015-0103-z.
- [38] **Lindon R.C.** Properties preserved under homomorphism // Pacific J. Math. Vol. 9, 1959. – P. 143-154.

Сведения об авторе

Васильев Станислав Николаевич, 1946 г. рождения. Окончил Казанский авиационный институт им. А.Н. Туполева в 1970 г., д.ф.-м.н. (1989). Г.н.с. ИПУ им. В.А. Трапезникова РАН. Академик Российской академии наук. Член Российской ассоциации искусственного интеллекта (ИИ). В списке научных трудов более 400 работ в области математической теории систем, теории управления и ИИ. Author ID (RSCI): 2467; Author ID (Scopus): 6603565134; Researcher ID (WoS): O-7577-2017. vassilyev_sn@mail.ru.



Поступила в редакцию 22.01.2023, Принята к публикации 15.02.2023.



Scientific article

DOI: 10.18287/2223-9537-2023-13-1-10-28

Interactive generation of new knowledge based on automatic means of logical inference

© 2023, S.N. Vassilyev

V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia

Abstract

The work relates to the field of knowledge representation and algorithmization of their processing. To describe the subject of consideration by a set of statements and their processing, logical means of the first order type are used. Unlike

the well-known achievements in the field of automatic proof (AP) in formal axiomatic theories, the automatic text synthesis of statements is complicated by the poor formalizability of the value of knowledge concept contained in the text. Based on the development of the calculus of positively-formed standardized formulas (pfs-formulas) the possibilities of analysis automation of the description of the subject matter and interactive generation of meaningful statements to improve this description are investigated. The proposed method of interactive knowledge generation integrates the capabilities of AP with algorithmized human-machine synthesis of new knowledge texts. AP in the calculus of pfs-formulas helps a person not only in the analysis of properties, but also in the choice of the required properties of the description of the subject matter. When the absence of a property desirable for a person in the analyzed knowledge is revealed by means of AP, the hypotheses is formed that logically guarantees the presence of the required property. Variants of hypotheses are automatically synthesized in a preliminary form, which are then refined interactively to improve the content of the generated text of new knowledge. The rules of synthesis are justified. Their application continues the inference in situations of inapplicability of the basic rule of inference in the calculus of pfs-formulas. The extension of calculus by the rules of synthesis transforms it as a first-order calculus of the deductive type into a means of deductive-abductive inference, which differs from the known works using abduction by the absence of a number of restrictions. Illustrative examples demonstrate the representation, analysis and generation of knowledge. These examples also show in interactive knowledge processing the importance of visualization of the processed text, provided by the large-block structure of pfs-formulas constructed from typical quantifiers. In conclusion, the results of the paper and possible directions for the development of the results are formulated.

Key words: *knowledge representation, knowledge generation, logical synthesis of hypotheses, abductive inference, automatic theorem proving.*

For citation: *Vassilyev SN. Interactive generation of new knowledge based on automatic means of logical inference [In Russian]. Ontology of designing. 2023; 13(1): 10-28. DOI: 10.18287/2223-9537-2023-13-1-10-28.*

Conflict of interest: The author declares no conflict of interest.

References

- [1] **Mendelson E.** Introduction to mathematical logic. 6th edition. New York: CRC Press; 2015. 472 p.
- [2] **Gentzen G.** Untersuchungen uber das logische schliessen. Mathematische Zeitschrift. 1935; 39: 176-210.
- [3] **Maslov SYu.** An inverse method of establishing deducibility in the classical predicate calculus [In Russian]. Dokl. Akad. Nauk SSSR. 1964; 159(1): 17-20.
- [4] **Shanin NA, Davydov GV, Maslov SYu, Mints GE, e.a.** The algorithm of machine search for natural logical inference in the calculus of statements [In Russian]. M.-L.: Science; 1965. 39 p.
- [5] **Robinson JA.** A Machine-oriented logic based on the resolution principle. J. ACM. 1965; 12: 23-41.
- [6] **D'Agostino M, Gabbay DM, Hahnle R, Posegga J** (eds.). Handbook of Tableau Methods. Dordrecht: Kluwer Academic Publishers; 1999. 680 p.
- [7] **Kovács L, Voronkov A.** First-order theorem proving and Vampire. In: Sharygina, N., Veith, H. (eds.). Proc. of the 25th Conf. on Computer Aided Verification, LNCS. Springer, Heidelberg, 2013; 8044: 1–35. DOI: 10.1007/978-3-642-39799-8_1.
- [8] **Otten J.** NanoCoP: a Non-clausal Connection Prover. Proc. of the Int. Joint Conference on Automated Reasoning. 2016: 302-312.
- [9] **Vassilyev SN, Zherlov AK.** On calculi of type-quantifier formulas [In Russian]. Dokl. Akad. Nauk. 1995; 343(5): 583–585.
- [10] **Vassilyev SN, Zherlov AK, Fedosov EA, Fedunov BE.** Intelligent Control of Dynamical Systems [In Russian]. Moscow: Fizmatlit Publ.; 2000. 352 p.
- [11] **Wang H.** Toward mechanical mathematics. IBM J. Res. Develop. 1960; 4(1): 2-22.
- [12] **Gödel K.** Die Vollständigkeit der Axiome des Logischen Funktionenkalküls. Monatsh. Math. Phys. 1930; 37: 349-360.
- [13] **Turing AM.** On Computable Numbers, with an Application to the Entscheidungsproblem // Proc. Lond. Math. Soc., Ser. 2. 1936; 42: 230-265; 1935; 43: 544-546.
- [14] **Church A.** A Note on the Entscheidungsproblem. J. Symbolic Logic. 1936; 1: 40-41.
- [15] **Chang C.-L, Lee RC-T.** Symbolic Logic and Mechanical Theorem Proving. N.-Y., San Francisco, London: Academic Press Inc., 1973. 358 p.
- [16] **Cherkashin EA.** KVANT/1 Program System for Automated Theorem Proving. PhD Thesis [In Russian]. Irkutsk: IDSTU SO RAN; 1999. 16 p.

- [17] **Larionov A, Davydov A, Cherkashin E.** The calculus of positively constructed formulas, its features, strategies and implementation. In: Proc. of the 36th Intern. Convention on Information and Communication Technology, Electronics and Microelectronics (2013, Opatija, Croatia). N.Y.: IEEE; 2013: 1023-1028.
- [18] **Davydov AV, Larionov AA, Cherkashin EA.** The method for translating first-order formulas into positively constructed formulas [In Russian]. *Software & Systems*. 2019; 32(4): 556–564. DOI: 10.15827/0236-235X.128.556-564.
- [19] **Bourbaki N.** Theory of Sets. Paris: Hermann; 1968. 424 p.
- [20] **Mal'cev AI.** Model correspondences [In Russian]. *Izv. of the USSR Academy of Sciences. Mathematics*. 1959; 23(3): 313-336.
- [21] **Mal'cev AI.** Algebraic systems. Springer Verlag; 1973. 317 p.
- [22] **Walthers C.** Many-sorted unification. *J. of the ACM*. 1988; 35(1): 1–17.
- [23] **Palacz W, Grabska E, Slusarczyk G.** Ontological Approach to Design Reasoning with the Use of Many-Sorted First Order Logic. In: Artificial Intelligence and Soft Computing: Proc. of the 15th Int. Conf. (Zakopane, Poland, 2016, June 12–16), Part II. 2016: 364–374. DOI:10.1007/978-3-319-39384-1_31.
- [24] **Nagul NV.** The method of logical-algebraic equations in the dynamics of systems. *S. P. Math. J.* 2013; 24(4): 645–662. DOI: 10.1090/S1061-0022-2013-01258-1.
- [25] **Nagul NV.** Generating conditions for preserving the properties of controlled discrete event systems. *Automation and Remote Control*. 2016; 77(4): 672–686. DOI:10.1134/S0005117916040111.
- [26] **Matrosov VM, Anapolsky LYu, Vassilyev S.N.** The comparison method in the mathematical theory of systems [In Russian]. Novosibirsk: Science; 1980. 481 p.
- [27] **Vassilyev SN.** The comparison method in the analysis of systems. I-IV [In Russian]. *Differential equations*. 1981; 17(9):1562-1573; 1981;17(11):1945-1954; 1982;18(2):197-205; 1982;18(6):938-947.
- [28] **Vassilyev SN.** Machine synthesis of mathematical theorems. *J. Logic Program.* 1990; 9(2–3): 235–266.
- [29] **Vassilyev SN.** On the Implication of Properties of Related Systems: The Method for Obtaining Implication Conditions and Application Examples. *J. of Computer and Systems Sciences International*. 2020; 59(4): 479-493. DOI: 10.1134/S1064230720040140.
- [30] **Kowalski R.** History of Logic Programming. In: Computational Logic, D. Gabbay and J. Woods (eds.). Elsevier. 2014. P.523-569.
- [31] **Horn A.** On sentences which are true on direct unions of algebras. *J. Symbolic Logic*. 1951; 16: 14-21.
- [32] **Calmerauer A, Kanoui H, Pasero R, Roussel P.** Un Systeme de Communication Homme-Machine en Francais [In French]. Rapport, Groupe d'Intelligence Artificielle, Universite d'Aix-Marseille II, 1973.
- [33] **Kowalski R.** Predicate Logic as a Programming Language. In: Proc. of the IFIP-74 Congress. 1974: 569-574.
- [34] **Poole D.** Representing Diagnosis Knowledge. *J. Annals of Mathematics and Artificial Intelligence*. 1994; 11: 33-50.
- [35] **Kobrinskii BA.** Images in Logical-and-Linguistic Artificial Intelligence. *J. Biomedical Engineering and Medical Imaging*. 2019; 6(1): 1-8. DOI: 10.14738/jbemi.61.61.61.
- [36] **Buzikov M, Galyaev A, Guryev Yu, Titov K, Yakushenko E, Vassilyev S.** Intelligent Control of Autonomous and Anthropocentric On-board Systems. *Procedia Computer Science*. 2019; 150: 10-18. DOI:10.1016/j.procs.2019.02.004.
- [37] **Kowalski R, Sadri F.** Reactive computing as model generation. *New Generat. Comput.* 2015; 33: 33-67. DOI:10.1007/s00354-015-0103-z.
- [38] **Lindon RC.** Properties preserved under homomorphism. *Pacific J. Math*. 1959; 9: 143-154.

About the author

Stanislav Nikolaevich Vassilyev (b. 1946) graduated from the Tupolev Aviation Institute (Kazan-city) in 1970, Doctor of Ph.D. (1989). Chief Researcher at the V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences (RAS). Academician of the RAS. Member of the Russian Association of Artificial Intelligence (AI). He is an author and co-author more than 400 publications in the field of mathematical systems theory, control theory and AI. Author ID (RSCI): 2467; Author ID (Scopus): 6603565134; Researcher ID (WoS): O-7577-2017. vassilyev_sn@mail.ru.

Received January 22, 2023. Accepted February 15, 2023.



Подход к оцениванию уровня интеллектуальности информационной системы

© 2023, С.В. Микони

Санкт-Петербургский Федеральный исследовательский Центр РАН,
Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, Россия

Аннотация

На основе анализа определений естественного и искусственного интеллекта (ИИ) предложены базисные функции, отражающие познавательную деятельность мозга. Их трактовка в широком смысле с использованием правил русской грамматики позволила описать весь спектр понятий, применяемых в ИИ. Показана возможность выражения базисных функций через более простые и объединения их в более сложные. Обоснована полнота, непротиворечивость и избыточность базиса познавательных функций мозга. 14 базисных функций разделены на четыре группы, отражающие различные стороны работы со знанием. Это позволило рассматривать систему ИИ как специализированную систему обработки информации. За признак, выделяющий подсистему ИИ в информационной системе, принят имитационный подход к моделированию функций человеческого мозга. Уровень интеллектуальности информационной системы характеризуется показателями долей аппаратного и программного обеспечения, использованных для реализации моделей ИИ, в общем объеме аппаратного и программного обеспечения информационной системы. Приведён пример оценивания уровня интеллектуальности информационной системы беспилотного летательного аппарата.

Ключевые слова: естественный интеллект, искусственный интеллект, ум, разум, уровень интеллектуальности, информационная технология, информационная система.

Цитирование: Микони С.В. Подход к оцениванию уровня интеллектуальности информационной системы // Онтология проектирования. 2023. Т.13, №1(47). С.29-43. DOI: 10.18287/2223-9537-2023-13-1-29-43.

Финансирование: исследования проводились в рамках бюджетной темы FFZF–2022–0004.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

В национальной стратегии развития искусственного интеллекта (ИИ) [1] понятие «ИИ» определено как «комплекс технологических решений, имитирующий когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и позволяющий при выполнении задач достигать результаты, как минимум сопоставимые с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает информационно-коммуникационную инфраструктуру, программное обеспечение, в том числе реализующее методы машинного обучения, процессы и сервисы (прежде всего в «облачной» архитектуре) по обработке данных и выработке решений». Технологии в области ИИ определены как «технологии, основанные на ИИ, включая, в том числе, компьютерное зрение, обработку естественного языка, распознавание и синтез речи, рекомендательные системы и интеллектуальные системы поддержки принятия решений...» [1].

Выход документа [1] оживил обсуждение в научной среде определения ИИ, его возможностей и границ. К одной из первых публикаций в этом направлении следует отнести [2], в которой приведён краткий обзор исследований по ИИ. С начала 90-х годов в России практически не разрабатывались, за некоторыми исключениями, собственные промышленные ин-

теллектуальные технологии. Поэтому вместе с западными технологиями пришли соответствующая терминология и взгляды на ИИ. Акцент сместился с изучения и моделирования умственных способностей человека на их машинную реализацию. В подтверждение технологического характера современного ИИ в [2] приведён пример так называемой периодической системы технологий ИИ [3]. В отличие от физических законов, которые отражены в периодической системе элементов Д.И. Менделеева, автор системы технологий ИИ представил схему связей между различными подходами и задачами ИИ. В ней не учитывается связь между объектами моделирования (функциями мозга) и применяемыми для моделирования моделями и методами ИИ, не учитываются отношения «Общее-частное», «Целое-часть», «Причина-следствие» между обсуждаемыми сущностями и т.д. Многообразие этих связей во всём их объёме невозможно представить таблицей на плоскости.

В работе [2] приведены следующие определения ИИ:

- ИИ – это междисциплинарная область исследований и набор технологий, позволяющий создавать технические системы, решающие задачи, ранее доступные только человеку;
- ИИ – это способность системы приобретать, обрабатывать и применять знания (знания – это факты, информация и навыки, приобретенные в результате опыта или обучения).

В первом определении, хотя и существует отсылка к человеку, но не указывается связь между объектом моделирования – естественным интеллектом (ЕИ) – и его искусственными моделями. Второе определение ИИ, данное в ISO / IEC 3WD 22989, на самом деле определяет не ИИ, а свойство системы ИИ (СИИ), и не охватывает всех решаемых ИИ задач.

В работе [2] не затронут важный вопрос, на который необходимо дать ответ в связи с появлением документа [1]: чем СИИ отличается от системы обработки информации (СОИ)? Практические соображения на этот счёт приведены, например, в обзоре [4, 5]: «на высоких уровнях автономности на бортовые управляющие системы БЛА возлагаются функции, традиционно ассоциирующиеся с интеллектуальной деятельностью человека» [4, стр. 3]. Этим высказыванием утверждается, что автономность БЛА обеспечивается не только средствами ИИ. Таким образом, актуальна задача выявления интеллектуальной составляющей информационной системы (ИС) с последующим определением уровня её интеллектуальности. Настоящая статья посвящена решению этой задачи.

1 Естественный интеллект

Известны различные определения интеллекта. В Энциклопедии Британии даётся следующее перечислительное определение интеллекта [6]: интеллект (от лат. *intellectus* «восприятие»; «разумение», «понимание»; «понятие», «рассудок» или «ум») — качество *психики*, состоящее из способности *приспосабливаться* к новым ситуациям, способности к *обучению* и *запоминанию* на основе опыта, *пониманию* и применению *абстрактных* концепций и использованию своих *знаний* для *управления* окружающей человека средой¹.

Наиболее общим понятием этого определения является психика. В [7] психика определена как *системное свойство высокоорганизованной живой материи, заключающееся в активном отражении субъектом объективного мира, в построении неотчуждаемой от него картины мира и регуляции на этой основе своего поведения и деятельности*¹. Это описательное определение охватывает все свойства психики, в том числе не названные в определении [6].

Существенные признаки (СП) понятия «интеллект» в определении [6] хотя и относятся к операциям психики, но различаются по уровню общности. Приспособление субъекта к новой ситуации осуществляется на основе его обучения, а обучение требует понимания и приобре-

¹ Перевод автора.

тения нового знания. Понимание требует, в свою очередь, запоминания ранее полученного знания, ибо смысл понимания заключается в перекодировании новой информации в коды известного знания, используемые его носителем [8].

Упомянутые в определении [6] свойства психики образуют следующую причинно-следственную цепочку: *восприятие* информации → *понимание* полученной информации → *сопоставление* понятой информации с известной → *запоминание* новой информации как результата *обучения* → *реагирование* на полученную информацию (управление действиями) по известным правилам. В случае новизны информации выполняется её *анализ* на предмет соответствия известным правилам поведения. При отсутствии правил поведения, обеспечивающих гармонию с внешней средой или противодействие ей, вырабатываются необходимые новые правила. Выработка и выполнение правил поведения, обеспечивающих гармонию с внешней средой, и есть приспособительная реакция на изменение внешней среды.

В русском языке наиболее близким термину «интеллект» является слово «ум». Неслучайно в СИИ используются такие словосочетания, как умный дом, умная машина и пр. Пользуясь грамматической системностью русского языка, интересно выявить бытовой смысл производных от слова «ум» – слов *уметь* и *разуметь*. Согласно [9] слово «уметь» означает «обладать навыками, необходимыми, чтобы сделать что-либо». Слово «разуметь» в смысле «понимать, постигать смысл вещей» отражает более высокий уровень мышления.

Более глубокий смысл разума (*разуметь*) отражает описательное определение интеллекта. «Интеллект включает в себя рациональный и логический аспекты человеческого разума. При изучении человеческого разума интеллект описывает и идентифицирует способность человеческого разума делать правильные выводы о том, что истинно, а что ложно в действительности, и как решать проблемы»¹ [10].

Из этого определения можно заключить, что для решения любой проблемы достаточно ума, а для того, чтобы делать правильные выводы о том, что истинно, а что ложно в действительности, необходим разум. Смысл этих слов сводится соответственно к ремеслу и творчеству. Под ремеслом понимается решение задачи по известному алгоритму. Творчество предполагает создание нового алгоритма для решения задачи². В общем случае мыслительная деятельность предполагает наличие обоих начал в различной пропорции.

Разум предполагает знание более общих закономерностей и законов, на основании которых можно делать заключения об истинности любых суждений. И чем большим числом уровней знания обладает разум, тем больше истин он может установить, что следует из теоремы Гёделя о неполноте [11], доказывающей необходимость метазнания для установления истинности некоторой теории предметной области (ПрО).

Информацию, которой оперирует мозг, можно разделить на обрабатываемую и обрабатывающую (инструментальную). Инструментальная информация нижнего уровня общности включает модели и методы, используемые для решения различных классов задач. Верхние уровни инструментальной информации содержат совокупности законов, закономерностей и правил, используемые для создания моделей и методов. Перечень правил вывода (способов решения задач) представляет собой не линейный список, а сложную структуру.

Набор решаемых задач зависит от объёма инструментальной информации, т.е. от объёма внешнего знания по отношению к рассматриваемой ПрО. Эта часть знания характеризует интеллект человека, как способность самостоятельно решать новые задачи.

² Автору это известно по личному опыту. При поступлении в институт после трёх лет службы в Советской Армии задачу на устном экзамене по математике решал в течение трёх часов. Экзаменатору-фронтовику было жаль ставить неуд абитуриенту в военной форме. Задача была решена только с третьей попытки. На вопрос экзаменатора, почему задача решена не школьному, был ответ: «Я не учился в старших классах школы, поскольку окончил Московский нефтяной техникум».

2 Измерение интеллекта человека

Сопоставление объектов по некоторому свойству требует применения количественного показателя. В этом смысле такое свойство человека, как интеллект, не является исключением. На значение показателя интеллекта влияет много факторов. К таковым, в частности, относятся: логическое мышление, пространственная память, визуализация, вербальные способности. С помощью факторного анализа в структуре интеллекта выделяется до 120 базовых факторов [13]. Очевидно, что большое число факторов затрудняет измерение уровня интеллекта.

У психологов получил применение фактор общего интеллекта g , приемлемо оценивающий элементарные познавательные способности испытуемых. Для оценки фактора общего интеллекта g предложен коэффициент интеллекта (англ. IQ — *intelligence quotient*) — количественная оценка уровня интеллекта человека (коэффициент умственного развития) [13]. Количественная оценка определяет уровень интеллекта индивида относительно уровня интеллекта среднестатистического человека. Важно подчеркнуть, что IQ оценивает мыслительные способности индивида (а не его уровень знаний).

Каждый тест IQ состоит из перечня заданий нарастающей сложности, характеризующих логическое и пространственное мышление индивида. В тесты входят логические и арифметические задания, проверка оперативной (кратковременной) памяти, ориентировка в практических ситуациях — умение самостоятельно сопоставлять, обобщать известные факты и т.п. По результатам теста подсчитывается IQ . Одним из наиболее известных тестов является тест Айзенка [14]. Более точными являются тесты Д. Векслера, Дж. Равена, Р. Амтхауэра, Р.Б. Кеттелла [15]. На данный момент не существует какого-либо единого стандарта на тесты IQ .

Чем более общего вида решается задача, тем менее конкретны её результаты. Если психолога интересует уровень интеллекта индивида, то на практике востребованы оценки компетенций применительно к конкретным областям знания. Два человека с близкими оценками по тестам IQ могут существенно различаться в способностях решать различные практические задачи. Здесь под задачей понимается любая проблема, с которой сталкивается индивид. На эти способности накладывается ещё и тип нервной системы, отражающей волевые качества, стрессоустойчивость и пр. Отсюда и следует практика применения специальных тестов на компетенции с учётом времени их решения. Такая практика применима и к системам ИИ, поскольку объектом оценивания является не общий уровень интеллекта, а реализация отдельных функций мозга.

3 Функциональность мозга

Доминирующие позиции английского языка в компьютерных науках отразились на отечественной терминологии. Многие англоязычные термины в этой области введены в русский язык в кириллоязычной записи либо в переводе, который не всегда соответствовал англоязычному оригиналу. Например, слово *integrity* в информационной безопасности было переведено на русский язык как целостность (информации). Между тем, в английском языке имеется в виду *нетронутость* информации, т.е. если субъект тронул информацию (проник к ней), ознакомился с ней в своих целях, но сохранил её целостность (ничего не пропало и не искажено), при этом информационная безопасность может быть нарушена.

Сходные проблемы были отмечены, например, при использовании в разных Про популярного англицизма «концепт», заменившего русское слово «понятие» [8]. Неправильное понимание смысла слова влечёт неправильные действия.

Функции мозга проявляются через акты мышления. Они образуют следующий базис актов мышления индивида: *воспринять* информацию, *преобразовать* её, *сохранить*, *понять*,

представить, присоединить её к имеющемуся знанию, *вспомнить* информацию, *распознать, выбрать* объект, *вообразить, сообразить, рассуждать, предвидеть, править*.

Восприятие мозгом внешней информации осуществляется через органы чувств, а внутренней информации (о состоянии организма) – через нервную систему.

Под *преобразованием* информации понимается как приведение её к форме, удобной для передачи и хранения, так и получение новой информации на основе имеющейся.

Сохранение информации означает её запоминание. Краткосрочное хранение информации осуществляется в оперативной (кратковременной) памяти. Психологи оценивают её объём у среднестатистического человека в 7 ± 2 слова [16].

Понять информацию означает найти её связь с имеющимся знанием, т.е. выразить мысль через известные субъекту понятия [8]. Понимание речи (текста) заключается в расшифровывании сообщения [17, с.105]. Источник сообщения шифрует (кодирует) свою мысль некоторыми словами, погружаемыми в выбранную им грамматическую форму, а получатель сообщения расшифровывает (перекодирует) сообщение в те слова и структуры, которые соответствуют его представлению передаваемой ему мысли [18].

Для себя и для других полученная информация *представляется* в форме речевого акта или изображения. В более широком смысле слово *представить* означает *представить* себя (рефлексия) или другого (другое).

Присоединить полученную информацию к имеющемуся знанию означает сделать единым (целостным) знание с его приращением и в таком виде сохранить его в долговременной памяти.

Вспомнить известную информацию означает извлечь её из долговременной памяти. Удобным способом нахождения нужной информации является использование ассоциации, как связи между отдельными представлениями, при которой одно из них вызывает другое.

Слово *распознать* имеет общий корень со словом «знание». Нельзя распознать то, чего не знаешь. Следовательно, объект уже познан и (его образ) хранится в памяти индивида. В задачу распознавания входит выделение объекта по некоторой совокупности признаков среди других известных или похожих на него объектов.

Под *выбором* понимается нахождение предпочтительного объекта из известного списка, либо класса (образца), которому он принадлежит.

Воображение определяется как «способность человека к спонтанному созданию или преднамеренному построению образов, представлений, идей объектов, которые в пережитом опыте воображающего в целостном виде ранее не воспринимались или же вообще не могут быть восприняты посредством органов чувств» [19]. Именно такое свойство мозга, как воображение, является источником создания искусственных (виртуальных) миров. Озарение (*insight*) является одной из разновидностей воображения.

Слово *сообразение*, кроме традиционного значения «обдумывание», обозначает и более широкий смысл. Он обусловлен наличием приставки «со» к корню «образ»³. Приставка «со» в русском языке обозначает в существительных и прилагательных наличие некоторой взаимосвязи и совместного действия нескольких объектов. Отсюда сообразить – значит сопоставить образы, а в более широком смысле – образовать коалицию. Это значение слова имеет прямое отношение к коллективному поведению индивидов и подразумевает их *согласие* на выполнение совместных действий. Слово *предвидеть* не требует особых пояснений.

Рассуждать означает строить умозаключения на основе исходной информации. В науке исходной информацией для умозаключений служат аксиомы – утверждения, которые при-

³ Образ является общим корнем слов «воображение» и «сообразение» (с заменой буквы з на ж при спряжении слов). В образуемое им словообразовательное гнездо входят также слова «образование», «преобразование» и пр. Все эти слова имеют прямое отношение к познавательному процессу, ибо образ, как изображение, является универсальным языком общения.

нимаются за истину, чтобы служить предпосылкой для дальнейших рассуждений и вывода заключения. Рассуждение включает такие операции, как *сопоставление* разных суждений, *вывод* следствия на основе обобщения известных фактов (от частного к общему) и конкретизации общих закономерностей (от общего к частному).

Править означает формировать сигналы на исполнительный орган для изменения режима его функционирования или направления движения.

На рисунке 1 перечисленные акты мышления сгруппированы по следующим признакам.

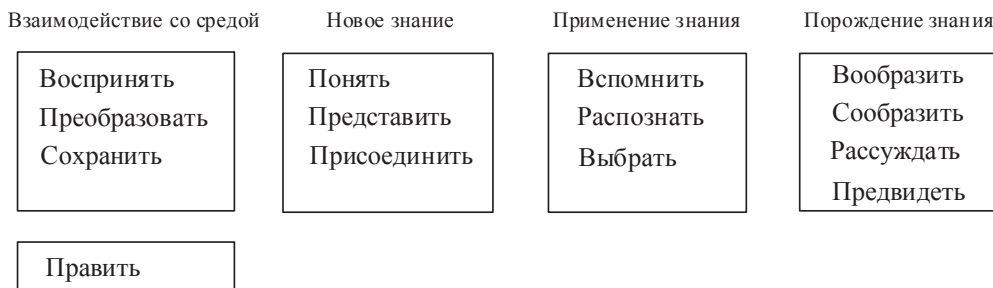


Рисунок 1 – Группы актов мышления

В левую часть рисунка вошли акты мышления, отражающие взаимодействие мозга с внешней средой. Первые три действия относятся к органам чувств и не обязательно связаны с мозгом. С точки зрения обработки информации названные функции соответствуют подготовке данных для обработки, т.е. являются предварительной ступенью познания. Действие *править* использует результаты обработки информации для воздействия на внешнюю или внутреннюю (управление организмом) среду, т.е. то, что в технике называется *управлением*.

Акты *понять*, *представить* и *присоединить* отражают процесс получения нового знания. Познание — совокупность процессов, процедур и методов **приобретения знаний** о явлениях и закономерностях *объективного мира*. Буквально, по-знание – это получение нового знания.

Действия *вспомнить* информацию, *распознать* и *выбрать* объект представляют собой варианты использования имеющегося знания.

Действия *вообразить*, *сообразить*, *рассуждать* и *предвидеть* используют имеющееся знание для *творчества* (получения оригинального знания). Из них действие *вообразить*, т.е. найти образ (войти в него) соответствует озарению. Эти действия следует отнести к варианту познания, заключающегося в приобретении нового знания путём вторичной обработки информации, как высшей ступени познания.

Акты мышления, вошедшие в предложенный базис, с одной стороны не являются элементарными, а с другой стороны могут использоваться в различных сочетаниях для выполнения более сложных познавательных процессов. Например, процесс *объяснить* (сделать ясным смысл) может включать в себя такие акты как *рассуждать* (делать вывод), *преобразовать* и *представить* – изобразить в доступной форме. Процесс *обучения* охватывает первые шесть действий, начиная с восприятия информации.

4 Информатика и ИИ

Назначением мозга является переработка информации [20]. Наука о методах и процессах сбора, хранения, обработки, передачи, анализа и оценки информации с применением компьютерных технологий, обеспечивающих возможность её использования для принятия решений, получила название информатика [21]. Информатика охватывает теоретические дисциплины (такие как теория алгоритмов, теория вычислений, теория информации и автоматиза-

ция) и практические дисциплины (включая проектирование и внедрение аппаратного и программного обеспечения СОИ).

Реализация искусственного мышления не имеет чётких границ между программной и аппаратной составляющими. Граница между этими составляющими мышления может сдвигаться как в одну, так и в другую сторону. Соотношение роли процессов и их носителя нашло отражение и в названии информатики, как науки. В англоязычном названии *computer science* (компьютерная наука) первичен компьютер как средство вычислений. В франко-немецком термине (фр. *informatique*; нем. *informatik*) первичны процессы обработки информации.

Компьютер используется не только для решения вычислительных задач, но и для задач классификации, обучения, распознавания, диагностики, управления, прогнозирования, принятия решений, машинного перевода, обработки текстов на естественном языке, анализа изображений, поиска зависимостей, кластерного анализа и т.д. [22, 23]. Исследования в этой области оформились в научное направление «ИИ». Из многих его определений можно привести описательное определение, связывающее ИИ с ЕИ: *ИИ – это раздел информатики, в рамках которого ставятся и решаются задачи моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными*. Возникновение ИИ связано с кибернетикой и ведёт свой отсчёт от Дартмутской Конференции (1956) [24].

Научное направление «ИИ» решает два класса задач: моделирование естественного мышления и применение принципов функционирования живого организма для решения задач математики. Это подтверждает терминология ИИ, опубликованная в словаре [25].

5 Моделирование мышления

Мозг человека представляет собой большую нейронную сеть. Она постоянно анализирует сигналы органов чувств и внутренних органов тела, накопленную ранее информацию, строит и корректирует модели окружающего мира и делает прогнозы на основе этих моделей. При взаимодействии мозга с органами чувств он работает в режиме преобразователя информации.

Способность мозга работать в двух режимах – как *преобразователь* и как *генератор* информации – позволяет отделить автономную работу мозга от его взаимодействия с внешней средой через органы чувств и с организмом через нервную систему. Органы чувств служат для накопления информации о внешней и внутренней среде в период обучения индивида и для решения задач в реальном времени по мере поступления информации, подлежащей обработке. Такое разделение функций мозга позволяет изучать мышление как самостоятельный процесс, хотя бы временно не зависящий от воздействия внешней среды.

Моделирование в совокупности всех актов естественного мышления, по крайней мере, на данном этапе развития науки и техники представляется невозможным. Мозг человека – универсальная ИС, и именно как система сложен для понимания и моделирования. Поэтому распространённый путь исследования – специализация ИС, моделирующих отдельные акты мышления и их совокупности. В моделях ИИ роль мозга, как универсального средства обработки информации, выполняет компьютер, либо сеть компьютеров.

Трудности моделирования заключаются в отсутствии ясной картины функционирования мозга. По энцефалограмме мозга можно только судить об активности нейронов различных участков мозга при решении задач. Частота и количество активных зон характеризует интенсивность процесса мышления. Психологи отмечают максимальную интенсивность мышления у переводчиков [26]. Во время синхронного перевода мозг *параллельно* выполняет четыре операции: *воспринимает* и *обрабатывает* текущий фрагмент сообщения на исходном

языке, *удерживает* в памяти предыдущие части и *порождает* эквивалентное сообщение на языке перевода. Однако при этом неизвестны процессы обработки сигналов, реализующие эти операции. По этим причинам моделирование процессов мышления выполняется на функциональном уровне.

Функциональная модель (Ф-модель) объекта моделирования не различает его функции и реализующие их действия [27]. Это различие проявляется при раскрытии Ф-модели как чёрного ящика. Модель акта мышления конкретизируется структурно-функциональной моделью (СФ-моделью) на принятом уровне общности. Выполняемая мозгом мысль представляется структурно-операционной моделью (СО-моделью). СФ- и СО- модели – это то, что в программировании называют структурой данных и методом соответственно. Поскольку оригинальные СФ- и СО-модели мозга и мышления неизвестны, создаются математические и имитационные модели различных функций мозга.

Моделирование восприятия информации в статье рассматривается в двух аспектах – восприятие внутренней и внешней информации. К внутренней информации относятся сигналы периферической нервной системы индивида. С учётом применения общего языка нейронов для передачи информации внутри организма правомерна аналогия восприятия внутренней информации с процессами сбора и обработки данных в компьютере. К подсистемам СОИ также следует отнести базы данных.

Примером моделирования органов чувств является персептрон, предложенный в 1958 году [28]. Разработанные на основе моделей персептрона и нейрона нейросетевые модели решают задачи классификации и распознавания образов.

Нейросетевые модели являются примером имитационного подхода к решению задачи распознавания. Помимо имитационного подхода задачи распознавания решаются путём классификации образов с применением алгебраических и логических моделей границ между смежными классами, а также с применением метрик таксономии.

В основе формализации понимания лежит *принцип подстановки* – неизвестное выражается через известное. Моделирование систем понятий осуществляется в рамках онтологических моделей. Моделирование понимания осуществляется с применением иерархических семантических сетей [8]. Для моделирования знания используются онтологии и базы знаний. Потребителю знание предоставляется в знаковой и/или образной форме, что предполагает обработку естественного языка и изображений. Вспоминание фактов предполагает применение ассоциативного поиска нужного знания.

Функция *воображения* моделируется моделями вымышленных (виртуальных) миров.

Функция *соображения*, как коллективного действия, моделируется многоагентной системой [29].

Для моделирования функции *предвидения* наряду с прогностическими моделями классической математики применяются, в частности, модели машинного обучения.

Для *объяснения* результатов, получаемых системой ИИ, разрабатываются подсистемы объяснения с применением логического вывода и когнитивной графики.

Характерной особенностью моделей ИИ является моделирование *неопределённости*, свойственной человеческому мышлению и поведению. В отличие от вероятности, как частотной характеристики событий, неопределённость человеческого мнения и поведения определяется экспертным способом (групповой экспертизой). Для решения соответствующих задач применяется, в том числе, нечёткая логика. Она используется и в моделировании рассуждений с оценкой степени истинности заключения. Экспертный характер имеют и модели многомерного оценивания ПрО. Качество моделей многомерного оценивания, как и экспертных систем, зависит от квалификации их разработчиков, что определяет степень доверия к ним.

Перечисленные модели и методы ИИ не претендуют на полноту и приведены как примеры моделирования познавательных функций мозга моделями ИИ.

Несмотря на родство некоторых предложенных слов относительно словообразующего гнезда [30] (например, с корнем образ), они отражают независимые действия. Например, слова *вообразить* и *сообразить* значительно различаются по сути выполняемых действий. В первом случае индивид конструирует искусственный мир, а во втором – решает задачу группового взаимодействия, сообразуя различные сущности.

Независимость базисных функций мозга означает невыводимость одной функции из другой. Это свойство доказывается попарным сопоставлением значений слов, обозначающих функции мозга. Таким же образом доказывается непротиворечивость системы. Ни одна из познавательных функций мозга не исключает другую.

Неизбыточность базиса следует из возможности трактовки слов в самом широком смысле. Это позволяет выводить из них вспомогательные, а, следовательно, зависимые действия мозга. В качестве примера можно рассмотреть распознавание. *Распознать* ситуацию, значит *сопоставить* её со всеми известными ситуациями, для того, чтобы *найти* ближайшую к ней. Подобным образом понятие *рассуждать* раскладывается на такие операции, как *сравнивать*, *обобщать*, *конкретизировать*, находящиеся к нему в отношении «часть-целое». Включение их в базис познавательных функций влечёт избыточность базиса и взаимозависимость познавательных функций.

6 Оценивание уровня интеллектуальности ИС

В состав современного сложного изделия входит ИС, предназначенная для обработки информации и управления исполнительными объектами. Разновидности ИС следует рассматривать как специализированные подсистемы СОИ. Любая из них нуждается в связях с внешней средой, в хранении и передаче данных. Специализация СИИ заключается в моделировании функций мозга. Следовательно, признаками, присущими интеллектуальной составляющей ИС, являются моделируемые ею познавательные функции мозга.

В целом СОИ можно именовать интеллектуальной, если она включает хотя бы одну подсистему ИИ. Для выделения интеллектуальной составляющей ИС можно сформулировать следующие принципы:

- поскольку назначением мозга является обработка информации, за его искусственный аналог принимается СОИ;
- ИС любого назначения является разновидностью СОИ;
- под интеллектуальной составляющей ИС следует понимать её подсистему, моделирующую одну или несколько познавательных функций мозга.

Обработка информации, не связанная с получением новой информации, не относится к познавательным действиям. Это же справедливо и для функции управления (править). Сложнее определить, все ли СОИ, назначением которых является получение новой информации, можно отнести к СИИ.

Например, назначением автопилота является ведение самолёта по заданному маршруту. Для решения этой задачи востребована познавательная функция *распознавание* ситуации. Необходимо обнаружить *расхождение* с заданным курсом (высотой и направлением полёта), при наличии расхождения определить его величину и сформировать управляющее воздействие на соответствующий исполнительный орган (рулевые поверхности). Автопилот, как программно-аппаратный комплекс, моделирует познавательную функцию мозга, используя модель автоматического регулирования, применяемую в различных ПрО. СИИ отличаются

от СОИ по типу используемых моделей. СОИ используют модели, отражающие физические законы. Автопилот и подобные ему системы не относят к СИИ [4].

В отличие от математических моделей, отражающих физические законы, модели ИИ имитируют функцию мозга. Имитационный характер моделей ИИ особенно проявляется в моделях с обучением. Выделение интеллектуальной составляющей ИС представляет собой нетривиальную процедуру. Для её выполнения следует привлекать системный анализ (СА) ИС. На первом этапе необходимо отделить подсистему обработки информации, не связанную с получением новой информации. К таким подсистемам относятся подсистемы сбора, передачи и хранения данных. На следующем этапе следует отделить подсистемы, реализующие математические модели, отражающие физические законы, от подсистем, имитирующих познавательные функции мозга.

Применение СА интеллектуальной составляющей ИС можно показать на примере беспилотного летательного аппарата (БПЛА). Одной из основных его функций является наблюдение за местностью. В нём видеокамера выполняет функцию человеческого зрения. Однако в СОИ она является источником данных для передачи их оператору и не выполняет интеллектуальных функций по анализу изображений. Сигналы, поступающие от видеокамеры, передаются оператору БПЛА и преобразуются в форму изображения на экране планшета.

Оператора интересует не вся картинка на планшете, а только её информативные участки. Например, для БПЛА военного назначения к ним могут относиться места расположения противника. Оператор должен обнаружить такие места на карте и дать сигнал БПЛА для выполнения очередной операции.

В терминологии ИИ обнаружение информативных признаков на карте называется обработкой изображений. Она связана с моделированием таких функций мозга, как *понимание* изображения и *распознавание* его частей. Моделирование этих познавательных функций требует привлечения дополнительных знаний о закономерностях, с применением которых выделяются информативные участки. Оно возлагается на подсистему обработки изображений. Подсказка в виде указания на информативные участки облегчает работу оператора по принятию решений.

Следующим шагом по повышению уровня интеллектуальности ИС является *выбор* очередности участков, предлагаемых вниманию оператора. Выполнение этой функции требует разработки модуля поддержки принятия решений, как подсистемы ИИ. В том случае, если выбор варианта будет не рекомендательным, а подлежащим исполнению, весь цикл операций по обработке изображений переносится на борт БПЛА.

Количественная оценка интеллектуальности ИС может быть рассчитана как отношение ресурсов, затраченных на реализацию СИИ, к ресурсам, выделенным на реализацию всей ИС. Для решения этой задачи необходимо определить виды ресурсов и единицы их измерения.

Достижения в области микроэлектроники позволяют реализовать модели и методы ИИ такие, например, как нейрокомпьютер, в аппаратном исполнении. Однако общим случаем является программно-аппаратная реализация СОИ. Учитывая это, требуется рассчитать доли аппаратной и программной составляющих интеллектуальной ИС. К аппаратной составляющей СИИ следует относить ту её часть, которая специализирована на реализации моделируемой функции мозга.

Уровень интеллектуальности ИС характеризуется двумя показателями – аппаратным и программным. Каждый из них оценивается как отношение соответствующего ресурса (аппаратного / программного), выделенного на реализацию функции мозга к общему ресурсу ИС. Объем аппаратуры может измеряться, например, в количестве элементов, а объем программ – в килобайтах. В случае реализации нескольких функций мозга в одной ИС показатель её

интеллектуальности рассчитывается как среднее арифметическое отдельных показателей. Аппаратный и программный показатели интеллектуальности ИС могут быть включены в её обозначение. Например, ИС, в которой на реализацию функции мозга затрачено 30% аппаратных и 40% программных ресурсов, может быть обозначена как ИИС-30/40 (интеллектуальная ИС с мерой 30/40).

Анализ функций БПЛА даёт необходимую информацию для расчёта показателя интеллектуальности его ИС. Общий объём аппаратуры и программного обеспечения БПЛА, реализующего все функции наблюдения, известен. В том случае, когда все операции по анализу изображений и управлению БПЛА выполняются вручную, ИС БПЛА не является интеллектуальной. В том случае, когда автоматизировано выделение информативных участков на карте, показатель интеллектуальности ИС определяется как отношение объёма затраченных на обработку изображений аппаратных и программных ресурсов к их общим объёмам в ИС. Если к функции обработки изображений добавляется автоматизация выбора, в числители соответствующих дробей добавляются объёмы ресурсов, востребованных подсистемой выбора. Уровень интеллектуальности ИС при этом возрастает.

Аналогичным образом оценивается соотношение ресурсов, выделяемых на другие информационные и интеллектуальные функции БПЛА, в частности, на функционирование органов «слуха» и «осознания», реагирующих соответственно на ультразвуковое и тепловое облучение БПЛА средствами противника.

Для работы автоматической системы управления требуется только два сигнала – пуск и останов. Примером автоматической системы управления, основанной на строгих математических моделях, является система автоматического регулирования параметров. В силу высокой степени неопределённости выбор на конечном множестве альтернатив и распределённое управление возможно с применением моделей ИИ. Задача автоматического выбора альтернативы сводится к переходу от системы поддержки принятия решений к системе принятия решений. Она может быть решена средствами ИИ для штатных ситуаций функционирования объекта. В нештатных ситуациях приходится прибегать к помощи человека.

Автоматизация распределённого управления опирается на достижения теории многоагентных систем. Коллективный разум – это общий или групповой интеллект, который возникает в результате сотрудничества, коллективных усилий и конкуренции многих людей и проявляется в принятии решений на основе согласия (консенсуса) [31]. Здесь также существует возможность автономного функционирования многоагентной системы в штатных ситуациях. Уровень интеллектуальности управления можно определить отношением штатных ситуаций ко всем возможным ситуациям. Трудность заключается в определении перечня нештатных ситуаций. Она преодолима, если под нештатными ситуациями понимать те, автоматизация которых невозможна или нецелесообразна по экономическим причинам. В этом случае они известны, и названное отношение может быть вычислено.

Заключение

Предложенный в работе базис познавательных функций мозга послужил основой для систематизации операций со знанием. Новое знание различается по способам его приобретения – из внешних источников и путём самостоятельного вывода. Действия по применению знания выделены в отдельную группу. Действиям мозга по взаимодействию с внешней средой поставлена в соответствие подсистема сбора, передачи и хранения данных, входящая в состав любой СОИ.

Трудность в выделении интеллектуальной составляющей ИС заключается в том, что, по крайней мере, некоторые функции мозга могут моделироваться математическими моделями,

отражающими физические законы. Поэтому за признак, выделяющий модели ИИ, принят имитационный подход к их созданию. Имитационное моделирование позволяет формализовать неопределённость, присущую мышлению, с применением экспертных методов, нечёткой логики и машинного обучения.

Для получения количественной оценки интеллектуальной составляющей ИС необходима информация о способах реализации познавательных функций и выделенных на них ресурсов. Эта информация известна создателям изделий. Предложенный подход может быть применён в организациях-разработчиках ИС.

Список источников

- [1] Национальная стратегия развития искусственного интеллекта на период до 2030 года. Указ Президента РФ от 10.10.2019 N 490 "О развитии искусственного интеллекта в Российской Федерации".
- [2] **Боргест Н.М.** Стратегии интеллекта и его онтологии: попытка разобраться // Онтология проектирования. 2019. Т.9, №4(34). С.407-428. DOI: 10.18287/2223-9537-2019-9-4-407-428.
- [3] **Душкин Р.В.** Искусственный интеллект. М.: ДМК Пресс, 2019. 380 с. <https://www.litres.ru/roman-dushkin/iskusstvennyy-intellekt/>.
- [4] **Самойлов Д.В.** Беспилотные комплексы с технологиями искусственного интеллекта как приоритетное направление развития вооружения и военной техники в ведущих зарубежных странах // Авиационные системы, 2022, № 8, с.2-29.
- [5] **Самойлов Д.В.** Беспилотные комплексы с технологиями искусственного интеллекта как приоритетное направление развития вооружения и военной техники в ведущих зарубежных странах (Продолжение) // Авиационные системы, 2022, № 9, с.35-56.
- [6] *Encyclopaedia Britannica*. Архивная копия от 19 октября 2014 на *Wayback Machine*.
- [7] **Colman Andrew M.** A Dictionary of Psychology (3rd ed.). Oxford [etc.]: Oxford University Press. 2008.
- [8] **Микони С.В.** Понятность онтологической модели как характеристика её качества // Онтология проектирования. 2021. Т.11, №1(39). С.20-34.
- [9] Большой толковый словарь русского языка / Гл. ред. С.А. Кузнецов. СПб.: Норинт, 1998.
- [10] APA Dictionary of Psychology by Gary R. VandenBos (Ed.) Washington, DC: American Psychological Association, 2007. 1024 p. ISBN 978-1-59147-380-0.
- [11] **Успенский В.А.** Теорема Гёделя о неполноте в элементарном изложении // Успехи математических наук. 1974. Т. XXIX. Вып. 1 (175). С.3-46.
- [12] **Anderson J.R.** Cognitive Psychology and Its Implications. 2005. NY: Worth Publishers. 519 p.
- [13] Коэффициент интеллектуальности // Большая Советская энциклопедия (в 30 т.) / А.М. Прохоров (гл. ред.). 3-е изд. М.: Сов. энциклопедия, 1973. Т. XIII. С.306. 608 с.
- [14] **Айзенк Г.** Новые IQ тесты. – М.: ЭКСМО, 2003. 192 с. ISBN 5-04-006713-5.
- [15] **Дружинин В.Н.** Психология общих способностей. 3-е издание. СПб.: Питер. 2007. 368 с.
- [16] **Miller G.A.** The Magical Number Seven, Plus or Minus Two// The Psychological Review, 1956. V.63, p.81–97.
- [17] **Полонников Р.И.** Избранные труды в двух томах. Том 1. СПб: "Анатолия". 2013. 495 с.
- [18] **Редозубов А.Д.** Формализация смысла. Часть 3. Формирование контекстов // Онтология проектирования. – 2021. Т.11, №4(42). С.437-449. DOI:10.18287/2223-9537-2021-11-4-437-449.
- [19] **Климов Е.А.** Основы психологии: Учебник для вузов. М.: Культура и спорт, ЮНИТИ, 1997. 295 с.
- [20] **Линдсей П., Норман Д.** Переработка информации у человека. Перевод с англ. по ред. Лурия А.Р. М.: Мир, 1974. 550 с.
- [21] **Михайлов А.И., Черный А.И., Гиляревский Р.С.** Основы информатики. 2-е изд., перераб. и доп. М.: Наука, 1968.
- [22] **Бир С.** Кибернетика и менеджмент: Пер. с англ. / Под ред. А.Б. Челюсткина. Изд. 2-е. М.: Комкнига. 2006. 280 с.
- [23] **Чень Ч., Ли Р.** Математическая логика и автоматическое доказательство теорем. М.: Наука. 1983. 360 с.
- [24] **Moore J.** The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. DOI:10.1609/aimag.v27i4.1911.
- [25] Толковый словарь по искусственному интеллекту (Авторы-составители А.Н. Аверкин, М.Г. Гаазе-Рапопорт, Д.А. Поспелов). Москва, Радио и связь, 1992. <http://www.raai.org/library/tolk/aivocpred.html>.
- [26] **Gile D.** Variability in the perception of fidelity in simultaneous interpretation. *Hermes* 1999. 22. P.51-79.
- [27] **Микони С.В., Соколов Б.В., Юсупов Р.М.** Квалиметрия моделей и полимодельных комплексов. М.: РАН. 2018. 314 с.

-
- [28] **Розенблатт Ф.** Перцептрон: вероятностная модель хранения и организации информации в мозге. *Корнельская авиационная лаборатория, Психологическое обозрение*, 1958. Т.65, №6, С.386-408. DOI:10.1037/h0042519.
- [29] **Городецкий В.И., Грушинский М.С., Хабалов А.В.** Многоагентные системы. Обзор. <https://textarchive.ru/c-1842736-p6.html>.
- [30] **Тихонов А.Н.** Новый словообразовательный словарь русского языка для всех, кто хочет быть грамотным. – М.: АСТ, 2014. 639 с.
- [31] **Скобелев П.О.** Ситуационное управление и мультиагентные технологии: коллективный поиск согласованных решений в диалоге // *Онтология проектирования*. 2013. №2. С.26-48.
-

Сведения об авторе



Микони Станислав Витальевич, 1936 г. рождения. Окончил Ленинградский институт инженеров железнодорожного транспорта им. Образцова в 1963 г., д.т.н. (1992), профессор (1994), ведущий научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН. Член Российской ассоциации искусственного интеллекта (1998). В списке публикаций 350 работ, из них 2 монографии и 7 учебных пособий в области технической диагностики, дискретной математики, системного анализа, теории принятия решений, искусственного интеллекта. AuthorID (РИНЦ): 100261; Author ID (Scopus): 57192370467; Researcher ID (WoS): W-3236-2019; <https://orcid.org/0000-0001-7153-6804>. smikoni@mail.ru.

Поступила в редакцию 21.01.2023, после рецензирования 02.02.2023. Принята к публикации 25.02.2023.



Approach to assessing the level of intelligence of an information system

© 2023, S.V. Mikoni

St. Petersburg Federal Research Center of the Russian Academy of Sciences

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg, Russia

Abstract

Based on the analysis of the definitions of natural and artificial intelligence (AI), basic functions are proposed that reflect the cognitive activity of the brain. Their interpretation in a broad sense, using the rules of Russian grammar, made it possible to describe the entire range of concepts used in AI. The possibility of expressing basic functions in terms of simpler ones and combining them into more complex ones is shown. The completeness, consistency and non-redundancy of the basis of the cognitive functions of the brain are substantiated. The 14 basic functions are divided into 4 groups, reflecting different aspects of working with knowledge. This made it possible to consider the AI system as a specialized information processing system. For a sign that distinguishes the AI subsystem in the information system, a simulation approach to modeling the functions of the human brain is adopted. The level of intelligence of an information system is characterized by indicators of the share of hardware and software required for the implementation of AI models in the total amount of hardware and software of the information system. An example of assessing the level of intelligence of an information system of an unmanned aerial vehicle is given.

Key words: natural intelligence, artificial intelligence, mind, intelligence, level of intelligence, information technology, information system.

Citation: Mikoni SV. Approach to assessing the level of intelligence of an information system [In Russian]. *Ontology of designing*. 2023; 13(1): 29-43. DOI: 10.18287/2223-9537-2023-13-1-29-43.

Financial support: Research carried out on this topic was carried out under budget topic FFZF-2022-0004.

Conflict of interest: The author declares no conflict of interest.

List of figures

Figure 1 - Groups of acts of thinking

References

- [1] National strategy for the development of artificial intelligence for the period up to 2030. Decree of the President of the Russian Federation of October 10, 2019 N 490 "On the development of artificial intelligence in the Russian Federation" [In Russian].
- [2] **Borgest NM.** Strategies of intelligence and its ontology: an attempt to understand [In Russian]. *Ontology of designing*. 2019; 9(4): 407-428. DOI: 10.18287/2223-9537-2019-9-4-407-428.
- [3] **Dushkin RV.** Artificial Intelligence [In Russian]. Moscow: DMK Press, 2019. 380 p. <https://www.litres.ru/romandushkin/iskusstvennyy-intellekt/>.
- [4] **Samoilov DV.** Unmanned systems with artificial intelligence technologies as a priority direction for the development of weapons and military equipment in leading foreign countries [In Russian]. *Aviation systems*. 2022; 8: 2-29.
- [5] **Samoilov DV.** Unmanned systems with artificial intelligence technologies as a priority direction for the development of weapons and military equipment in leading foreign countries (Continuation) [In Russian]. *Aviation systems*. 2022; 9: 35-56.
- [6] Encyclopaedia Britannica Archived October 19, 2014 at the Wayback Machine.
- [7] **Colman Andrew M.** A Dictionary of Psychology (3rd ed.). 2008. Oxford [etc.]: Oxford University Press. ISBN 9780191726828.
- [8] **Mikoni SV.** Comprehensibility of an ontological model as a characteristic of its quality [In Russian]. *Ontology of designing*. 2021; 11(1): 20-34. DOI: 10.18287/2223-9537-2021-11-1-20-34.

- [9] Big explanatory dictionary of the Russian language [In Russian]. Ch. ed. S. A. Kuznetsov. - St. Petersburg: Nor-int, 1998.
- [10] **Gary R.** VandenBos (ed.), APA Dictionary of Psychology (Washington, DC: American Psychological Association, 2007, ISBN 978-1591473800).
- [11] **Uspensky VA.** Gödel's incompleteness theorem in an elementary presentation [In Russian]. *Uspekhi matematicheskikh nauk.* 1974. T. XXIX. Issue. 1 (175). pp. 3-46.
- [12] **Anderson JR.** Cognitive Psychology and Its Implications. 2005. NY: Worth Publishers. 519 p.
- [13] The coefficient of intelligence [In Russian]. Great Soviet Encyclopedia (in 30 volumes) / A. M. Prokhorov (editor-in-chief). 3rd ed. Moscow: Sov. Encyclopedia, 1973. T. XIII. P.306. 608 p.
- [14] **Eysenck G.** New IQ tests. Moscow: EKSMO, 2003. 192 p. ISBN 5-04-006713-5.
- [15] **Druzhinin VN.** Psychology of general abilities. 3rd edition. St. Petersburg: Peter. 2007. 368 p.
- [16] **Miller GA.** The Magical Number Seven, Plus or Minus Two. *The Psychological Review*, 1956; 63: 81—97.
- [17] **Polonnikov RI.** Selected works in two volumes [In Russian]. Volume 1. St. Petersburg ITMO University. Anatolia. 2013. 495 p.
- [18] **Redozubov AD.** Formalization of the meaning. Part 3. Formation of contexts [In Russian]. *Ontology of Designing.* 2021; 11(4): 437-449. DOI: 10.18287/2223-9537-2021-11-4-437-449.
- [19] **Klimov EA.** Fundamentals of Psychology: A textbook for universities [In Russian]. Moscow: Culture and sport, UNITI, 1997. 295 p.
- [20] **Lindsay P, Norman D.** Information processing in humans [In Russian]. Translation from English. by red. Luria A.R. Moscow: Mir, 1974. 550 p.
- [21] **Mikhailov AI, Cherny AI, Gilyarevsky RS.** Fundamentals of Informatics [In Russian]. 2nd ed., revised. and additional. Moscow: Nauka, 1968.
- [22] **Beer S.** Cybernetics and Management [In Russian]. Per. from English. / Ed. A.B. Chelyustkina. Ed. 2nd. [In Russian]. Moscow: Komkniga. 2006. 280 p.
- [23] **Chen Ch, Li R.** Mathematical logic and automatic theorem proving [In Russian]. Moscow. Nauka. 1983. 360 p.
- [24] **Moor J.** The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years [In Russian]. DOI:10.1609/aimag.v27i4.1911
- [25] Explanatory Dictionary of Artificial Intelligence [In Russian]. Authors-compilers A.N. Averkin, M.G. Gaaze-Rapoport, D.A. Pospelov. Moscow, Radio and communication, 1992. <http://www.raai.org/library/tolk/aivopred.html>.
- [26] **Gile D.** Variability in the perception of fidelity in simultaneous interpretation. *Hermes* 1999. 22. P.51-79.
- [27] **Mikoni SV, Sokolov BV, Yusupov RM.** Qualimetry of models and multi-model complexes [In Russian]. Moscow: RAN, 2018. 314 p.
- [28] **Rosenblatt F.** Perceptron: a probabilistic model of information storage and organization in the brain [In Russian]. Cornell Aviation Laboratory, *Psychological Review*, 1958; 65(6): 386-408. DOI: 10.1037/h0042519.
- [29] **Gorodetsky VI, Grushinsky MS, Khabalov AV.** Multi-agent systems. Review [In Russian]. <https://textarchive.ru/c-1842736-p6.html>.
- [30] **Tikhonov AN.** A new word-building dictionary of the Russian language for everyone who wants to be literate. Moscow: AST, 2014. 639 p. ISBN 978 5 17 082826 5.
- [31] **Skobelev PO.** Situation-driven decision making and multi-agent technology: finding solutions in dialogue [In Russian]. *Ontology of Designing.* 2013; 3(2): 26-48.

About the author

Stanislav Vitalievich Mikoni (b. 1936) graduated from the Obraztsov Institute of Engineers of Railway Transport (Leningrad) in 1963, D. Sc. Eng. (1992), Professor (1994), Leading Researcher of the St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences. He is a member of the Russian Association of Artificial Intelligence (1998). He is the author and a co-author of 350 publications in the field of technical diagnostic, discrete mathematic, system analyses, artificial intelligence, and decision making theory. AuthorID (РИИЦ): 100261; Author ID (Scopus): 57192370467; Researcher ID (WoS): W-3236-2019; <https://orcid.org/0000-0001-7153-6804>. smikoni@mail.ru.

Received January 21, 2023, Revised February 2, 2023. Accepted February 25, 2023.

ПРИКЛАДНЫЕ ОНТОЛОГИИ ПРОЕКТИРОВАНИЯ

УДК 004.82: 620.9

Научная статья

DOI: 10.18287/2223-9537-2023-13-1-44-54



Семантическое моделирование при построении цифровых двойников энергетических объектов и систем

© 2023, Л.В. Массель ✉, А.Г. Массель

Институт систем энергетики им. Мелентьева (ИСЭМ) СО РАН, Иркутск, Россия

Аннотация

Рассматривается задача построения цифровых двойников и умных цифровых двойников для моделирования и управления в энергосистемах. Под энергосистемой понимается совокупность энергетических ресурсов всех видов, методы их получения (добычи), преобразования, распределения и использования, а также технические средства и организационные комплексы, обеспечивающие снабжение потребителей различными видами энергии. Анализируются интегрированные интеллектуальные энергосистемы и основные направления цифровизации энергетики. Рассмотрены понятие «цифровые двойники» в технических сферах, онтологический подход к построению цифровых двойников и семантические модели построения умных цифровых двойников. Предложено использование фрактального подхода при выполнении онтологического инжиниринга, который даёт возможность формализации понятий предметной области и позволяет строить разномасштабные онтологии с использованием метауровней онтологий. Представлены модели цифровых двойников и умных цифровых двойников. Разработанные подходы проиллюстрированы на примерах построения цифрового двойника солнечной электростанции и умного цифрового двойника топливно-энергетического комплекса. Описанный в статье подход даёт возможность интеграции цифровых и умных цифровых двойников разных уровней в единое цифровое решение при моделировании объектов энергетики и энергосистем.

Ключевые слова: интегрированные интеллектуальные энергосистемы, цифровой двойник, фрактальный подход, онтологии, семантические модели.

Цитирование: Массель Л.В., Массель А.Г. Семантическое моделирование при построении цифровых двойников энергетических объектов и систем // Онтология проектирования. 2023. Т.13, №1(47). С.44-54. DOI:10.18287/2223-9537-2023-13-1-44-54.

Финансирование: работа выполнена в рамках проекта по госзаказу ИСЭМ СО РАН, № АААА-А21-121012090007-7, тема № FNEU-2021-0007.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

Разработка интегрированных интеллектуальных энергетических систем (ИИЭС) – одна из важных тенденций в российской энергетике. Энергетика является инфраструктурной отраслью, основная задача которой – обеспечение энергоснабжения потребителей с требуемой надёжностью и приемлемым качеством энергоносителя [1-5].

Инфраструктурная роль присуща системам электро-, газо- и теплоснабжения, имеющим развитую транспортную и распределительную сетевую инфраструктуру. Эти энергетические системы принято подразделять на производственные, транспортные и энергоснабжающие [1]. Производственные и транспортные энергетические системы интегрируются при использовании энергоносителя одной системы в другой (например, газ в качестве топлива на элек-

тростанциях и котельных, электроэнергия на газоперекачивающих агрегатах и т.д.). Эта интеграция определяет ведущую роль энергетических систем в топливно-энергетическом комплексе (ТЭК).

До недавнего времени интеграция систем электро-, тепло- и газоснабжения рассматривалась на уровне источников электроэнергии и тепла (тепловые электростанции). С появлением у потребителей возможностей выбора устройств использования энергии, в качестве обобщения этих возможностей было введено понятие «энергохаб» [2, 3]. Такая централизация требует реализации новых принципов построения ИИЭС и создания интеллектуальных систем управления (ИСУ) с развитым информационно-коммуникационным обеспечением.

В [2, 3] структура ИИЭС представлена в трёх уровнях. На первом уровне – системы электро-, газо- и теплоснабжения, на втором – уровни микро-, мини- и макросистемы; на третьем, функциональном уровне, – энергетические функции (производство, транспорт, распределение и потребление энергоресурсов), коммуникационные и управляющие функции (получение информации, её обработка, передача и представление, системы управления режимами и развития ИИЭС) и функции принятия решений (модели и методы обоснования решений по разработке ИИЭС).

Наряду с использованием методов искусственного интеллекта (ИИ) в ИИЭС всё большее внимание энергетиков привлекают так называемые «системы целеполагания» [4, 5]. Это ИСУ, имеющие несколько целей функционирования или способные генерировать эти цели, выбирать наиболее подходящую цель в зависимости от среды, прогнозировать поведение среды и собственное состояние. Для систем с неполной информацией и высокой сложностью объектов управления, таких как энергосистемы, всё более актуально использование ситуационных методов управления на основе интеллектуальных технологий [6], а также развитие семиотического подхода к построению интеллектуальных системы поддержки принятия решений [7].

Принятие в России программ «Цифровая экономика» [8] и «Цифровая энергетика» [9] потребовало цифровой трансформации энергетики. Цифровая трансформация, как переход к цифровому обществу, включает следующие направления: формирование нового варианта экономических отношений (цифровая экономика); построение нового уровня отношений общества и государства (цифровое правительство); создание высокотехнологичной инфраструктуры (цифровое пространство) [10]. Данная статья посвящена работам, проводимым в третьем направлении.

Цифровые двойники (ЦД) – важный тренд цифровизации. Технологии ЦД сегодня активно используются в различных областях техники, в том числе и в энергетике в России [11–14] и за рубежом [15–21]. В настоящее время ЦД представляет собой реальное отображение всех компонентов жизненного цикла объекта с использованием физических данных, виртуальных данных и данных о взаимодействии между ними [15]. ЦД объединяет информацию о показателях эффективности объекта и его детальную математическую модель (ММ), параметры которой уточняются с использованием реальных данных [17].

ЦД включает пять программных компонентов [22]:

- ядро ЦД – математические, имитационные и информационные модели;
- системы наблюдения, распознавания и сбора данных, системы мониторинга и управления, обеспечивающие сбор данных с физического объекта, мониторинг и управление физическим объектом (физический элемент/физическая часть ЦД);
- системы, обеспечивающие хранение собранных данных, – реляционные СУБД, СУБД с открытым исходным кодом, облачные хранилища, распределённые файловые системы и др.;

- сервисные элементы, предоставляющие сервисы и интерфейс для клиентов – инструменты для поддержки сервисов оптимизации, математического моделирования, прогнозирования и т.д.;
- платформа Интернета вещей, обеспечивающая связи между компонентами.

ЦД хорошо зарекомендовали себя на разных этапах жизненного цикла объектов и систем. Основная проблема при построении ЦД – это отсутствие единой концепции взаимосвязи и общего взаимодействия различных уровней управления. В статье предложено использовать ЦД для моделирования и интеллектуального управления энергетическими объектами и системами, которые объединяются единой концепцией.

1 Предлагаемые подходы и методы

Вариант онтологического подхода к построению ЦД предложен в [13, 14]. В этих работах в отличие от работ [23-25] рассматривается обобщённый подход к построению не только ЦД отдельных объектов, но и ЦД систем. На основе анализа [11, 12] построена обобщённая архитектура ЦД энергосистемы (подробно рассмотрена в [14]). ЦД получает данные от датчиков, измерительных и автоматизированных систем, а также из баз нормативных и справочных данных. Ядром ЦД являются математические (ММ)¹ и имитационные, информационные и онтологические модели. Взаимодействие всех типов моделей в ЦД основывается на онтологических моделях.

В соответствии с тремя уровнями управления: оперативным, тактическим и стратегическим [26] в энергетике выделяют следующие уровни: объектное управление (функционирование объектов), системное управление (функционирование систем) и управление развитием ТЭК (стратегическое управление). Наиболее сложным является стратегическое управление.

При решении задач управления функционированием энергосистем могут найти применение интеллектуальные модели, причём не только для формирования рекомендаций оперативному персоналу, принимающему решения, но и как самостоятельные элементы многоагентного управления. Интеллектуальные модели позволяют проводить необходимые расчёты в режиме реального времени и адаптировать устройства автоматики энергосистем для изменения уставок релейной защиты и противоаварийной автоматики, специального автоматического отключения нагрузки и отбора резервов мощности у потребителя и в генерирующих установках [27].

При управлении развитием ИИЭС возрастает роль прогнозирования. В [27] выделены три типа прогнозирования.

- *Количественное прогнозирование*, когда на основе ММ многосвязной системы (включая взаимосвязь энергетики и экономики, социальной и технологической сфер) при сохранении структуры и тенденций общего развития получают новые оценки и формулируют будущие параметры экономического и энергетического развития.
- *Качественное прогнозирование*, когда экспертно выбираются несколько сценариев будущего развития, и из этих сценариев выбирается тот, который наиболее соответствует представлениям заказчика и исполнителя этого прогноза.
- *Когнитивное прогнозирование*, когда у эксперта формируется целевое видение будущей системы и путей достижения этого будущего.

Для анализа тенденций в стратегическом управлении предлагается использовать семантические модели, которые включают онтологические, когнитивные, событийные и вероят-

¹ Представляется, что ММ энергетических объектов и систем, широкий спектр которых разработан в Институте систем энергетики им. Мелентьева СО РАН, могут быть использованы при построении ЦД энергетических объектов и систем.

ностные модели [6, 7]. Онтологическому инжинирингу посвящены работы [28, 29], применение когнитивных моделей описано, например, в [30, 31].

Для управления энергетическими объектами достаточно использовать традиционные ЦД. При управлении энергосистемами могут потребоваться элементы интеллектуальных технологий, а при стратегическом управлении (управлении развитием ТЭК) потребуется интеллектуальный, или умный, ЦД (УЦД). В отличие от [32], где УЦД понимается как виртуальная модель, которая может имитировать планирование на основе методологии глубокого обучения с подкреплением, предлагается расширить содержание данного понятия и использование методов машинного обучения дополнить методами работы с экспертными знаниями.

Модель ЦД (*Digital Twin – DT*) может быть представлена следующим образом:

$$DT = (IM, MM, DB, VT, SE, RC),$$

где *IM* – информационные модели; *MM* – ММ; *DB* – базы данных (БД); *VT* – средства визуализации, *SE* – необходимые сервисные компоненты; *RC* – средства связи между компонентами в режиме реального времени.

Методы ИИ и перечисленные программные компоненты, необходимые для построения ЦД, интегрируются в УЦД.

Модель УЦД (*Smart Digital Twin – SDT*) может быть представлена следующим образом:

$$SDT = (IM, MM, DB, VT, SE, RC, SM, ES, KB),$$

где *SM* – семантические модели, *ES* – экспертные системы, *KB* – базы знаний.

В свою очередь *SM = (OM, CM, EM, PM)*, где *OM* – онтологические модели, *CM* – когнитивные модели, *EM* – событийные модели, *PM* – вероятностные модели.

Построение ЦД включает шесть этапов:

- 1) анализ существующих ММ и программ для ЭВМ, их реализующих;
- 2) онтологический инжиниринг предметной области (соответствующей энергосистемы) и построение онтологических моделей;
- 3) определение исходных данных (состав, источники поступления, возможность получения оперативных данных, БД и т.д.) и характера их взаимодействия с ММ;
- 4) модификация, при необходимости, ММ и реинжиниринг программ и программных комплексов;
- 5) разработка на основе реинжиниринга программных комплексов веб-приложений и веб-сервисов для реализации ЦД;
- 6) внедрение прототипов ЦД для энергетических объектов и систем.

При разработке УЦД добавляется этап анализа интеллектуальных технологий и средств ИИ и их интеграции в УЦД.

При построении онтологий используется фрактальный подход [33], где под фрактальностью понимается возможность использования одни и те же технологии построения для онтологий разного масштаба. Графически фрактальную стратифицированную модель (ФС-модель), описывающую информационное пространство, можно представить в виде набора вложенных сферических оболочек (страт или слоёв), где каждая страта (слой) объединяет информационные модели одного типа. На практике не работают со всем информационным пространством, а только с его частью (сектором), которую можно представить, например, в виде конуса. На рисунке 1 показана ФС-модель, в которой представлены конусы, описывающие четыре типа стратификации: 1) *по типу управления* (оперативное, тактическое, стратегическое); 2) *по функциональным уровням* построения ЦД (ТЭК, энергетические системы, объекты энергетики, где степень агрегации возрастает с каждым слоем); 3) *по типам ЦД* (УЦД на уровне ТЭК, ЦД и УЦД на уровне энергосистем, ЦД на уровне объекта); 4) *по типам моделей* (модели знаний в экспертных системах, семантические модели, ММ). Используется отображение между слоями (от любого слоя к каждому) [33].

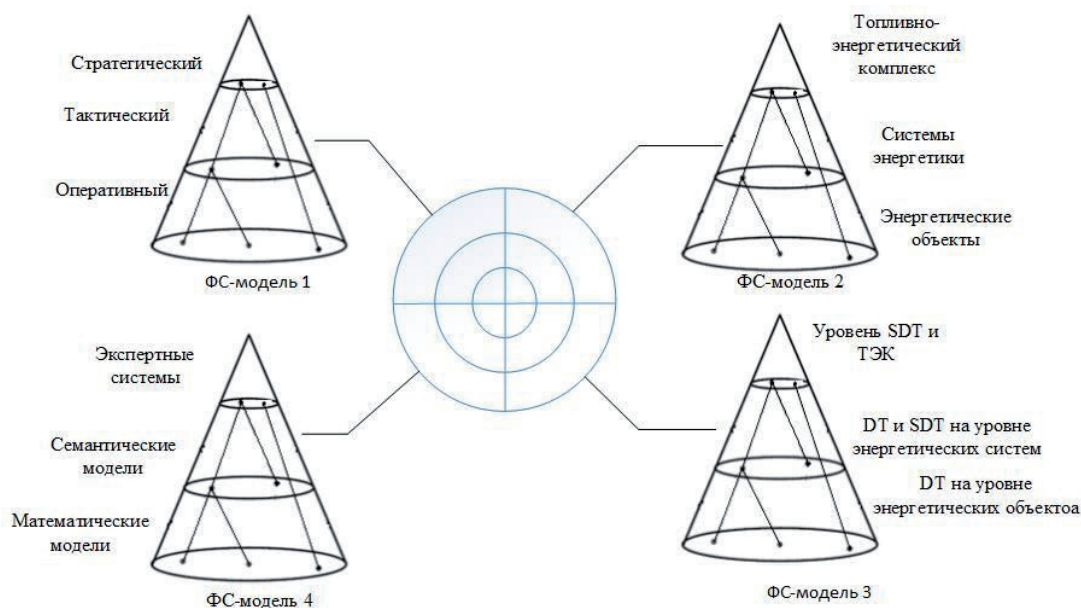


Рисунок 1 – Фрактальная стратифицированная модель, объединяющая ФС-модели для четырёх секторов

Например, для слоя типов моделей (4): $F_S^M: D_M \rightarrow D_S$ – отображение из слоя ММ в слой семантических моделей; $F_K^S: D_S \rightarrow D_K$ – отображение из слоя семантических моделей в слой моделей знаний; $F_K^M: D_M \rightarrow D_K$ – отображение из слоя ММ в слой моделей знаний. Допускаются обратные отображения, например: $F_M^S: D_S \rightarrow D_M$ – отображение из слоя семантических моделей в слой ММ. Для каждого слоя есть инструменты построения моделей разработаны программные продукты, поддерживающие онтологическое, когнитивное, событийное и вероятностное моделирование, а также соответствующие ММ. Отображения поддерживаются соответствующими инструментами. В частности, используется онтологическое пространство знаний, в котором присутствуют связи между основными терминами и их отображения в различных моделях.

2 Полученные результаты

Результаты применения предложенных подходов представлены для двух уровней управления: управления объектами энергетики и стратегического управления развитием ТЭК.

Используя фрактальный подход [33], строится иерархическая система онтологий, на верхнем уровне которой - метаонтология энергетической системы, определяющая её метаконцепты и отношения между ними, на втором уровне она детализируется. Например, метаконцепт «Электроэнергетические системы» детализируется понятиями выработки электроэнергии: гидро-, тепло- электростанции, возобновляемые источники энергии (ВИЭ) и т.д.

На рисунке 2 представлена онтология архитектуры ЦД солнечной электростанции, интегрирующая цифровую модель, цифровую тень, ММ, методы машинного обучения для прогнозирования погодных условий и др. Интеграция моделей осуществляется за счёт использования единого онтологического пространства знаний, в котором представлены отображения уровней систем, уровней управления, различных моделей и связей концептов в этих моделях. Использование онтологий на уровне проектирования ЦД позволило выявить основные информационные потоки, получаемые с объектов, формализовать основные факторы, которые необходимы для использования в ММ. На основе онтологического инжиниринга спроектированы БД, в которые поступают оперативные данные с датчиков, на основании этих данных строится цифровая модель.

Оперативные данные используются также вычислителем на основе ММ. При формировании онтологий объектов ВИЭ и их ЦД удалось сгруппировать агрегированные данные, необходимые для применения на следующих уровнях управления [34].

Для стратегического управления развитием ТЭК предлагается реализация УЦД, ядром которого является программный комплекс (ПК) INTEC-A (результат реинжиниринга ПК для исследований направлений развития ТЭК с учётом требований энергобезопасности). Описание архитектуры ПК приводится в [35].

УЦД реализуется за счёт автоматизации сбора входной информации с уровней энергетических систем и добавления средств семантического моделирования для внесения и моделирования различных угроз и мероприятий экспертами-энергетиками. Совместное использование ММ и экспертных знаний позволяет рассматривать ситуации с



Рисунок 2 – Онтология архитектуры ЦД солнечной электростанции

большой степенью неопределённости. Так, на данном этапе исследований осуществлена интеграция когнитивных моделей с ММ ТЭК. Это позволяет эксперту взаимодействовать с ММ, используя графическое представление когнитивных моделей в виде когнитивных карт с отображением концептов, характеристики которых автоматически собираются из ММ.

За счёт использования единой методологии построения ЦД и УЦД решается задача получения актуальных данных непосредственно с объектов при условии встраивания объектов в систему ЦД. На рисунке 3 отображена онтология верхнего уровня, описывающая взаимодействия ЦД и УЦД в ТЭК.

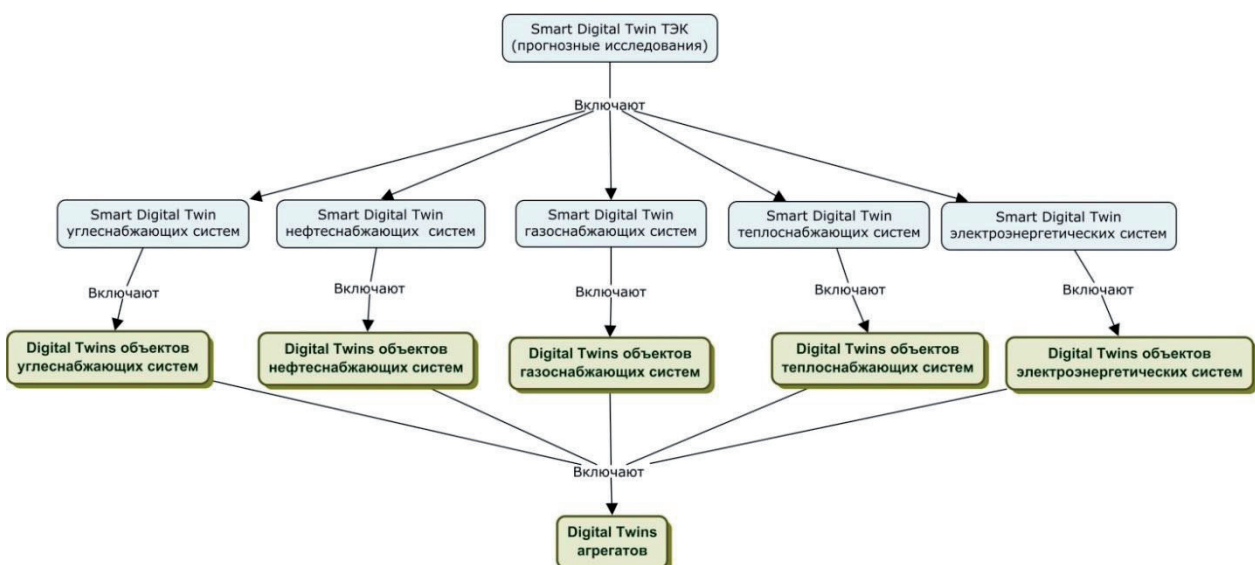


Рисунок 3 – Онтология верхнего уровня, описывающая взаимодействия ЦД и УЦД в ТЭК

Взаимодействие ЦД начинается с уровня агрегатов, которые объединяются сначала на уровне объектов энергетических систем, затем учитываются в УЦД энергетических систем, и потом агрегируются на уровне УЦД ТЭК в целом. Данный подход позволит решить проблемы с актуализацией необходимой информации.

Заключение

В статье рассмотрена концепция ИИЭС и связанная с ней проблема интеллектуального управления энергосистемами.

На основе онтологического подхода к построению ЦД в энергетике предложена обобщённая архитектура УЦД энергетических систем, включающая математические, информационные и онтологические модели. При построении ЦД в электроэнергетике планируется использовать ММ энергетических объектов и систем, которые разработаны в ИСЭМ СО РАН.

Описаны предложенные модели ЦД и УЦД. Модель ЦД включает: информационные модели, ММ, БД; инструменты визуализации, необходимые сервисные компоненты, средства связи между компонентами. Модель УЦД дополнительно включает семантические модели, базы знаний и экспертные системы.

При разработке онтологий использован фрактальный подход, позволяющий строить разномасштабные онтологии с использованием одних и тех же методов.

Предложенные подходы построения ЦД и УЦД иллюстрируются их применением при разработке ЦД солнечной электростанции и УЦД ТЭК. Предложена архитектура ЦД солнечной электростанции, объединяющая ММ и реализующую её цифровую модель, а также цифровую тень и реализующую её БД.

Список источников

- [1] *Бушуев В.В., Каменев А.С., Кобец Б.Б.* Энергетика как инфраструктурная «система систем» / Энергетическая политика. 2012. № 5. С.3-14.
- [2] *Воропай Н.И., Стенников В.А., Сендеров С.М., Барахтенко Е.А. и др.* Интегрированные инфраструктура энергетических систем регионального и межрегионального уровня / Энергетическая политика. 2015. Т.3. С.24-32.
- [3] *Воропай Н.И., Стенников В.А., Сендеров С.М.* Интегрированные интеллектуальные системы в энергетике России / Системные исследования в энергетике: методология и результаты. Под редакцией А.А. Макарова и Н.И. Воропая. М.: ИНЭИ РАН. 2018. С.87-101.
- [4] *Воропай Н.И., Стенников В.А.* Интегрированные интеллектуальные энергетические системы / Известия РАН. Энергетика. 2014. № 1. С.64-73.
- [5] *Воропай Н.И., Стенников В.А.* Интегрированные энергетические системы / Инновационная энергетика-21. Под редакцией В.М. Батенина, В.В. Бушуева, Н.И. Воропая. М.: ПК «Энергия». 2017. С.181-193.
- [6] *Массель Л.В., Массель А.Г.* Интеллектуальные вычисления в исследованиях направлений развития энергетики // Известия Томского политехнического университета. 2012. № 5. Управление, вычислительная техника и информатика. С.135-141.
- [7] *Massel L.V., Massel A.G.* Intelligent system of semiotic type for decision-making support in Russia energy sector based on situational management conception / Proceedings of IV International scientific conference "Information technologies in science, management, social sphere and medicine" (ITSMSSM 2017) // Advances in Computer Science Research (ACSR). Volume 72, 2017. P.423-429. <https://doi.org/10.1051/e3sconf/202020902019>.
- [8] «Цифровая экономика РФ». Минцифры России. Последнее обновление: 9 августа 2022. https://digital.gov.ru/ru/activity/directions/858/?utm_referrer=https%3a%2f%2fwww.google.com%2f.
- [9] Ведомственный проект «Цифровая энергетика». <https://minenergo.gov.ru/node/14559>.
- [10] *Сойфер В.А.* Human fActor // Онтология проектирования. 2021. Т.11, №1(39). С.8-19. DOI:10.18287/2223-9537-2021-11-1-8-19.
- [11] *Андрюшкевич С.К., Ковалев С.П., Нефедов Е.А.* Подходы к разработке и применению цифровых двойников энергосистем // Цифровая подстанция. 2019. № 12. С.38-43.

- [12] **Ковалев С.П.** Проектирование информационного обеспечения цифровых двойников энергосистем // Системы и средства информатики. 2020. Т.30. №1. С.66–81.
- [13] **Massel L., Massel A.** Development of Digital Twins and Digital Shadows of Energy Objects and Systems Using Scientific Tools for Energy Research // ENERGY-21 – Sustainable Development & Smart Management: proceedings. E3S Web of Conferences, 2020. Volume 209. P.1-7. DOI: 10.1051/e3sconf/202020902019.
- [14] **Воропай Н.И., Массель Л.В., Колосок И.Н., Массель А.Г.** ИТ-инфраструктура для построения интеллектуальных систем управления развитием и функционированием энергосистем на основе цифровых двойников и цифровых образов / Известия РАН. Энергетика. 2021. №1. С.3-13. DOI: 10.31857/S0002331021010180.
- [15] **Grieves M.W.** Digital Twin: Manufacturing Excellence through Virtual Factory Replication–LLC, 2014, 7 p.
- [16] **Fei Tao, Fangyuan Sui, Ang Liu, Qinglin Qi, Meng Zhang, Boyang Song, Zirong Guo, Stephen C.-Y. Lu & A. Y. C. Nee.** Digital twin-driven product design framework, International Journal of Production Research, 2019; 57:12, pp. 1-19. DOI: 10.1080/00207543.2018.1443229.
- [17] **Lim, Kendrik Yan Hong & Zheng, Pai & Chen, Chun-Hsien.** A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives. Journal of Intelligent Manufacturing. 2020. P.1313-1337. DOI: 10.1007/S10845-019-01512-W.
- [18] **Ferguson S., Bennett E., Ivashchenko A.** Digital twin tackles design challenges. World Pumps. 2017 (4). P.26–28. DOI: 10.1016/S0262-1762(17)30139-6.
- [19] **Karanjkar N., Joglekar A., Mohanty S., Prabhu V., Raghunath D., Sundaresan R.** Digital Twin for Energy Optimization in an SMT-PCB Assembly Line, IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), Bali, 2018. P.85-89. DOI:10.1109/IOTAIS.2018.8600830.
- [20] **Pileggi P., Verriet J., Broekhuijsen J., C. van Leeuwen, Wijbrandi W., Konsman M.** A Digital Twin for Cyber-Physical Energy Systems, 7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), Montreal, QC, Canada, 2019. P.1-6. DOI: 10.1109/MSCPES.2019.8738792.
- [21] **Tao F., Zhang H., Liu A., Nee A.Y.C.** Digital Twin in Industry: State-of-the-Art, in IEEE Transactions on Industrial Informatics, vol.15, no.4. P.2405-2415, April 2019. DOI: 10.1109/TII.2018.2873186.
- [22] **Прохоров А., Лысачев М.** Цифровой двойник. Анализ, тренды, мировой опыт. М.: ООО «АльянсПринт». 2020. 401 с.
- [23] **Марьясин О.Ю.** Разработка онтологий для цифрового двойника зданий // Онтология проектирования. 2019. Т.9, №4(34). С.480-495. DOI: 10.18287/2223-9537-2019-9-4-480-495.
- [24] **Stennikov V., Barakhtenko E., Sokolov D., Mayorov G.** Principles of Building Digital Twins to Design Integrated Energy Systems. *Computation*, vol.10, no.12, Dec. 2022. P.222. DOI:10.3390/COMPUTATION10120222.
- [25] **Steinmetz C., Rettberg A., Ribeiro F.G.C, Schroeder G., Pereira C.E.** Internet of things ontology for digital twin in cyber physical systems. Brazilian Symp Comput Syst Eng SBESC, vol. 2018-November, P.154–159, Jul. 2018. DOI:10.1109/SBESC.2018.00030.
- [26] **Anthony R. N.** Planning and Control: a Framework for Analysis. Cambridge MA: Harvard University Press, 1965.
- [27] **Макаров А.А.** Подходы к оценке устойчивости и рисков долгосрочного развития российской энергетики / Системные исследования в энергетике: методология и результаты. М.: ИНЭИ РАН. 2018. С.87-101.
- [28] **Euzenat J., Shvaiko P.** Ontology matching: Second edition. Springer-Verlag Berlin Heidelberg. 2013. P.511. DOI:10.1007/978-3-642-38721-0/COVER.
- [29] **Suarez-Figueroa M.C., Gomez-Perez A., Motta E., Gangemi A. (Eds.).** Ontology engineering in a networked world. Springer Science & Business Media, 2012. P.444. DOI:10.1007/978-3-642-24794-1.
- [30] **Groumpos P., Stylios C.** Modelling supervisory control systems using fuzzy cognitive maps / Chaos, Solitons & Fractals. 2000, Vol.11, №1-3. P.329-336.
- [31] **Papageorgiou E., Stylios C., Groumpos P.** An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps / IEEE Transactions on Biomedical Engineering. 2003, Vol. 50. № 12. P.1326-1339.
- [32] **Ruiz J.C.S., Bru J.M., Escoto R.P.** Smart Digital Twin for ZDM-based job-shop scheduling. 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), Rome, Italy, 2021, P.510-515, DOI: 10.1109/MetroInd4.0IoT51437.2021.9488473.
- [33] **Массель Л.В.** Фрактальный подход к структурированию знаний и примеры его применения // Онтология проектирования. 2016. Т.6, №2(20). С.149-161. DOI: 10.18287/2223-9537-2016-6-2-149-161.
- [34] **Массель Л.В., Массель А.Г., Шукин Н.И., Цыбилов А.Р., Лосев А.С.** Построение цифровых двойников ветровой и солнечной электростанций на основе онтологического подхода // Автоматизация в промышленности, 2022. № 7. С.28-32. DOI: 10.25728/avtprom.2022.07.04.
- [35] **Мамедов Т.Г., Массель А.Г.** Адаптация методики реинжиниринга унаследованного программного обеспечения // Информационные и математические технологии в науке и управлении. 2021. №4(24). С.88-99. DOI: 10.38028/ESI.2021.24.4.009.

Сведения об авторах



Массель Людмила Васильевна, 1949 г. рождения. Окончила Томский политехнический институт (1971). Д.т.н. (1995), профессор (1999). Главный научный сотрудник ИСЭМ СО РАН, профессор Института информационных технологий и анализа данных Иркутского национального технического университета. В списке научных трудов около 300 статей и глав монографий в области проектирования информационных систем и технологий, семантического моделирования, разработки систем интеллектуальной поддержки принятия решений в области энергетики. Author ID (РИНЦ): 8466; ORCID: 0000-0002-9088-9012, Author ID (Scopus): 56440157300; Researcher ID (WoS): B-4793-2014.lvmassel@isem.irk.ru

Массель Алексей Геннадьевич, 1985 г. рождения. Окончил Иркутский государственный университет в 2007 г., к.т.н. (2011). Старший научный сотрудник ИСЭМ СО РАН, доцент Института информационных технологий и анализа данных Иркутского национального технического университета. В списке научных трудов более 70 работ в области семантического моделирования, проектирования информационных систем и технологий, разработки систем интеллектуальной поддержки принятия решений в области энергетики. ORCID: 0000-0002-0351-0415; Author ID (РИНЦ): 173012; Author ID (Scopus): 57220804764. amassel@gmail.com.



Поступила в редакцию 01.11.2022, после рецензирования 26.01.2023. Принята к публикации 02.02.2023.



Scientific article

DOI: 10.18287/2223-9537-2023-13-1-44-54

Semantic modeling in the construction of digital twins of energy objects and systems

© 2023, L.V. Massel ✉, A.G. Massel

L.A. Melentiev Energy Systems Institute, Siberian Branch of the Russian Academy of Science, Irkutsk, Russia

Abstract

The article deals with the problem of building Digital Twins and Smart Digital Twins for control and management in power systems. The energy system is understood as a set of energy resources of all types, methods for their production (extraction), transformation, distribution and use, as well as technical means and organizational complexes that ensure the supply of consumers with all types of energy. Integrated intelligent energy systems are analyzed as one of the important trends in the Russian energy sector, and the main directions of digitization of the energy sector are considered. The concept of "digital twins" in technical fields is considered as one of the main digitalization trends, an ontological approach to building digital twins and semantic models for building smart digital twins are proposed. It is proposed to use a fractal approach when performing ontological engineering, which makes it possible to formalize the concepts of the subject area and allows you to build different-scale ontologies using metalevels of ontologies. Formalized models of digital twins and smart digital twins are presented. The developed approaches are illustrated by the example of construction of digital twins of a solar power plant and smart digital twins of a fuel and energy complex. The approach described in the article makes it possible to integrate different levels of digital and smart digital twins into a single digital solution when modeling energy facilities and power systems.

Key words: Integrated intelligent energy systems, digital twin, ontologies, semantic models, fractal approach.

Citation: Massel LV, Massel AG. Semantic modeling in the construction of digital twins of energy objects and systems [In Russian]. *Ontology of designing*. 2023; 13(1): 44-54. DOI: 10.18287/2223-9537-2023-13-1-44-54.

Financial Support: The results were obtained within the framework of the project under the state order of MESI SB RAS, № AAAA-A21-121012090007-7, topic № FWEU-2021-0007.

Conflict of interest: The authors declare no conflict of interest.

List of figures

Figure 1 – Fractal stratified model linking FS-models for four sectors

Figure 2 – Ontology of digital twin architecture

Figure 3 – Top-level ontology describing the interactions of DT and SDT in energy

References

- [1] **Bushuev VV, Kamenev AS, Kobec BB.** Energy as an infrastructure "system of systems" [In Russian]. *Energy policy*. 2012; 5: 3-14.
- [2] **Voropaj NI, Stennikov VA, Senderov SM, Barakhtenko EA and etc.** Integrated infrastructure energy systems of the regional and interregional level [In Russian]. *Energy policy*. 2015; 3: 24-32.
- [3] **Voropaj NI, Stennikov VA, Senderov SM.** Integrated intelligent systems in the energy sector of Russia [In Russian]. *Systems Research in Energy: Methodology and Results* Edited by Makarov A.A. and Voropaj N.I. Moscow: ERI RAS. 2018. P.87-101.
- [4] **Voropaj NI, Stennikov VA.** Integrated intelligent energy systems [In Russian]. *News of the RAS. Energy*. 2014; 1: 64-73.
- [5] **Voropaj NI, Stennikov VA.** Integrated energy systems [In Russian]. *Innovative power industry-21*. Edited by Batenin V.M., Bushuev V.V., Voropaj N.I. Moscow: PC "Energy". 2017. P.181-193.
- [6] **Massel LV, Massel AG.** Intelligent computing in studies of energy development trends [In Russian]. *Bulletin of the Tomsk Polytechnic University. Management, computer technology and informatics*. 2012; 321(5): 135-141.
- [7] **Massel LV, Massel AG.** Intelligent system of semiotic type for decision-making support in Russia energy sector based on situational management conception /Proceedings of IV International scientific conference "Information technologies in science, management, social sphere and medicine" (ITSMSSM 2017)) // *Advances in Computer Science Research (ACSR)*. 2017; 72: 423-429. DOI: 10.2991/itsmssm-17.2017.89.
- [8] Digital economy of the Russian Federation. Ministry of Digital Development of Russia. Last updated: August 9, 2022. https://digital.gov.ru/ru/activity/directions/858/?utm_referrer=https%3a%2f%2fwww.google.com%2f.
- [9] Departmental project "Digital Energy". <https://minenergo.gov.ru/node/14559>.
- [10] **Soifer VA.** Human fActor [In Russian]. *Ontology of Designing*. 2020; 11(1): 8-19. DOI: 10.18287/2223-9537-2021-11-1-8-19..
- [11] **Andryushkevich SK, Kovalev SP, Nefedov EA.** Approaches to the development and application of digital twins of energy systems [In Russian]. *Digital substation*. 2019; 12: 38-43.
- [12] **Kovalev SP.** Designing information support for digital twins of energy systems [In Russian]. *Systems and means of informatics*. 2020; 30(1): 66–81. DOI: 10.14357/08696527200106.
- [13] **Massel L, Massel A.** Development of Digital Twins and Digital Shadows of Energy Objects and Systems Using Scientific Tools for Energy Research // *ENERGY-21 – Sustainable Development & Smart Management: proceedings*. E3S Web of Conferences, 2020. Volume 209. DOI:10.1051/e3sconf/202020902019.
- [14] **Voropaj NI, Massel LV, Kolosok IN, Massel AG.** IT-infrastructure for construction of intelligent management systems of development and functioning of energy systems based on digital twins and digital images [In Russian]. *News of the Russian Academy of Sciences. Energy*. 2021; 1: 3-13. DOI:10.31857/S0002331021010180.
- [15] **Grieves MW.** Digital Twin: Manufacturing Excellence through Virtual Factory Replication–LLC, 2014, 7 p.
- [16] **Fei Tao, Fangyuan Sui, Ang Liu, Qinglin Qi, Meng Zhang, Boyang Song, Zirong Guo, Stephen C.-Y. Lu & A. Y. C. Nee.** Digital twin-driven product design framework, *International Journal of Production Research*, 2019; 57:12, pp. 1-19. DOI:10.1080/00207543.2018.1443229.
- [17] **Lim, Kendrik Yan Hong & Zheng, Pai & Chen, Chun-Hsien.** A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*. 2020. P.1313-1337. DOI:10.1007/S10845-019-01512-W.
- [18] **Ferguson S, Bennett E, Ivashchenko A.** Digital twin tackles design challenges. *World Pumps*. 2017; 4: 26–28. DOI:10.1016/S0262-1762(17)30139-6.
- [19] **Karanjkar N, Joglekar A, Mohanty S, Prabhu V, Raghunath D, Sundaresan R.** Digital Twin for Energy Optimization in an SMT-PCB Assembly Line, *IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*, Bali, 2018. P.85-89. DOI:10.1109/IOTAIS.2018.8600830.
- [20] **Pileggi P, Verriet J, Broekhuijsen J, C. van Leeuwen, Wijbrandi W, Konsman M.** A Digital Twin for Cyber-Physical Energy Systems, *7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, Montreal, QC, Canada, 2019. P.1-6. DOI:10.1109/MSCPES.2019.8738792.
- [21] **Tao F, Zhang H, Liu A, Nee AYC.** Digital Twin in Industry: State-of-the-Art, in *IEEE Transactions on Industrial Informatics*, April 2019; 15(4): 2405-2415. DOI:10.1109/TII.2018.2873186.

- [22] **Prohorov A, Lysachev M.** Digital twin. Analysis, trends, world experience [In Russian]. Moscow: Al'yansPrint. 2020. 401 p.
- [23] **Maryasin OYu.** Development of ontologies for the digital twin of buildings [In Russian]. *Ontology of designing*. 2019; 9(4): 480-495. DOI: 10.18287/2223-9537-2019-9-4-480-495.
- [24] **Stennikov V, Barakhtenko E, Sokolov D, Mayorov G.** Principles of Building Digital Twins to Design Integrated Energy Systems. *Computation*, Dec. 2022; 10(12): 222. DOI: 10.3390/COMPUTATION10120222.
- [25] **Steinmetz C, Rettberg A, Ribeiro FGC, Schroeder G, Pereira CE.** Internet of things ontology for digital twin in cyber physical systems. Brazilian Symp Comput Syst Eng SBESC, 2018. P.154–159. DOI:10.1109/SBESC.2018.00030.
- [26] **Anthony RN.** Planning and Control: a Framework for Analysis. Cambridge MA: Harvard University Press, 1965.
- [27] **Makarov AA.** Approaches to assessing the sustainability and risks of long-term development of the Russian energy sector [In Russian]. Systems Research in Energy: Methodology and Results. M.: ERI RAS. 2018. P.87-101.
- [28] **Euzenat J, Shvaiko P.** Ontology matching: Second edition. Springer-Verlag Berlin Heidelberg. 2013. 511 p. DOI:10.1007/978-3-642-38721-0/COVER.
- [29] **Suarez-Figueroa MC, Gomez-Perez A, Motta E, Gangemi A.** (Eds.). Ontology engineering in a networked world. Springer Science & Business Media, 2012. P.444. DOI:10.1007/978-3-642-24794-1.
- [30] **Groumpos P, Stylios C.** Modelling supervisory control systems using fuzzy cognitive maps / Chaos, Solitons & Fractals. 2000; 11(1-3): 329-336.
- [31] **Papageorgiou E, Stylios C, Groumpos P.** An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps / IEEE Transactions on Biomedical Engineering. 2003; 50(12): 1326-1339.
- [32] **Ruiz JCS, Bru JM, Escoto RP.** Smart Digital Twin for ZDM-based job-shop scheduling. 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), Rome, Italy, 2021. P.510-515, DOI: 10.1109/MetroInd4.0IoT51437.2021.9488473.
- [33] **Massel LV.** Fractal approach to knowledge structuring and examples of its application [In Russian]. *Ontology of designing*. 2016; 6(2): 149-161. DOI: 10.18287/2223-9537-2016-6-2-149-161.
- [34] **Massel LV, Massel AG, Shchukin NI, Tsybikov AR, Losev AS.** Construction of digital twins of wind and solar power plants based on an ontological approach [In Russian]. Automation in Industry, 2022; 7: 28-32. DOI: 10.25728/avtprom.2022.07.04.
- [35] **Mammadov TG, Massel AG.** Adaptation of the methodology of reengineering of legacy software [In Russian]. Information and mathematical technologies in science and management. 2021; 4(24): 88-99. DOI: 10.38028/ESI.2021.24.4.009.

About the authors

Liudmila Vasilievna Massel (b. 1949) graduated from the Tomsk Polytechnic Institute (1971), Doctor of Technical Sciences (1995), professor (1999). Chief Researcher at the Melentiev Energy Systems Institute SB RAS. Professor at the Institute of Information Technologies and Data Analysis of the Irkutsk National Research Technical University. The list of scientific papers includes about 300 articles and chapters of monographs in the field of information systems and technology design, semantic modeling, and the development of intelligent decision support systems in the field of energy. ORCID: 0000-0002-9088-9012, Author ID (Scopus): 56440157300; Researcher ID (WoS): B-4793-2014. lvmassel@isem.irk.ru ✉.

Aleksei Gennadievich Massel (b. 1985) graduated from the Irkutsk State University (2007), PhD in Engineering Science (2011). Senior researcher at the Melentiev Energy Systems Institute SB RAS. Associate Professor at the Institute of Information Technologies and Data Analysis of the Irkutsk National Technical University. The list of scientific works includes more than 70 articles in the field of semantic modeling, design of information systems and technologies, and the development of intelligent decision support systems in the field of energy solutions. ORCID: 0000-0002-0351-0415; Author ID (РИИЦ): 173012; Author ID (Scopus): 57220804764. amassel@gmail.com.

Received November 1, 2022. Revised January 26, 2023. Accepted February 2, 2023.



Онтология идентификации человека по движениям тела и лицу в видеонаблюдениях

© 2023, А.Е. Колоденкова

Самарский государственный технический университет, Самара, Россия

Аннотация

В настоящее время разработка моделей и методов распознавания по движениям тела и лицу в видеонаблюдениях является актуальной задачей. Особенно это важно в обеспечении безопасности на объектах с массовым скоплением людей для противодействия преступлениям террористической направленности. В статье приведена классификация основных биометрических признаков и параметров, характеризующих потенциального нарушителя, разработанная для систем контроля безопасности, пропускных систем предприятий. Предложена структурная схема слияния биометрических данных и распознавания нарушителя, которая может лежать в основе разработки систем контроля безопасности. Рассмотрены виды систем и методы распознавания человека по движениям тела и лицу, выявлены их достоинства и недостатки. Отмечено, что для распознавания нарушителя в условиях множества биометрических признаков целесообразно использовать комбинацию методов распознавания. Это позволит принимать правильные решения относительно выявления потенциального нарушителя. В статье сделана попытка рассмотреть основные аспекты, касающиеся распознавания человека по движениям тела и лицу в видеонаблюдениях в целом, в отличие от известных работ, посвящённых отдельным биометрическим признакам.

Ключевые слова: методы распознавания, биометрические признаки, потенциальный нарушитель, системы контроля безопасности, видеонаблюдения, онтология.

Цитирование: Колоденкова А.Е. Онтология идентификации человека по движениям тела и лицу в видеонаблюдениях // *Онтология проектирования*. 2023. Т.13, №1(47). С.55-74. DOI: 10.18287/2223-9537-2023-13-1-55-74.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

Согласно статистике в Российской Федерации в 2022 г. зарегистрировано 2233 преступления террористического характера и 1566 – экстремистской направленности [1]. С увеличением преступлений террористического характера на стратегических объектах, в учебных заведениях, в различных организациях с массовым скоплением людей главная роль отводится сотрудникам, обеспечивающим внутриобъектовый и пропускной режим [2, 3].

С развитием информационных технологий обеспечение надёжной защиты объектов невозможно без применения систем контроля безопасности (СКБ), основанных на биометрии. СКБ основываются на уникальных характеристиках человека, которые сложно фальсифицировать, и позволяют однозначно распознать конкретного человека. В СКБ существенными биометрическими признаками являются движения тела (походка, жестикуляция, осанка) и мимика (лицо), поскольку их можно наблюдать издали по видео без прямого контакта с человеком.

Целью настоящей статьи является анализ подходов к распознаванию человека по движениям тела и лицу при выявлении потенциального нарушителя из общей массы людей, а также предложить рекомендации разработчикам систем распознавания нарушителя по видеонаблюдению.

1 Онтология предметной области

Онтология (онтологическая модель) предметной области (ПрО) представляет собой систему, состоящую из набора понятий и отношений между ними. Формальную модель онтологии ПрО можно представить в виде кортежа:

$$O = \langle X, A, R, F \rangle,$$

где X – конечное множество понятий ПрО; A – конечное множество атрибутов понятий X ; R – конечное множество отношений между понятиями, заданной ПрО; F – конечное множество функций интерпретации, заданных отношениями онтологии.

Фрагмент онтологии процесса идентификации нарушителя в виде семантической сети представлен на рисунке 1.

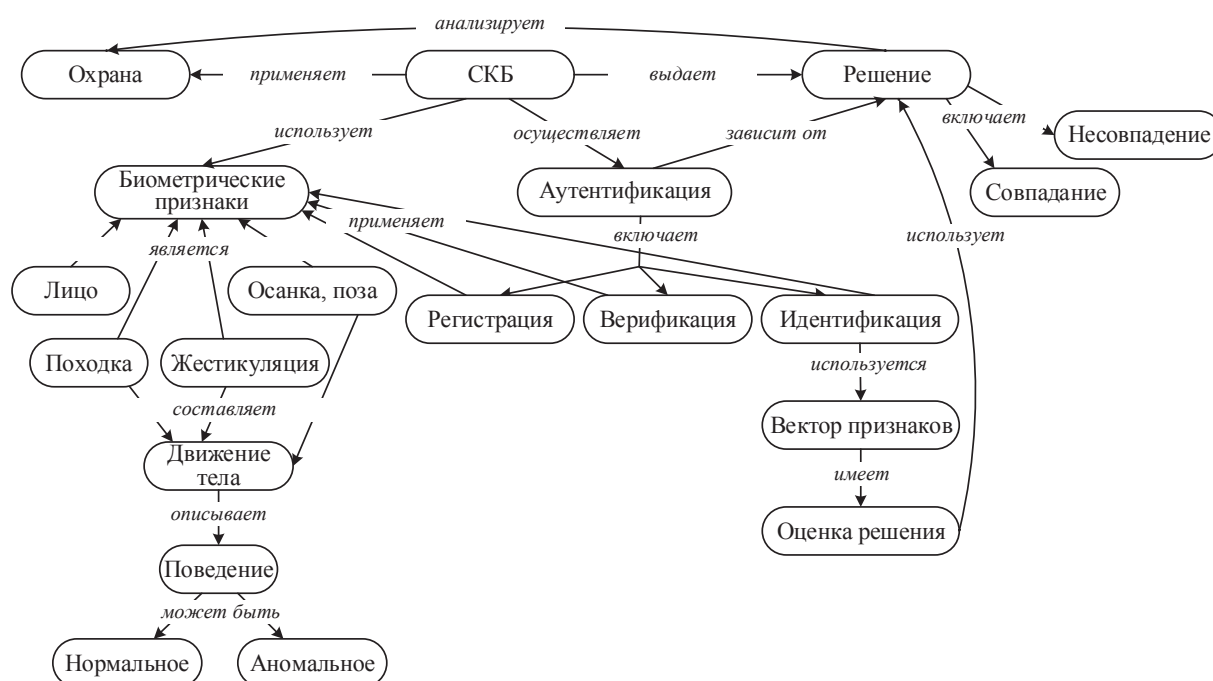


Рисунок 1 – Фрагмент онтологии процесса идентификации нарушителя

ПрО «процесса идентификации нарушителя» включает список понятий, для которых определены атрибуты и отношения. Фрагмент спецификации понятий онтологии идентификации нарушителя представлен в таблице 1. Онтологии позволяют структурировать знания при идентификации нарушителя, а также обеспечить формальное представление системы понятий ПрО и поддержку требуемой функциональности.

Таблица 1 – Фрагмент спецификации понятий онтологии процесса идентификации нарушителя

Понятия	Атрибуты	Отношения с атрибутами
Охрана	ФИО. Серия и номер удостоверения	Решение. СКБ
Вектор признаков	Имя. Номер. Дополнительная информация	Идентификация
Лицо	Положение губ. Положение бровей. Положение глаз	Биометрический признак
Походка	Размер шага. Темп. Вид	Биометрический признак. Движение тела

2 Классификация биометрических признаков при идентификации

Существует множество определений понятия и типологии нарушителя [4-6]. В работе под потенциальным нарушителем понимается человек, который совершает или может со-

вершить преступление, направленное на устрашение и насилие, добивается своих целей путём захвата заложников с угрозой их уничтожения (террорист). Для выработки эффективных мер, направленных на предотвращение преступлений террористического характера, необходимо иметь представление о личности потенциального нарушителя, что невозможно без изучения её особенностей (структуры, биометрических признаков, психологического портрета и др.) [7-10].

Признаки человеческого тела, поведенческие (походка, голос и др.), либо физиологические (лицо, отпечаток пальца, ДНК и др.) могут быть использованы в качестве биометрических признаков для распознавания личности. Данные признаки должны удовлетворять следующим свойствам: уникальность, универсальность (биометрические данные могут быть получены у любого человека независимо от их пола, возраста, местоположения), постоянство (например, рукописная подпись лица, выполненная в документе, должна визуально соответствовать подписи этого лица, содержащейся в иных документах), неуязвимость, собираемость [11].

В литературе рассматривается большое количество психологических портретов террористов. Личность нарушителя, совершившего преступление террористического характера, отличается от других категорий преступников (убийцы, воры, наркоманы и др.), во-первых, устойчивыми психологическими и психическими нарушениями, нарушениями социальной адаптации; во-вторых, преступлениями против человечества больше, чем против отдельных лиц (обычные преступники делают это больше для собственной выгоды) [12].

Деятельность террористов сложно распознать. Основными признаками поведения террориста являются: отрешённый взгляд, отсутствие визуального контакта с окружающими (жёсткий и сфокусированный взгляд вперёд), подчинённая поза, тяжёлая и скованная походка, определённая заторможенность реакций, внешне имеет вид собранного и умиротворенного человека [13-15]. Возможен и другой вариант их поведения, например, когда террорист может быть заметно возбуждён, агрессивен. Тогда у него может проявляться обильное выделение пота; он резко двигается, постоянно оглядывается, боясь слежки [14].

В работах [16-18] описаны основные черты лица террориста: неестественная бледность; скованное, каменное лицо, которое не выражает никаких эмоций; крупные и блестящие глаза; губы могут быть сильно сжаты либо чуть заметно двигаться (возможно чтение молитвы); возможные шрамы; и многие др.

Составить и выделить универсальный тип террориста невозможно, а можно лишь условно выделить некоторые биометрические признаки, характеризующие его [12, 13]. Сложность задачи обусловлена многоаспектностью такого явления как терроризм, а также многогранностью личности человека.

Основные типы биометрических признаков для выявления потенциального нарушителя изображены на рисунке 2 (адаптирован по материалам [19-21]).

Биометрические признаки человеческого тела



Рисунок 2 – Типы биометрических признаков для выявления потенциального нарушителя

Физиологическими признаками для криминологии, сотрудников полиции и охраны являются движение тела и лицо [22, 23].

Физиологические (статические) биометрические признаки – признаки, присутствующие с рождением человека. Их получают в результате измерения части человеческого тела. Достоинства данных признаков в относительно ограниченной возможности их изменить и скорости процесса распознавания с использованием компьютеров. Недостатки физиологических признаков: дороговизна биометрических сканеров; возможность повреждения биометрических идентификаторов; использование лицевых аксессуаров (кепки, очки, платки, маски и т.п.); необходимость определённых условий окружающей среды (освещение и т.п.).

Поведенческие (динамические) биометрические признаки – признаки, основанные на определенном шаблоне поведения (действия), выполняемого человеком (косвенно измеряют признаки облика человека). Главной особенностью данных признаков является использование времени в качестве метрики. Достоинствами поведенческих признаков являются: присутствие индивидуального набора анализируемых признаков; повышение точности распознавания в СКБ многофакторной идентификации. Недостатками являются: отсутствие метода и модели для оценки точности распознавания человека по движению тела; необходимость большого количества личных данных для верного определения поведения нарушителя, поскольку оно может меняться (усталость, опьянение, плохое самочувствие и т.п.).

Биометрические признаки имеют следующие общие достоинства: уникальные особенности для каждого человека; неизменные черты с течением времени [24]. Существуют ограничения, к которым биометрия уязвима, например, компьютерные атаки, изменение БД [25, 26].

Выявление потенциального нарушителя осуществляется в условиях множества биометрических признаков. Предлагается следующая классификация основных биометрических признаков, характеризующих потенциального нарушителя (рисунок 3, адаптирован по материалам [23, 27-29]).

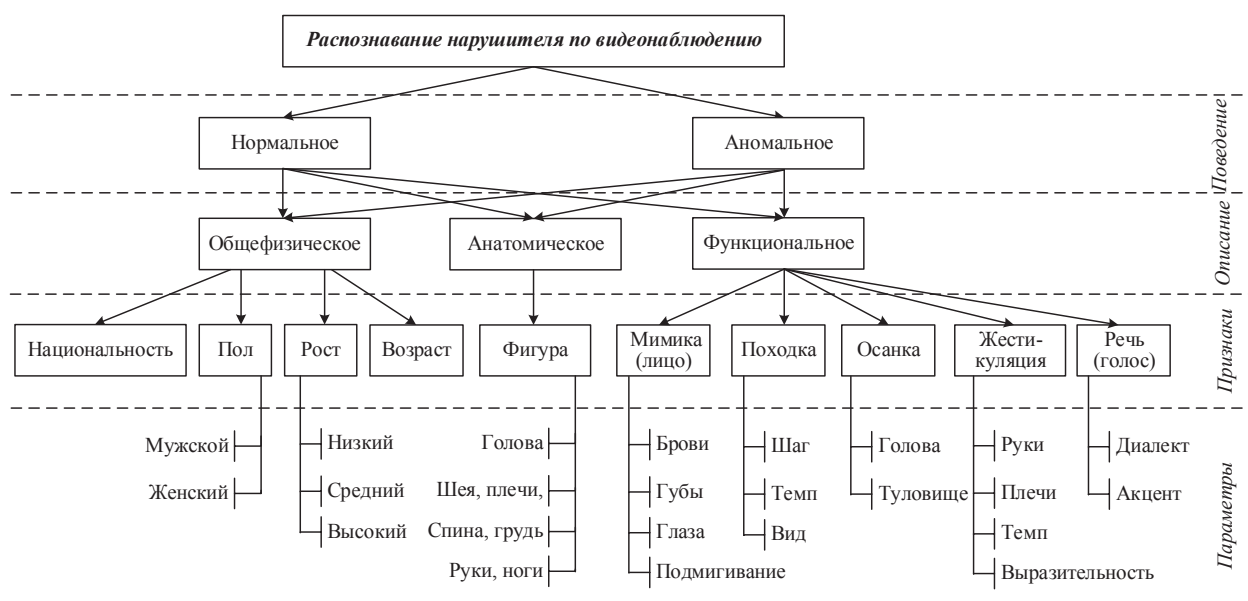


Рисунок 3 – Классификация основных биометрических признаков, характеризующих потенциального нарушителя

Поведение человека может быть «нормальным» и «аномальным (девиантным)». В настоящей работе под аномальным поведением понимается последовательность действий человека, которые не соответствуют модели типичного поведения в конкретной ситуации, имеют отклонение от нормы поведения, склонность к нарушениям, агрессии [28-30].

Описание внешности человека. Общефизическое описание – элементы внешности; анатомическое – описание всей фигуры в целом либо отдельных областей тела; функциональное – особенности, которые проявляются в движении. Каждый из анатомических признаков характеризуется по форме, размеру, положению, цвету [29]. При этом некоторые параметры могут быть представлены в виде словесного (нечёткого) описания, например, телосложение – слабое, очень слабое, среднее, коренастое, атлетическое; упитанность человека – худой, худощавый, средней упитанности, полный. Функциональное описание: мимика – движение мышц и элементов лица, которые могут меняться в зависимости от эмоционального состояния человека; походка – совокупность индивидуальных телодвижений при ходьбе; жестикуляция – комплекс движений рук, плеч, головы человека, которыми он сопровождает свою речь; речь – данные речевого механизма.

3 Виды систем распознавания

В настоящее время системы распознавания человека подразделяются на несколько видов (рисунок 4, адаптирован по материалам [24, 31-33]).

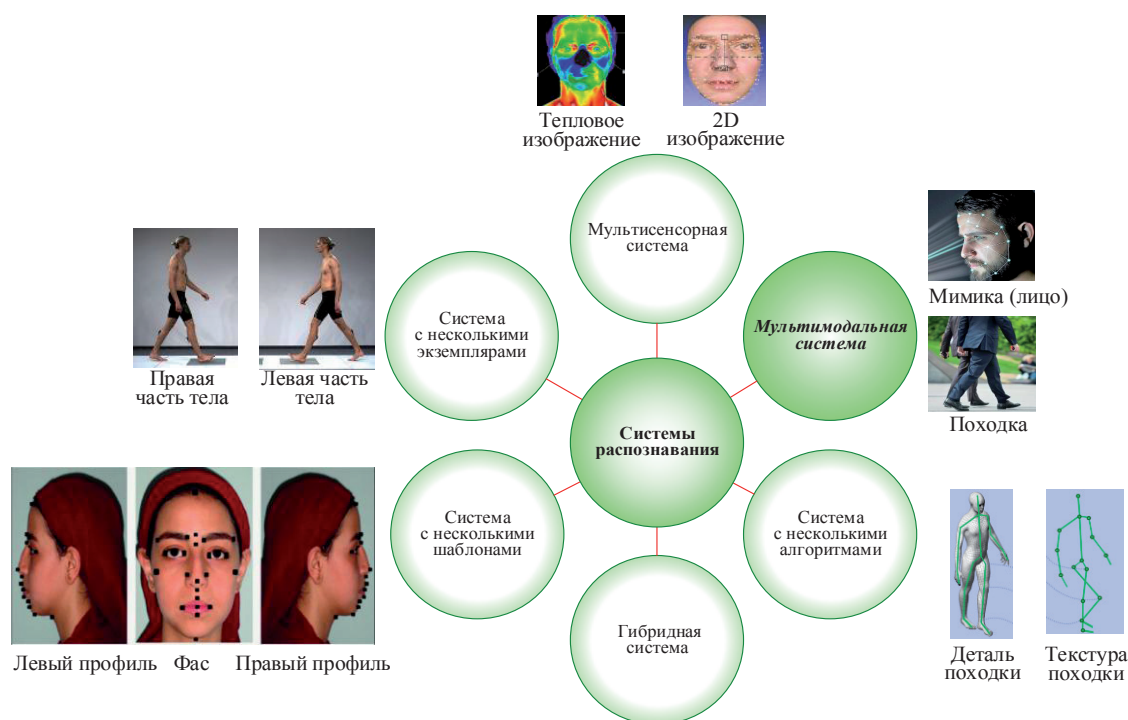


Рисунок 4 – Виды систем распознавания человека

Мультисенсорная система позволяет объединять полученные данные с различных датчиков (сканеров) для одного и того же биометрического признака для извлечения разнообразной информации (несколько датчиков – один биометрический признак). Например, данная система может использовать 2D, 3D или тепловое изображение лица для распознавания.

Мультимодальная система позволяет объединять более одного биометрического признака, что повышает точность распознавания человека. Например, данная система может использовать одновременно лицо и голос для распознавания человека. Такие системы являются дорогостоящими, поскольку для их работы требуется несколько датчиков, каждый из которых воспринимает различные биометрические признаки.

Система с несколькими алгоритмами позволяет обрабатывать один биометрический признак, полученный с одного датчика с использованием различных подходов к извлечению признаков и различных алгоритмов сопоставления (несколько алгоритмов – один биометрический признак). Например, система распознавания отпечатков пальцев, походки может использовать как детали, так и особенности текстуры для сопоставления отпечатков пальцев и походки.

Гибридная система позволяет объединять более одной из вышеупомянутых биометрических систем для надёжного распознавания. Например, если два алгоритма распознавания лица будут объединены с двумя алгоритмами распознавания отпечатков пальцев, то такая система будет мультимодальной с несколькими алгоритмами.

Система с несколькими шаблонами позволяет использовать несколько шаблонов одного и того же биометрического признака, полученных с помощью одного датчика (несколько образцов с одним датчиком – один биометрический признак). Например, при распознавании человека по снимкам лица, которые сделаны при разных точках доступа и освещении.

Система с несколькими экземплярами позволяет фиксировать несколько экземпляров одного и того же биометрического признака (несколько экземпляров – один биометрический признак). Например, изображения левой и правой частей тела, отпечатки двух или более пальцев или ладоней человека могут быть объединены по одному изображению одного и того же человека.

При распознавании человека встречаются *унимодальные системы*, позволяющие использовать один биометрический признак человека для распознавания и проверки. Данные системы предлагают надёжное решение для приложений идентификации и верификации, однако необходимо учитывать ограничения их использования, которые могут привести к ошибкам распознавания.

- Шумы в измеряемых данных (изображение отпечатка пальца со шрамом; образец голоса, изменённый холодом; плохая освещённость лица и др.).
- Неуниверсальность – отсутствие получения значимых биометрических данных (система распознавания радужной оболочки не может получить информацию потенциального нарушителя с длинными ресницами, опущенными веками, патологией глаз).
- Поддельные атаки, которые встречаются при использовании поведенческих признаков (подпись, голос, походка и др.). В этом случае система, основанная только на анализе одного признака, может привести к неверному распознаванию. Если система включает дополнительный признак, то это приведёт к увеличению вероятности распознавания (маловероятно, что у двух разных людей одинаковая походка и лицо).

При распознавании человека СКБ, основанные на слиянии многих биометрических признаков, способны эффективно обрабатывать зашумлённые или некачественные данные (характерно для систем, основанных на распознавании лица и голоса) [33].

Интеграция разнородной информации, поступающей из разных источников, является одним из основных приёмов проектирования СКБ по распознаванию человека, основанных на слиянии множества биометрических признаков [33-36].

СКБ должны соответствовать следующим критериям [31, 32, 37-39].

- *Универсальность* – каждый нарушитель обладает своими уникальными биологическими признаками. Люди без рук, пальцев, с нарушением походки, осанки или с повреждёнными глазами также должны быть зарегистрированы и учтены.
- *Уникальность, отличительность* – никакие два любых нарушителя не могут быть одинаковыми с точки зрения биометрических признаков. Уникальность может измеряться частотой ложного сопоставления.
- *Надёжность* – биометрические признаки должны быть неизменными в течение определённого периода времени. Радужная оболочка глаза обычно остаётся неизменной на протяжении десятилетий, лицо человека со временем значительно меняется, палец часто подвергается травмам, подпись и её динамика также может измениться. Надёжность может измеряться частотой ложных несоответствий.
- *Представление* – достижение точности и скорости распознавания с учетом задействованных эксплуатационных факторов и факторов окружающей среды, необходимых для достижения приемлемой точности.
- *Приемлемость* – у пользователей и общественности не должно быть возражений против измерения/сбора биометрических признаков.
- *Доступность, собираемость* – биометрический признак может быть измерен количественно с помощью какого-либо сенсорного устройства и легко визуализирован. Доступность может быть количественно оценена по пропускной способности системы.
- *Устойчивость к обходу* – тесты и доказательства того, что разработанная система противостоит мошенническим методам.

Сравнение биометрических признаков для выявления потенциального нарушителя с использованием рассмотренных критериев СКБ представлено в таблице 2, адаптированной по материалам [32, 38].

Таблица 2 – Сравнение биометрических признаков

Критерии СКБ Биометрические признаки	Универсальность	Уникальность	Надёжность	Доступность	Представление	Приемлемость	Устойчивость к обходу
Мимика (лицо)	высокая	низкая	средняя	высокая	низкое	высокая	средняя
Походка	высокая	низкая	средняя	высокая	низкое	высокая	высокая
Жестикуляция	средняя	низкая	средняя	средняя	низкое	высокая	средняя
Осанка	средняя	низкая	низкая	средняя	низкое	высокая	средняя

Какой биометрический признак является наилучшим, однозначно ответить нельзя, поскольку каждый признак имеет свои сильные и слабые стороны, и выбор обычно зависит от постановки задачи.

4 Слияние биометрических данных

Схема слияния биометрических данных и распознавания нарушителя по движениям тела и лицу, которая может лежать в основе разработки СКБ, представлена на рисунке 5.

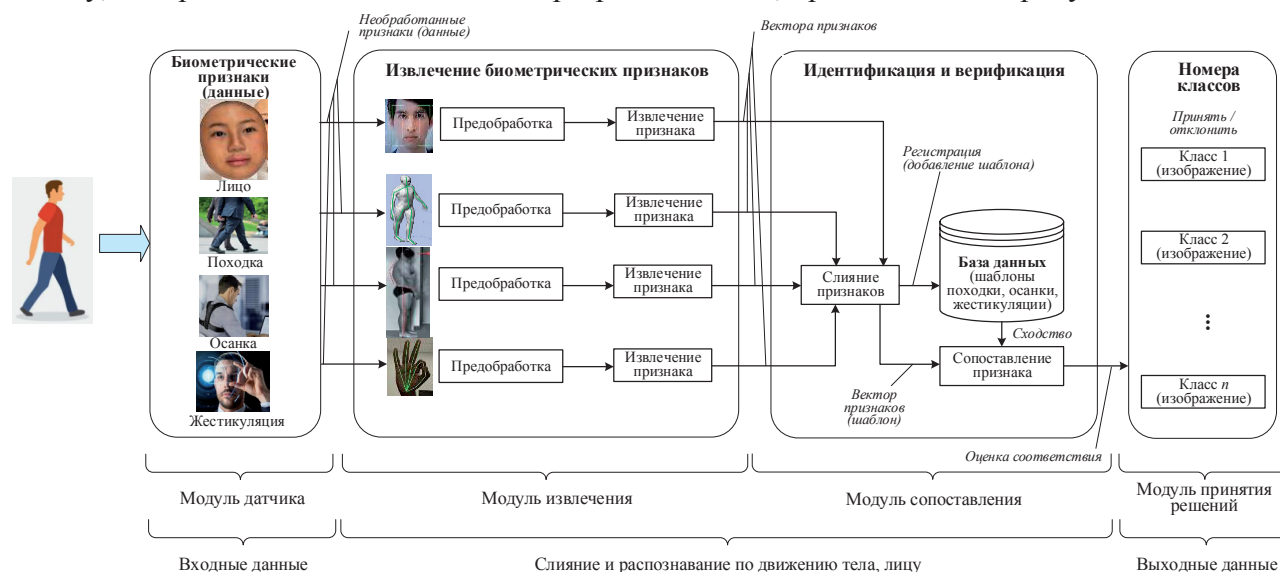


Рисунок 5 – Структурная схема слияния биометрических данных и распознавания нарушителя по движениям тела и лицу

СКБ, работающие с множеством биометрических признаков, состоят из модулей [31, 40, 41]:

- *модуль датчика* собирает биометрические данные человека (например, датчик отпечатков пальцев), которые задаются в качестве входных данных и служат входом для модуля извлечения признаков;
- *модуль извлечения (выделения) признаков* извлекает значения признаков после предобработки (например, положение и ориентация мелких точек на изображении отпечатка пальца), которые дают компактное представление этих признаков;
- *модуль сопоставления (соответствия)* сравнивает значения признаков со значениями в шаблоне, которые хранятся в БД, путём создания соответствующей оценки (степень сходства или расхождения между двумя биометрическими векторами признаков), которая передаётся в модуль принятия решений;

- *модуль принятия решений* устанавливает (выявляет) нарушителя, который либо подтверждается, либо отклоняется на основе оценки соответствия, сгенерированной в модуле сопоставления, либо распознаёт личность человека.

Важный вопрос, возникающий при слиянии биометрических признаков, - определение типа данных, которые должны быть объединены и выбор метода слияния [31]. Классификация уровней слияния в СКБ, работающих с множеством биометрических данных, которые могут быть объединены на различных уровнях показана на рисунок 6 (адаптирован по материалам [31, 40-42]).



Рисунок 6 – Классификация уровней слияния в СКБ

Согласно категории *до сопоставления признаков* слияние осуществляется на уровне датчиков, на уровне признаков до того, как будет выполнено сопоставление биометрических данных [37, 38, 40-43].

Слияние на уровне датчиков – объединение необработанных данных, полученных с различных биометрических датчиков, в один вектор. Информация одного и того же биометрического признака, полученного с различных датчиков, объединяется для получения единого биометрического признака. Т.е. дополнительная информация, соответствующая, например, отпечаткам пальцев, которая может быть получена с использованием различных типов датчиков, интегрируется с использованием метода слияния на уровне датчиков. Другой пример – изображения лиц, полученные с нескольких камер, могут быть объединены для формирования единого изображения лица.

Структура *слияния на уровне датчиков* представлена на рисунке 7. Данные поступают с разных датчиков, которые должны быть совместимы, что не всегда возможно (например, может оказаться невозможным объединить изображения лиц, полученные с камер с разным разрешением); все методы должны быть совместимы с исходными данными и известны заранее [44]; один датчик или различные совместимые датчики (отпечаток пальца, сканер радужной оболочки глаза и т.д.) представляют шаблоны (образцы) одного обнаруженного биометрического признака. Данное слияние рекомендуется для систем с множеством датчиков и несколькими выборками, которые делают несколько снимков одной и той же биометрической информации [45]. Предполагается, что это повысит точность распознавания.

Слияние на уровне признаков – слияние векторов признаков, которые получены либо с использованием нескольких датчиков, либо с использованием нескольких алгоритмов извле-

чения признаков из одних и тех же датчиков (векторы признаков одного биометрического признака, полученные от разных датчиков; векторы признаков одного биометрического признака, полученные от разных объектов (векторы признаков отпечатков пальцев левой и правой руки); векторы признаков, сгенерированные из нескольких биометрических признаков) (рисунок 8).

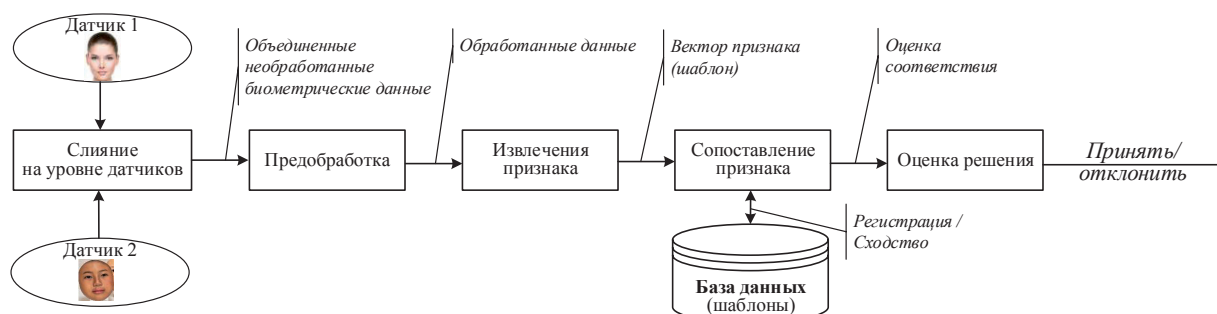


Рисунок 7 – Слияние данных на уровне датчиков

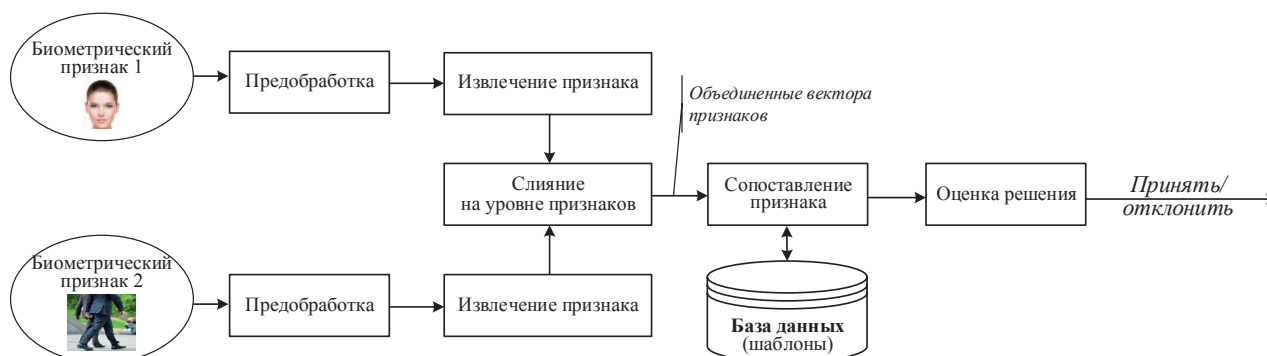


Рисунок 8 – Слияние данных на уровне признаков

Если признаки, извлечённые из одного биометрического признака, не зависят от признаков, извлечённых из другого, можно объединить два вектора в один новый вектор, который может быть как однородным, так и гетерогенным. Когда векторы признаков однородны (например, множественные отпечатки пальцев нарушителя), один результирующий вектор признаков может быть вычислен как средневзвешенное значение отдельных векторов. Когда векторы признаков неоднородны (например, векторы признаков, полученные с использованием различных методов извлечения признаков, или векторы признаков различных биометрических признаков), то происходит объединение с целью формирования единого вектора признаков. Объединение невозможно, когда наборы признаков несовместимы (например, признаки пальцев и лица). Новый вектор признаков имеет более высокую размерность и представляет личность человека в подробном виде, который сравнивается с шаблоном регистрации (объединённым вектором признаков, хранящимся в БД).

При слиянии векторов признаков необходимо учитывать трудности, возникающие по следующим причинам [37, 40]:

- объединение двух векторов признаков может привести к получению вектора признаков с очень большой размерностью;
- большая размерность вектора признаков приводит к увеличению вычислительных ресурсов и ресурсов хранения;
- большинство коммерческих биометрических систем не предоставляют доступ к векторам признаков, которые они используют в своих продуктах.

Сдерживающим фактором при распознавании нарушителя является отсутствие единой общедоступной БД биометрических образцов преступников.

Согласно категории *после сопоставления признаков* слияние осуществляется на уровнях оценки сопоставления, ранга и принятия решений (см. рисунок 6) [32, 40, 45].

Слияние на уровне оценки сопоставления (достоверности) – векторы признаков, создаваемые независимо для каждого датчика, сравниваются с шаблонами, которые хранятся в БД отдельно для каждого биометрического признака, и объединяются для оценки соответствия в виде одиночного скалярного балла (показателя). Данные оценки могут быть объединены для подтверждения подлинности заявленной личности. Структура слияния на уровне оценки сопоставления представлена на рисунке 9.

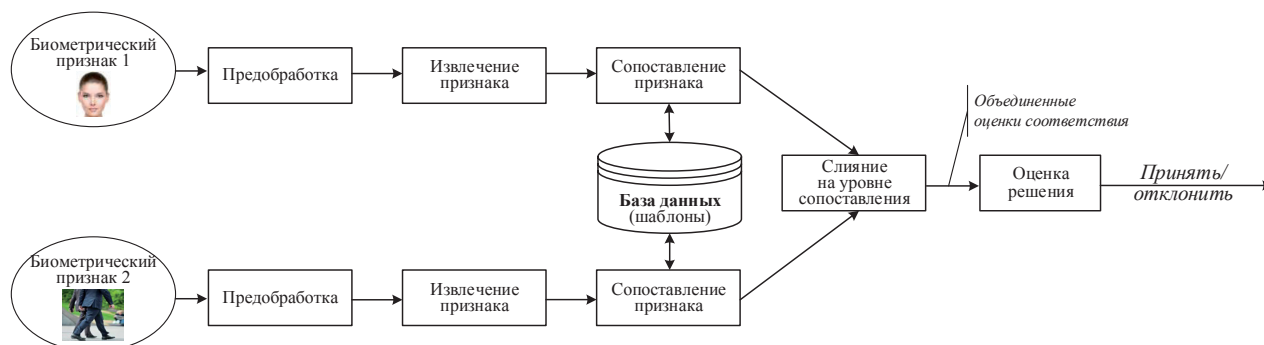


Рисунок 9 – Слияние данных на уровне оценки сопоставления

Соответствующий балл указывает на близость входного вектора признаков к вектору шаблона. Совпадающие баллы не могут напрямую быть использованы или объединены, поскольку оценки получены из разных датчиков и основаны на разных методах масштабирования. Для решения данной задачи введены три варианта схемы слияния: на основе плотности, преобразования, классификатора [46, 47].

Схема на основе плотности основана на оценке распределения баллов. Данная схема включает фильтр Калмана, расширенный фильтр Калмана и методы слияния фильтров частиц. Схема обеспечивает точную оценку, но требует большого количества тренировочных образцов, а также больше времени и усилий для настройки работы по сравнению с другими схемами.

Схема на основе преобразования обычно применяется для нормализации баллов. Этот процесс необходим для изменения масштаба параметров с целью обеспечения совместимости между несколькими переменными оценки [48]. Эта схема может быть применена с использованием различных методов (правило суммы, правило произведения, минимальное правило и максимальное правило).

В схеме на основе классификатора оценки соответствия, полученные с помощью нескольких сопоставителей, объединяются для построения единого вектора признаков, который затем передается в подходящий классификатор с целью получения заключительной метки: является ли пользователь законным или самозванцем. Для классификации вектора соответствия в этой схеме применяются: метод опорных векторов, байесовский вывод, теория Демпстера-Шефера, модель максимальной энтропии, нейронная сеть.

Слияние на уровне ранга – объединение рангов, полученных каждым отдельным биометрическим сопоставлением, и определение нового ранга, который будет использоваться при принятии окончательного решения. Высокий ранг указывает на хорошее соответствие. Слияние на уровне ранга используется для идентификации, а не для верификации [49].

Слияние на уровне принятия решений – формирование окончательного решения из полученных индивидуально отдельных решений о личности нарушителя по различным биомет-

рическим признакам (рисунок 10) с использованием различных методов, например, метода голосования большинством.

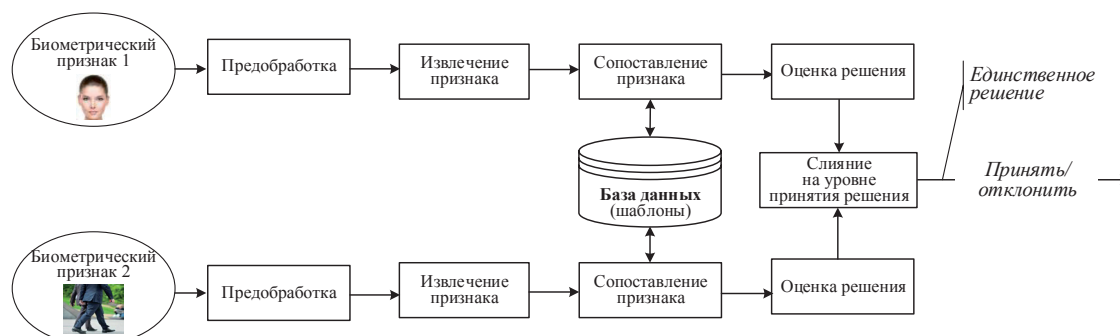


Рисунок 10 – Слияние данных на уровне принятия решения

Главное преимущество метода голосования большинством заключается в том, что он не требует предварительных знаний о сопоставителе, а также подготовки для принятия окончательного решения.

СКБ могут работать либо в режиме верификации, либо в режиме идентификации (см. рисунок 5). Процесс биометрической аутентификации (процесс проверки подлинности личности) разделен на три основных этапа (рисунок 11 адаптирован по материалам [37, 38]).

Регистрация – процесс регистрации нового человека в системе. На данном шаге сначала биометрический признак личности фиксируется датчиком для генерирования цифрового представления признака. Датчик должен получать всю важную информацию о личности, которая обычно представлена в виде изображения. Затем осуществляется предварительная обработка биометрических данных X (улучшение, нормализация, сегментация и удаление шума). Целью нормализации элемента является изменение местоположения (среднего) и масштаба (дисперсии) значения элемента с помощью функции преобразования, чтобы отобразить их в общую область. Далее происходит извлечение набора признаков из обработанных биометрических данных X' для создания эталонного шаблона Z_X . Этот шаг является решающим, поскольку для успешного распознавания потенциального нарушителя необходимо извлечь и выбрать правильные признаки. Шаблон Z_X , представляющий личность потенциального нарушителя, который будет использоваться для последующего сравнения, сохраняется в БД. Как видно из рисунка 5 этап регистрации является общим, как для этапа верификации, так и для этапа идентификации.

Верификация – процесс проверки подлинности личности потенциального нарушителя путём сравнения (1:1 соответствие) предоставленной биометрическим признаком Z_Q (биометрический признак запроса, полученный на этапах распознавания) с хранимым в БД шаблоном, который соответствуют заявленному нарушителю Z_X . Т.е. необходимо определить, является ли заявленная личность действительной? (является ли этот человек тем, за кого он себя выдаёт?). Результирующая оценка соответствия S (балл) сравнивается с сохранённым пороговым значением, вычисленным для заявленного потенциального нарушителя, или общим пороговым значением. Оценка соответствия определяется в диапазоне от 0 до 100 %.

Идентификация – процесс распознавания личности потенциального нарушителя путём сравнения биометрического признака Z_Q со всей БД ($Z_X = \{X_1, X_2, \dots, X_m\}$, где m – количество нарушителей, зарегистрированных в БД) путём «один ко многим», (1:N) с заданной степенью сходства. Т.е. требуется ответить на вопрос: «Кто этот человек?». Далее неизвестному лицу может быть присвоен идентификатор, соответствующий наиболее похожему профилю, найденному в БД, либо отклоняется это лицо. Если совпадение не найдено, это означает, что данное лицо не зарегистрировано в БД. Решение о принятии или отклонении человека принимается путём сравнения ответа системы с пороговым значением (называемым порогом принятия решения). В результате будет выбрана наиболее похожая личность по используемой геометрии лица искомого человека, а не идентичная, как в процессе верификации. В этом случае будет идентификация или ответ, что данное лицо есть в БД. Если лицо отсутствует в БД, или степень сходства является меньше заданной для всех сравнений, то результата не будет. Ошибка системы возрастает за счёт сравнения «один ко многим», и этап идентификации становится критическим для системы распознавания.

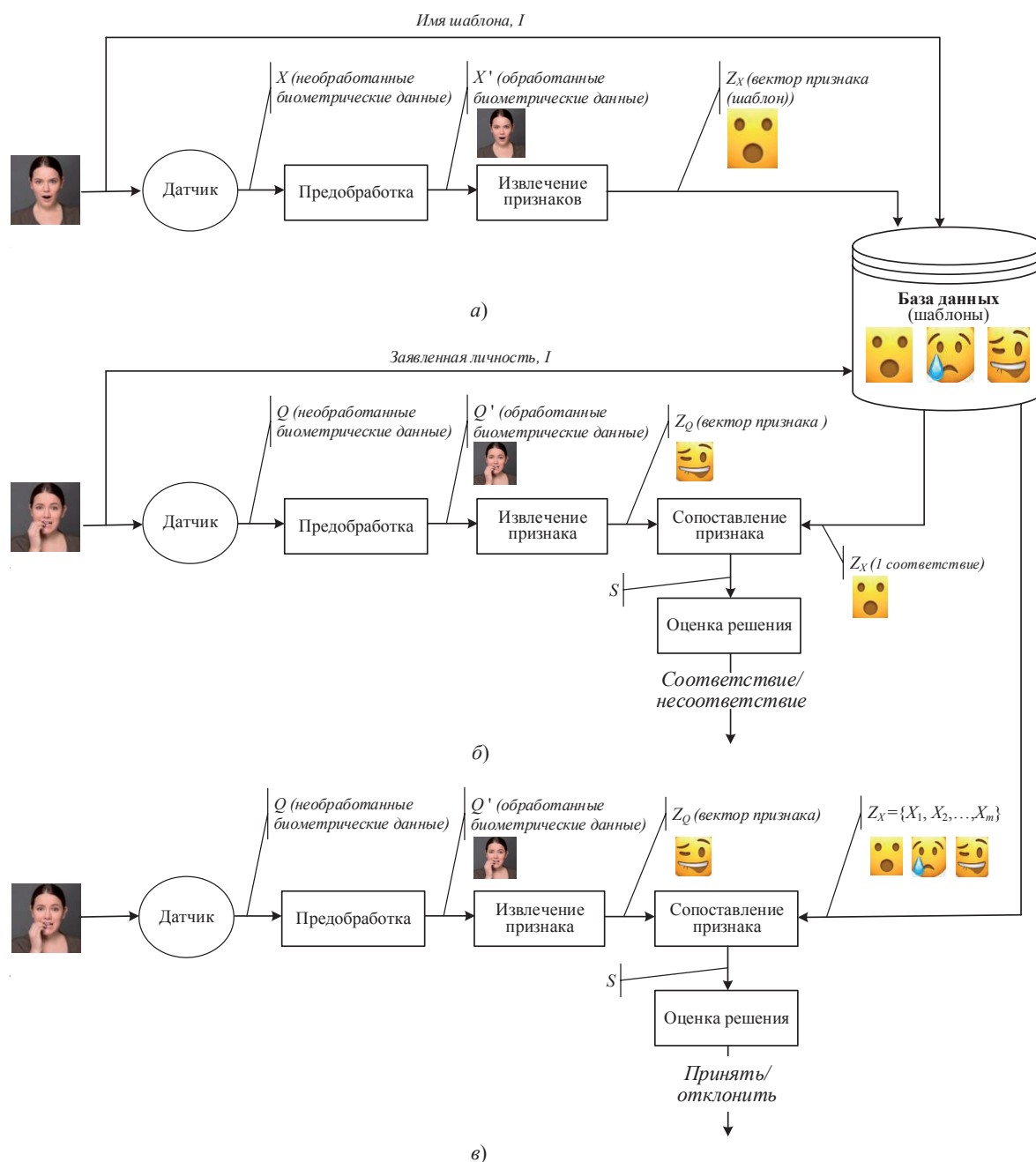


Рисунок 11 – Схемы процесса биометрической аутентификации:

а) этап регистрации; б) этап верификации; в) этап идентификации

X – признак, полученный во время регистрации; Z_X – набор признаков шаблона; Z_Q – набор признаков запроса; Q – биометрический признак запроса (выборка), полученный на этапах распознавания; S – оценка соответствия; M – количество нарушителей, зарегистрированных в БД

5 Методы и технологии распознавания человека по движению тела и лицу

Для решения задачи распознавания личности на видео по движению тела и лицу разработано большое количество различных методов. Классификация методов и технологий, которые используются при распознавании человека по движениям тела и лицу в видеонаблюдениях, показана на рисунке 12 (адаптирован по материалам [50-59]).

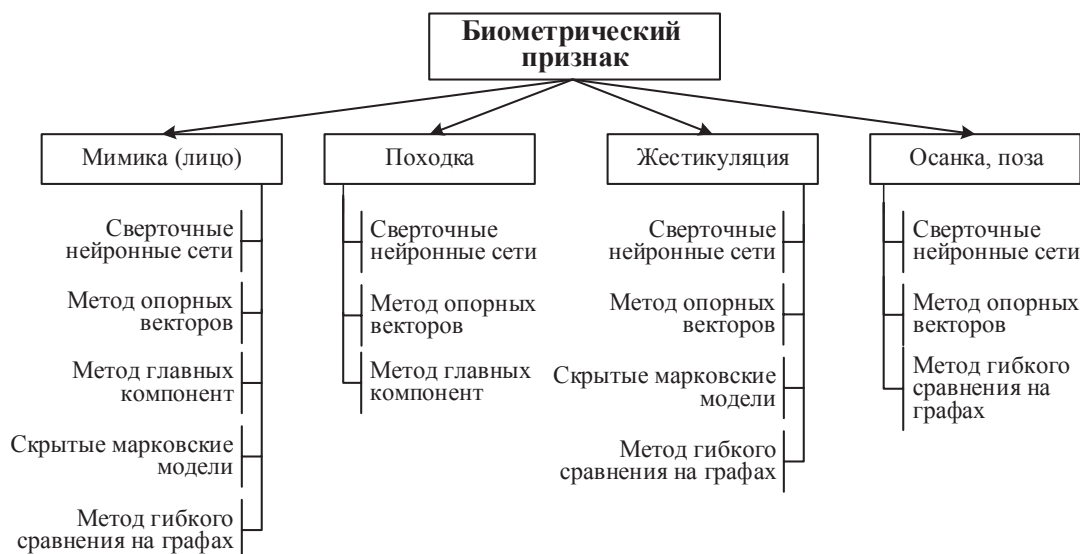


Рисунок 12 – Классификация методов и технологий распознавания человека по движению тела и лицу

Приведённое разделение методов и технологий, применяющихся при распознавании личности, носит условный характер, поскольку на практике они пересекаются и взаимодействуют между собой. В этом случае применяются гибридные подходы, сочетающие в себе различные методы и технологии распознавания человека. Ключевым моментом при их разработке является то, что они не должны конфликтовать между собой.

Наиболее применяемыми являются следующие методы:

Свёрточные нейронные сети (*Convolutional Neural Networks, CNN*) – это класс искусственных нейронных сетей, чаще всего применяемых для задач классификации, обнаружения и анализа объектов на изображениях, сконцентрированных на небольших участках изображения, и выделения в них важных особенностей. В работе [50] представлена модифицированная архитектура *CNN* для извлечения отличительных черт лица путём добавления двух операций нормализации к двум слоям. Операция нормализации представляет собой пакетную нормализацию, которая обеспечивает ускорение работы *CNN*. Для классификации лиц использовался классификатор *Softmax*.

В работе [51] предложен подход, основанный на идее многоклассовой классификации на видеопоследовательностях. Оценка качества предлагаемого подхода проводилась на основе набора данных, который включает более 15000 видеопоследовательностей. В качестве классификаторов были апробированы пять архитектур нейронных сетей. Результаты исследований показали, что предлагаемый подход может осуществлять идентификацию человека в режиме реального времени без использования специализированного оборудования с точностью около 80 %.

В работе [52] предложена новая модель *CNN* для распознавания походки на основе позы. Данная модель учитывает движение точек в областях вокруг суставов человека. Для извлечения информации о движении оценивается оптический поток между последовательными кадрами.

Метод опорных векторов (*Support Vector Machines, SVM*) – один из самых известных алгоритмов обучения с учителем, применяемый для задачи классификации и регрессии в машинном обучении. В работе [53] предложен метод распознавания лиц на основе комбинирования анализа основных компонент ядра (*Kernel Principal Component Analysis, KPCA*) и метода опорных векторов. Сначала используется метод *KPCA* для извлечения признаков из входных изображений, а затем применяется метод *SVM* к извлечённым признакам для классификации входных изображений.

В работе [54] описан прототип системы реального времени, способной распознавать четыре жеста, которые соотносятся с человеческими эмоциями на основе движений рук. Предложена структура для использования датчика походки *Kinect* для идентификации жеста. Объекты, извлечённые из 3D-скелета с помощью датчика *Kinect v2*, классифицируются с использованием метода *SVM*.

В работе [55] предложен метод распознавания поз человека с использованием классификатора *SVM*. Для одновременной съёмки двух наборов последовательностей изображений используются две камеры. После захвата последовательностей изображений используется алгоритм сегментации движущихся объектов, чтобы от-

личить человеческое тело от фона. Экспериментальные результаты показали, что предложенный метод обеспечивает высокую скорость и уровень распознавания.

Недостатками метода опорных векторов являются существенные временные затраты при настройке и необходимость большого объема памяти.

Метод главных компонент (*Principal Components Method, PCA*) обеспечивает уменьшение размерности пространства биологических признаков с наименьшими потерями информации. Например, в работе [56] рассмотрены некоторые аспекты применения метода главных компонент для решения задачи распознавания изображений. Предложен алгоритм многоуровневой линейной конденсации для вычисления главных компонент больших наборов изображений. Данный алгоритм использует аппроксимацию, которая позволяет сократить порядок матриц с сохранением собственных значений в заданном диапазоне.

В работе [57] предлагается метод распознавания личности по походке, регистрируемой с использованием видеосъемки в оптическом диапазоне, состоящий в выделении движущегося человека на видеоряде с последующей нормализацией размера и снижением размерности с использованием метода главных компонент и классификацией с использованием метода опорных векторов. Экспериментальные исследования показали высокую точность распознавания личности (не менее 90 %).

Недостатками метода главных компонент являются: потеря некоторой информации, стандартизация данных по единичной шкале; ковариационную матрицу трудно оценить точным образом, даже простейшая инвариантность не может быть зафиксирована *PCA* [58], чувствительность к входным данным.

Скрытые марковские модели (*Hidden Markov Models, HMM*) учитывают пространственно-временные характеристики сигналов, поэтому получили широкое применение в распознавании изображений лиц. Например, в работе [59] предлагается метод распознавания лиц на основе скрытой марковской модели. Предложенный метод снижает вычислительную сложность распознавания лиц на основе *HMM*, при этом немного повышая скорость распознавания.

Недостатками *HMM* являются: ограниченные применения при не очень большом объеме БД; необходимость подбирать параметры модели для каждой конкретной БД.

Метод гибкого сравнения на графах (*Elastic Bunch Graph Matching*) – метод компьютерного зрения для распознавания объектов или классов объектов в изображении на основе графического представления, извлеченного из других изображений. Данный метод использовался для распознавания и анализа лиц, а также для жестов и других классов объектов.

Недостатками метода гибкого сравнения на графах являются: вычислительная сложность процесса распознавания, а также низкая технологичность при запоминании новых эталонов.

Заключение

Распознавание человека по движениям тела и лицу в видеонаблюдениях играет важную роль в обеспечении безопасности на объектах с массовым скоплением людей, т.к. позволяет раскрыть личность потенциального нарушителя, совершающего преступления, а также предупредить преступления.

Рассмотрены структурные схемы слияния биометрических признаков и принципы их работы, приведено сравнение методов и технологий распознавания человека по движению тела и лицу, показаны их достоинства и недостатки.

На основании проведенного анализа могут быть сформулированы общие рекомендации для разработчиков СКБ по видеонаблюдению:

- СКБ необходимо проектировать как систему поддержки принятия решений;
- необходимо разработать базовую архитектуру СКБ и процесс аутентификации по движению тела и лицу (при комбинировании технологически отработанных методов);
- информационной основой проектируемой СКБ должны быть значения различных биометрических признаков;
- рекомендуется применять мультимодальные системы, которые позволяют объединять несколько биометрических признаков;
- необходимо определить, как будет осуществляться распознавание нарушителя - по одному или по нескольким признакам, как они получены (с одного или нескольких датчиков), и выбрать уровень слияния биометрических признаков.

Список источников

- [1] Состояние преступности в России за январь - ноябрь 2022 года. Москва. Генеральная прокуратура Российской Федерации. 60 с. <http://crimestat.ru/analytics>.
- [2] Типовая модель действий нарушителя, совершающего на объекте образования преступление террористической направленности в формах вооруженного нападения, размещения взрывного устройства, захвата заложников, 41 с. <https://minobrnauki.gov.ru>.
- [3] **Зенов А.Ю.** Комплексный подход к обнаружению, классификации и распознаванию нарушителя на охраняемой территории // *Известия высших учебных заведений. Поволжский регион. Технические науки*. 2012. № 2 (22). С.23-32.
- [4] **Смирнов А.М.** К вопросу о фундаментально-теоретической модели изучения личности преступника // *Гуманитарные, социально-экономические и общественные науки*. 2020. № 1. С.1-5. DOI: 10.23672/SAE.2020.1.53256.
- [5] **Ким Е.В., Ру П.Г.** Личность преступника: криминологический анализ // *Ученые заметки ТОГУ*. 2013. Т.4. № 4. С.402-407.
- [6] **Абельцев С.Н.** О личности преступника и практической значимости ее изучения // *Вестник Тамбовского университета. Гуманитарные науки*. 2000. № 3 (19). С.83-85.
- [7] **Копылова Г.К., Прозоров А.В.** Психология в деятельности органов внутренних дел. М.: ЦОКР МВД России, 2006. 236 с.
- [8] **Герасименко В.А.** Основы защиты информации в АС. М.: Наука, 2001. 178 с.
- [9] **Ольшанский Д.В.** Психология терроризма. М.: Юрайт, 2015. 194 с.
- [10] **Стуколова Л.С., Закирова Д.А.** Психология современного терроризма // *Аллея науки*. 2018. Т.5. № 6 (22). С.546-549.
- [11] **Jain, A.K., Ross A., Prabhakar S.** An introduction to biometric recognition // *IEEE Trans. Circuits Syst. Video Technol.* 2004. Vol.14. No.1. P.4-20. DOI: 10.1109/TCSVT.2003.818349.
- [12] **Газизулин А.И.** Криминологическая характеристика личности террориста нашего времени // *NovaInfo.Ru*. 2018. Т.1. № 84. С.161-165. DOI: 10.24411/2312-0444-2021-4-292-297.
- [13] **Аренова Л.К., Набиева Е.А.** Личность лица, совершившего акт терроризма // *Актуальные проблемы права и государства в XXI веке*. 2018. Т. 10. № 1. С.38-46.
- [14] **Пимакова О.Г.** Личность преступника террориста // *Виктимология*. 2018. № 4 (18). С. 54-58.
- [15] **Тарчоков Б.А.** Мотивационные особенности вовлечения молодежи в террористическую деятельность // *Историческая и социально-образовательная мысль*. 2015. № 6. С.211-213. DOI: 10.17748/2075-9908-2015-7-6/1-211-213.
- [16] **Шендра С.Е., Хонин А.А., Войлошников А.Д.** Психологический портрет личности террориста // *Молодой ученый*. 2022. № 13.1 (408.1). С.32-33.
- [17] **Малеева М., Кленникова Е., Мартынова Я.** Психологический портрет террориста, 7 с. <https://scienceforum.ru/2017/article/2017038036>.
- [18] **Лепешкин Н.Я., Василин В.Г., Обирин А.И., Талынев В.Е.** Психологические основы терроризма и анти-террористической деятельности в современных условиях. Хабаровск: Хабаровский пограничный институт Федеральной службы безопасности Российской Федерации, 2008, 348 с.
- [19] **Bouchrika I., Jain S., Arora S., Singh U.P.** A Survey of using biometrics for smart visual surveillance: gait recognition // *Advanced Sciences and Technologies for Security Applications*. 2018. P.3-23. DOI: 10.1007/978-3-319-68533-5_1.
- [20] **Singh J.P.** Vision-based gait recognition: a survey // *IEEE Access*. 2018. Vol.6. P.70497-70527. DOI: 10.1109/ACCESS.2018.2879896.
- [21] **Priyanka S, Kaur M.** Classification in pattern recognition: a review // *IJARCSSE All Rights Reserved*. 2013. Vol. 3. P.298-306.
- [22] **Кобец П.Н.** О комплексном изучении личности преступника в отечественной криминологии // *Проблемы развития личности: матер. междунаrod. науч.-практ. конф. Прага, 2013*. С.93-95.
- [23] **Аверьянова Т.В., Белкин Р.С., Корухов Ю.Г., Россинская Е.Р.** Криминалистика: учебник для вузов. М.: НОРМА (НОРМА-ИНФРА М), 2001. 908 с.
- [24] **Ryszard C.** Multimodal biometrics for person authentication. 2020, 516 p. https://www.researchgate.net/publication/345480873_Multimodal_Biometrics_for_Person_Authentication. DOI: 10.5772/intechopen.85003.
- [25] **Evans N., Marcel S., Ross A., Teoh ABJ.** Biometrics security and privacy protection // *IEEE Signal Process Mag.* 2015. Vol.32(5). P.17-18. DOI: 10.1109/MSP.2015.2443271.
- [26] **Anil K.Jain, Nandakumar Karthik, Ross Arun.** 50 years of biometric research: Accomplishments, challenges, and opportunities // *Pattern Recognition Letters*. 2016. Vol.79. P.80-105. DOI:10.1016/j.patrec.2015.12.013.

- [27] Криминалистическое исследование внешних признаков человека (габитоскопия). Физиогномика убийцы или как определить преступника по внешности: словесный портрет с полным описанием, 17.04.2021. <https://lehre.ru/do/kriminalisticheskoe-issledovanie-vneshnih-priznakov-cheloveka-gabitoskopiya.html>.
- [28] **Федюнина А.П.** Выявление характерологических признаков и составление психологического портрета возможного нарушителя и лояльного сотрудника в сфере информационной безопасности // *Вестник АГТУ*. 2007. № 4. С. 231-236.
- [29] Словесный портрет. Описание внешности человека по методу словесного портрета. 05.11.2021. <https://goaravetisyan.ru/slovesnyi-portret-opisanie-vneshnosti-cheloveka-po-metodu-slovesnogo>.
- [30] **Акимов А.А., Мустафина С.А.** Обзор современных методов искусственного интеллекта по распознаванию девиантного поведения индивида // *Вестник Технологического университета*. 2020. Т. 23. № 8. С.69-79.
- [31] **Siddiqui A.M.N., Telgad R., Deshmukh P.D.** Multimodal biometric systems: study to improve accuracy and performance // *International Journal of Current Engineering and Technology*. 2014. Vol.4. No.1. P.165-171.
- [32] **Gad R., Nawal El-Fishawy, Ayman El-Sayed, Zorkany M.** Multi-biometric systems: a state of the art survey and research directions // *International Journal of Advanced Computer Science and Applications*. 2015. Vol. 6(6). P. 128-138. DOI:10.14569/IJACSA.2015.060618.
- [33] **Ayodele Oloyede, Aderonke Adegbenjo.** Current practices in information fusion for multimodal biometrics // *American Journal of Engineering Research (AJER)*. 2017. Vol. 6. P.148-154.
- [34] **Ковалев С.М., Колоденкова А.Е., Шнасель В.** Интеллектуальные технологии слияния данных при диагностировании технических объектов // *Онтология проектирования*. 2019. Т.9. №1(31). С.152-168. DOI: 10.18287/2223-9537-2019-9-1-152-168.
- [35] **Долгий А.И., Колоденкова А.Е., Ковалев С.М.** Проблемы и методы слияния разнородных данных в гибридных интеллектуальных системах // *Гибридные и синергетические интеллектуальные системы*: матер. IV Всерос. Пospelовской конференции с междунар. участием. 2018. С.181-187.
- [36] **Ailon N., Charikar M., Newman A.** Aggregating inconsistent information: ranking and clustering // In *Proceedings of 37th Annual ACM Symposium on Theory of Computing (STOC)*. 2015. P.684-693.
- [37] **AlMahafzah H., AlRwashdeh M.Z.** A Survey of multibiometric systems // *International Journal of Computer Applications*. 2012. Vol.43. No.15. P.36-43.
- [38] **Soltane M., Bakhti M.** Multi-modal biometric authentications: concept issues and applications strategies // *International Journal of Advanced Science and Technology*. 2012. Vol. 48. P.1-38.
- [39] **Sathish G., Saravanan S.V., Narmadha S., Maheswari S.U.** Multi-algorithmic iris recognition // *International Journal of Computer Applications*. 2012. Vol.38. No.11. P.13-21.
- [40] **Mwaura G.W., Mwangi W., Otieno C.** Multimodal biometric system: fusion of face and fingerprint biometrics at match score fusion level. *International Journal of Scientific & Technology Research*. 2017. Vol. 6. P.41-49.
- [41] **Ross A., Jain A.** Information fusion in biometrics. *Pattern Recognition Letters*. 2003. P.2115-2125.
- [42] **Delac K., Grgic M.** A Survey of biometric recognition methods // *46th International Symposium, ELMAR-2004*, 2004. P.184-193.
- [43] **Aly O.M., Salama G.I., Mahmoud T.A., Onsi H.M.** A multimodal biometric recognition system using feature fusion based on PSO // *International Journal of Advanced Research in Computer and Communication Engineering*. 2013. Vol.2. P.4336-4343. DOI:10.1007/s10916-019-1391-5.
- [44] **Ghayoumi M.** A review of multimodal biometric systems: fusion methods and their applications. *IEEE/ACIS 14th International Conference Computer and Information Science (ICIS)*. 2015. P.131-136. DOI: 10.1109/ICIS.2015.7166582.
- [45] **Radha N., Kavitha A.** Rank level fusion using fingerprint and iris biometrics // *Indian Journal of Computer Science and Engineering (IJCSE)*. 2011. Vol.2. No.6. P.917-923.
- [46] **Haryati Jaafar, Dzati Athiar Ramli.** A review of multibiometric system with fusion strategies and weighting factor // *International Journal of Computer Science Engineering (IJCSE)*. 2013. Vol.2. No.4. P.158-165.
- [47] **Abderrahmane H., Noubel G., Ziet L., Zahid A., Dipankar D.** Weighted quasi-arithmetic mean based score level fusion for multi-biometric systems // *IET Biometrics*. 2020. Vol. 9(3). P.91-99. DOI: 10.1049/iet-bmt.2018.5265.
- [48] **Nandakumar K., Chen Y., Dass C., Jain A.K.** Likelihood ratio based biometric score fusion // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2007. P.1-9.
- [49] **Ross A., Jain A.K.** Fusion techniques in multibiometric systems // *Face Biometrics for Personal Identification*. 2007. P.185-212.
- [50] **Coşkun M., Uçar A., Yildirim Ö., Demir Y.** Face recognition based on convolutional neural network // *2017 International Conference on Modern Electrical and Energy Systems (MEES)*. 2017. P.376-379.
- [51] **Уздяев М.Ю., Яковлев Р.Н., Дударенко Д.М., Жебрун А.Д.** Идентификация человека по походке в видеопотоке // *Известия Юго-Западного государственного университета*. 2020. №24(4). С.57-75. DOI: 10.21869/2223-1560-2020-24-4-57-75.

- [52] **Sokolova A., Konushin A.** Pose-based deep gait recognition // IET Biometrics. 2019. Vol.8. P.134-143. DOI: 10.48550/arXiv.1710.06512.
- [53] **Ivanna T., Iwan S., Andreas F.** Face recognition between two person using kernel principal component analysis and support vector machines. *International Journal on Electrical Engineering and Informatics*. 2010. Vol.2. P.53-61. DOI:10.15676/ijeei.2010.2.1.5.
- [54] **Maret Y., Oberson D., Gavrilova M.** Real-time embedded system for gesture recognition // 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2018. P.30-34. DOI: 10.1109/SMC.2018.00014.
- [55] **Chia-Feng, Chung-Wei Liang, Lee Chiung-Ling, Chung I-Fang** Vision-based human body posture recognition using support vector machines // Proceedings: 4th International Conference on Awareness Science and Technology. 2012. P.150-155. DOI:10.1109/iCAwST.2012.6469605.
- [56] **Мокеев А.В., Мокеев В.В.** Об эффективности распознавания лиц с помощью линейного дискриминантного анализа и метода главных компонент // *Вестник Южно-Уральского государственного университета. Компьютерные технологии, автоматическое управление, радиоэлектроника*. 2013. Vol.13. No.3. P.61-70.
- [57] **Струкова О.В., Ширипова Л.В., Мясников Е.В.** Распознавание личности по походке: опыт использования метода главных компонент и машины опорных векторов // *Информационные технологии и нанотехнологии*: сб. тр. IV междунар. конф. и молодеж. школы (ИТНТ-2018). Самара: Новая техника, 2018. С.822-832.
- [58] **Li C., Diao Y., Ma H., Li Y.** A Statistical PCA method for face recognition // *Intelligent Information Technology Application*. 2008. P.376-380.
- [59] **Nefian Ara, Hayes Monson.** Face detection and recognition using hidden Markov models. 1998. Vol.1. P.141-145.

Сведения об авторе



Колоденкова Анна Евгеньевна, 1982 г. рождения. Окончила Уфимский государственный авиационный технический университет в 2004 г., д.т.н. (2017). Заведующая кафедрой «Информационные технологии» Самарского государственного технического университета. Член Российской ассоциации искусственного интеллекта. В списке научных трудов более 180 работ в области атомной энергетики, программной инженерии, системного анализа, интеллектуальных и биометрических систем, мягких вычислений, экспертной поддержки принятия решений, технической диагностики и мониторинга состояния промышленного оборудования. AuthorID (РИНЦ): 175446; ORCID: 0000-0002-9784-1871; Author ID (Scopus): 57190670136;

Researcher ID (WoS): F-1341-2018. anna82_42@mail.ru.

Поступила в редакцию 09.01.2023, после рецензирования 05.02.2023. Принята к публикации 16.02.2023.



Scientific article

DOI: 10.18287/2223-9537-2023-13-1-55-74

Ontology of human identification by face and body motions in video surveillance systems

© 2023, A.E. Kolodenkova

Samara State Technical University, Samara, Russia

Abstract

At the present stage of advancing information technology, the development of models and recognition methods by body movements and faces in video surveillance systems is a topical problem. This task is essential for security issues, especially at facilities with mass gatherings to counter a terrorism-related crime. The paper presents a classification of the main biometric features and parameters that characterize a potential violator. This classification has been developed for security control systems and access systems of enterprises. A block diagram of merging biometric data and violator recognition by body motions and face which can be used as the basis for the development of security control systems is

proposed. The types of systems and methods of human recognition by body movements and face are considered, their advantages and disadvantages are revealed. It is noted that for accurate violator recognition under a set of biometric features, it is reasonable to use a combination of recognition methods which will allow to make the right decisions regarding the identification of a potential violator. This paper attempts to consider the main aspects related to human recognition by body movements and face in video surveillance in general, in contrast to well-known works devoted to individual biometric features.

Key words: *recognition methods, biometric features, potential violator, security control systems, video surveillance systems, ontology.*

Citation: Kolodenkova AE. Ontology of human identification by face and body motions in video surveillance systems [In Russian]. *Ontology of designing*. 2023. 13(1): 55-74. DOI: 10.18287/2223-9537-2023-13-1-55-74.

Conflict of interest: The author declares no conflict of interest.

List of figures and tables

- Figure 1 - Ontology of the violator identification process
- Figure 2 - Types of biometric features to identify potential violators
- Figure 3 - Classification of the main biometric features and indicators characterizing a potential violator
- Figure 4 - Types of recognition systems
- Figure 5 - Block diagram of merging biometric data and recognition of the violator by body and face movements
- Figure 6 - Merging level classification in recognition systems
- Figure 7 - Sensor-level merge structure
- Figure 8 - Feature level merge structure
- Figure 9 - Merge structure at the mapping level
- Figure 10 - Merge structure at the decision-making level
- Figure 11 - Biometric authentication process
- Figure 12 - Classification of methods and technologies of human recognition by body movement and face
- Table 1 - Fragment of the concepts specification of the ontology of the violator identification
- Table 2 - Comparison of biometric features

References

- [1] Portal of legal statistics of the Prosecutor General's Office of the Russian Federation, 60 p. <http://crimestat.ru/analytics>.
- [2] A typical model of the actions of a violator who commits a terrorist-oriented crime at an educational facility in the forms of an armed attack, placement of an explosive device, hostage-taking [In Russian]. 41 p. <https://minobrnauki.gov.ru>.
- [3] **Zenov AYu.** An integrated approach to the detection, classification and recognition of an intruder in a protected area [In Russian]. *News of higher educational institutions. Volga region. Technical sciences*. 2012; 2 (22): 23-32.
- [4] **Smirnov AM.** On the question of a fundamental theoretical model for studying the personality of a criminal [In Russian]. *Humanities, socio-economic and social sciences* 2020; 1: 1-5. DOI: 10.23672/SAE.2020.1.53256.
- [5] **Kim EV, Ri PG.** Criminal identity: criminological analysis [In Russian]. *Scientific notes of TOGU*. 2013; 4:402-407.
- [6] **Abeltsev SN.** On the identity of the criminal and the practical significance of its study [In Russian]. *Bulletin of the Tambov University. Humanities*. 2000; 3 (19): 83-85.
- [7] **Kopylova GK, Prozorov AV.** Psychology in the activities of internal affairs bodies [In Russian]. Moscow: Central Committee of the Ministry of Internal Affairs of Russia, 2006. 236 p.
- [8] **Gerasimenko VA.** Fundamentals of information protection in AS [In Russian]. Moscow: Nauka, 2001. 178 p.
- [9] **Olshansky DV.** Psychology of terrorism [In Russian]. Moscow: Yurayt, 2015. 194 p.
- [10] **Stukolova LS, Zakirova DA.** Psychology of modern terrorism [In Russian]. *Alley of Science*. 2018; 5: 546-549.
- [11] **Jain AK, Ross A, Prabhakar S.** An introduction to biometric recognition // *IEEE Trans. Circuits Syst. Video Technol* 2004; 14: 4-20. DOI: 10.1109/TCSVT.2003.818349.
- [12] **Gazizullin AI.** Criminological characteristics of the personality of a terrorist of our time [In Russian]. *NovaIn-fo.Ru* 2018; 1: 161-165. DOI: 10.24411/2312-0444-2021-4-292-297.
- [13] **Arenova LK, Nabieva EA.** Personality of a person who committed an act of terrorism [In Russian]. *Actual problems of law and the state in the XXI century*. 2018; 10: 38-46.
- [14] **Primakova OG.** The identity of the criminal terrorist [In Russian] *Victimology*. 2018; 4 (18): 54-58.

- [15] **Tarchokov BA.** Motivational features of youth involvement in terrorist activities [In Russian]. *Historical and socio-educational thought*. 2015; 6: 211-213 DOI: 10.17748/2075-9908-2015-7-6/1-211-213.
- [16] **Shendra SE, Khanin AA, Voyloshnikov AD.** Psychological portrait of a terrorist's personality [In Russian]. *Young scientist*. 2022; 13.1 (408.1): 32-33.
- [17] **Maleeva M, Klenikova E, Martynova Ya.** Psychological portrait of a terrorist [In Russian]. 7 p. <https://scienceforum.ru/2017/article/2017038036>.
- [18] **Lepeshkin NYa, Vasilin VG, Obirin AI, Talynev VE.** Psychological foundations of terrorism and anti-terrorist activity in modern conditions [In Russian]. Educational and methodological manual. - Khabarovsk: Khabarovsk Border Institute of the Federal Security Service of the Russian Federation, 2008, 348 p.
- [19] **Bouchrika I, Jain S, Arora S, Singh UP.** A Survey of using biometrics for smart visual surveillance: gait recognition. *Advanced Sciences and Technologies for Security Applications*. 2018: 3-23. DOI:10.1007/978-3-319-68533-5_1.
- [20] **Singh JP.** Vision-based gait recognition: a survey. *IEEE Access*. 2018; 6: 70497-70527. DOI: 10.1109/ACCESS.2018.2879896.
- [21] **Priyanka S, Kaur M.** Classification in pattern recognition: a review. *IJARCSSE All Rights Reserved* 2013; 3: 298-306.
- [22] **Kobets PN.** On the complex study of the personality of a criminal in domestic criminology [In Russian]. Problems of personality development: mater. international scientific-practical. conf. Prague, 2013: 93-95.
- [23] **Averyanova TV, Belkin RS, Koruhov YG, Rossinskaya ER.** Criminalistics: textbook for universities [In Russian]. - Moscow: NORMA, 2001. 908 p.
- [24] **Ryszard C.** Multimodal biometrics for person authentication. 2020, 516 p. https://www.researchgate.net/publication/345480873_Multimodal_Biometrics_for_Person_Authentication. DOI: 10.5772/intechopen.85003.
- [25] **Evans N., Marcel S., Ross A., Teoh ABJ** Biometrics security and privacy protection // *IEEE Signal Process Mag.* 2015; 32(5): 17-18. DOI: 10.1109/MSP.2015.2443271.
- [26] **Anil K.Jain, Nandakumar Karthik, Ross Arun** 50 years of biometric research: Accomplishments, challenges, and opportunities // *Pattern Recognition Letters* 2016; 79: 80-105. DOI:10.1016/j.patrec.2015.12.013.
- [27] Forensic examination of external signs of a person (habitoscopy). Physiognomy of a murderer or how to identify a criminal by appearance: a verbal portrait with a full description [In Russian]. 31 p. <https://lehr.ru/do/kriminalisticheskoe-issledovanie-vneshnih-priznakov-cheloveka-gabitoskopiya.html>.
- [28] **Fedyunina AP.** Identification of characterological signs and drawing up a psychological portrait of a possible violator and a loyal employee in the field of information security [In Russian]. *Bulletin of AGTU* 2007; 4: 231-236.
- [29] Verbal portrait. Description of a person's appearance by the method of verbal portrait. Verbal description of a person's appearance criminology [In Russian]. <https://goaravetisyan.ru/slovesnyi-portret-opisanie-vneshnosti-cheloveka-po-metodu-slovesnogo>.
- [30] **Akimov AA, Mustafina SA.** Review of modern artificial intelligence methods for recognizing deviant behavior of an individual [In Russian]. *Bulletin of the Technological University* 2020; 8: 69-79.
- [31] **Almas M. N. Siddiqui, Rupali Telgad, Prapti D. Deshmukh** Multimodal biometric systems: study to improve accuracy and performance // *International Journal of Current Engineering and Technology* 2014; 4: 165-171.
- [32] **Gad R., Nawal El-Fishawy, Ayman El-Sayed, Zorkany M.** Multi-biometric systems: a state of the art survey and research directions. *International Journal of Advanced Computer Science and Applications* 2015; 6(6): 128-138. DOI: 10.14569/IJACSA.2015.060618.
- [33] **Ayodele Oloyede, Aderonke Adegbenjo.** Current practices in information fusion for multimodal biometrics // *American Journal of Engineering Research (AJER)* 2017; 6: 148-154.
- [34] **Kovalev SM, Kolodenkova AE, Snasel V.** Intelligent data fusion technologies in the diagnosis of technical objects [In Russian]. *Design Ontology* 2019; 1(31): 152-168. DOI: 10.18287/2223-9537-2019-9-1-152-168.
- [35] **Dolgiy AI, Kolodenkova AE, Kovalev SM.** Problems and methods of merging heterogeneous data in hybrid intelligent systems [In Russian]. *Hybrid and synergetic intelligent systems* 2018: 181-187.
- [36] **Ailon N., Charikar M., Newman A.** Aggregating inconsistent information: ranking and clustering // In *Proceedings of 37th Annual ACM Symposium on Theory of Computing (STOC)* 2015: 684-693.
- [37] **AlMahafzah H, AlRwashdeh MZ.** A Survey of multibiometric systems // *International Journal of Computer Applications* 2012; 43: 36-43.
- [38] **Mohamed Soltane, Mimen Bakhti.** Multi-modal biometric authentications: concept issues and applications strategies. *International Journal of Advanced Science and Technology* 2012; 48: 1-38.
- [39] **Sathish G., Saravanan SV, Narmadha S, Maheswari SU.** Multi-algorithmic iris recognition. *International Journal of Computer Applications* 2012; 38: 13-21.
- [40] **Mwaura GW, Mwangi W, Otieno C.** Multimodal biometric system: fusion of face and fingerprint biometrics at match score fusion level // *International Journal of Scientific & Technology Research* 2017; 6: 41-49.

- [41] **Arun Ross, Anil Jain** Information fusion in biometrics // Pattern Recognition Letters 2003: 2115-2125.
- [42] **Delac K, Grgic M.** A Survey of biometric recognition methods // 46th International Symposium, ELMAR-2004, 2004: 184-193.
- [43] **Aly OM, Salama GI, Mahmoud TA, Onsi HM.** A multimodal biometric recognition system using feature fusion based on PSO // International Journal of Advanced Research in Computer and Communication Engineering 2013; 2: 4336-4343. DOI:10.1007/s10916-019-1391-5.
- [44] **Ghayoumi M.** A review of multimodal biometric systems: fusion methods and their applications // IEEE/ACIS 14th International Conference Computer and Information Science (ICIS) 2015: 131-136. DOI: 10.1109/ICIS.2015.7166582.
- [45] **Radha N, Kavitha A.** Rank level fusion using fingerprint and iris biometrics // Indian Journal of Computer Science and Engineering (IJCE) 2011; 2: 917-923.
- [46] **Haryati Jaafar, Dzati Athiar Ramli** A review of multibiometric system with fusion strategies and weighting factor // International Journal of Computer Science Engineering (IJCE) 2013; 2: 158-165.
- [47] **Abderrahmane H, Noubel G, Ziet L, Zahid A, Dipankar D.** Weighted quasi-arithmetic mean based score level fusion for multi-biometric systems. *IET Biometrics* 2020; 9(3): 91-99. DOI: 10.1049/iet-bmt.2018.5265.
- [48] **Nandakumar K., Chen Y., Dass C., Jain A.K.** Likelihood ratio based biometric score fusion // IEEE Transactions on Pattern Analysis and Machine Intelligence 2007: 1-9.
- [49] **Ross A., Jain A.K.** Fusion techniques in multibiometric systems // Face Biometrics for Personal Identification 2007: 185-212.
- [50] **Coşkun M, Uçar A, Yildirim Ö, Demir Y.** Face recognition based on convolutional neural network // 2017 International Conference on Modern Electrical and Energy Systems (MEES) 2017: 376-379.
- [51] **Uzdayev MYu, Yakovlev RN, Dudarenko DM, Zhebrun AD.** Identification of a person by gait in a video stream [In Russian]. Proceedings of the Southwestern State University 2020; 24(4): 57-75. DOI: 10.21869/2223-1560-2020-24-4-57-75.
- [52] **Sokolova A, Konushin A.** Pose-based deep gait recognition // IET Biometrics 2019; 8: 134-143. DOI: 10.48550/arXiv.1710.06512.
- [53] **Timotius Ivanna, Setyawan Iwan, Febrianto Andreas** Face recognition between two person using kernel principal component analysis and support vector machines // International Journal on Electrical Engineering and Informatics 2010; 2: 53-61. DOI:10.15676/ijeei.2010.2.1.5.
- [54] **Maret Y., Oberson D., Gavrilova M.** Real-time embedded system for gesture recognition // 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) 2018: 30-34. DOI: 10.1109/SMC.2018.00014.
- [55] **Chia-Feng, Chung-Wei Liang, Lee Chiung-Ling, Chung I-Fang** Vision-based human body posture recognition using support vector machines // Proceedings: 4th International Conference on Awareness Science and Technology 2012: 150-155. DOI: 10.1109/iCAwST.2012.6469605.
- [56] **Mokeyev AV, Mokeyev VV.** On the effectiveness of facial recognition using linear discriminant analysis and the method of principal components [In Russian]. *Bulletin of the South Ural State University. Computer technology, automatic control, radio electronics* 2013; 13: 61-70.
- [57] **Strukova OV, Shiripova LV, Myasnikov EV.** Personality recognition by gait: the experience of using the principal component method and the support vector machine [In Russian]. *Information technologies and nanotechnologies* 2018: 822-832.
- [58] **Li C, Diao Y, Ma H, Li Y.** A Statistical PCA Method for Face Recognition. *Intelligent Information Technology Application* 2008: 376-380.
- [59] **Nefian Ara, Hayes Monson** Face detection and recognition using hidden Markov models 1998; 1: 141-145.

About the author

Anna Evgenievna Kolodenkova (b. 1982) graduated from the Ufa State Aviation Technical University (Ufa-city) in 2004, D. Sc. Eng. (2017). She is an Associate Professor and the Head of «Information technologies» Department at Samara State Technical University. She is a member of Russian Association of Artificial Intelligence. She is a co-author of about 180 scientific articles and abstracts in the field of nuclear energy, software engineering, system analysis, intelligent and biometric systems, soft computing, expert decision support and technical diagnostics and monitoring of industrial equipment condition. AuthorID (RCI): 175446. ORCID: 0000-0002-9784-1871. Author ID (Scopus): 57190670136; Researcher ID (WoS): F-1341-2018. anna82_42@mail.ru.

Received January 9, 2023. Revised February 05, 2023. Accepted February 16, 2023.



Информационная система поддержки предварительного проектирования малых космических аппаратов

© 2023, И.С. Ткаченко ✉, С.Л. Сафронов, М.Д. Коровин, М.А. Иванушкин, А.В. Крестина

Самарский национальный исследовательский университет имени академика С.П. Королёва, Самара, Россия

Аннотация

С развитием информационных технологий значительно выросла роль систем поддержки процесса проектирования технических изделий. Возможности систем управления знаниями позволяют обеспечивать общность терминологии, определений объектов и отношений между ними. Использование формальной семантики необходимо для поддержки многократного использования знаний. В статье рассматриваются пути повышения эффективности процесса проектирования малых космических аппаратов и создания средств для обмена знаниями о функциональных характеристиках типовых подсистем таких аппаратов. Рассматривается создание информационной системы поддержки предварительного проектирования малых космических аппаратов, которая содержит информацию, связанную с отношениями между функциями и компонентами подсистем и их характеристиками. Созданная система позволит упростить синтез различных конфигураций изделий. При разработке информационной системы использовались современные подходы в области структурных и функциональных классификаций систем малых космических аппаратов, существующие в отрасли таксономии, подходы к созданию структур баз данных. База данных информационной системы содержит набор характеристик широкой номенклатуры оборудования малых космических аппаратов. В качестве примера применения информационной системы поддержки проектирования малых космических аппаратов рассмотрен процесс выбора характеристик системы электропитания для перспективного малого космического аппарата, предназначенного для обеспечения связи между подвижными наземными терминалами

Ключевые слова: информационная система, проектирование, малый космический аппарат, космическая платформа, база данных, бортовая система, пользовательский интерфейс.

Цитирование: Ткаченко И.С., Сафронов С.Л., Коровин М.Д., Иванушкин М.А., Крестина А.В. Информационная система поддержки предварительного проектирования малых космических аппаратов // Онтология проектирования. 2023. Т.13, №1(47). С.75-89. DOI:10.18287/2223-9537-2023-13-1-75-89.

Финансирование: работа выполнена при финансовой поддержке Программы развития Самарского национального исследовательского университета имени академика С.П. Королёва на 2021-2030 годы в рамках Программы стратегического академического лидерства “Приоритет 2030”, соглашение ПР-НУ/2.1-17-2022 от 11.07.2022 г.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

Исследования Земли из космоса с применением космических аппаратов (КА) проводятся уже более 60 лет, предоставляя людям всё большее количество необходимой и полезной информации. Без актуальной и качественной информации о состоянии объектов и явлений, связанных с поверхностью Земли и атмосферой, становится невозможным ведение эффективной деятельности в области метеорологии, экологии, экономики и других областях.

Всё большую популярность при решении актуальных задач приобретает использование малых КА (МКА) и космических группировок на их основе. По сравнению с относительно

большими КА они быстрее проходят стадии жизненного цикла [1], требуют значительно меньших затрат, что достигается информационной поддержкой проектно-конструкторской деятельности, начиная от технических предложений и заканчивая функционированием на орбите. МКА максимально используют коммерчески доступные современные технологии, обеспечивающие требуемую функциональность.

Разработка новых способов и развитие технологий проектирования и конструирования на базе подходов системной инженерии позволит не только решить актуальную проблему создания лёгких, эффективных и надёжных МКА, но и способствует снижению стоимости серийных изделий. За последние десятилетия было проведено много исследований в области повышения эффективности системной инженерии [2]. Системы управления знаниями позволяют упростить представление природы и структуры инженерных систем и их компонентов, помочь различным заинтересованным сторонам лучше понять проблемы, характерные для сложных технических систем и их социотехнического окружения [3].

1 Контекст задачи проектирования

С формальной точки зрения контекст представляет собой определённую систему отсчёта, пространство имён. В более широком значении контекст — среда, в которой существует объект [4]. В большинстве моделей систем в области инженерного проектирования контекст оценивается редко или просто определяется как нечто, окружающее результаты проектной деятельности и придающее им смысл [5]. В настоящее время компании развиваются в условиях жёсткой конкуренции и постоянно меняющейся среды, и организациям приходится непрерывно адаптировать свои структуры, процессы и продукты для достижения успеха. Следовательно, сложность моделей, необходимых для представления систем, растёт, понятие контекста становится решающим.

Известны стандарты и обзоры, посвящённые различным аспектам системной инженерии: задачи и инструменты на основе моделей для сервис-ориентированной системной инженерии [6, 7]; подходы к архитектуре типа «система из систем» [8] и представлению знаний [9, 10], атрибутам качества [11], системной интеграции [12] и разработке требований [13, 14].

По результатам анализа приведённых работ можно заключить, что общие подходы к информационной поддержке проектирования при применении в конкретной области должны быть адаптированы под её специфику. Для этого необходимы разработка как подробных моделей предметной области (ПрО), так и описаний подходов в целом.

Хотя методы системной инженерии и своды лучших практик формализованы и широко представлены в различных монографиях и стандартах, на сегодняшний момент всё ещё не преодолены два существенных вызова: трудности с разработкой систем в рамках бюджета и в срок [15] и значительные потери ресурсов, связанные с исправлением ошибок [16-18].

В [15] выделены причины отсутствия успеха в преодолении этих вызовов, среди которых: ограниченность описаний лучших практик в форме метамоделей, отсутствие общепринятой и согласованной терминологии, неэффективное сотрудничество из-за непонимания и неправильной интерпретации.

В частности, отмечается необходимость гармонизации в интерпретации многочисленных стандартов и различных моделей [19]. Онтологии используются для обеспечения семантической совместимости как между людьми, так и между людьми и компьютерами. В [20] отмечается, что «системная наука и инженерия нуждаются в чётко определённой фундаментальной, универсальной, общей, необходимой и достаточной онтологии, которая бы лежала в основе понятий и терминов, которые она использует, чтобы они были точными и однозначными».

ми». В [21] подчёркивается ключевая роль единой информационной среды в контексте проектирования сложных технических систем.

Целью создания информационной системы (ИС) поддержки предварительного проектирования МКА является формирование единого информационного пространства, которое содержит сведения о работе МКА, его бортовых систем (БС) и об эффективности их совместного функционирования.

Настоящая работа посвящена обеспечению общего *пользовательского* контекста процессов предварительного проектирования МКА при помощи традиционных ИС на основе таксономических моделей Про.

2 Модель Про

МКА разрабатываются в соответствии с техническими заданиями, в которых определены их тактико-технические характеристики. Эти характеристики определяются по результатам проектирования системы, состоящей из наземного и космического компонентов.

Одним из направлений в проектировании, обеспечивающих эффективное решение задач создания МКА в короткие сроки, является использование платформенных принципов. Переход к построению МКА на базе крупных компонентов позволяет создавать семейство МКА, отличающихся по составу целевой аппаратуры, но имеющих общую платформу, включающую в себя набор обеспечивающих систем.

Платформа является базовой частью создаваемого на её основе МКА и представляет собой несущую конструкцию, снабжённую служебными системами и устройствами для функционирования целевой аппаратуры различного назначения, но с близкими эксплуатационными требованиями. Платформа должна состоять из базового и целевого состава обеспечивающих БС. Базовый состав одинаков для всего семейства МКА, целевой формируется на этапе адаптации платформы под конкретную задачу. Создание МКА на базе универсальных платформ широко применяется отечественными и зарубежными проектными организациями. Перечень основных технических требований, которые должны быть учтены при составлении задания на проектирование МКА, включает [22]: назначение; живучесть и стойкость к внешним воздействиям; надёжность; удобство технического обслуживания; транспортабельность; безопасность; технологичность; конструктивные требования; технико-экономические характеристики и др.

На рисунке 1 представлена таксономическая схема подсистем МКА, использованная при создании ИС. На схеме элементы без рамок представляют собой группы конкретных компонентов, информация о которых содержится в базе данных (БД) разработанной ИС. Блоки в рамках представляют абстрактные системы, необходимые для наглядности схемы.

Обеспечивающими системами являются системы, общие для МКА различного назначения: система управления, система ориентации и стабилизации, система электропитания, система обеспечения теплового режима, двигательная установка, ряд механических систем и элементов, таких как средства отделения от ракеты-носителя, антенные устройства, механизмы фиксации и раскрытия подвижных устройств, экранно-вакуумная теплоизоляция (ЭВТИ) и т.д. Солнечная батарея и её конструкция, аккумуляторная батарея, высокоскоростная радиоперелиния, исполнительные органы системы управления движением и др. элементы выбираются в зависимости от характеристик и требований целевой аппаратуры [1].

При выборе параметров для описания каждого из типов подсистем МКА в качестве основы использовался приведённый выше перечень требований из [22]. Информация для БД собиралась из каталогов продукции предприятий, выпускающих оборудование для изделий космической техники [23-27].

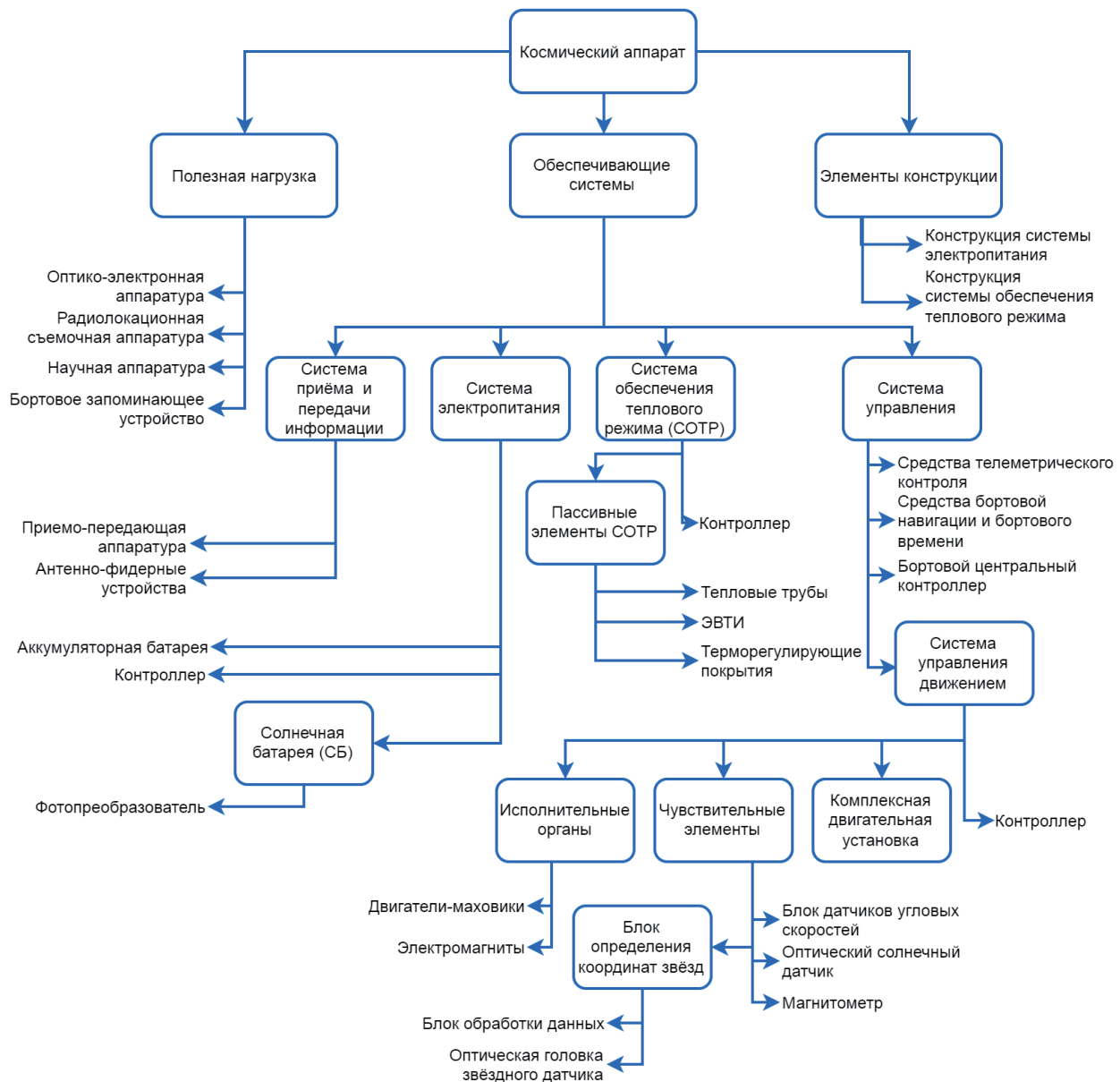


Рисунок 1 – Таксономическая схема подсистем МКА

3 Информационная система

Большинство БД, созданных для обмена проектной информацией и поддержки процесса проектирования, не содержат данные о функциональных характеристиках изделий, их стоимости, совместимости и надёжности, которые играют решающую роль в получении успешного продукта [28].

Этап проектирования характерен тем, что рассматривается множество возможных вариантов изделия. На ранних стадиях важно исключить очевидно неоптимальные варианты и сосредоточить ресурсы разработчиков на тех вариантах, которые могут получить дальнейшее развитие. Время, затраченное на разработку того или иного варианта, является показателем эффективности проектного подразделения и предприятия в целом. Для повышения эффективности проектных работ предприятия повсеместно применяют современные информационные технологии, такие как автоматизированное управление инженерными данными,

управление и работа с большими сборками, параллельное проектирование при реализации комплексных проектов с участием большого количества исполнителей [29]. БД, содержащая информацию о функциональности изделий, может стать важным подспорьем в повышении эффективности проектной деятельности.

В большинстве случаев пользователи не взаимодействуют напрямую с БД и пользуются дополнительным слоем абстракции над БД в виде пользовательского приложения. Такой подход позволяет упростить взаимодействие пользователей с БД, например, за счёт того, что при запросах через клиентский интерфейс от пользователя как правило не требуются навыки применения языка запросов (например, *SQL*) или знания информации о схеме БД. Приложение может представлять пользователю информацию в наглядной форме.

Использование клиент-серверной архитектуры позволяет пользователям ИС взаимодействовать с БД (*Postgres*) без необходимости установки СУБД на локальный компьютер. Кроме того, отделение клиентской части от приложения и БД позволит в будущем реализовать интерфейсы, адаптированные под мобильные устройства.

Клиентский интерфейс приложения представлен на рисунке 2. Условия для поиска интересующих пользователя объектов задаются взаимодействием пользователя с графическими компонентами экрана. Из описания конкретного изделия существует возможность перейти на сайт его производителя.

Каждая страница клиентского интерфейса ИС содержит обязательные элементы: название системы; логотипы организаций, участвующих в разработке; навигационное меню по системам МКА; блок для задания параметров поиска; поле для выдачи результатов поиска.

Блок для задания параметров поиска включает в себя: фильтры поиска в БД; поле поиска по запросу; кнопки для запуска процесса поиска в БД и для очистки заданных фильтров поиска.

БАЗА ДАННЫХ СИСТЕМ МАЛЫХ КОСМИЧЕСКИХ АППАРАТОВ

СИСТЕМА ОБЕСПЕЧЕНИЯ ТЕПЛОВОВОГО РЕЖИМА
СИСТЕМА ЭНЕРГОПИТАНИЯ
СИСТЕМА УПРАВЛЕНИЯ
СИСТЕМА ПРИЕМА И ПЕРЕДАЧИ ИНФОРМАЦИИ
ЦЕЛЕВАЯ АППАРАТУРА
ЭЛЕМЕНТЫ КОНСТРУКЦИИ

Система управления движением

☐ Контроллер

☐ Чувствительные элементы

☐ Блок датчиков угловых скоростей
☐ Блок обработки данных звёздного датчика
☐ Оптическая головка звёздного датчика
☐ Оптический солнечный датчик
☐ Магнитометр

☐ Исполнительные органы

☐ Двигатели-манипуляторы
☐ Электромагниты

Производитель

☐ ОАО «Пеленг»
☐ АО «Корпорация «НИИИЗМ»
☐ ООО «НИЛАКТ ДОСААФ»
☐ НПП «АНТАРЕС»
☐ ООО «СПУТНИКС»
☐ ОАО «Рамenskoe ГКБ»
☐ ИКИ РАН
☐ АО НПП «Геофизика-Космос»
☐ ООО «Азмерит»
☐ НПК «Оптолинк»

Поиск

ПРИМЕНИТЬ
ОЧИСТИТЬ

Подсистема	Наименование	Производитель
Блок датчиков угловых скоростей	Датчик угловой скорости (НИЛАКТ)	ООО «НИЛАКТ ДОСААФ»
Блок датчиков угловых скоростей	SX-VMS-03 (Спутникс)	ООО «СПУТНИКС»
Блок датчиков угловых скоростей	SX-WR-03 (Спутникс)	ООО «СПУТНИКС»
Блок датчиков угловых скоростей	Одноосный измеритель угловой скорости (ОИУС)	НПП «АНТАРЕС»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС200	НПК «Оптолинк»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС501	НПК «Оптолинк»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС1000	НПК «Оптолинк»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС2000	НПК «Оптолинк»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС5000	НПК «Оптолинк»
Блок датчиков угловых скоростей	Одноосный волоконно-оптический гироскоп ОИУС1001	НПК «Оптолинк»
Блок датчиков угловых скоростей	Прецизионный измеритель угла поворота ВЕ 112 (ФЦПУ)	АО НПП «Геофизика-Космос»
Блок датчиков угловых скоростей	Оптико-электронный датчик угла ОЗДУ-2М	АО НПП «Геофизика-Космос»
Блок датчиков угловых скоростей	Унифицированный ряд приборов БИУС ВОА	НПП «АНТАРЕС»
Блок датчиков угловых скоростей	Унифицированный ряд приборов БИУС ВО	НПП «АНТАРЕС»
Блок обработки данных звёздного датчика	мБОКЗ-2 Блок обработки данных (БОД-2)	ИКИ РАН
Блок обработки данных звёздного датчика	Блок управления и синхронизации	ОАО «Пеленг»
Оптическая головка звёздного датчика	АЗДК-1 (Азмерит)	ООО «Азмерит»
Оптическая головка звёздного датчика	SX-MICROBOKZ-01 (Спутникс)	ООО «СПУТНИКС»
Оптическая головка звёздного датчика	мБОКЗ-2 Оптическая головка (ОГ 32/2000)	ИКИ РАН
Оптическая головка звёздного датчика	Прибор звездной ориентации БОКЗ-М60	ИКИ РАН

© Названный комплекс разработан Самарского университета 2022

Рисунок 2 – Клиентский интерфейс приложения ИС поддержки проектирования МКА

4 База данных

БД обеспечивающих БС и полезных нагрузок МКА различного целевого назначения разработана как часть веб-приложения. Клиентская часть выполнена на *JavaScript* и *HTML5*¹, работает в веб-браузере и взаимодействует с БД через *Django Backend*².

Семантическая структура ПрО описана в разделе 2. В дополнение к схеме ПрО, приведённой на рисунке 1 в БД добавлены описания аппаратных интерфейсов, обеспечивающих работу подсистем в составе МКА. Под аппаратными интерфейсами подразумеваются способы физического соединения элементов МКА, например, шины и разъёмы. Соответствие интерфейсов используется в качестве одного из критериев совместимости компонентов. В случае если отобранные системы не совместимы, пользователь будет об этом уведомлен.

Фрагмент *ER*³-диаграммы разработанной БД ИС поддержки проектирования приведён на рисунке 3.

В таблицах, описывающих аппаратуру и компоненты КА, используются суррогатные первичные ключи. В промежуточных таблицах, отражающих связи вида «многие ко многим», используются составные первичные ключи.

В силу относительно малых размеров некоторых таблиц, промежуточные таблицы универсальных интерфейсов используются для связи всех подсистем, т.е. отдельные промежуточные таблицы «подсистема-интерфейс» не создавались. Вместо этого в промежуточной таблице *interface* введено дополнительное поле *subsystem*, являющееся частью составного ключа и содержащее имя соответствующей таблицы, что позволяет обеспечивать условие уникальности первичного ключа. Подробная информация об интерфейсах содержится в отдельной таблице.

Ввиду относительно малого количества записей в БД, обусловленного спецификой ПрО, при поиске каждая таблица может быть обработана системой управления БД (СУБД) за один проход. В связи с этим индексирование таблиц не проводилось (за исключением индексов, автоматически создаваемых СУБД для первичных ключей), так как в рассматриваемом случае они не способны обеспечить значимый прирост производительности.

При анализе планов выполнения запросов было отмечено, что СУБД выполняет поиск по таблицам за сопоставимое время вне зависимости от использования фильтров по индексированным и неиндексированным полям.

5 Работа с системой

В качестве примера применения ИС поддержки проектирования МКА рассмотрен процесс выбора характеристик системы электропитания (СЭП) для перспективного МКА, предназначенного для обеспечения связи между подвижными наземными терминалами. В первом приближении МКА представляется как совокупность платформы обеспечивающих систем и модуля полезной нагрузки (МПН).

Одним из подходов к предварительному проектированию МКА является метод проектирования с использованием заимствованных элементов.

¹ *HTML5* (англ. *HyperText Markup Language, version 5*) — язык для структурирования и представления содержимого всемирной паутины (англ. *World Wide Web*). *HTML5* предназначен для поддержки мультимедиа-технологий с одновременным сохранением обратной совместимости, удобочитаемости кода для человека и простоты анализа для парсеров.

² *Django* — свободный полнофункциональный фреймворк для создания веб-приложений, написанный на *Python*.

³ *ER* от англ. *Entity-Relationship model*, модель «сущность — связь». В программной инженерии *ER*-модель часто выступает абстрактной моделью ПрО, определяющей структуру данных или информации, которая может быть реализована в БД, обычно реляционной.

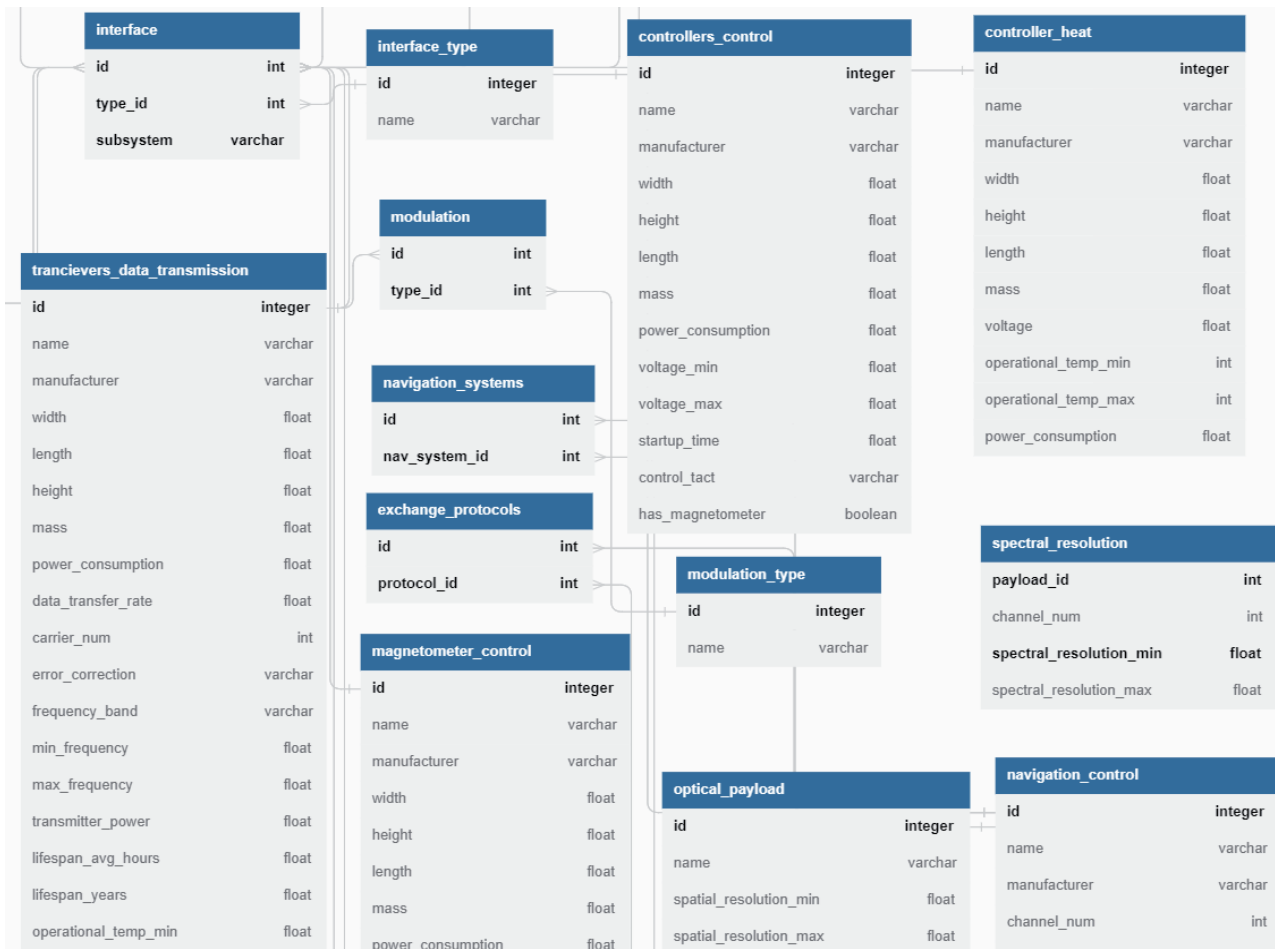


Рисунок 3 – Фрагмент ER-диаграммы БД подсистем МКА

При формировании проектного облика МКА с использованием заимствованных элементов, характерного для МКА на базе платформ, задача определения проектных параметров МКА может быть сведена к поиску множества компонентов подсистем МКА, характеристики которых обеспечивают функционирование МПН.

В общем случае задача проектирования может формулироваться в терминах нелинейного программирования [30, 31]. Необходимо найти вектор проектных переменных X^* , при котором выполняется условие:

$$f(X^*) \leq f(X) \forall X \in \Omega$$

$$\Omega = \{X : g_j(X) \leq 0, j = 1, 2, \dots, m\},$$

где $f(X)$ – целевая функция, $X = \{x_1, x_2, \dots, x_n\}$ – вектор проектных параметров, Ω – область допустимых значений, определяемая комплексом ограничений – $g_j(X) \leq 0$.

Для определения основных характеристик СЭП в качестве проектных параметров принимаются: масса аккумуляторной батареи (m_{AB}), масса солнечных батарей (m_{CB}), масса контроллера заряда-разряда (m_K).

В качестве критерия оптимальности в настоящей работе используется минимум массы СЭП, обеспечивающей функционирование МПН. Таким образом, целевая функция: $M_{СЭП} = m_{AB} + m_{CB} + m_K \rightarrow \min$.

Условия функционирования МПН определяются из режимов работы, приведённых в таблице 1.

Таблица 1 – Исходные данные по электропотреблению МПН

Режимы работы МПН	Длительность на рабочем витке, мин	Электропотребление, Вт		Допустимое время работы в сутки, мин
		при работе	среднесуточное	
1 Режим включения (РВ)	3	450	3,75	12
2 Режим ретрансляции (РР)	13	2090	75,5	52
3 Режим выключения (РВЫК)	3	450	3,75	12
4 Дежурный режим (ДР)	77,7	100	94,7	1364

Согласно исходным данным, в сутки предполагается до четырёх рабочих сеансов с максимальной длительностью сеанса ретрансляции до 13 минут при пролёте региона интереса: два сеанса на двух последовательных восходящих витках и через ~ 12 часов на двух последовательных нисходящих витках. График электропотребления МПН для типовых суток полёта МКА приведён на рисунке 4.

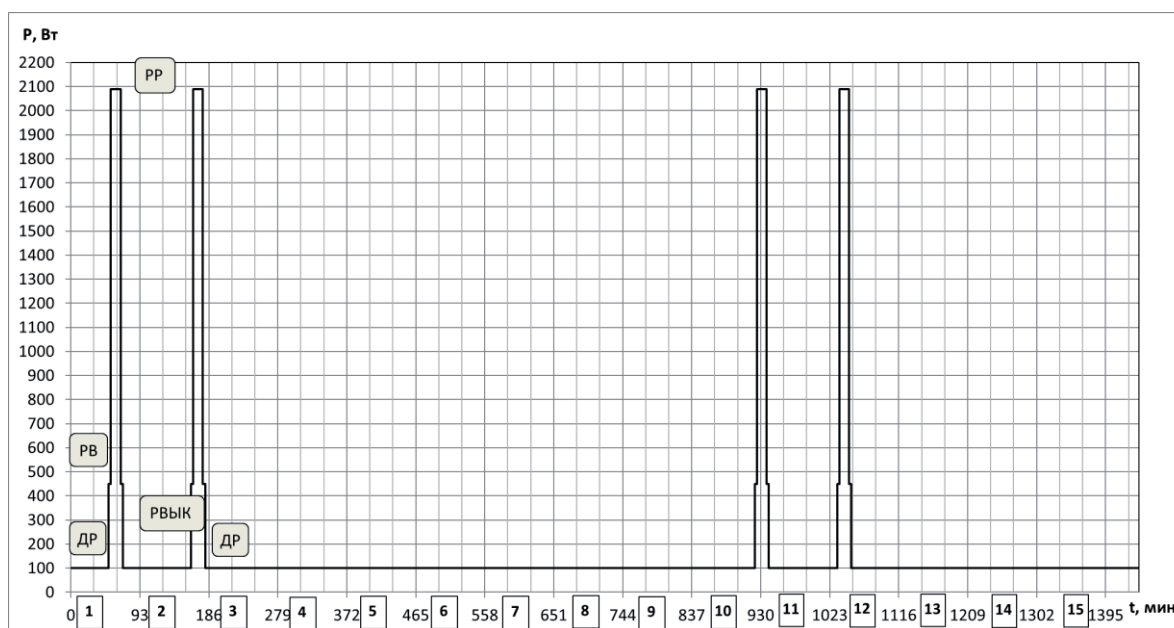


Рисунок 4 – График электропотребления МПН

График электропотребления бортовой аппаратуры МКА и уровня освещённости солнечных панелей приведён на рисунке 5.

В качестве исходных данных, необходимых для определения параметров СЭП используются сводка электропотребления и характеристики компонентов СЭП. Возможности разработанной ИС поддержки проектирования позволяют вместо усреднённых значений использовать характеристики реально существующих изделий, что позволяет снизить время на поиск информации и повысить точность первых проектных приближений.

В таблице 2 приведены исходные данные для определения параметров СЭП, полученных как в результате разработанной циклограммы работы МКА, так и из ИС. В таблице характеристики удельной мощности и деградации фотоэлектрического преобразователя (ФЭП), а также значения коэффициентов полезного действия (КПД) аккумуляторной батареи и разрядного/зарядного устройства соответствуют значениям, хранящимся в разработанной БД ИС поддержки проектирования.

Схема рабочего процесса использования системы представлена на рисунке 6. При работе с ИС пользователь задаёт значения или диапазоны значений интересующих его параметров, используя интерфейс приложения (1). Приложение формирует запрос к БД (2) и представляет полученную информацию пользователю (3).

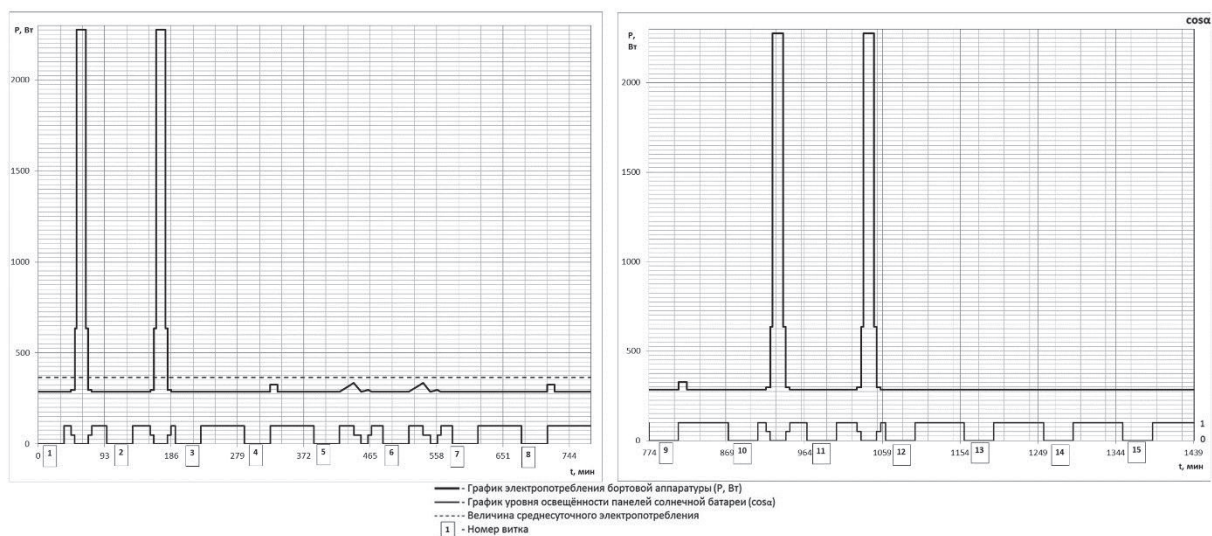


Рисунок 5 – График электропотребления бортовой аппаратуры МКА и уровня освещённости

Таблица 2 – Исходные данные для определения параметров СЭП

Характеристика	Значение
Среднесуточное электропотребление бортовой аппаратуры, Вт	418,7
1 МПН	177,7
2 БС контроля и управления	55
3 Система управления движением	55
4 Система электропитания	20
5 Система обеспечения теплового режима	96
Резерв	15
Среднее напряжение бортовой сети, В	27
Среднесуточный уровень освещённости панелей БС ($\cos \alpha$)	0,51
Удельная мощность ФЭП (арсенид-галлиевые), Вт/м ²	255
Удельная мощность ФЭП (кремний), Вт/м ²	150
Среднее значение КПД аккумуляторной батареи	0,9
Среднее значение КПД разрядного/зарядного устройства	0,95
Коэффициент деградации ФЭП	1,2

Дальнейший расчёт параметров СЭП проводится с учётом данных, приведённых в таблице 2. Пример результатов расчёта требуемых для обеспечения работы МПН параметров СЭП приведён в нижней таблице на рисунке 6. Полученные значения параметров СЭП позволяют определить границы областей, в которых может находиться искомое проектное решение для МКА.

Заключение

В работе описана ИС поддержки проектирования и представлена структура БД обеспечивающих БС и полезных нагрузок МКА. ИС может быть использована для повышения эффективности проектирования МКА на базе платформ.

Набор требований к описанию ПрО разработан на основе существующих таксономий. Элементы БД идентифицируются уникальными кодами для каждого элемента. Конструкторская информация группируется в функциональные блоки. Развитие проекта предполагает работы по дальнейшему наполнению конструктивными решениями всех таблиц БД и расширению возможностей веб-платформы.



Характеристика	Значение
Среднесуточное электропотребление БА, Вт	418,7
Среднее напряжение бортовой сети, В	27
Требуемая мощность ФЭП, Вт	1116,5
Площадь под размещение ФЭП (арсенид-галлиевые), м ²	4,4
Площадь под размещение ФЭП (кремний), м ²	7,4
Ёмкость АБ, А·ч	51,6
Собственное потребление СЭП, Вт	20

Рисунок 6 – Схема работы ИС поддержки проектирования МКА

СПИСОК ИСТОЧНИКОВ

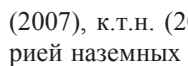
- [1] **Ткаченко И.С.** Анализ ключевых технологий создания многоспутниковых орбитальных группировок малых космических аппаратов // *Онтология проектирования*. 2021. Т.11, №4(42). С.478-499. DOI: 10.18287/2223-9537-2021-11-4-478-499.
- [2] **Hallberg N., Jungert E., Pilemalm S.** Ontology for Systems Development // *International Journal of Software Engineering and Knowledge Engineering*. 2014. P.329-345. DOI:10.1142/S0218194014500132.
- [3] **Mezhuyev V.** Ontology Based Development of Domain Specific Languages for Systems Engineering // 2014 International Conference on Computer and Information Sciences (ICCOINS). Kuala Lumpur, Malaysia, 2015. P.1-6. DOI:10.1109/ICCOINS.2014.6868825.
- [4] The MOKA Consortium, "Managing engineering knowledge: MOKA: methodology for knowledge-based engineering applications", ASME Press, New York, 2001.
- [5] **Gu Q., Lago P.** Exploring service-oriented system engineering challenges: a systematic literature review // *Service Oriented Computing and Applications (SOCA)*. 2009. №3. P.171-188. DOI:10.1007/s11761-009-0046-7.
- [6] ГОСТ Р ИСО 9001-2015. Система менеджмента качества – Требования.
- [7] **Ma J., Wang G., Jinzhi L., Vangheluwe H., Kiritsis D., Yan Y.** Systematic Literature Review of MBSE Tool-Chains // *Applied Sciences*. 2022. №12. P.3431. DOI:10.3390/app12073431.
- [8] **Klein J.** A systematic review of system-of-systems architecture research // *QoSA 2013 - Proceedings of the 9th International ACM Sigsoft Conference on the Quality of Software Architectures*. 2013. P.13-22. DOI:10.1145/2465478.2465490.
- [9] **Abdalla M., Bellare M., Neven G.** Robust encryption. TCC 2010, LNCS 5978, Springer, 2010. P.480-497.
- [10] **O'Neill A.** Definitional issues in functional encryption. Cryptology ePrint Archive, Report 2010/556, 2010. <http://eprint.iacr.org/2010/556>.
- [11] **Bianchi T., Santos D.S., Felizardo K.R.** Quality Attributes of Systems-of-Systems: A Systematic Literature Review // 2015 IEEE/ACM 3rd International Workshop on Software Engineering for Systems-of-Systems. 2015. P.23-30, DOI:10.1109/SESoS.2015.12.
- [12] **Vargas I., Gottardi T., Braga R.** Approaches for integration in system of systems: a systematic review. // In *Proceedings of the 4th International Workshop on Software Engineering for Systems-of-Systems (SESoS '16)*. Association for Computing Machinery. New York, NY, USA, 2016. P.32–38. DOI:10.1145/2897829.2897835.
- [13] **Vierhauser, M., Rabiser, R., Grünbacher, P.** Requirements monitoring frameworks: a systematic review // *Inf. Softw. Technol.* 2016. №80. P.89-109. DOI:10.1016/j.infsof.2016.08.005.
- [14] **de Lima R.M., de Vargas D., Fontoura L.M.** System of systems requirements: a systematic literature review using snowballing. Knowledge Systems Institute Graduate School. 2017. P.97-100. DOI:10.18293/SEKE2017-114.
- [15] **Lan Yang, Kathryn Cormican, Ming Yu,** Ontology-based systems engineering: A state-of-the-art review, *Computers in Industry*, Volume 111, 2019, P.148-171, DOI:10.1016/j.compind.2019.05.003.
- [16] **Dwivedi Y.K., Ravichandran K., Williams M.D., Miller S., Lal B., Antony G.V., Kartik M.** IS/IT project failures: a review of the extant literature for deriving a taxonomy of failure factors // In: Dwivedi, Y.K., Henriksen, H.Z., Wastell, D., Dé, R. (Eds.), *Gd. Successes Fail. IT. Public Priv. Sect. TDIT 2013. IFIP Adv. Inf. Commun. Technol.*. Springer, Berlin, Heidelberg. P.73–88. DOI:10.1007/978-3-642-38862-0_5.
- [17] **Hallberg N., Pilemalm S., Timpka T.** Quality driven requirements engineering for development of crisis management systems // *Int. J. Inf. Syst. Cris. Response Manag.* 2012. №4. P.35-52. DOI:10.4018/jiscrm.2012040103.
- [18] **Гвоздев В.Е., Бежаева О.Я., Насырова Р.А.** Модели возникновения ошибок на предпроектной стадии разработки компонент информационно-вычислительных систем // *Онтология проектирования*. 2020. Т.10, №1. С.73-86. DOI: 10.18287/2223-9537-2020-10-1-73-86.
- [19] **Pardo C., Pino F.J., García F., Piattini M., Baldassarre M.T.** An ontology for the harmonization of multiple standards and models // *Comput. Stand. Interfaces*. 2012. №34. P.48-59. DOI:10.1016/j.csi.2011.05.005.
- [20] **Dori D., Sillitto H.** What is a system? An ontological framework // *Syst. Eng.* 2017. №20. P.207–219. DOI:10.1002/sys.21383.
- [21] **Axelsson J.** Achieving System-of-Systems Interoperability Levels Using Linked Data and Ontologies // *INCOSE International Symposium*. 2020. P.651-665. DOI:10.1002/j.2334-5837.2020.00746.x.
- [22] **Сафронов С.Л., Ткаченко И.С., Иванушкин М.А., Волгин С.С.** Современные подходы к созданию малых космических аппаратов дистанционного зондирования Земли на базе унифицированных платформ. Самара: Изд-во Самарского университета, 2019. 276 с.
- [23] Бортовая аппаратура. Сайт Российские космические системы. <https://russianspacesystems.ru/bussines/cosmostroy/bortovaya-apparatura/>.
- [24] Навигационная аппаратура. Сайт АО «ИРЗ». <https://www.irz.ru/products/12/index.htm>.
- [25] Отдел оптико-физических исследований ИКИ РАН. <http://ofo.ikiweb.ru/>.
- [26] АО «НПЦ «Полус». <https://polus-tomsk.ru/?id=212>.

- [27] Артериальные тепловые трубы. Сайт ООО НПП «Тепловые агрегаты и системы». <http://npptais.ru/artt/>.
- [28] **Zhang L, Tan R, Peng Q, Shao P, Dong Y, Wang K.** Construction and Application of Enterprise Knowledge Base for Product Innovation Design // Applied Sciences. 2022. №12(13). DOI:10.3390/app12136358.
- [29] **Комарова Л.А., Филатов А.Н.** Применение технологии нисходящего проектирования, основанной на решениях Windchill PDMLink и САПР Pro/ENGINEER, для разработки изделий ракетно-космической техники // Известия Самарского научного центра РАН. 2011. №1-2.
- [30] **Есипов Б.А.** Методы оптимизации и исследование операций. Самара: Изд-во Самарского университета. 2007. 180 с.
- [31] **Salmin V.V. et al** Determination of the main design parameters of cost-effective remote sensing satellite systems at the stage of preliminary design // Journal of Physics: Conference Series. 2021. Vol. 1745. Issue 1. 012089. DOI: 10.1088/1742-6596/1745/1/012089.

Сведения об авторах



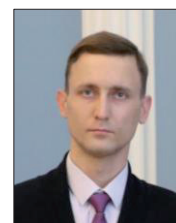
Ткаченко Иван Сергеевич, 1985 г. рождения. Окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (2008), к.т.н. (2012). Доцент кафедры космического машиностроения Самарского университета, директор института авиационной и ракетно-космической техники, В списке научных трудов более 70 работ в области исследования и создания МКА. Author ID (РИНЦ): 602393; Author ID (Scopus): 12645515700, ORCID: 0000-0001-8892-7975. Tkachenko.is@ssau.ru. ✉.



Сафронов Сергей Львович, 1984 г. рождения. Окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (2007), к.т.н. (2012), доцент кафедры космического машиностроения, заведующий лабораторией наземных испытаний летательных аппаратов. В списке научных трудов более 60 публикаций в области проектирования космической техники. Научные интересы: платформы МКА, системный анализ. AuthorID (РИНЦ): 723511. Author ID (Scopus): 56940457200, ORCID:0000-0002-7530-0636. safronov@ssau.ru.



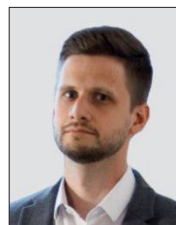
Коровин Максим Дмитриевич, 1988 г. рождения. Окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (2013). Автор более 30 публикаций. Author ID (РИНЦ): 746456; Author ID (Scopus): 56566556100, ORCID: 0000-0002-2298-8457. korovin.md@ssau.ru.



Иванушкин Максим Александрович, 1991 г. рождения. Окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (2015), ассистент кафедры космического машиностроения, младший научный сотрудник научно-исследовательского института космического машиностроения. Автор более 25 публикаций. Научные интересы: проектирование МКА, многоспутниковые космические системы. Author ID (РИНЦ): 882454. Author ID (Scopus): 57188855864, ORCID: 0000-0001-9525-0229. ivanushkin.ma@ssau.ru.



Крестина Анастасия Владимировна, 1995 г. рождения. Окончила Самарский государственный аэрокосмический университет имени академика С.П. Королёва (2017), ассистент кафедры космического машиностроения. Автор более 15 научных работ в области малых космических аппаратов. Научные интересы: космический мусор, уход МКА с орбиты. Author-ID (РИНЦ): 1047564. Author-ID (Scopus): 57210124218, ORCID: 0000-0002-6892-9903. stasy2403@yandex.ru.



Поступила в редакцию 19.11.2022, после рецензирования 30.01.2023. Принята к публикации 11.02.2023.



An information system to support preliminary design of small spacecraft

© 2023, I.S. Tkachenko ✉, S.L. Safronov, M.D. Korovin, M.A. Ivanushkin, A.V. Krestina

Samara University (Samara National Research University named after academician S.P. Korolev), Samara, Russia

Abstract

With the development of information technology, the role of support systems for the technical products design process has increased significantly. The capabilities of knowledge management systems make it possible to ensure the commonality of terminology, object definitions and relations between them. In addition, the use of formal semantics is necessary to support the reuse of knowledge. The article discusses ways to improve the efficiency of the small spacecraft design process and a creation of tools for the knowledge exchange about the functional characteristics of typical subsystems of such vehicles. The creation of an information system to support the preliminary design of small spacecraft, which contains information linked to the relationship between the functions and components of subsystems and their characteristics, is considered. The created system facilitates the synthesis of various product configurations. When developing the information system, modern approaches in the field of structural and functional classifications of small spacecraft systems and approaches existing in the field of taxonomy to create database structures were used. The information system database contains a set of characteristics for a wide range of small spacecraft equipment. As an example of the use of an information system to support the design of small spacecraft, the process of choosing the characteristics of a power supply system for a promising small spacecraft designed to provide communication between mobile ground terminals is considered.

Key words: information system, small spacecraft, space platform, database, onboard system, design, user interface.

For citation: Tkachenko IS, Safronov SL, Korovin MD, Ivanushkin MA, Krestina AV. An information system to support preliminary design of small spacecraft [In Russian]. *Ontology of designing*. 2023; 13(1): 75-89. DOI:10.18287/2223-9537-2023-13-1-75-89.

Financial Support: The work was carried out with the financial support of the Program of Samara National Research University named after Academician S.P. Korolev for 2021-2030 within the Program of Strategic Academic Leadership «Priority 2030», Agreement № 30/22B from 15.06.2022.

Conflict of interest: The authors declare no conflict of interest.

List of figures and tables

Figure 1 – Taxonomic scheme of small spacecraft subsystems

Figure 2 – Application user interface

Figure 3 – An ER-diagram fragment of the spacecraft subsystems database

Figure 4 – Payload module power graph

Figure 5 – Power consumption graph of the onboard equipment of the small spacecraft and the level of illumination

Figure 6 – Block diagram of the design support information system

Table 1 – Input data of the payload module power consumption

Table 2 – Initial data for determining the power supply system parameters

References

- [1] **Tkachenko IS.** Analysis of key technologies for creating multisatellite orbital constellations of small spacecraft [In Russian]. *Ontology of designing*. 2021; 11(4): 478-499. DOI: 10.18287/2223-9537-2021-11-4-478-499.
- [2] **Hallberg N, Jungert E, Pilemalm S.** Ontology for Systems Development // *International Journal of Software Engineering and Knowledge Engineering*. 2014. pp.329-345. DOI:10.1142/S0218194014500132.

- [3] **Mezhuyev V.** Ontology Based Development of Domain Specific Languages for Systems Engineering // 2014 International Conference on Computer and Information Sciences (ICCOINS). Kuala Lumpur, Malaysia, 2015. p.1-6. DOI:10.1109/ICCOINS.2014.6868825.
- [4] The MOKA Consortium, "Managing engineering knowledge: MOKA: methodology for knowledge-based engineering applications", ASME Press, New York, 2001.
- [5] **Gu Q, Lago P.** Exploring service-oriented system engineering challenges: a systematic literature review // Service Oriented Computing and Applications (SOCA). 2009; 3: 171-188. DOI:10.1007/s11761-009-0046-7.
- [6] ГОСТ Р ИСО 9001-2015 Quality Management System – Requirements [In Russian].
- [7] **Ma J, Wang G, Jinzhi L, Vangheluwe H, Kiritsis D, Yan Y.** Systematic Literature Review of MBSE Tool-Chains // Applied Sciences. 2022; 12: 3431. DOI:10.3390/app12073431.
- [8] **Klein J.** A systematic review of system-of-systems architecture research // QoSA 2013 - Proceedings of the 9th International ACM Sigsoft Conference on the Quality of Software Architectures. 2013. pp. 13-22. DOI:10.1145/2465478.2465490.
- [9] **Abdalla M, Bellare M, Neven G.** Robust encryption. TCC 2010, LNCS 5978, Springer, 2010. pp. 480-497.
- [10] **O'Neill A.** Definitional issues in functional encryption. Cryptology ePrint Archive, Report 2010/556, 2010. <http://eprint.iacr.org/2010/556>.
- [11] **Bianchi T, Santos DS, Felizardo KR.** Quality Attributes of Systems-of-Systems: A Systematic Literature Review // 2015 IEEE/ACM 3rd International Workshop on Software Engineering for Systems-of-Systems. 2015. pp.23-30, DOI:10.1109/SESOS.2015.12.
- [12] **Vargas I, Gottardi T, Braga R.** Approaches for integration in system of systems: a systematic review. // In Proceedings of the 4th International Workshop on Software Engineering for Systems-of-Systems (SESOS '16). Association for Computing Machinery. New York, NY, USA, 2016. pp. 32–38. DOI:10.1145/2897829.2897835.
- [13] **Vierhauser M, Rabiser R, Grünbacher P.** Requirements monitoring frameworks: a systematic review // Inf. Softw. Technol. 2016. №80, pp. 89-109. DOI:10.1016/j.infsof.2016.08.005.
- [14] **de Lima RM, de Vargas D, Fontoura LM.** System of systems requirements: a systematic literature review using snowballing. Knowledge Systems Institute Graduate School. 2017. pp.97-100. DOI:10.18293/SEKE2017-114.
- [15] **Lan Yang, Kathryn Cormican, Ming Yu,** Ontology-based systems engineering: A state-of-the-art review, Computers in Industry, Volume 111, 2019, P.148-171, DOI:10.1016/j.compind.2019.05.003.
- [16] **Dwivedi YK, Ravichandran K, Williams MD, Miller S, Lal B, Antony GV, Kartik M.** IS/IT project failures: a review of the extant literature for deriving a taxonomy of failure factors // In: Dwivedi, Y.K., Henriksen, H.Z., Wastell, D., Dé, R. (Eds.), Gd. Successes Fail. IT. Public Priv. Sect. TDIT 2013. IFIP Adv. Inf. Commun. Technol.. Springer, Berlin, Heidelberg, pp.73–88. DOI:10.1007/978-3-642-38862-0_5.
- [17] **Hallberg N, Pilemalm S, Timpka T.** Quality driven requirements engineering for development of crisis management systems // Int. J. Inf. Syst. Cris. Response Manag. 2012; 4: 35-52. DOI:10.4018/jiscrm.2012040103.
- [18] **Gvozdev VE, Bezhaeva OY, Nasyrova RA.** Models of errors at the pre-design stage of the development of information and computing systems components [In Russian]. *Ontology of designing*. 2020; 10(1): 73-86. DOI: 10.18287/2223-9537-2020-10-1-73-86.
- [19] **Pardo C, Pino FJ, García F, Piattini M, Baldassarre MT.** An ontology for the harmonization of multiple standards and models // Comput. Stand. Interfaces. 2012; 34: 48-59. DOI:10.1016/j.csi.2011.05.005.
- [20] **Dori D, Sillitto H.** What is a system? An ontological framework // Syst. Eng. 2017; 20: 207–219. DOI:10.1002/sys.21383.
- [21] **Axelsson J.** Achieving System-of-Systems Interoperability Levels Using Linked Data and Ontologies // INCOSE International Symposium. 2020. pp. 651-665. DOI:10.1002/j.2334-5837.2020.00746.x.
- [22] **Safronov SL, Tkachenko IS, Ivanushkin MA, Volgin SS.** Modern approaches to the creation of small satellites for remote sensing of the Earth on the basis of unified platforms [In Russian]: monograph. Samara: Samara University publ.; 2019. 276 p.
- [23] Onboard equipment. The site of Russian space systems. <https://russianspacesystems.ru/bussines/cosmostroy/bortovaya-pparatura/>.
- [24] Navigation equipment. Website of JSC "Irz". <https://www.irz.ru/products/12/index.htm>.
- [25] Department of Optical and Physical Research of IKI RAS. <http://ofo.ikiweb.ru/>.
- [26] JSC "NPC" Polus ". <https://polus-tomsk.ru/?id=212>.
- [27] Arterial heat pipes. The site of the NPP LLC "Heat Aggregates and Systems". <http://npptais.ru/artt/>.
- [28] **Zhang L, Tan R, Peng Q, Shao P, Dong Y, Wang K.** Construction and Application of Enterprise Knowledge Base for Product Innovation Design // Applied Sciences. 2022. №12(13). DOI:10.3390/app12136358.
- [29] **Komarova LA, Filatov AN.** The application of the technology of downward design based on the solutions of Windchill PdmLink and CAD Pro/Engineer for the development of products of rocket and space technology [In Russian] // News of the Samara Scientific Center of the Russian Academy of Sciences. 2011. No. 1-2.

- [30] **Esipov BA.** Optimization methods and operations research. Samara: Publishing House of the Samara University. 2007. 180 p.
- [31] **Salmin VV et al.** Determination of the main design parameters of cost-effective remote sensing satellite systems at the stage of preliminary design // Journal of Physics: Conference Series. 2021. Vol. 1745. Issue 1. 012089 DOI: 10.1088/1742-6596/1745/1/012089.
-

About the authors

Ivan Sergeevich Tkachenko (b. 1985), Graduated from Samara State Aerospace University named after Academician S.P. Korolev (2008), Candidate of Technical Sciences (2012). Associate Professor of the Department of Space Engineering of Samara University, Director of Institute of Aviation and Rocket-Space Engineering. He has more than 70 works in the field of research and creation of small spacecrafts. Author ID (RSCI): 602393; Author ID (Scopus): 12645515700, ORCID: 0000-0001-8892-7975. tkachenko.is@ssau.ru. ✉

Sergey Lvovich Safronov, (b. 1984), Graduated from Samara State Aerospace University named after Academician S.P. Korolev (2007), Candidate of Technical Sciences (2012), Associate Professor of Space Engineering Department, Head of Aircraft Ground Test Laboratory. He has more than 60 publications in the field of space technology design. Research interests are: small spacecraft platforms, systems analysis. AuthorID (RSCI): 723511. Author ID (Scopus): 56940457200, ORCID:0000-0002-7530-0636. safronov@ssau.ru.

Maxim Dmitrievich Korovin, (b. 1988), Graduated from Samara State Aerospace University named after Academician S.P. Korolev (2013). He is the author of more than 30 publications. Research interests are: database design, information processing, small satellite design. Author ID (RSCI): 746456; Author ID (Scopus): 56566556100, ORCID: 0000-0002-2298-8457. korovin.md@ssau.ru.

Maxim Alexandrovich Ivanushkin, (b. 1991), Graduated from Samara State Aerospace University named after Academician S.P. Korolev (2015), Assistant of Space Engineering Department, Junior Researcher of Space Engineering Research Institute. He is the author of more than 25 publications. Research interests are: design of small spacecraft, multi-satellite space systems. Author ID (RSCI): 882454. Author ID (Scopus): 57188855864, ORCID: 0000-0001-9525-0229. ivanushkin.ma@ssau.ru.

Anastasiya Vladimirovna Krestina, (b. 1995), Graduated from Samara State Aerospace University named after Academician S.P. Korolev (2017), Assistant of Space Engineering Department. She is the author of more than 15 scientific works in the field of small spacecraft. Scientific interests are: space debris, small spacecraft orbital insertion. Author-ID (RSCI): 1047564. Author-ID (Scopus): 57210124218, ORCID: 0000-0002-6892-9903. stasy2403@yandex.ru.

Received November 19, 2022. Revised January 30, 2023. Accepted February 11, 2023.



Проектирование системы сбора полётной информации для расчёта аэродинамических характеристик БПЛА

©2023, Хамза Мазин Абдулаали Хамза

Казанский национальный исследовательский технический университет им. А.Н. Туполева - КАИ,
Казань, Россия

Технологический университет, Багдад, Ирак

Аннотация

В статье описана разработка бортовой системы сбора информации для расчёта аэродинамических характеристик беспилотных летательных аппаратов (БПЛА), которые возможно получить только в условиях свободного полёта. Представлена функциональная схема, описан алгоритм работы системы регистрации параметров полёта. Предложенные решения позволяют собрать большой объём информации и повысить эффективность системы за счёт отбора и сжатия информации. Собранные данные используются для определения аэродинамических характеристик, обеспечивая отказ от испытаний моделей БПЛА в аэродинамических трубах, за счёт чего может быть достигнута экономия средств и времени на разработку новых БПЛА. Рассмотрены особенности информационного обмена между элементами системы, позволяющие ускорить передачу данных и повысить точность измерений за счёт увеличения частоты опроса измерительных устройств. Полученные результаты могут представлять интерес для разработчиков автономных измерительных систем. Особое значение для расчёта аэродинамических характеристик БПЛА по предлагаемому методу имеет высокая точность измерения параметров траектории летательного аппарата, включая возникающие в процессе движения линейные ускорения и угловые скорости. В статье рассматривается вариант аппаратной реализации предложенной системы для БПЛА.

Ключевые слова: аэродинамические характеристики, параметры полёта, программно-аппаратный комплекс, сбор информации, передача данных, БПЛА.

Цитирование: Хамза Мазин Абдулаали Хамза. Проектирование системы сбора полётной информации для расчёта аэродинамических характеристик БПЛА // Онтология проектирования. 2023. Т.13, №1(47). С.90-98. DOI:10.18287/2223-9537-2023-13-1-90-98.

Благодарности: автор выражает признательность членам редколлегии журнала «Онтология проектирования» за замечания и рекомендации по усовершенствованию данной статьи.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

Достижения в аэродинамике, навигационных системах, электронике, робототехнике, компьютерных технологиях и др. позволили выйти на новый уровень в разработке беспилотных летательных аппаратов (БПЛА) [1]. В числе приоритетных направлений применения БПЛА: военные, спасательные, правоохранительные и природоохранительные операции [2-4], академические исследования [5-8].

При разработке крупногабаритных БПЛА (взлётная масса более 100 кг) исследования в аэродинамических трубах можно производить на небольших масштабных моделях, а полученные на них результаты можно экстраполировать на полноразмерный аппарат с помощью теории подобия.

Целью исследования является разработка системы сбора полётной информации высокой точности для получения аэродинамических характеристик летательного аппарата (ЛА) в

условиях свободного полёта. Большинство аэродинамических характеристик при разработке БПЛА определяется путём продувок в аэродинамической трубе. Некоторые аэродинамические характеристики могут быть определены только в полёте. Разрабатываемая система предназначена для определения аэродинамических характеристик БПЛА в полёте.

По совокупности регистрируемых системой данных путём вычислений можно получить направление осей ЛА и направление вектора скорости. Например, есть возможность получить угол наклона траектории θ - угол между направлением земной скорости $\vec{V}_k$ и горизонтальной плоскостью OX_gZ_g нормальной системы координат. Это позволит в дальнейшем рассчитать зависимости аэродинамических коэффициентов ($c_x, c_y, c_z, m_x, m_y, m_z$) от углов атаки и скольжения, а также от скорости полёта и от скоростей вращения вокруг осей.

Система сбора информации должна регистрировать проекции вектора истинной скорости по осям траекторной системы координат. В большинстве случаев бортовой самописец БПЛА регистрирует данные об угловых скоростях в системе координат ЛА, но не имеет возможности регистрировать истинную скорость.

Для получения данных о положении ЛА в траекторной системе координат используется приёмник аппаратуры *RTK* (*Real Time Kinematic* – «кинематика реального времени») [10] работающий с навигационными спутниковыми системами (*GPS*, США; *ГЛОНАСС*, Россия; *Бэйдоу*, Китай) по протоколу Национальной ассоциации морской электроники (*National Marine Electronics Association - NMEA*), формирующий пакет с информацией о положении БПЛА и векторе его скорости.

1 Структура системы сбора полётной информации

Система сбора информации для расчёта аэродинамических характеристик имеет следующую структуру (см. рисунок 1) и включает:

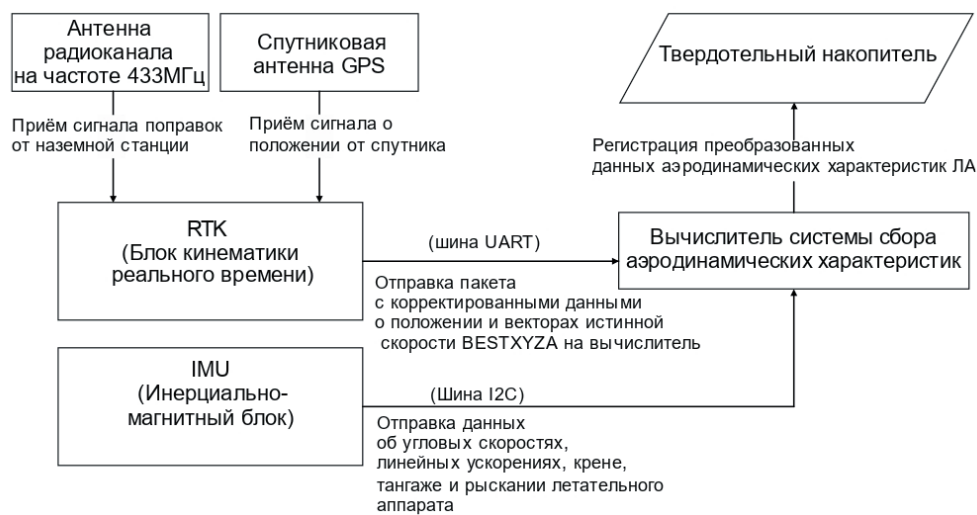


Рисунок 1 – Структура системы сбора информации для расчёта аэродинамических характеристик БПЛА

- инерциально-магнитный (измерительный) блок (*Inertial Measurement Unit - IMU*) для определения ориентации ЛА в земной системе координат;
- приёмник сигнала *RTK* для определения положения ЛА в земной системе координат, в том числе для получения проекций векторов истинной скорости;
- вычислительный блок для реализации алгоритма регистрации получаемых данных от блоков *IMU* и *RTK* и для преобразования этих данных в формат, удобный для последующей обработки записи на накопитель.

В БПЛА в качестве бортовой электроники часто используются микроконтроллеры и межмашинные интерфейсы встраиваемых систем, такие как универсальный асинхронный приёмопередатчик (*Universal Asynchronous Receiver Transmitter - UART*) или последовательный асимметричный интерфейс для соединения низкоскоростных периферийных компонентов с микроконтроллерами (*Inter-Integrated Circuit - I2C*). *IMU* блок представляет собой микроэлектромеханический датчик (*Microelectromechanical systems - MEMS*), с встроенными трёхосевым акселерометром, магнитометром, барометром и вычислителем, способным возвращать данные в цифровом виде [11-13]. *IMU* передаёт следующий набор данных:

- ускорения ЛА по соответствующим трём осям системы координат БПЛА (в испытываемой модели оси направлены согласно обозначениям на датчике);
- угловые скорости ЛА относительно трёх осей системы координат БПЛА;
- три угловых параметра положения ЛА: тангаж ϑ , крен γ , рыскание ψ .

Для записи данных в память устройства используется программная структура из девяти указанных параметров согласно документации на микросхему, используемую в нём. В случае выполнения задачи расчёта аэродинамических характеристик из блока *RTK* считывается пакет *BESTXYZA* (*Best available cartesian position and velocity* - наилучшее доступное декартово положение и скорость) [14].

- Пакет *BESTXYZA* содержит полный набор данных, содержащих позицию (в метрах), скорость (в м/с) и время (в секундах) регистрации (*position, velocity, time - PVT*): статус пакета решения по точке; тип позиционирования: координата позиции X - $P-X$; координата позиции Y - $P-Y$; координата позиции Z - $P-Z$; стандартное отклонение от точки $P-X$ - $P-X\sigma$; стандартное отклонение от точки $P-Y$ - $P-Y\sigma$; стандартное отклонение от точки $P-Z$ - $P-Z\sigma$.
- Пакет *BESTXYZA* содержит информацию об истинных скоростях по осям OX , OY и OZ земной системы координат, а также статус считываемого сообщения для коррекции и соответствующие среднеквадратические отклонения от наземной станции: статус пакета решения по скорости; тип скорости; вектор скорости по оси $O-X$ - $V-X$; вектор скорости по оси $O-Y$ - $V-Y$; вектор скорости по оси $O-Z$ - $V-Z$; стандартное отклонение $V-X$ - $V-X\sigma$; стандартное отклонение $V-Y$ - $V-Y\sigma$; стандартное отклонение $V-Z$ - $V-Z\sigma$; измерение задержки получаемого значения скорости по временной метке, его следует вычесть из текущего времени, чтобы получить улучшенные результаты; разность времени полученного пакета; время решения.
- Пакет *BESTXYZA* отправляется от корректирующей станции бортовой системе и содержит: идентификатор соответствующей станции; общее количество отслеживаемых спутников; количество спутников, используемых в решении; количество спутников первого уровня, используемых при генерации в решении; количество спутников первого и второго уровней, используемых при генерации в решении [15].

Изначально пакет представляет собой сборку данных для определения лучших значений по скоростям (в текущем случае - скорость БПЛА). При формировании пакета для записи на накопитель из получаемого пакета берутся значения $P-X$, $P-Y$, $P-Z$, а так же значения среднеквадратических отклонений $P-X\sigma$, $P-Y\sigma$ и $P-Z\sigma$ для получения информации об угловых скоростях БПЛА на момент регистрации состояния объекта.

2 Алгоритм сбора и регистрации полётной информации БПЛА

Алгоритм сбора и регистрации полётной информации о пространственном положении и движении БПЛА для расчёта его аэродинамических характеристик реализован в виде программы на языке *C* для микроконтроллера, являющегося вычислителем системы сбора полётной информации (см. рисунок 2). Программа разделяется на две части. В первой части создаются три структуры с полями, соответствующими принимаемым данным от *IMU* и *RTK* блоков и полям выходного пакета для формирования записи переработанных данных на внешний накопитель. Эти структуры являются буферами для чтения и записи данных. Вторая часть программы представляет из себя бесконечный цикл, в каждом проходе по циклу производится опрос шин, на которых находятся *IMU* и *RTK* блоки. Так как время регистра-

ции данных от блоков не детерминировано, то принятые данные записываются в буферы, объявленные в первой части программы.

В структуре буферов во избежание наложения записей прописаны дополнительные поля для временной метки. Когда оба буфера принимаемой информации заполнены, программа переходит к считыванию необходимых полей для записи на внешний накопитель и формирует запись буфера выходного пакета. После формирования выходного пакета буферы, принимающие информацию, очищаются для следующей записи, а функция записи считывает данные из буфера отправки и записывает данные на внешний накопитель.

В бесконечном цикле последовательно считываются данные из шин *I2C* и *UART* (см. рисунок 1) с помощью функций *cbfuncrtk* и *cbfuncimu* (см. рисунок 2). Результатом является запись в буферные структуры. Если обе структуры были успешно записаны, то данные передаются функции *rec_proc* для преобразования в выходной формат, после чего производится запись на внешний носитель.

В результате получена схема, состоящая из двух структур, описывающих блок данных, полученных от инерциальной системы, и блок данных, полученный от *RTK* на борту ЛА. Из полученных структур формируется пакет для регистрации, выгрузки и последующего изучения полученных параметров. Пакет формируется в текстовом формате. Универсальная текстовая форма позволяет конвертировать полученные данные в форматы, удобные для машинной обработки или чтения их в текстовом редакторе (см. рисунок 3). На способ определения аэродинамических характеристик летательного аппарата по результатам лётного эксперимента получен патент на изобретение (см. рисунок 4).

Аппаратный блок (см. рисунок 5) выполнен в виде микроконтроллера, который выполняет функцию вычислителя разрабатываемой системы, и группы дополнительного периферийного оборудования для сопряжения с блоками *RTK* и *IMU*.

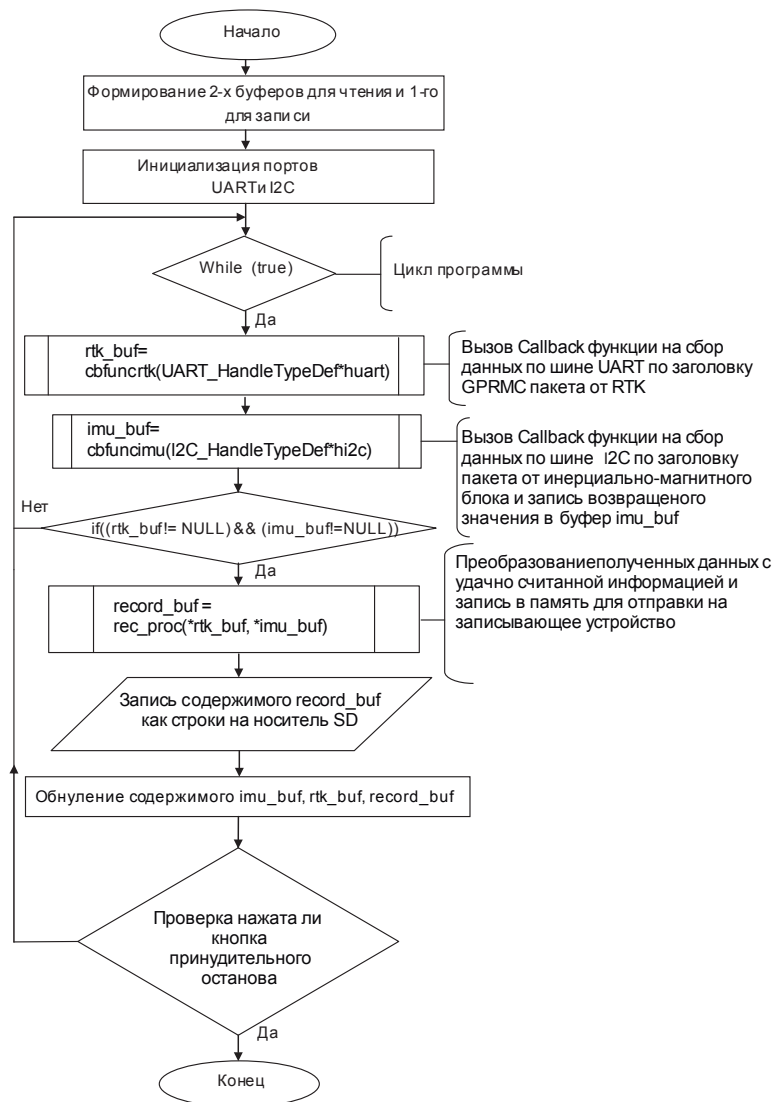


Рисунок 2 – Алгоритм сбора и регистрации полётной информации для расчёта аэродинамических характеристик БПЛА

Аппаратный блок устанавливается на платформу в корпус испытательного БПЛА (см. рисунок 6), где он сопрягается с информационной сетью (*UART*) устройства автоматического пилотирования.

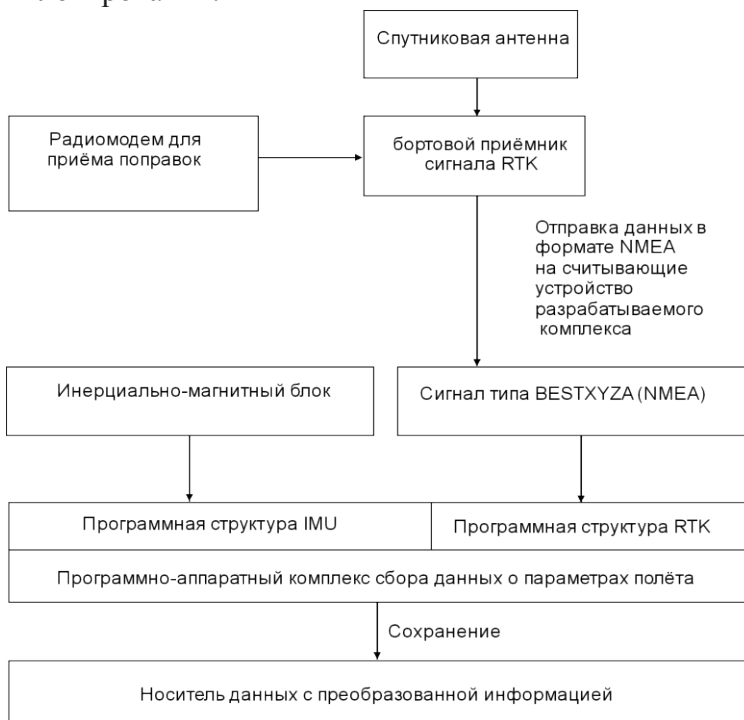


Рисунок 3 – Структурная схема получения и сохранения полётных данных



Рисунок 4 – Патент на изобретение № 2790358 от 17.02.2023

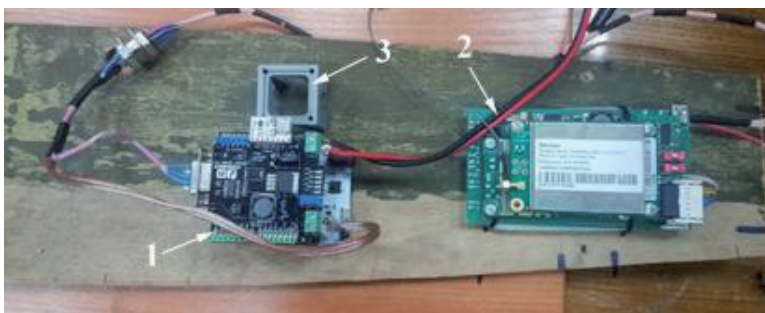


Рисунок 5 – Общий вид аппаратного блока: контроллер (1), блоки *RTK* (2) и *IMU* (3)



Рисунок 6 – Аппаратный блок, установленный в корпус перед монтажом основной системы управления БПЛА

3 Использование накопленной информации для определения характеристик БПЛА

По накопленным данным можно определить параметры движения БПЛА, необходимые для вычисления аэродинамических характеристик:

$$\Psi = \arctg \frac{V_N}{V_E}, \quad \Theta = \arctg \frac{V_H}{\sqrt{V_N^2 + V_E^2}}, \quad (1)$$

где θ – угол наклона траектории, Ψ – угол направления горизонтального движения, V_N – проекция истинной скорости на северное направление, V_E – проекция истинной скорости на восточное направление, V_H – вертикальная проекция истинной скорости.

Чтобы определить аэродинамические характеристики, необходимо построить модель пространственного движения БПЛА, которая имеет вид [16]:

$$m\ddot{X} = mg + M_\gamma^{-1}M_\vartheta^{-1}M_\psi^{-1}P + M_\gamma^{-1}M_\Theta^{-1}M_\Psi^{-1}R_A, \quad (2)$$

где $\ddot{X}$ – вектор вторых производных от перемещений в неподвижной системе координат, g – вектор ускорения силы тяжести, P – вектор силы тяги двигателя и других сил, связанных с планером БПЛА, R_A – вектор аэродинамических сил, m – масса БПЛА, M^{-1} – обратные матрицы поворота, необходимые для определения проекций сил, заданных в системе координат БПЛА или скоростной системе координат, на земную систему координат для углов поворота: γ – угол крена, ϑ – угол тангажа, ψ – угол курса БПЛА, Θ – угол наклона траектории, Ψ – угол направления горизонтального движения.

Зная уравнение движения, можно построить систему линейных уравнений [16]:

$$F = A^{-1}B, \quad (3)$$

где A – матрица коэффициентов, B – вектор правых частей, содержащий результаты измерений. Матрицы A и B имеют следующий вид:

$$A = M_\gamma M_\vartheta M_\psi M_\gamma^{-1} M_\Theta^{-1} M_\Psi^{-1},$$

$$B = m\ddot{X}_{\text{ст}} - M_\gamma M_\vartheta M_\psi mg - M_\gamma M_\vartheta M_\psi M_\gamma^{-1} M_\Theta^{-1} M_\Psi^{-1} P. \quad (4)$$

Прямые и обратные матрицы поворотов не являются перестановочными, поэтому заменить присутствующие в произведениях пары прямых и обратных матриц поворота вокруг одной и той же оси на единичную матрицу не получится. В системе уравнений не присутствуют в явном виде углы атаки и скольжения, что с одной стороны требует дополнительно их вычисления при обработке данных, с другой стороны облегчает использование в процессе лётного эксперимента высокоточных навигационных систем *РТК*. В процессе штатной работы определяются только проекции истинной скорости на земную систему координат, что позволяет легко определять углы наклона траектории и горизонтального направления движения, а не углы атаки и скольжения [16].

После определения аэродинамических сил можно вычислить их коэффициенты.

Пример результата вычислительного эксперимента показан на рисунке 7. Этап первый –

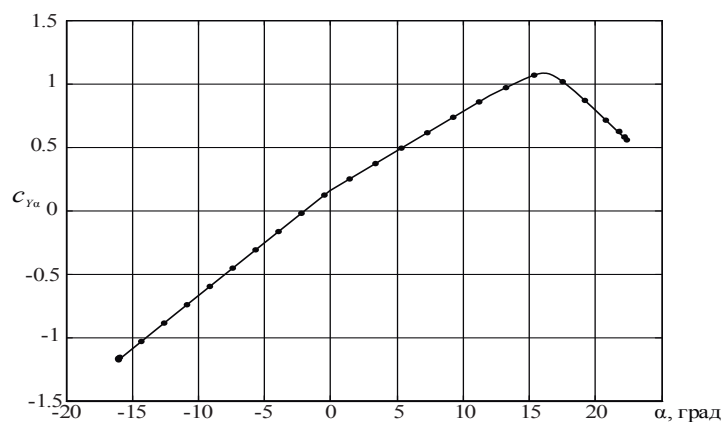


Рисунок 7 – Пример определения аэродинамического коэффициента подъемной силы по результатам эксперимента

имитация движения с помощью динамической модели движения ЛА с заданными аэродинамическими коэффициентами (линия на рисунке 7). Этап второй – решение обратной задачи согласно предложенному алгоритму для получения аэродинамических коэффициентов ЛА (точки на рисунке 7), движение которого моделировалось на первом этапе. Высокая точность совпадения в данном случае объясняется тем, что исследовались результаты численного эксперимента, в котором все случайные факторы были заранее известны. Для проведения

реального, лётного эксперимента необходима специальная измерительная система, проектированию которой и посвящена данная статья.

Заключение

Разработана программа, решающая две задачи - измерения и регистрации параметров полёта БПЛА. Задача записи на внешний носитель решается сразу после считывания, пошаговое преобразование внутри задач не вызывает задержек во время выполнения.

Разработанная система и программное обеспечение позволяют собирать информацию, необходимую для определения аэродинамических характеристик БПЛА с помощью метода, предложенного в статье [16].

Список источников

- [1] **Просвирина Н.В.** Анализ и перспективы развития беспилотных летательных аппаратов // Московский экономический журнал. 2021. №10. С.560-575. DOI: 10.24412/2413-046X-2021-10619.
- [2] **Маммадов А.З.** Модель инерциальной навигации для беспилотных летательных аппаратов // Universum: технические науки : электрон.научн. журн. 2021. 5(86) С.5-9. DOI:10.32743/UniTech.2021.86.5.11683: <https://7universum.com/ru/tech/archive/item/11683>.
- [3] **Кузнецов Г.А., Кудрявцев И.В., Крылов Е.Д.** Ретроспективный анализ, современное состояние и тенденции развития отечественных беспилотных летательных аппаратов // Инженерный журнал: наука и инновации. 2018. № 9. С.1-22. DOI: 10.18698/2308-6033-2018-9-1801.
- [4] **Guo Hongtao, Yu Yan, Hongya Xia, Li Yu, Binbin Lv.** The Prediction and Correction Method of Aircraft Static Aeroelastic Effects: A Review of Recent Progress. *Actuators* 2022, no. 11: 309. P.1-20 DOI:10.3390/act11110309.
- [5] **Mazin Abdulaali Hamzah, Mikhaylov S., Makhanko A.** Algorithm for Determining Aerodynamic Characteristics based on the Results of a Flight Experiment. *Design Engineering*, 2022, 14547-14559. <http://thedesigengineering.com/index.php/DE/article/view/8638>.
- [6] **Солдаткин В.В., Солдаткин В.М.** Построение модели и обработка сигналов многофункционального приемника в аэротрических системах самолета // Изв. вузов. Приборостроение. 2021. т. 64, № 7. С. 551–558 DOI: 10.17586/0021-3454-2021-64-7-551-558.
- [7] **Рендал У. Биард, Тимоти У.Маклэйн.** Малые беспилотные летательные аппараты: теория и практика: ТЕХНОСФЕРА, 2015. 312 с. ISBN 978-5-94836-393-6.
- [8] **Lukyanov O.E., Barcenas O.U.E., Zolotov D.V.** Experimental Model of an Electric Power Plant for Small UAV's Automatic Control Systems. *2021 International Scientific and Technical Engine Conference (EC)*, Samara, Russian Federation, 2021. P.1-4, DOI: 10.1109/EC52789.2021.10016802.
- [9] **Лукьянов О.Е., Золотов Д.В.** Методологическое обеспечение подготовки проектантов и операторов беспилотных летательных аппаратов // Вестник Самарского университета. Т.20, №1, 2021, С.14-28, DOI:10.18287/2541-7533-2021-20-1-14-28.
- [10] **Mohsan S.A.H., Khan M.A., Noor F., Ullah, I., Alsharif M.H.** Towards the Unmanned Aerial Vehicles (UAVs): A Comprehensive Review. *Drones* 2022, 6, 147. (1-27). DOI:10.3390/drones6060147.
- [11] **Pflüger J, Von Langsdorff M., Breitsamter C.** Efficacy of an elasto-flexible morphing wing at high lift using fluid-structure-interaction simulations. *Front. Aerosp. Eng.* 2022 1:975600. (1-14). DOI:10.3389/fpace.2022.9756004.
- [12] **Stöcker C., Bennett R., Nex F., Gerke M., Zevenbergen J.** Review of the current state of UAV regulations. In: *RemoteSensing* 9.5. 2017. P.33–35. DOI:10.3390/rs9050459.
- [13] **Михайловский К.В., Барановски С.В.** Учёт обледенения поверхности крыла из полимерных композиционных материалов при проектных расчётах. *Известия высших учебных заведений. Машиностроение*, 2019, № 3. С.61–70, DOI:10.18698/0536-1044-2019-3-61-70.
- [14] LIS331DLH. MEMS digital output motion sensor ultra low-power high performance 3-axes “nano” accelerometer. July 2009 Doc ID 15094 Rev 3. 38 p. <https://www.st.com/resource/en/datasheet/lis331dlh.pdf>.
- [15] BESTXYZ. Best available cartesian position and velocity <https://docs.novatel.com/OEM7/Content/Logs/BESTXYZ.htm>.
- [16] **Михайлов С.А., Маханько А.А., Хамза М.А.Х.** Алгоритм определения аэродинамических характеристик по результатам лётного эксперимента // Изв. вузов. Авиационная техника. 2022. № 1. С.123–130.

Сведения об авторе

Хамза Мазин Абдулаали Хамза, 1981 г.рождения. Окончил Технологический университет Багдада (UOT), Ирак (2004), магистратуру Тамбовского государственного технического университета (2014), аспирант 4-го курса кафедры аэрогидродинамики Казанского национального исследовательского технического университета им. А.Н. Туполева - КАИ. ORCID:0000-0003-2782-7000, ResearcherID (WoS): HKO-3541-2023, (SPIN-код): 8990-0688, AuthorID (РИНЦ) 1178112. mazinhamza641@gmail.com.



Поступила в редакцию 09.12.2022, после рецензирования 20.02.2023. Принята к публикации 28.02.2023.



Scientific article

DOI: 10.18287/2223-9537-2023-13-1-90-98

Designing a flight information collection system for calculating the aerodynamic characteristics of a UAV

© 2023, Hamzah Mazin Abdulaali Hamzah

Kazan National Research Technical University named after A.N. Tupolev – KAI, Kazan, Russia
University of Technology (UOT), Baghdad, Iraq

Abstract

The article describes the development of an onboard information collection system for calculating the aerodynamic characteristics of unmanned aerial vehicles (UAVs), which can only be obtained under free flight conditions. A functional diagram is presented, a flight parameters registration system operation algorithm is described. The proposed solutions make it possible to collect a large amount of information and improve system efficiency through the selection and compression of information. The collected information is used to determine the aerodynamic characteristics, eliminating the need to test UAVs models in wind tunnels, which can save money and time on the development of new UAVs. The features of the information exchange between the system elements are considered, which make it possible to speed up the data transfer, and therefore improve the accuracy of measurements by increasing the frequency of polling measuring devices. The results obtained may be of interest to developers of autonomous measuring systems. Of particular importance for calculating the aerodynamic characteristics of an UAV using the proposed method is the high accuracy of measuring the parameters of the aircraft trajectory, including the linear accelerations and angular velocities that occur in the course of motion. In addition, the article considers a variant of the hardware implementation of the proposed system for UAVs.

Key words: aerodynamic characteristics, flight parameters, software and hardware complex, collection of information, data transfer, UAV.

For citation: Hamzah Mazin Abdulaali Hamzah. Designing a flight information collection system for calculating the aerodynamic characteristics of a UAV [In Russian]. *Ontology of designing*. 2023; 13(1): 90-98. DOI:10.18287/2223-9537-2023-13-1-90-98.

Acknowledgment: The author is grateful to the members of the editorial board of the journal «*Ontology of designing*» for their comments and recommendations for improving this article.

Conflict of interest: The author declares no conflict of interest.

List of figures

Figure 1 - The structure of the information collection system for determining flight parameters

Figure 2 - Algorithm for collecting and registering flight information to determine flight parameters

Figure 3 - Block diagram of receiving and saving flight data

Figure 4 - Invention Patent No. 2790358 of 02/17/2023

Figure 5 - General view of the hardware unit: controller (1), RTK (2) and IMU (3) units

Figure 6 - The hardware unit installed in the aircraft body before mounting the main UAV control system

Figure 7 - An example of determining the aerodynamic lift coefficient based on the results of the experiment

References

- [1] **Prosvirina HV.** Analysis and prospects for the development of unmanned aerial vehicles [In Russian]. *Moscow Economic Journal*. 2021; 10: 560-575. DOI: 10.24412/2413-046X-2021-10619.
- [2] **Mammadov AZ.** Model of inertial navigation for unmanned aerial vehicles [In Russian]. *Universum: technical sciences: electron. scientific magazine* 2021; 5(86): 5-9. DOI:10.32743/UniTech.2021.86.5.11683.
- [3] **Kuznetsov GA, Kudryavtsev IV, Krylov ED.** Retrospective analysis, current state and development trends of domestic unmanned aerial vehicles [In Russian]. *Engineering Journal: Science and Innovations*. 2018; 9: 1-22. DOI: 10.18698/2308-6033-2018-9-1801.
- [4] **Guo Hongtao, Yu Yan, Hongya Xia, Li Yu, Binbin Lv.** The Prediction and Correction Method of Aircraft Static Aeroelastic Effects: A Review of Recent Progress. *Actuators* 2022 11, no.11: 309. (1-20) DOI:10.3390/act11110309.
- [5] **Hamzah MA, Mikhaylov S, Makhanko A.** Algorithm for Determining Aerodynamic Characteristics based on the Results of a Flight Experiment. *Design Engineering*, 202214547-14559. <http://thedesigengineering.com/index.php/DE/article/view/8638>.
- [6] **Soldatkin VV, Soldatkin VM.** Building a model and signal processing of a multifunctional receiver in aircraft aerometric systems [In Russian]. *Izv. universities.instrumentation*. 2021; 64(7): 551–558. DOI:10.17586/0021-3454-2021-64-7-551-558.
- [7] **Rendal W. Beard and Timothy W. McLain.** Small unmanned aerial vehicles: theory and practice. TECHNO-SPHERE, 2014. 312 p. ISBN 978-5-94836-393-6.
- [8] **Lukyanov OE, Barcenas OUE, Zolotov DV.** Experimental Model of an Electric Power Plant for Small UAV's Automatic Control Systems, 2021 International Scientific and Technical Engine Conference (EC), Samara, Russian Federation, 2021, pp.1-4, DOI:10.1109/EC52789.2021.10016802.
- [9] **Lukyanov OE, Zolotov DV.** Methodological support for the training of designers and operators of unmanned aerial vehicles [In Russian]. *Bulletin of the Samara University*. 2021; 20(1): 14-28, DOI: 10.18287/2541-7533-2021-20-1-14-28.
- [10] **Mohsan SAH, Khan MA, Noor F, Ullah I, Alsharif MH.** Towards the Unmanned Aerial Vehicles (UAVs): A Comprehensive Review. *Drones* 2022; 6, 147(1-27). DOI:10.3390/drones6060147.
- [11] **Pflüger J, Von Langsdorff M, Breitsamter C.** Efficacy of an elasto-flexible morphing wing at high lift using fluid-structure-interaction simulations. *Front. Aerosp. Eng.* 2022. 1:975600. P.1-14. DOI:10.3389/fpace.2022.9756004.
- [12] **Stöcker C, Bennett R, NexF, Gerke M, Zevenbergen J.** Review of the current state of UAV regulations. In: *Remote Sensing* 9.5 (2017). P.33–35. DOI:10.3390/rs9050459.
- [13] **Mikhailovsky KV, Baranovsky SV.** Accounting for icing of the wing surface made of polymer composite materials in design calculations [In Russian]. *News of higher educational institutions. Engineering*. 2019; 3: 61–70. DOI: 10.18698/0536-1044-2019-3-61-70.
- [14] LIS331DLH.MEMS digital output motion sensor ultra low-power high performance 3-axes “nano” accelerometer. July 2009 Doc ID 15094 Rev 3. 38 p. <https://www.st.com/resource/en/datasheet/lis331dlh.pdf>.
- [15] BESTXYZ. Best available cartesian position and velocity <https://docs.novatel.com/OEM7/Content/Logs/BESTXYZ.htm>.
- [16] **Mikhailov SA, Makhanko AA, Hamza MAH.** Algorithm for determining aerodynamic characteristics based on the results of a flight experiment [In Russian]. *Izv. universities. Aviation technology*. 2022; 1: 123–130.

About the author

Hamzah Mazin Abdulaali Hamzah (b.1981) Graduated from the University of Technology (UOT), Baghdad, Iraq (2004), received Master's degree in Tambov State Technical University (2014), 4th year postgraduate student at the Department of Aerohydrodynamics in Kazan National Research Technical University named after A.N. Tupolev - KAI. ORCID:0000-0003-2782-7000, Researcher ID (WoS): HKO-3541-2023, (SPIN-код): 8990-0688, Author ID (RSCI) 1178112. mazinhamza641@gmail.com

Received December 9, 2022. Revised February 20, 2023. Accepted February 28, 2023.



Designing an ontology of the e-learning course content

© 2023, A.V. Solovov ✉, A.A. Menshikova

Samara National Research University, Samara, Russia

Abstract

Knowledge modeling, closely related to ontologies, is an important semantic technology and research area. The article deals with the e-learning course content model concept. The content model is based on structuring the content into separate fragments, called learning elements. These learning elements integrate into a tree directed graph. The content model is defined as a combination of such a graph and a table of attributes of educational elements with requirements for didactic indicators of their study. The rules for building models of the electronic educational content are formulated. The mathematical properties of these models are discussed and their integral characteristics are introduced. The proposed approach to content modeling is in line with the SCORM specifications for international e-learning, complements them with targets, didactic design algorithms and analysis of educational materials. Formation algorithms and methods of presenting the content model make it possible to automate the process of its construction and didactic analysis in the form of a visual interactive dialogue between developers of electronic educational resources in instrumental author's environments.

Key words: *e-learning, electronic educational resources, structuring of educational material, content model, tree-oriented graphs, SCORM.*

For citation: Solovov A.V., Menshikova A.A. Designing an ontology of the e-learning course content. *Ontology of designing*. 2023; 13(1): 99-112. DOI:10.18287/2223-9537-2023-13-1-99-112.

Conflict of interest: The authors declare no conflict of interest.

Introduction

In the past few years, knowledge modeling, closely related to ontologies, has been an important semantic technology and research area [1]. However, "while ontologies have become the de facto standard in the field of knowledge base development, the processes of extracting and especially structuring knowledge still remain a kind of "blank spot" in the modern literature on knowledge engineering" [2, p.88].

Structuring knowledge is essential in learning. It is the structural "dissection" of knowledge for their presentation in the framework of lectures, in various types of textbooks (printed or electronic) that has always been and continues to be one of the main functions of the teacher. Structuring of educational material can be defined as the process of organizing information to improve its understanding and memorization. As a result of this process, fragments of the studied material are connected in meaning into an integral group or several such groups.

The design of e-learning also begins with the structuring of educational material. There are different, including ontological approaches to the formalization of this process. In the work [3] as a formal basis for individualized e-learning, it is proposed to use semantic models that include the apparatus of vector representations of knowledge graphs, which has the flexibility and expressiveness of the ontological approach. Modeling of e-learning processes using directed graphs is offered in the work [4]. The article [5] presents a formal description of the structural-hierarchical didactic model of e-learning. A distinctive feature of this model is the support for dividing educational objects into didactic components. The book [6] proposes the concept of electronic educational resources integrated into a multimedia system open for development.

The development of the problems of e-learning makes it possible to transfer the educational process to the industrial "rails", to introduce specialization and division of labor into it. Industrialization entails the unification and standardization of various educational procedures. The most famous are the standards of international organizations AICC [7], IMS [8], ADL [9]. An overview of the various standards is given in [10, 11].

The basis of international unified procedures for structuring educational materials since the late 90s are the SCORM (The Sharable Content Object Reference Model) specifications [12] and its development in the xAPI (Tin Can) and cmi5 specifications [13]. One of the basic ideas of SCORM is the compilation of electronic educational resources from blocks of educational material, called Sharable Content Objects (SCOs). Such objects may include semantically local text fragments, graphic illustrations, computer programs, video clips, any other typical elements of hypermedia or their combinations.

SCORM does not impose restrictions on the size of SCOs and contact training time with them. At the same time, it is assumed that the object represents a relatively small part of the content of the studied educational material. The content developer should determine the size of the SCO based, first, on the amount of information needed to achieve the learning outcome, and second, on the degree of multiple use that the developer wants to obtain.

Various SCOs are placed in network depositories (corporate or global), which provides access to them to users of these networks. Developers of training materials, using metadata about SCOs, find suitable objects and arrange from various SCOs their aggregation in the form of electronic textbooks, computer courses, etc. The developer does not always copy the selected SCOs. You can specify only their network URLs. The collected aggregation is hosted in a Learning Management System (LMS) that supports the SCORM specifications. Any such LMS can run and execute SCOs, regardless of the technology platform on which these learning objects were created.

However, the SCORM specifications do not contain specific structuring techniques and models, making them difficult to apply in practice. The Russian School of Didactics has advanced research experience in the field of structuring educational materials. The most famous in this regard are the didactic developments of V.P. Bespalko [14] and E.L. Belkin [15]. In our research, these developments have been adapted and developed in relation to the design of electronic educational resources (EER) [16]. The models for structuring training materials proposed in the works [14-16] are adequate to the basic concepts of SCORM and complement them in terms of didactic goal-setting of SCOs. However, these models do not have a mathematical justification, and the methodology for their construction is focused on the usual, non-automated procedures for designing educational material.

The purpose of this research is to provide a mathematical justification for structuring models [14-16], to investigate the properties and introduce integral characteristics of these models, allowing for didactic analysis and construction of automated procedures for designing the structure of educational material. The research is based on methods of system analysis, discrete mathematics, pedagogical psychology and didactics, many years of experience of the authors in the field of education, theory and technologies of e-learning.

1 Content model

In accordance with [16], the educational material planned for study is divided into separate learning elements (LE). LE is understood as objects, phenomena, concepts, methods of activity selected from the relevant science and included in the curriculum of the academic discipline or section of the academic discipline for their study. The set of LE is presented in the form of a structural scheme, which is called the content graph (CG) of the educational material. The nodes (vertices) of

the graph are LE, the edges are hierarchical connections between them. Note that the concept of LE and the presentation of the structure of the educational material in the form of CG are equivalent to the concept of SCOs and their aggregations in SCORM.

In parallel with the construction of the CG, the LE attribute specification (table) is compiled, in which the LE names are entered. An analogue of this process is the compilation of the table of contents of the textbook, when its content is preliminarily divided into sections, subsections or chapters and paragraphs. However, when constructing CG educational material, unlike compiling a table of contents, there is no need to care about the sequence of presentation of LE. It is important to display only the hierarchical structure of the educational material. After structuring and selecting the content of the educational material for each LE, didactic requirements are formulated for the level of assimilation α ($\alpha \in 0, 1, 2, 3, 4$), the level of presentation β ($\beta \in 1, 2, 3, 4$) and the level of awareness γ ($\gamma \in 1, 2, 3$) of the educational material, which are included in the specification of the LE [16, p.12]. At the same time, for each indicator, one or two columns of the LE table are filled.

In the first column, which is not always included in the specification, the "starting" value of the indicator (the estimated level before training) is indicated, in the second column, which is mandatory for inclusion in the specification, the "finish" value of the indicator (the required level after training). Note that the first versions of SCORM (SCORM-2) did not contain such elements of didactic goal-setting. In the latest version (SCORM-4), this gap was partially filled by the inclusion in the characteristics of SCOs of didactic goals based on the taxonomy of the Bloom-Anderson level of knowledge [17].

The totality of the CG and the specification of the attributes of the LE is called the model of the content of the educational material of the EER [16]. As an illustrative example, this article discusses the content model prepared for a fragment of educational material on the theory of orgraphs from the book [18] (Figure 1). Here, in the specification of the attributes of the LE, approximate didactic requirements for the level of knowledge of students of a technical university studying a course of discrete mathematics are indicated.

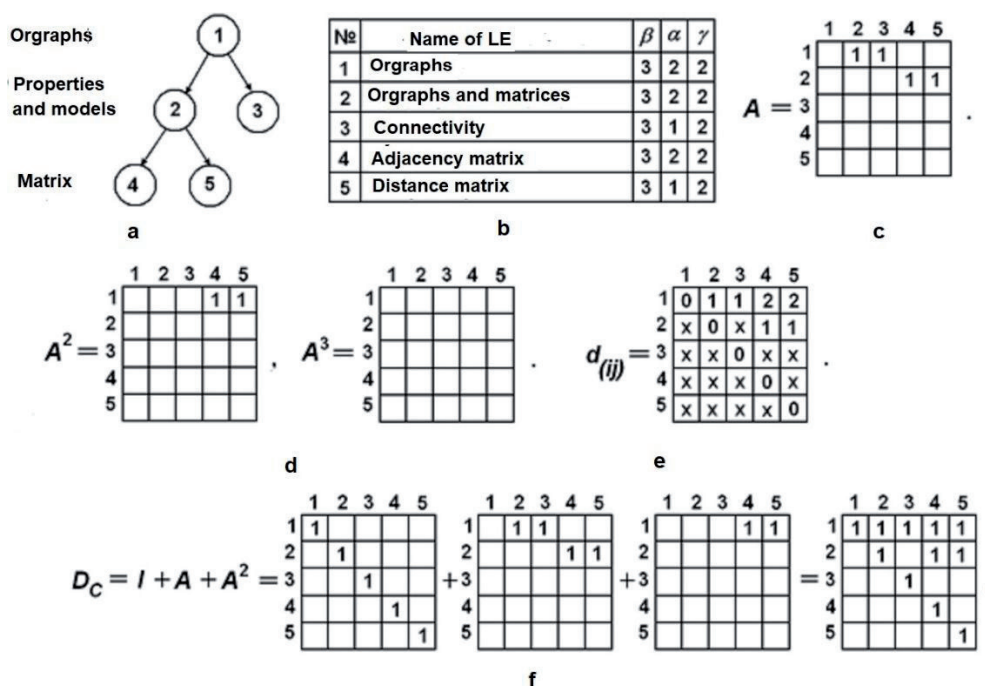


Figure 1 – Example of a content model:

a – content graph; b – specification of LE attributes; c – the matrix of adjacency of CG; d – the degree of the adjacency matrix CG; e – the distance matrix of the CG; f – CG achievability matrix

2 Definition and rules for constructing a content graph

We will represent the CG as a oriented graph of the tree structure $D = (V, Y)$, where V is the finite set of n vertices (the set of LE), and Y is the finite set m of the oriented edges (hierarchical connections between the LE) of the orgraph. When constructing the CG, we will observe the following rules (see Figure 1, a):

- 1) the graph has the form of an inverted tree with one root vertex - one LE corresponding to the name of the topic being structured;
- 2) communication (orientation of the edges) is carried out only in the direction from the root (from top to bottom);
- 3) there are no separate (hanging) vertices to which there is no connection (arc) from the higher LE, except for the root;
- 4) only one arc from higher LE can approach a lower LE in the hierarchy;
- 5) higher LE should be associated with at least two lower LE, otherwise the lower LE is included in the higher LE;
- 6) grouping of LE at the same level is carried out on any common basis (general basis);
- 7) the numbering of the vertices of the CG begins at the root and continues sequentially along the levels of grouping of the LE from top to bottom and from left to right. Sometimes it is convenient to number the vertices of the CG in the same way as the table of contents of printed materials. Then the root vertex of the CG is assigned the number 0, the vertices of the first level - 1, 2, 3, ..., the vertices of the second level - 1.1, 1.2, 1.3, ..., 2.1, 2.2, 2.3 ... etc.

We will also assume that the content of lower LE is not a simple decomposition (fragmentation) of the content of the associated higher LE. In particular, the content of lower LE can detail, disclose the individual components of the content of the associated higher LE. Conversely, the content of the higher LE, although it integrates the content of the associated lower LE, is not a simple unification of them.

The mathematical model of CG is its adjacency matrix A (see Figure 1, c). When it is filled, the rows and columns of the matrix are put in accordance with the LE numbers, which are located on the left and top of the matrix. The cells in this matrix can contain zeros or ones. Zero means that there is no hierarchical relationship between the LE specified in the row number and the LE specified in the column number (there is no edge in the CG). Zeros, as a rule, are not put, since the matrix of adjacency of the CG is usually weakly filled. One is placed in the cell of the matrix when there is a hierarchical relationship between the LE. For example, the units in cells 1-3 and 2-5 indicate the presence of corresponding edges in the CG between LE 1 and LE 3, between the LE 2 and the LE 5 (see Figure 1, c).

3 Content graph properties

Property 1. The number of CG arcs is one less than the number of its vertices, $m=n-1$, with $n \geq 1$ and $n \neq 2$.

The CG can be constructed by starting with the root vertex and sequentially adding typical fragments in the form of one vertex and an arc entering it (Figure 2). It follows that the number of arcs of the CG will be one less than the number of its vertices. An exception is the case of $n=2$, in which the CG cannot be constructed, since according to Rule 5 of the CG construc-

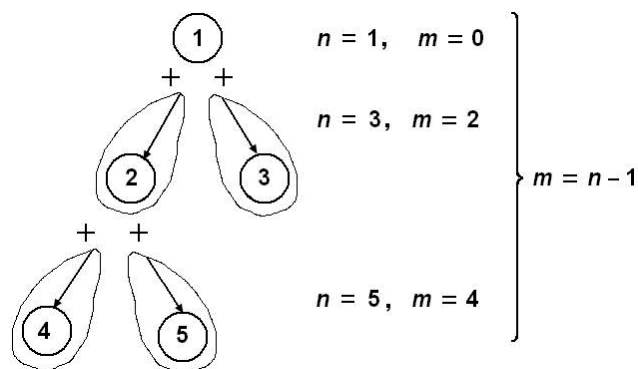


Figure 2 – To property 1 of the content graph

tion, the higher LE must be associated with at least two lower LE.

Property 2. The columns of the vertex adjacency matrix CG $A=(a_{ij})$ contain only one unit except for the column corresponding to the root vertex, which contains only zeros. This property is determined by the fact that, according to the rules of construction of CG 3-4, there is only one incoming arc in any vertex of the CG, except the root (see Figure 1, c).

Property 3. For CG with an adjacency matrix $A=(a_{ij})$, the element $a_{ij}^{(t)}$ in the matrix A^t , where t is the power, may be 0 or 1. The unit defines a single simple (without repeating vertices) path from vertex v_i to vertex v_j of length t .

If $t=1$, then the result is obvious – the adjacency matrix A indicates the presence of single-length paths (see Figure 1, c). Let be $t=2$. To go from vertex v_i to v_j in two steps, you need to go from v_i to some vertex v_k in one step and then from v_k to v_j in the next step. The transition from v_i to v_k is determined by the coefficient a_{ik} of the matrix A , the transition from v_k to v_j is determined by the coefficient a_{kj} . The transition from v_i to v_j via v_k is determined by the sum of $\sum_{k=1}^n a_{ik} a_{kj}$.

This sum is the coefficient $a_{ij}^{(2)}$ of the matrix A^2 . From property 2 of the CG (see above) it follows that in column k of coefficients a_{ik} and in column j of coefficients a_{kj} , only one coefficient can be equal to one, and the remaining coefficients are zero. Therefore, each column j of the matrix A^2 can be either completely zero or contain one unit, i.e. the path from v_i to v_j , if any, is the only and simple. By making similar reasoning, it is possible to show the validity of this property for A^3 , etc. for A^t (see Figure 1, d).

Property 4. All paths in the CG are simple (with no repeating vertices). According to property 3, each column of the matrix A^t can either be completely zero or contain one unit, i.e. the path from any vertex v_i to another vertex v_j , if any, is the only and simple (without repeating vertices).

Property 5. Let the CG have an adjacency matrix A and a distance matrix (d_{ij}) . Then, if the value d_{ij} ($i \neq j$) is defined, then it is equal to t , for which the coefficient $a_{ij}^{(t)}$ in A^t is 1. For $i=j$ $d_{ii}=0$.

The proof follows from property 3 (see above), according to which the coefficients of the matrix A^t indicate all simple paths of length t in CG. Zeros on the principal diagonal of the matrix (d_{ij}) determine the path length of the corresponding vertex to itself (see Figure 1, d, e).

Property 6. Any vertex of the CG is reachable from its root, and to each vertex there is a single and simple path from the root.

Let's start moving from any vertex towards the root in the direction opposite to the orientation of the edges. On this path, there will be only one possible direction in each branch (vertex) of the orgraph, since any vertex of the CG, except for the root vertex, has only one incoming edge. Given that the CG has no higher hanging vertices other than the root, such an advance will have only one trajectory necessarily leading to the root, and therefore, conversely, from the root to any vertex there is necessarily a single and simple path, i.e. all vertices are reachable from the root. Note that any vertex is considered a path, so the root top is achievable for itself.

Property 7. The achievability matrix D_c of the CG is determined through its adjacency matrix A by the formula

$$D_c = I + A + A^2 + \dots + A^{(n-1)/2}. \quad (1)$$

The first term of this formula, unit matrix I , determines the fact that each vertex of the CG is achievable for itself. The subsequent terms indicate all possible paths in the CG of length 1, 2, ..., $(n-1)/2$, the units in the columns of the matrices $A, A^2, \dots, A^{(n-1)/2}$ indicating these paths are at different positions and do not coincide. The last term corresponds to the longest (potentially) simple path to the CG. Its length is $m/2 = (n-1)/2$, since according to rule 5 of the CG construction, the higher vertex must be adjacent to at least two lower vertices. Consequently, the summation result of for-

mula (1) indicates all simple paths in the CG and thus determines the achievability matrix (see Figure 1, f).

Property 8. Any two vertices of the CG are connected.

The proof follows from the consideration that from any vertex of the CG there is a simple half-way to the root, and from the root any vertex of the CG is achievable. Therefore, any two vertices of the CG are connected at least through the root.

Property 9. The CG is a weakly connected (weak) orgraph with a degree (category) of connectivity equal to 1.

This property is defined by the fact that any pair of CG vertices is conjugated (see property 8), but has neither the properties of a strongly connected orgraph (i.e., the two-way reachability of all vertices) nor a one-way connected orgraph (i.e., the one-way reachability of all vertices) [18].

4 Integral characteristics of the content model

Let's introduce some characteristics that allow you to analyze the structure of educational materials.

1) **Number of learning elements n .** This characteristic determines the number of vertices of the CG and characterizes, in a certain, but, of course, not fully the amount of educational material. The value of $n \geq 1$ and $n \neq 2$ (see property 1 CG).

2) **The number of levels (bases) of structuring U .** The value of U shows the number of levels (the depth of structuring of the educational material), the degree of hierarchical nesting of some educational elements into others. It is defined by the following two theorems.

Theorem 1. For CG with adjacency matrix A , the exponent of degree t in the series of matrices $A, A^2, \dots, A^t, A^{t+1}, \dots$ determines the number of levels of structuring U if there are at least two ones in addition to zeros among the A^t coefficients, and in the A^{t+1} matrix all coefficients are zero.

Proof. In accordance with CG property 3 (see above), the exponent of the t in the matrix A^t determines the presence in the CG of paths of length t , and this length corresponds to the longest tracks. All paths to the CG are simple (see CG property 4) and only the lower level of the CG structure can be moved from any vertex. Therefore, the magnitude of the longest path t is equal to the number of levels of structuring U . At the last level of structuring, there must be at least two vertices (which corresponds to two units in the matrix A^t), since according to rule 5 of the construction of the CG, the higher vertex must be adjacent to at least two lower vertices.

Theorem 2. The maximum possible depth of structuring of the CG of the U_{max} training material depends on the number of LE n ($n \geq 1$, $n \neq 2$) and is determined by the following ratios:

$$U_{max} = (n-1)/2 \text{ for odd } n = 1, 3, 5, 7, \dots; \quad (2)$$

$$U_{max} = (n-2)/2 \text{ for even } n = 4, 6, 8, \dots. \quad (3)$$

Proof. The increment n from 1 or from 4 in steps 2 gives the maximum increment of U per unit if the structuring is performed according to the schemes shown in Figure 3. Summarizing these schemes, we get expressions (2, 3). The value of $n=2$ is excluded from consideration in accordance with rule 5 of the construction of the CG.

3) **Relative depth of structuring of educational material**

$$\bar{U} = U/U_{max}. \quad (4)$$

It is always useful to determine the value $\bar{U}$ and its proximity to a limit value equal to one to assess the use of the potential of hierarchical structuring. Thus, for the above example of CG (see Figure 1, a) $U=U_{max}=2$, and $\bar{U}=1$, which means the maximum possible degree of hierarchical structuring.

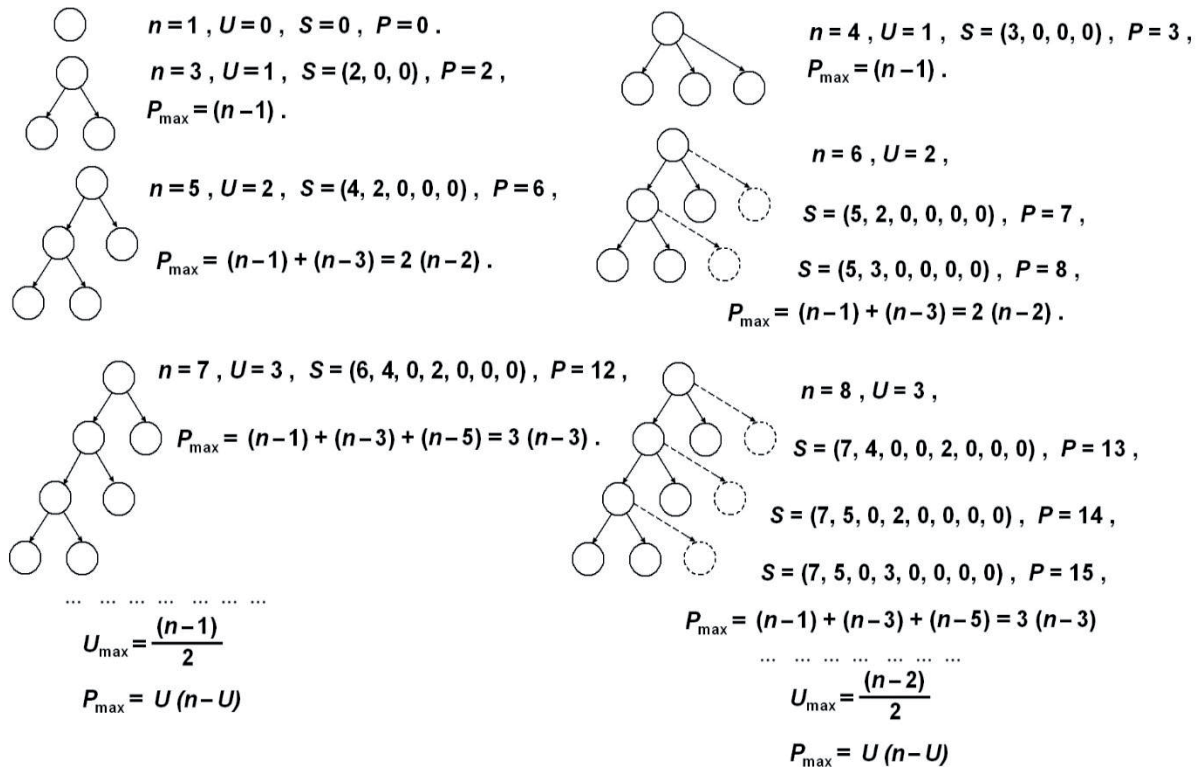


Figure 3 – To theorems 2 and 3

4) Vector of structuring of educational material

$$S = (D_c - I)E, \quad (5)$$

Where is D_c – achievability matrix; I – unit matrix; E – column vector of n units.

Vector S allows you to assess the degree of structuring of all LE. Each S_i coefficient of the vector S determines the scalar value - the degree of structuring of the LE with the number i (i.e., the number of lower LE included in it). Thus, for the above example of CG (see Figure 1, a) $S = (4, 2, 0, 0, 0)$. The analysis of the vector S makes it possible to clearly distinguish local, independent LE, the value of S_i for which is zero, and integrated LE, which generalize, hierarchically include other LE (the S_i value for such LE is greater than zero). Thus, LE with $S_i = 0$ can be used according to the SCORM ideology as local independent learning objects - SCOs. They can be prepared independently of other learning facilities and placed together with the appropriate meta description in EER repositories for repeated reuse.

5) **The degree of branching of the model of the content** of the educational material. Let's denote this characteristic P and define it by the formula:

$$P = E^T S = E^T (D_c - I) E. \quad (6)$$

The P value characterizes the branching of the CG of the educational material. It is related to the number of LE and the number of levels of structuring by the following theorem.

Theorem 3. The degree of branching P of the content model of the educational material depends on the number of levels of structuring U and the number of LE n ($n \geq 1, n \neq 2$) and is associated with them by inequalities:

$$n-1 \leq P \leq U(n-U); \quad (7)$$

$$n-1 \leq P \leq (n^2-1)/4 \text{ for odd } n = 1, 3, 5, 7, \dots; \quad (8)$$

$$n-1 \leq P \leq (n^2-2)/4 \text{ for even } n = 4, 6, 8, \dots. \quad (9)$$

Proof. The minimum level of branching at any $n \geq 3$ can be obtained if the number of structuring levels $U=1$ and all LE are directly related to the root. Then $P_{\min} = n-1$, which is also true for $n=1$.

Analyzing the structuring schemes in Figure 3, you can get a general formula for determining the P value that is maximum possible for the given values of U and n : $P_{max} = U(n-U)$. Thus, inequality follows from the above (7). Further, substituting in the expression (7) inequality (5, 6), we get inequalities (8, 9), respectively.

6) **Relative degree of branching** of the model of the content of the educational material

$$\bar{P} = P / P_{max} = P / (U(n-U)). \quad (10)$$

7) **Average level of presentation of educational material**

$$\beta_{cp.} = \sum_{i=1}^n \beta_i / n. \quad (11)$$

8) **Average level of assimilation of educational material**

$$\alpha_{cp.} = \sum_{i=1}^n \alpha_i / n. \quad (12)$$

9) **Average level of awareness of educational material**

$$\gamma_{cp.} = \sum_{i=1}^n \gamma_i / n. \quad (13)$$

Averaged target indicators determined by formulas (11-13) allow you to compare various training materials with each other, predict the complexity of their presentation during development, the laboriousness of preparing exercises for training and control. The greater the value of these indicators, the higher the labor intensity. For example, if $1 < \alpha_{cp.} < 2$, then the exercises for training and control should include two blocks: the first at the level of acquaintance ($\alpha = 1$), the second at the level of knowledge reproduction ($\alpha = 2$).

For the above content model example (see Figure 1), integral characteristics: $n = 5$, $U = 2$, $\bar{U} = 1$, $S = (4, 2, 0, 0, 0)$, $P = 6$, $\bar{P} = 1$, $\beta_{cp.} = 3$, $\alpha_{cp.} = 1.6$, $\gamma_{cp.} = 2$.

Thus, using the integral characteristics of the content model, it is possible to analyze and compare various educational materials with each other, to assess the complexity of preparing EER already at the stage of their design.

5 Content model design automation

The algorithms discussed above make it possible to automate the process of preparing a content model [19]. The EER developer creates a set of LE in dialogue with the computer and establishes hierarchical relationships between them, filling in the values of the target indicators in the specification of the LE attributes. The computer program controls the structure of the CG, according to the rules of its construction, visualizes the CG, forms matrices of adjacency, reachability and distances, calculates the integral characteristics of the content model, forms the table of contents of the educational material for its export to the EER layout tool program (Figures 4, 5)¹.

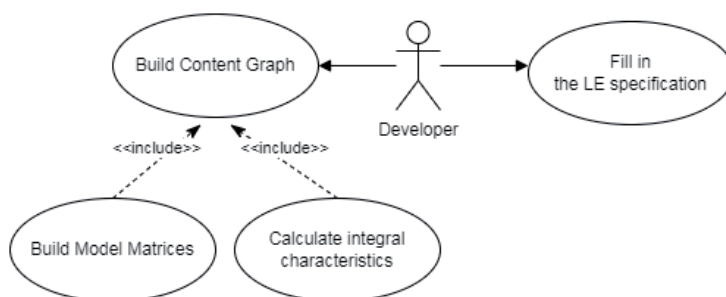


Figure 4 – Variant UML-diagram for using the computer program for the formation of the content model

¹ Here, when describing computer program scripts, unified modeling language (UML) diagrams are used <https://www.uml.org/>.

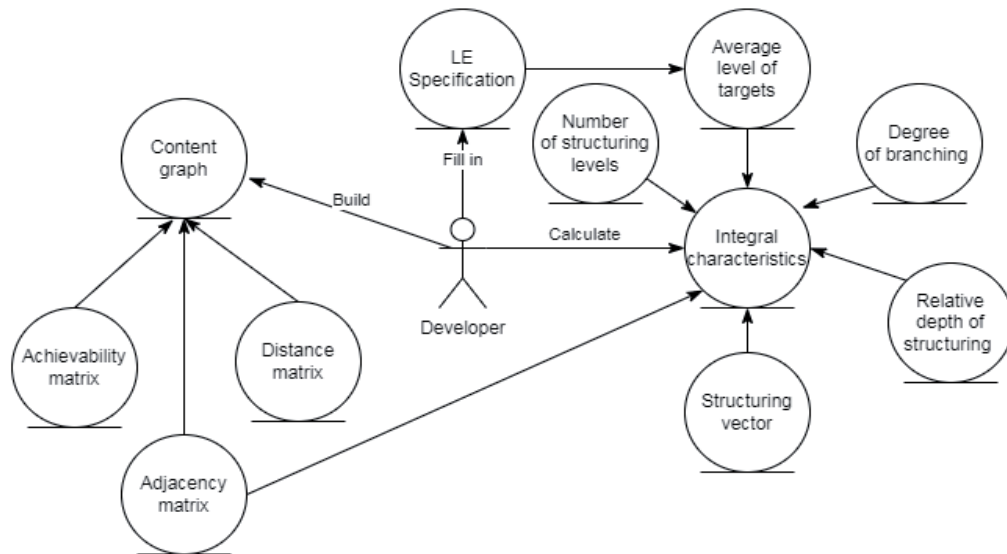


Figure 5 – UML interaction diagram between users and objects of the computer program for the formation of the content model

6 Example of structuring education material

Consider the content model of one of the modules of our course for graduate students on methods and technologies of e-learning [20]. The topic of the module: "Electronic information and educational environment of an educational institution (EIEE EI)". The purpose of studying the module is to get acquainted with the typical functionality of EIEE EI. The education material of the module is based on the article [21] with some additions from other sources (Figure 6). The structure of the module is based on the presentation of EIEE EI as an organizational and technical system [21, p.147].

The content graph of the educational material has two levels of structuring: the first level is the subsystems of the EIEE EI, the second is the components of these subsystems (see Figure 6, a). Didactic attributes of LE are determined based on the contingent of students. These are graduate students who have experience with some components of EIEE EI, but do not have a complete system understanding of such systems.

For each LE, the didactic parameters are chosen to be the same (see Figure 6, b). The level of presentation of educational material is adopted by the analytical-synthetic $\beta = 2$ [22, p. 59] in accordance with its basic source [21]. The required level of assimilation is minimal – "Acquaintance" $\alpha = 1$ [22, p.60], taking into account the user nature of the potential interaction of students with the services of EIEE EI. But the level of awareness is maximum $\gamma = 3$, since students study the basic concepts of EIEE EI in this course, based on the experience of using the services of the system in different academic disciplines [22, p.62].

Integral characteristics of this content model: the number of LE $n = 18$, the number of levels of structuring of educational material $U = 2$, the maximum possible depth of structuring $U_{max} = 8$, the relative depth of structuring $\bar{U} = 0.25$, the structuring vector $S = (17, 5, 0, 3, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$, the degree of branching of the content model $P = 30$, the relative degree of branching $\bar{P} = 0.94$, the average level of didactic indicators $\beta_{cp.} = 2$, $\alpha_{cp.} = 1$, $\gamma_{cp.} = 3$.

Based on the requirements for the level of assimilation for each LE, the corresponding block of EER modulo, in addition to the information description, contains 3-5 exercises for comprehension and consolidation of the educational material. In total, about 65 such exercises of the first level of mastering the $\alpha = 1$ have been developed modulo, taking into account the level of presentation of

the educational material $\beta = 2$ and the level of awareness $\gamma = 3$. The same set of test tasks is used to sample tests for the final control of knowledge by module.

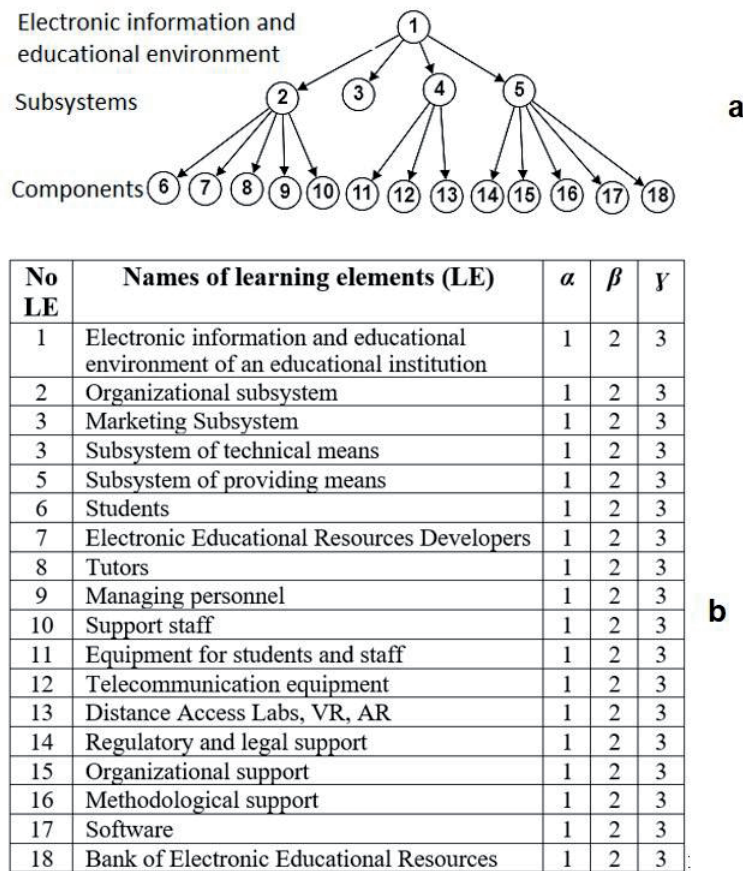


Figure 6 – Content model of the Electronic information and educational environment of an educational institution topic: a – content graph; b – specification of learning elements attributes

7 Discussion of the results

Some elements of the above process of modeling the structure of EER were proposed by us earlier in [16, 23]. Models of this kind are useful for rational structuring of the content of the educational resource in the form of a set of hierarchically organized LEs. For many years, the authors have been using the concept and methods of building a content model when designing EER in various academic disciplines [22]. A number of colleagues in other educational institutions apply our developments in the design of their own EER [24].

The accumulated experience allows us to recommend starting to apply the proposed models with "manual" design using a pencil and paper. And only then proceed to automate this process. The use of a computer allows you to work with detailed content models consisting of several dozen LE, see the example in [22, p.75], which is practically unrealistic when preparing models manually.

The presentation of the structure of the electronic educational resource in the form of the content model considered in this article allows:

- to allocate the necessary material from the studied academic discipline, to divide it into separate educational elements, to present it in the form of a visual and observable scheme, to clearly define the didactic requirements for its presentation and study;
- to involve experts and customers of EER to discuss the completeness of the content and targets for its presentation and study already at the initial stage of EER design;

- to form a systematic (holistic) representation of the content of the EER, both among developers and users of EER (teachers and students);
- evaluate and compare various training materials in terms of volume, degree of structure, branching, give a forecast on the labor intensity, number and type of required exercises for training and control;
- develop EER in accordance with international SCORM specifications.

It is also important to emphasize that the process of building a content model allows even experienced teachers to take a fresh look at their educational material in terms of structure, form of presentation and requirements for its assimilation.

It is very useful in a guide to the study of any EER to give a model of content with structure and didactic requirements. This allows students to form a holistic visual representation of the structure of the educational material, motivate and orient them in terms of the thoroughness of its study.

Conclusion

The rules for constructing models of content of electronic educational content have been formulated. Models of this kind are useful for rational structuring of the content of the educational resource in the form of a set of hierarchically organized fragments of educational material. The mathematical properties of these models are discussed, their integral characteristics are introduced and strictly substantiated. The proposed approach to content modeling is well consistent with scorm's international e-learning specifications, complementing them with didactic targets, didactic design algorithms and analysis of educational materials. Algorithms for the formation and methods of representation of the content model allow to automate the process of its construction and didactic analysis in the form of a visual interactive dialogue of developers of electronic educational resources in instrumental author's environments.

References

- [1] Ontology Summit 2020: Knowledge Graphs [In Russian]. *Ontology of Designing*. 2020; 10(2): 246-248.
- [2] **Gavrilova TA, Strakhovich EV**. Visual analytical thinking and mind maps for ontology engineering [In Russian]. *Ontology of designing*. 2020; 10(1): 87-99. DOI: 10.18287/2223-9537-2020-10-1-87-99.
- [3] **Mouromtsev D**. Models and methods of e-learning individualization in the context of ontological approach [In Russian]. *Ontology of designing*. 2020; 10(1): 34-49. DOI: 10.18287/2223-9537-2020-10-1-34-49.
- [4] **Solovov AV, Menshikova AA**. Discrete mathematical models in the study of automated learning processes. [In Russian]. *Information technologies*. 2001; 12: 43-48.
- [5] **Silkina N.S., Sokolinsky L.B.** Structural-Hierarchical Didactic Model of E-Learning. [In Russian]. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*. 2019; 8(4): 56–83. DOI: 10.14529/cmse190405.
- [6] **Osina AV**. Multimedia in Education: The Context of Informatization [In Russian]. Moscow: Agency «Izdatel'skii service». 2004. 420 p.
- [7] **AICC** – Aviation Industry CBT Committee. https://en.wikipedia.org/wiki/Aviation_Industry_Computer-Based_Training_Committee.
- [8] **IMS** – Instructional Management System. <https://www.1edtech.org/>.
- [9] **ADL** – Advanced Distributed Learning. <https://adlnet.gov/>.
- [10] **Silkina NS, Sokolinsky LB**. Models and standards of e-learning [In Russian]. *Bulletin of South Ural State University. Series: Computational Mathematics and Informatics*. 2014; 3(4): 5-35. DOI 10.14529/cmse140401.
- [11] **Zimin IV**. Recommendations for the preparation of content and structuring of an electronic training resource. [In Russian]. National Open University "INTUIT". 2016. 102 p.
- [12] **SCORM** – Sharable Content Object Reference Model. SCORM solved and explained. <https://scorm.com/>.
- [13] **1EDTECH**. <https://www.1edtech.org/>.
- [14] **Bespalko VP**. Fundamentals of the theory of pedagogical systems [In Russian]. Voronezh: Voronezh University Press. 1977. 303 p.

- [15] **Belkin EL.** Didactic foundations of control of cognitive activity in the conditions of application of technical means of teaching. [In Russian]. Yaroslavl: Verkhne-Volzhskoe knizhnoe izdatelstvo. 1982. 107 p.
- [16] **Solovov AV.** Design of computer systems for educational purpose: a textbook. [In Russian]. Samara: SSAU. 1995. 140 p.
- [17] **Anderson LW, Krathwohl DR.** A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman. 2001. 352 p.
- [18] **Roberts FS.** Discrete mathematical models with applications to social, biological and environmental problems. Englewood Cliffs, NJ: Prentice-Hall, Inc.; 1976. 560 p. (Russ. ed.: Roberts F.S. Diskretnye matematicheskie modeli s prilozheniyami k sotsial'nym, biologicheskim i ekologicheskim zadacham. Moscow: Nauka; 1986. 494 p.).
- [19] **Menshikova AA, Solovov AV.** Author's toolkit of the CADIS system [In Russian]. Unified educational information environment: electronic periodical. 2003. № 1.
- [20] **Solovov AV.** «Gold Cells» of Virtual Learning Environments. [In Russian]. *Vysshee obrazovanie v Rossii = Higher Education in Russia*. 2012; 11: 133-137.
- [21] **Solovov AV, Menshikova AA.** Models for the Design and Operation of Digital Educational Environments. [In Russian]. *Vysshee obrazovanie v Rossii = Higher Education in Russia*. 2021; 30(1): 144-155. DOI: 10.31992/0869-3617-2021-30-1-144-155.
- [22] **Solovov AV.** E-learning: problems, didactics, technology. [In Russian]. Samara: New Engineering, 2006. 464 p.
- [23] **Solovov AV.** Modeling the structure of electronic educational resources. [In Russian]. *Information technologies*. 2007; 3: 43-48.
- [24] **Kukharensko SP, Dzyubenko OL, Solovov AV, Menshikova AA.** Formation of an electronic information and educational environment for engineering analysis in a military university. Part 2. 2nd edition, corrected and enlarged. [In Russian]. Moscow: Limited Liability Company "Rusins". 2022. 170 p.

About the authors

Alexander Vasilevich Solovov (b. 1948) graduated from the Kuibyshev Aviation Institute named after S.P. Korolev (Kuibyshev, USSR) in 1972, PhD (1977). Professor at the Department of Technical Cybernetics (2006). Full member of the Russian Academy of Informatization of Education (1996). The list of scientific works includes more than 300 works in the field of CAD, theory and technologies of e-learning. ORSID: 0000-0001-6288-820X; Author ID (RSCI) : 560817; Author ID (Scopus): 57222040521. a_solovov@mail.ru. ✉.

Anastasia Alexandrovna Menshikova (b. 1972) graduated from the Samara State Aerospace University named after S.P. Korolev in 1996, Ph.D. (2004). Associate Professor of the Department of Supercomputers and General Informatics of Samara University. The list of scientific works includes more than 40 works. ORSID: 0000-0001-8201-7065; Author ID (RSCI): 382400; Author ID (Scopus): 57222036809; Researcher ID (WoS): H-6847-2017. nastya.menshikova@gmail.com.

Received February 20, 2023. Revised March 1, 2023. Accepted March 3, 2023.

УДК 519.673

Научная статья

DOI: 10.18287/2223-9537-2023-13-1-99-112



Проектирование онтологии содержания электронного учебного курса

© 2023, А.В. Соловов✉, А.А. Меньшикова

Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия

Аннотация

Моделирование знаний, тесно связанное с онтологиями, является важной семантической технологией и областью исследований. В статье рассматривается понятие модели содержания электронного учебного курса. В ос-

нову модели содержания положено структурирование содержание курса на отдельные фрагменты, называемые учебными элементами. Эти учебные элементы интегрируются в древовидный ориентированный граф. Модель содержания определена как совокупность такого графа и таблицы атрибутов учебных элементов с требованиями к дидактическим показателям их изучения. Формулируются правила построения моделей содержания электронного учебного курса. Обсуждаются математические свойства этих моделей и вводятся их интегральные характеристики. Предлагаемый подход к моделированию содержания хорошо согласуется с международными спецификациями электронного обучения *SCORM*, дополняет их целевыми показателями, алгоритмами дидактического проектирования и анализа учебных материалов. Алгоритмы формирования и способы представления модели содержания позволяют автоматизировать процесс её построения и дидактического анализа в форме визуального интерактивного диалога разработчиков электронных образовательных ресурсов в инструментальных авторских средах.

Ключевые слова: электронное обучение, электронные образовательные ресурсы, структуризация учебного материала, модель содержания, древовидные ориентированные графы, *SCORM*.

Цитирование: Соловов А.В., Меньшикова А.А. Проектирование онтологии содержания электронного учебного курса // Онтология проектирования. 2023. Т.13, №1(47). С.99-112. DOI:10.18287/2223-9537-2023-13-1-99-112.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Рисунки

Рисунок 1 – Пример модели содержания: а – граф содержания (ГС); б – спецификация атрибутов учебного элемента (УЭ); в – матрица смежности ГС; г – степени матрицы смежности ГС; д – матрица расстояний ГС; е – матрица достижимости ГС

Рисунок 2 – К свойству 1 графа содержания

Рисунок 3 – К теоремам 2 и 3

Рисунок 4 – *UML*-диаграмма вариантов использования компьютерной программы формирования модели содержания

Рисунок 5 – *UML*-диаграмма взаимодействия пользователей и объектов компьютерной программы формирования модели содержания

Рисунок 6 – Модель содержания темы «Электронная информационно-образовательная среда образовательного учреждения»: а – граф содержания; б – спецификация атрибутов учебных элементов

Список источников

- [1] Онтологический Саммит 2020: Графы знаний. *Онтология проектирования*. 2020. Т.10, №2(36). С.246-248.
- [2] Гаврилова Т.А., Страхович Э.В. Визуально-аналитическое мышление и интеллект-карты в онтологическом инжиниринге. *Онтология проектирования*. 2020. Т.10, №1(35). С.87-99. DOI 10.18287/2223-9537-2020-10-1-87-99.
- [3] Муромцев Д.И. Модели и методы индивидуализации электронного обучения в контексте онтологического подхода. *Онтология проектирования*. 2020. Т.10, №1(35). С.34-49. DOI 10.18287/2223-9537-2020-10-1-34-49.
- [4] Соловов А.В., Меньшикова А.А. Дискретные математические модели в исследовании процессов автоматизированного обучения. *Информационные технологии*. 2001. №12. С.43-48.
- [5] Силкина Н.С., Соколинский Л.Б. Структурно-иерархическая дидактическая модель электронного обучения. *Вестник ЮУрГУ. Серия: Вычислительная математика и информатика*. 2019. Т.8, №4. С.56–83. DOI: 10.14529/cmse190405.
- [6] Осин А.В. Мультимедиа в образовании: контекст информатизации. М.: Агентство «Издательский сервис». 2004. 420 с.
- [7] *AICC* – Aviation Industry CBT Committee (Международный комитет по компьютерному обучению в авиации). https://en.wikipedia.org/wiki/Aviation_Industry_Computer-Based_Training_Committee.
- [8] *IMS* – Instructional Management System (Системы организации обучения). <https://www.ledtech.org/>.
- [9] *ADL* – Advanced Distributed Learning (Расширенное распределённое обучение). <https://adlnet.gov/>.
- [10] Силкина Н.С., Соколинский Л.Б. Модели и стандарты электронного обучения. *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*. 2014. Т.3, №4. С.5-35. DOI 10.14529/cmse140401.

- [11] **Зимин И.В.** Рекомендации по подготовке содержания и структурирования электронного обучающего ресурса. Национальный Открытый Университет "ИНТУИТ". 2016. 102 с.
 - [12] **SCORM** – Sharable Content Object Reference Model (Модель обмена учебными материалами). SCORM solved and explained. <https://scorm.com/>.
 - [13] **IEDTECH**. <https://www.ledtech.org/>.
 - [14] **Беспалько В.П.** Основы теории педагогических систем. Воронеж: Изд-во Воронеж. ун-та. 1977. 303 с.
 - [15] **Белкин Е.Л.** Дидактические основы управления познавательной деятельностью в условиях применения технических средств обучения. Ярославль: Верхне-Волжское книжное издательство. 1982. 107 с.
 - [16] **Соловов А.В.** Проектирование компьютерных систем учебного назначения: учебное пособие. Самара: СГАУ. 1995. 140 с.
 - [17] **Anderson L.W., Krathwohl D.R.** A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman. 2001. 352 p.
 - [18] **Робертс Ф.С.** Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. Пер. с англ. М.: Наука. 1986. 494 с.
 - [19] **Меньшикова А.А., Соловов А.В.** Авторский инструментальный системы КАДИС. *Единая образовательная информационная среда: электронное периодическое издание*. 2003. № 1.
 - [20] **Соловов А.В.** "Золотые клетки" виртуальных учебных сред. *Высшее образование в России*. 2012. № 11. С.133-137.
 - [21] **Соловов А.В., Меньшикова А.А.** Модели проектирования и функционирования цифровых образовательных сред. *Высшее образование в России*. 2021. Т. 30. № 1. С.144-155. DOI: 10.31992/0869-3617-2021-30-1-144-155.
 - [22] **Соловов А.В.** Электронное обучение: проблематика, дидактика, технология. Самара: Новая техника. 2006. 464 с.
 - [23] **Соловов А.В.** Моделирование структуры электронных образовательных ресурсов. *Информационные технологии*. 2007. № 3. С.43-48.
 - [24] **Кухаренко С.П., Дзюбенко О.Л., Соловов А.В., Меньшикова А.А.** Формирование электронной информационно-образовательной среды для инженерного анализа в военном вузе. Часть 2. 2-е издание, исправленное и дополненное. М.: ООО "Русайнс". 2022. 170 с.
-

Сведения об авторах



Соловов Александр Васильевич, 1948 г. рождения. Окончил Куйбышевский авиационный институт имени С.П. Королёва в 1972 г., к.т.н. (1977). Профессор по кафедре технической кибернетики (2006). Действительный член Российской академии информатизации образования (1996). В списке научных трудов более 300 работ в области САПР, теории и технологий электронного обучения. ORSID: 0000-0001-6288-820X; Author ID (РИНЦ): 560817; Author ID (Scopus): 57222040521. a_solovov@mail.ru. ✉.



Меньшикова Анастасия Александровна, 1972 г. рождения. Окончила Самарский государственный аэрокосмический университет имени академика С.П. Королёва в 1996 г., к.т.н. (2004). Доцент кафедры суперкомпьютеров и общей информатики Самарского университета. В списке научных трудов более 40 работ. ORSID: 0000-0001-8201-7065; Author ID (РИНЦ): 382400; Author ID (Scopus): 57222036809; Researcher ID (WoS): H-6847-2017. nastya.menshikova@gmail.com.

Поступила в редакцию 20.02.2023, после рецензирования 1.03.2023. Принята к публикации 3.03.2023

ИНЖИНИРИНГ ОНТОЛОГИЙ

УДК 004.853:004.82

Научная статья

DOI: 10.18287/2223-9537-2023-13-1-113-124



Доверие к данным при пополнении онтологий и графов знаний

© 2023, А.С. Серый

Институт систем информатики им. А.П. Еришова СО РАН, Новосибирск, Россия

Аннотация

Рассматривается задача оценки доверия к информации, извлекаемой из текстовых источников для пополнения онтологий или графов знаний. За единицу информации или факт, принимается минимальное знание об экземпляре предметной области, выражаемое единичным *RDF*-триплетом. Приведено описание вероятностной модели оценки доверия, основанной на марковских случайных процессах. При оценке модель строится на основании доступной информации об источниках с учётом ранее извлечённых данных. Предложен метод оценки доверия к информации с параллельным взвешиванием источников. Подобный подход востребован в ситуациях, когда качественные характеристики источников неизвестны или недоступны. В рамках тестирования модели были автоматически сгенерированы наборы численных данных различных объёмов, проведены эксперименты по взвешиванию источников и оценке доверия к извлекаемой из них информации. Результаты экспериментов показали, что в большинстве случаев веса источников, вычисляемые на основе предлагаемой модели, тем больше, чем меньше среднее отклонение предоставленной ими информации от истинной, доверие к фактам увеличивается с уменьшением расстояния до истинных данных. Выполнено сравнение с моделями агрегации данных. В большинстве случаев агрегация, выполненная на основе оценки доверия, демонстрировала наименьшее среднее отклонение от истинных данных среди рассмотренных моделей. Полученные результаты показывают, что предлагаемая модель эффективна в сравнении с другими аналогичными моделями и может применяться в задачах оценки доверия к фактам, представляемым вещественными числами.

Ключевые слова: онтология, граф знаний, извлечение данных, доверие к информации, марковский процесс.

Цитирование: Серый А.С. Доверие к данным при пополнении онтологий и графов знаний // Онтология проектирования. 2023. Т.13, №1(47). С.113-124. DOI:10.18287/2223-9537-2023-13-1-113-124.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

Современное глобальное информационное пространство невозможно представить и проанализировать усилиями человека, даже если речь идёт об экспертах в конкретных предметных областях (Про). Одним из путей решения данной проблемы стало использование методов автоматического анализа данных, которые широко применяются во всех сферах человеческой деятельности, связанной с обработкой информации. В первую очередь это методы обработки неструктурированных источников, например текстов, изображений и веб-страниц, позволяющие извлекать определённую информацию и представлять её в структурированном виде — в базах данных, онтологиях и графах знаний. Последние являются основным, на данный момент, способом интеграции больших структурированных данных [1]. Для извлечения информации применяется множество различных методов и подходов: от конвейерных

процессов на основе технологии *Apache NiFi* [2] до трансформерных нейросетей типа *BERT* [3]. Полученные графы затем применяются в интеллектуальных информационных системах (ИС) как источники знаний и основа логического вывода [4].

Обработка большого числа источников почти неизбежно приводит к появлению противоречивых знаний, т.е. нескольких альтернативных утверждений относительно одной и той же сущности. Это могут быть, к примеру, разные прогнозы погоды, цены акций, ожидаемое время прибытия авиарейсов, противоречивая информация о местах жительства или работы людей и т.д. Противоречия возникают как вследствие ошибок, так и потому, что информация, предоставленная источником, давно не обновлялась и устарела¹. Таким образом, требуется не просто извлечь знания из источника, но и оценить их надёжность или уровень доверия к ним. В данной работе предлагается метод оценки доверия к информации, извлекаемой из различных источников для пополнения базы знаний (БЗ) ИС, основанной на онтологии.

1 Обзор предшествующих работ

Проблема оценки надёжности знаний, особенно в тех случаях, когда знания, полученные из разных источников, противоречат друг другу, исследуется давно. В работах [5, 6] надёжность источников и извлекаемых данных оценивались по заранее заданным правилам. Информация зачастую предполагалась статичной, т.е. представленной в виде завершённой таблицы соответствия фактов и источников [7, 8], которая затем не изменяется. В работах [9, 10] рассматриваются ситуации, когда информация из источников поступает последовательно, а истинные знания изменяются со временем. Исследования проводились на численных данных, в качестве примеров были выбраны прогнозы среднесуточной температуры, прогнозы капитализаций на фондовых рынках и время прибытия авиарейсов. Проведённый анализ результатов показал эффективность предлагаемых решений.

Более сложной задачей является оценка текстовых данных. Предметом исследования в данном направлении являются социальные сети: с одной стороны как источник большого количества противоречивой информации, с другой — как средство, оказывающее значительное влияние на образ мыслей и мнение людей. В работах [11, 12] исследовались способы и пути распространения слухов внутри социальных сетей. В [11] рассмотрены механизмы распространения слухов, проанализированы их жизненные циклы и зависимость таких показателей, как уровень поддержки и обсуждаемость, от типов пользователей, вовлекаемых в их распространение. Исследование [12] сосредоточено на верификации слухов. В работе [13] собран набор данных и разработана мультимодальная модель машинного обучения для решения задачи обнаружения и верификации слухов, касающихся девяти различных событий. Каждый элемент набора данных был аннотирован одной из трёх меток в зависимости от степени надёжности: **Правда** (*True*), **Неправда** (*False*) и **Не подтверждено** (*Unverified*). Схожая задача оценки высказываний пользователей решалась в [14] при помощи серии известных методов машинного обучения: наивный байесовский классификатор, логистическая регрессия, метод опорных векторов, деревья решений и др.

Сложность задач исследования в области анализа надёжности информации возрастает с ростом объёмов доступной информации и стремительным распространением ложной информации. Методы глубокого обучения применяются как современное и мощное средство. Разработанный в [12, 13] набор данных применялся для анализа новостей в социальной сети *Twitter* [15]. В [16] использована нейросетевая модель на основе свёрточных и рекуррентных нейронных сетей для распознавания ложной информации.

¹ В современном информационном пространстве особое значение приобретает проблема выявления заведомо ложной или умышленно искажённой информации. *Прим. ред.*

2 Модель доверия

2.1 Факты в онтологии

Пусть БЗ ИС построена на основе онтологии $\mathcal{O}$ некоторой ПрО, где $\mathcal{O} = \{C_O, D_O, Dat_O, Rel_O\}$. Конечное непустое множество C_O представляет совокупность концептов ПрО, конечные непустые множества D_O , Dat_O и Rel_O — соответственно доменов, атрибутов и отношений. Каждый атрибут из Dat_O имеет область значений $d \in D_O$, а элементы множества $Rel_O \subseteq C_O \times C_O$ — это бинарные отношения между концептами из C_O . Объединение $Dat_O \cup Rel_O$ атрибутов и отношений называется множеством свойств онтологии $\mathcal{O}$. Класс можно определить в виде тройки (c, Dat_c, Rel_c) , где через c обозначено имя класса, $Dat_c \subseteq Dat_O$, $Rel_c \subseteq Rel_O$ — его свойства. Каждый атрибут $\alpha^c \in Dat_c$ имеет область значений $d_{\alpha^c} \in D_O$, а каждое отношение $\rho^c \in Rel_c$ связывает класс c некоторым классом $c_{\rho^c} \in C_O$. Класс или множество классов c_{ρ^c} образуют область значений отношения ρ^c .

Пусть $a \in c_a$, если a является экземпляром класса $c_a \in C_O$. Экземпляр представляется тройкой вида $a = (c_a, Dat_a, Rel_a)$ такой, что $Dat_a = \{(\alpha, V_{\alpha_a}) | \alpha \in Dat_{c_a}, V_{\alpha_a} \subseteq d_{\alpha} \in D_O\}$ — атрибуты экземпляра a , и $Rel_a = \{(\rho, V_{\rho_a}) | \rho \in Rel_{c_a}\}$ — его связи с другими экземплярами. Здесь V_{ρ_a} — множество экземпляров, с которыми a связан отношением ρ .

В данной работе задача пополнения БЗ ИС рассматривается как задача пополнения онтологии, т.е. как добавление, удаление и изменение экземпляров в соответствии с данными, полученными извне. При этом за область рассмотрения остаётся редактирование ядра онтологии — множества $\mathcal{O}$. В терминах ИС, БЗ которой построена на основе онтологии, единицей информации считается минимальное знание об экземпляре ПрО — значение его атрибута или его связь с другим экземпляром. Можно называть такое знание единичным фактом или просто фактом. Автоматическая обработка текстовых источников позволяет извлекать факты и добавлять их в БЗ ИС. Информация, полученная из разных источников, может оказаться противоречивой, порождая множества конфликтных фактов. Требуется ранжировать эти множества по уровню доверия таким образом, чтобы предоставить пользователям ИС наиболее надёжную информацию.

При оценке доверия к фактам предлагаемая в данной работе модель основывается на доступной информации об источниках, из которых факты были получены. Под источниками здесь понимаются общедоступные электронные ресурсы, из которых извлекаются численные или текстовые данные. Предполагается, что в ИС используются некоторые качественные показатели источников, например рейтинг, если таковые доступны, или создаются собственные оценки, опираясь на всю информацию, доступную на текущий момент. Это означает, что модель доверия учитывает характеристики источников данных, но не включает описание методов их получения. Необходимо только, чтобы данные характеристики принимали значения из множества $\mathbb{R}^+$. В любой момент времени множество источников, из которых извлекаются факты для пополнения БЗ ИС, конечно. Пусть это будет множество S , а для любого источника $s \in S$ искомая качественная характеристика — μ^s . В дальнейшем индекс s в обозначении μ^s будет опускаться в тех случаях, когда не имеет значения, из какого конкретного источника была получена информация.

2.2 Доверие как случайный процесс

В каждый момент времени для факта F , являющегося частью экземпляра a , должна быть определена величина, показывающая насколько надёжным является данный факт по сравнению с другими фактами в ИС, Tr^F - трастовая метрика. БЗ не является статичной, и с появле-

нием новых источников, содержащих другие факты об a , доверие к F может изменяться. История изменений Tr^F представляет собой последовательность, в которой каждый следующий член зависит только от предыдущего, а также от поступившего в обработку источника $s \in S$. ИС может считать или не считать факт F достоверным, т.е. F имеет два потенциальных состояния: **Ненадёжный** (*Unreliable, U*) и **Надёжный** (*Reliable, R*). Информация о том, в каком из состояний на данный момент находится F , эквивалентна его Tr^F . Последовательность значений Tr^F становится эквивалентной последовательности вида $(X_t, t = 0, 1, \dots)$ дискретных случайных величин, принимающих значения из бинарного множества состояний $\{U, R\}$. Тогда для последовательности (X_t) выполняется условие $P(X_k = x_k | X_{k_1} = x_{k_1}, X_{k_2} = x_{k_2}, \dots, X_{k_r} = x_{k_r}) = P(X_k = x_k | X_{k_r} = x_{k_r})$, для любых $k_1 < k_2 < k_3 < \dots < k_r < k$, т.е. она удовлетворяет определению марковских случайных процессов.

Следующий член случайного процесса вычисляется всякий раз, когда в систему поступает новая информация об F . Пусть T - множество моментов времени, соответствующих членам случайного процесса X_t , т.е. $t \in T$. Значение Tr^F в момент t оценивается как вероятность того, что факт F является надёжным, т.е. $Tr^F = P(X_t^F = R)$.

Для любого t величина X_t распределена как (π_U^t, π_R^t) , где $\pi_x^t = P(X_t = x)$. Вектор распределения $\bar{\pi} = (\pi_U, \pi_R)$ можно назвать вектором распределения доверия (*Trust Distribution Vector, TDV*). TDV показывает вероятность факта оказаться достоверным или недостоверным. Очевидно, что $\pi_U + \pi_R = 1$. Распределение $\bar{\pi} = (\frac{1}{2}, \frac{1}{2})$ и близкие к нему соответствуют состоянию неопределённости, когда судить о достоверности F невозможно.

Согласно теории случайных процессов, вектор $\bar{\pi}^{t+1} = (\pi_U^{t+1}, \pi_R^{t+1})$ получается умножением вектора предыдущего шага $\bar{\pi}^t$ на 2×2 стохастическую матрицу перехода $P(t, t+1)$. Элементы $p_{ij}(t, t+1)$ матрицы $P(t, t+1)$ — это вероятности перехода из i -го состояния в j -е на шаге $(t+1)$, при этом $p_{i1} + p_{i2} = 1, i = 1, 2$. Здесь и далее предполагается, что состояния 1 и 2 — это состояния U и R соответственно. В таких обозначениях $\pi_1 = \pi_U, \pi_2 = \pi_R$.

2.3 Переходная матрица случайного процесса

В рамках модели матрица перехода $P(t, t+1)$ представляется как функция от μ и вектора $\bar{\pi}^t$.

$$P(t, t+1) = \text{softmax} \left(\pi^T \begin{pmatrix} \frac{\pi_1}{\mu} & \mu\pi_2 \end{pmatrix} \right) = \begin{pmatrix} \text{softmax} \left(\frac{\pi_1^2}{\mu}, \mu\pi_1\pi_2 \right) \\ \text{softmax} \left(\frac{\pi_1\pi_2}{\mu}, \mu\pi_2^2 \right) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{1 + e^{\pi_1(\mu\pi_2 - \pi_1/\mu)}} & \frac{1}{1 + e^{\pi_1(\pi_1/\mu - \mu\pi_2)}} \\ \frac{1}{1 + e^{\pi_2(\mu\pi_2 - \pi_1/\mu)}} & \frac{1}{1 + e^{\pi_2(\pi_1/\mu - \mu\pi_2)}} \end{pmatrix}. \quad (1)$$

Из формулы (1) следует, что увеличение рейтинга μ источника информации ведёт к росту вероятности перехода в состояние R и наоборот — информация из источников с низким рейтингом способствует переходу в состояние U . Функция $\text{softmax}(\bar{x}) = \left(\frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \right)_i, \bar{x} = (x_1, x_2, \dots, x_n)$ необходима для представления категориального распределения по строкам матрицы P .

Матрица P обладает несколькими полезными свойствами как функция $P(\mu, \bar{\pi}^t)$. На основании формулы (1) можно заключить, что с уменьшением рейтинга источника доверие к получаемой из него информации также уменьшается вплоть до нуля. Заведомо ложный источник с $\mu = 0$ приводит любой TDV , кроме $(0, 1)$, к вектору $(1, 0)$ за один шаг. Вектор $(0, 1)$,

т.е. такой, где $\pi_1 = 0$, приводит к $P = \begin{pmatrix} 1/2 & 1/2 \\ \frac{1}{1+e^\mu} & \frac{1}{1+e^{-\mu}} \end{pmatrix} \xrightarrow{\mu \rightarrow 0} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$. Таким образом, заведомо ложный источник переводит идеальный $TDV(0, 1)$ в $(1/2, 1/2)$.

Аналогично, в случае заведомо надёжного источника, при $\mu \rightarrow \infty$, матрица P переводит любой TDV , кроме $(1, 0)$, в идеальный вектор $(0, 1)$. В случае $\pi_2 = 0$ $P \xrightarrow{\mu \rightarrow \infty} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$. Это означает, что заведомо надёжный источник переводит распределение $(1, 0)$ в $(1/2, 1/2)$.

Третьим полезным следствием формулы (1) является тот факт, что для любого распределения $\bar{\pi}$ $(\bar{\pi}P)_2(\mu_1) < (\bar{\pi}P)_2(\mu_2)$ при $\mu_1 < \mu_2$, т.е. доверие Tr^F монотонно как функция от μ .

3 Экспериментальные исследования

3.1 Параметры и обозначения

ИС, реализующая модель доверия для оценки поступающих данных, полагается на качественные показатели источников, из которых они получены. Модель не содержит описания методов оценки источников. При доступности достаточного количества альтернативных фактов существуют методы, позволяющие параллельно оцениванию доверия к информации, «взвесить» её источники. В данной работе предлагается метод параллельного оценивания, основанный на [9, 10]. В этих работах описанные модели использовались для численных данных, таких как прогнозы среднесуточной температуры, количество пешеходов на улице, капитализация компаний на фондовых рынках и т.п. Основная решаемая с использованием моделей задача состоит в вычислении по имеющейся альтернативной информации из источников единственного агрегированного значения, наиболее близкого к истинному. Предлагаемая в данной работе модель предназначена для ранжирования всех полученных альтернативных значений по степени доверия к ним.

Входными параметрами оригинальных моделей [9, 10] являются множество моментов времени T , источников S и множество O объектов, информация о которых извлекается из источников. Источник $s \in S$ имеет вес $w_s \in \mathbb{R}^+$, и в момент времени t из него извлекается информация об $c_t^s \geq 0$ объектах. Решение задачи заключается в минимизации потерь вида (2).

$$L_t = \theta \sum_{s=1}^S w_s \sum_{o=1}^{c_t^s} (v_{o,t}^s - v_{o,t}^*)^2 - \sum_{s=1}^S c_t^s \log(w_s). \quad (2)$$

Здесь $v_{o,t}^s$ и $v_{o,t}^*$ — это, соответственно, значение, полученное для объекта o из источника s в момент времени t и агрегированное значение, вычисленное для o в момент t . Для минимизации потерь необходимо, с одной стороны, уменьшить вес источников, дающих информацию, сильно отличающуюся от агрегированного значения, с другой — избегать приближения весов к нулю. Значение параметра θ задаётся заранее.

В данном случае объектами выступают факты, т.е. единицы информации об экземплярах концептов онтологии ПрО. Количество источников заранее неизвестно, однако в любой момент времени оно конечно, поэтому, без ограничения общности, можно считать множество источников аналогичным таковому в оригинальной модели в ситуации, когда источник s может не содержать информации о конкретном объекте в момент времени t . Это означает, что в момент t учитываются только те источники, из которых удалось извлечь требуемую информацию. Множеством, аналогичным множеству объектов O , будет $\mathcal{F} = \{f | f = (a, \alpha), a \in c_a \in C_O, \alpha \in Dat_{c_a} \cup Rel_{c_a}\}$ — множество пар экземпляр/свойство, которое можно

назвать множеством типов извлекаемых фактов. В этом случае $v_{f,t}^s$ дополняет пару f до законченного триплета, т.е. факта. Время T при этом соответствует определению, данному в разделе 2.2. Полный список обозначений приведён в таблице 1.

Таблица 1 – Принятые обозначения

Обозначение	Определение
S	Количество источников, из которых извлекается информация. $S < +\infty$.
w_s	Вес источника s .
$\mathcal{F}$	Множество типов извлекаемых фактов.
T	Множество моментов времени.
$v_{f,t}^s$	Факт типа $f \in \mathcal{F}$, полученный из источника s в момент t .
$v_{f,t}^*$	Агрегированное значение f в момент t .
c_t^s	Количество фактов, полученных из источника s в момент t .
D	База данных ИС.
$D_{f,t}$	Количество альтернативных значений f в базе данных в момент t .
v_f^j	j -е альтернативное значение f в D .
$e_{f,t}^s$	Ошибка источника s в момент t на факте типа f .

В работе [9] рассмотрены ситуации, когда истинное знание постоянно меняется, то есть истинное значение факта зависит от t . Это соответствует задачам оценки прогнозов погоды или количества товаров на складе. Существуют задачи, где истинное знание не изменяется длительное время, но могут появляться источники, распространяющие неточную информацию и порождающие новые альтернативные значения. В качестве примера можно привести количество сотрудников в организации, данные статистических исследований, место работы или жительства персоны и т.п. В подобных ситуациях альтернативных значений сравнительно немного, они распространяются разными источниками и поэтому могут быть извлечены ИС многократно. Для таких случаев предлагается использовать функцию потерь (3), которая, наряду с новой информацией, учитывает и ту, что уже содержится в БЗ ИС.

$$L_t = \theta \left(\sum_{s=1}^S w_s \sum_{f=1}^{c_t^s} (v_{f,t}^s - v_{f,t}^*)^2 + \lambda \sum_{f \in \mathcal{F}} \sum_{j=1}^{D_{f,t}} Tr(v_f^j) (v_f^j - v_{f,t}^*)^2 \right) - \sum_{s=1}^S c_t^s \log(w_s). \quad (3)$$

В формуле (3) гиперпараметры θ и λ также должны быть заданы предварительно. Агрегированные значения в рамках построенной модели доверия носят вспомогательный характер и необходимы только для взвешивания источников и вычисления рейтинговых характеристик $\mu_{f,t}^s = 1/e_{f,t}^s$, вычисляемых как величина, обратная ошибке $v_{f,t}^s$ по отношению к $v_{f,t}^*$. Величина $\mu_{f,t}^s$ используется при оценке доверия к $v_{f,t}^s$. В зависимости от вида функции потерь агрегированные значения $v_{f,t}^*$ вычисляются по одной из формул (4):

$$\begin{aligned} v_{f,t}^* &= \frac{\sum_{s=1}^S w_s \cdot v_{f,t}^s}{\sum_{s=1}^S w_s}; \\ v_{f,t}^* &= \frac{\sum_{s=1}^S w_s \cdot v_{f,t}^s + \lambda \sum_{j=1}^{D_{f,t}} Tr(v_f^j) v_f^j}{\sum_{s=1}^S w_s + \lambda \sum_{j=1}^{D_{f,t}} Tr(v_f^j)}. \end{aligned} \quad (4)$$

Вид формулы для $v_{f,t}^*$ не влияет на способ взвешивания источников. Их веса вычисляются по формуле (5), применяемой в работе [9].

$$w_s = \frac{2\alpha - 2 + \sum_{t=1}^T c_t^s}{2\beta + \theta \sum_{t=1}^T \sum_{f=1}^{c_t^s} (e_{f,t}^s)^2}. \quad (5)$$

В [9] была доказана сходимость процесса взвешивания, т.е. веса источников при такой оценке сходятся к определённым значениям. Начальные веса определяются случайно и подчиняются гамма-распределению с параметрами α и β .

3.2 Результаты на наборах численных данных

Оригинальные модели предназначены для работы с численными данными, соответственно, все $v_{f,t}^s$ — это целые или вещественные числа. Предлагаемая модель оценки доверия, обозначенная *Markov Trust Evaluation model (MarkTE)*, сравнивалась с моделью агрегации [9] *DYNAMIC Truth Discovery (DynaTD)* на случайно сгенерированных массивах данных разных размеров, представленных в таблице 2.

Таблица 2 – Искусственные наборы данных

	S	T	F
Small	10	15	50
Medium	65	45	250
Large	150	100	500

Приведены результаты решения двух задач, обозначенных как *NF* и *FX*. Задача *NF* (*Not Fixed truth*) соответствует ситуации, когда из каждого источника в каждый момент времени извлекается значение $v_{f,t}^s$, а истинные значения зависят от t . Это означает, что для каждого $f \in \mathcal{F}$ и каждого t существует отдельное истинное значение $v_{f,t}^a$. Условия задачи *NF* соответствуют таковым в экспериментах, приведённых в работах [9, 10]. Так как истинные значения постоянно меняются, то в функции (3) не имеет смысла учитывать информацию, полученную ранее, поэтому в задаче *NF* в формуле (4) принята $\lambda = 0$. Задача *FX* (*Fixed truth*) соответствует ситуации, когда значение вида $v_{f,t}^s$ может быть извлечено из источника s не гарантированно, но с некоторой вероятностью, распределённой как $U_{(0.4,0.6)}$, а истинные значения v_f^a фиксированы и не зависят от времени. При этом количество альтернатив для f также фиксировано, а одни и те же значения могут встречаться в разных источниках в разное время. Данная ситуация схожа с теми, на которые ориентирована модель *MarkTE*, и при оценке учитывалась информация из БЗ с $\lambda > 0$. Путём экспериментов в условиях задач *NF* и *FX*, предложенная модель доверия была протестирована в разных условиях.

Для каждой задачи были сгенерированы отдельные наборы данных. Каждый источник $s \in S$ получил истинный вес w_s^a , и для каждого $f \in \mathcal{F}$ были определены истинные значения. Вероятность того, что источник предоставил информацию, не соответствующую действительности, обратно пропорциональна его весу. Как показано на рисунке 1, максимальная, средняя и медианная ошибки источников тем больше, чем меньше назначенный им истинный вес. Ошибки вычислялись как $|v_{f,t}^s - v_{f,t}^a|$ в задаче *NF* и $|v_{f,t}^s - v_f^a|$ в задаче *FX*.

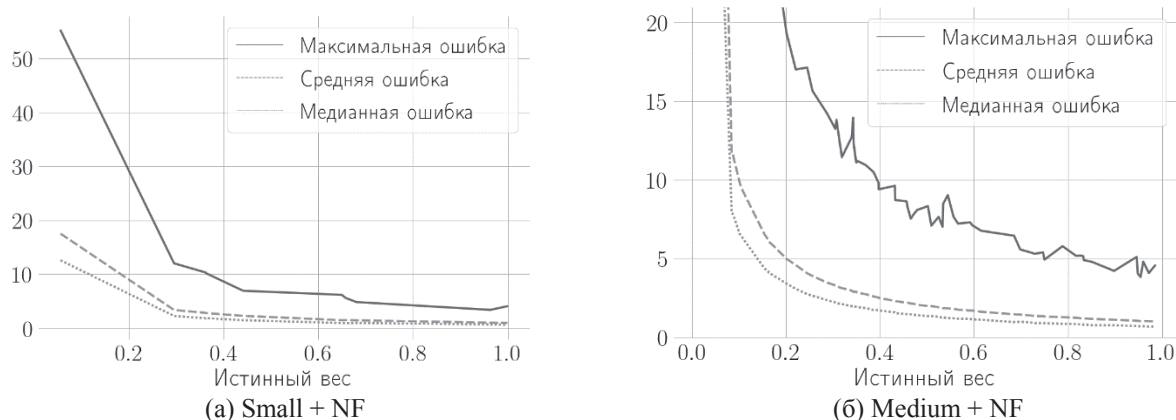


Рисунок 1 – Ошибки источников в зависимости от их истинного веса

Характеристики $\mu_{f,t}^s$ для переходных матриц (1) вычислялись как $\mu_{f,t}^s = 1/e_{f,t}^s$, где $e_{f,t}^s = |v_{f,t}^s - v_{f,t}^*|$. Зная вектор распределения доверия $\bar{\pi}^{t-1}$ значения $v_{f,t}^s$, распределение на шаге t вычислялось как $\bar{\pi}^t = \bar{\pi}^{t-1} \cdot P(\mu_{f,t}^s)$. Веса источников последовательно пересчитывались по формуле (5). Зависимость весов, вычисленных на основе модели *MarkTE*, от ошибок источника на истинных значениях показана на рисунке 2.

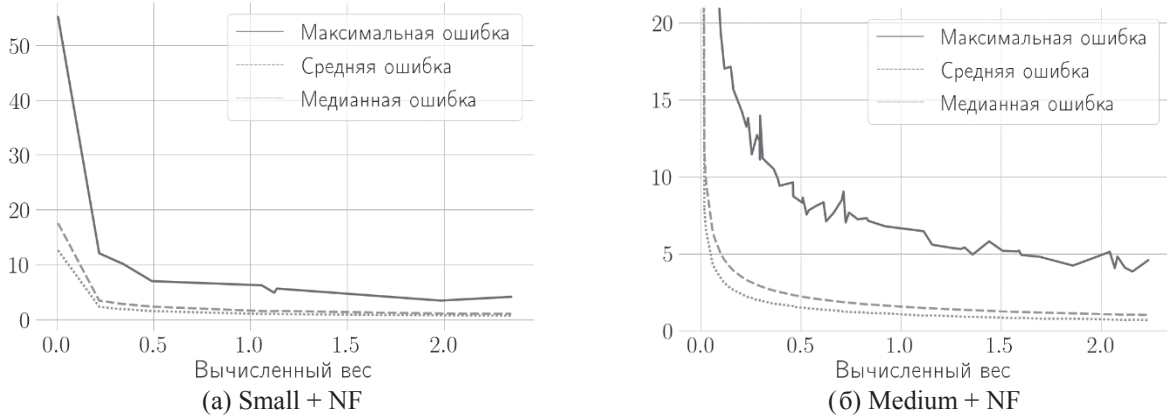


Рисунок 2 – Ошибки источников в зависимости от их вычисляемого веса

Сходство графиков на рисунках 1 и 2 говорит о том, что веса источников, предоставивших информацию, близкую к истинной, сходятся к величинам бóльшим, нежели те, к которым сходятся веса источников, содержащих информацию, далекую от истины. Таким образом, источники были взвешены корректно, и более надёжные получили бóльшие веса по сравнению с менее надёжными.

Модель оценки доверия служит для ранжирования полученных знаний, тогда как модель *DynaTD* выполняет агрегацию. Для сравнения полученных результатов была введена операция агрегации (6) для всех s , таких что $v_{f,t}^s$ существует, выполняемая на основе оценок доверия полученных знаний, а не их источников, как в (5).

$$v_{f,t}^m = \frac{\sum_{s=1}^S v_{f,t}^s \cdot Tr(v_{f,t}^s)}{\sum_{s=1}^S Tr(v_{f,t}^s)}. \quad (6)$$

В качестве показателей эффективности были использованы результаты сравнения полученных агрегированных значений с истинными: средняя абсолютная ошибка (*Mean Absolute Error*, *MAE*) и средняя квадратичная ошибка (*Root Mean Squared Error*, *RMSE*).

$$V = \{(t, s, f) \in T \times S \times \mathcal{F} | \exists v_{f,t}^s\}$$

$$MAE = \frac{\sum_{(t,s,f) \in V} |v_{f,t}^s - v_{f,t}^a|}{|V|}$$

$$RMSE = \sqrt{\frac{\sum_{(t,s,f) \in V} (v_{f,t}^s - v_{f,t}^a)^2}{|V|}}.$$

Базисный уровень был реализован двумя дополнительными моделями *Mean* и *Median*. Агрегированные значения $v_{f,t}^*$ в модели *Mean* вычисляются как среднее по всем $v_{f,t}^s$, в модели *Median* агрегацию выполняет функция медианы.

Результаты оценки *MAE* и *RMSE* для всех моделей и наборов данных на каждой задаче приведены в таблице 3. Видно, что агрегированные значения, полученные по формуле (6) на основе оценок доверия в соответствии с моделью *MarkTE*, в большинстве случаев оказались ближе к истинным. Отсюда можно заключить, что эффективность модели *MarkTE* находится на конкурентном уровне по сравнению с моделью *DynaTD* и другими, сравнения с которыми

были проведены в [9]. Полученные оценки доверия оказались более точными весовыми коэффициентами при агрегации.

Таблица 3 – Сравнение результатов моделей *MarkTE*, *DynaTD*, *Mean* и *Median*

Задача		Мера	<i>MarkTE</i>	<i>DynaTD</i>	<i>Mean</i>	<i>Median</i>
Small	NF	MAE	0,4367	0,5448	2,8096	0,5601
		RMSE	0,5765	0,9925	3,5978	0,7068
	FX	MAE	0,3841	0,5306	0,5449	0,6093
		RMSE	0,542	0,673	0,6888	0,8243
Medium	NF	MAE	0,2093	0,2087	1,2755	0,2507
		RMSE	0,2674	0,3603	1,5951	0,3155
	FX	MAE	0,3297	0,6205	0,6377	0,8149
		RMSE	0,4289	0,7759	0,7969	1,0263
Large	NF	MAE	0,1389	0,132	1,9844	0,1688
		RMSE	0,1814	0,2727	2,4887	0,2117
	FX	MAE	0,2791	0,582	0,6018	0,7967
		RMSE	0,3599	0,7325	0,757	1,0124

Каждая единица информации, выраженная значением $v_{f,t}^s$, получила оценку доверия $Tr(v_{f,t}^s)$. На рисунке 3 показана зависимость $Tr(v_{f,t}^s)$ от абсолютной ошибки $|v_{f,t}^s - v_{f,t}^a|$ на примере задачи *Medium+NF*.

Набор данных *Medium* (см. таблицу 2) содержал $65 \cdot 45 \cdot 250 = 731250$ значений. В целях наглядности на рисунке 3 представлены результаты для случайной выборки из 100 значений. На графике видна очевидная тенденция уменьшения доверия с ростом ошибки, из чего можно заключить, что предлагаемая модель *MarkTE* корректно оценивает доверие к поступающим в ИС данным, назначая наибольшие показатели доверия значениям с минимальной ошибкой.

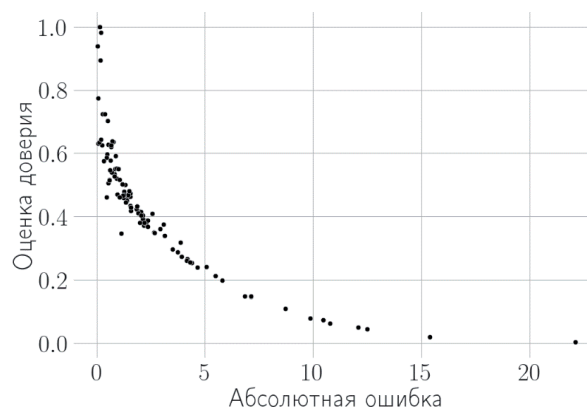


Рисунок 3 – Оценка доверия к фактам в зависимости от абсолютной ошибки

Заключение

Предлагаемая модель оценки доверия к информации, извлекаемой из внешних источников для пополнения БЗ ИС, построенной на основе онтологии некоторой ПрО, способна демонстрировать эффективные показатели по сравнению с другими моделями на задачах оценки численных данных. В качестве данных могут выступать показатели, например, стоимость акций и капитализация компаний, информация о наличии товаров на складе и др. В общем случае модель *MarkTE* способна оценить доверие к текстовым данным или данным, представленным в виде *RDF*-триплетов, что соответствует их представлению в онтологиях.

Список источников

- [1] **Баклавски К.** Онтологический Саммит 2020. Коммюнике: Графы знаний / К. Баклавски, М. Беннет, Г. Берг-Кросс, Т. Шнайдер, Р. Шарма, Д. Сингер. Перевод с англ. Д. Боргест // Онтология проектирования. 2020. Т.10, №4(38). С.540–555. DOI: 10.18287/2223-9537-2020-10-4-540-555.

- [2] **Simsek U., Umbrich J., Fensel D.** Towards a Knowledge Graph Lifecycle: A pipeline for the population of a commercial Knowledge Graph. In: A. Paschke, C. Neudecker, G. Rehm, J.A. Qundus and L. Pintscher (eds): Proceedings of the Conference on Digital Curation Technologies Qurator-2020 (Berlin, Germany, 2020, January 20-21). CEUR Workshop Proceedings, vol. 2535, CEUR-WS.org. https://ceur-ws.org/Vol-2535/paper_10.pdf.
- [3] **Fernández-Cañellas D. et al.** Enhancing Online Knowledge Graph Population with Semantic Knowledge. In: The Semantic Web ISWC 2020. Lecture Notes in Computer Science, vol 12506. Springer, Cham. 2020. P.183–200. DOI: 10.1007/978-3-030-62419-4_11.
- [4] **Cimmino A., García-Castro R.** Helio: a framework for implementing the life cycle of knowledge graphs. Semantic Web. Preprint 2022. P.1–27. DOI: 10.3233/SW-233224.
- [5] **Galland A., Abiteboul S., Marian A., and Senellart P.** Corroborating information from disagreeing views. In: Proceedings of the third ACM international conference on Web search and data mining WSDM-2010. (New York, USA, 2010, February 4–6). 2010. P.131–140. DOI: 10.1145/1718487.1718504.
- [6] **Li X., Dong X.L., Lyons K.B., Meng W., Srivastava D.** Truth finding on the deep web: Is the problem solved? In: Proceedings of the VLDB Endowment. vol. 6(2). 2012. P.97–108. DOI: 10.14778/2535568.2448943.
- [7] **Pochampally R. et al.** Fusing data with correlations. In: Proceedings of the 2014 ACM SIGMOD international conference on Management of data SIGMOD-2014 (Snowbird, Utah, USA, 2014, June 22–27). 2014. P.433–444. DOI: 10.1145/2588555.2593674.
- [8] **Dong X.L., Gabrilovich E., Murphy K., Dang V., Horn W., Lugaresi C., Sun S., Zhang W.** Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources. In: Proceedings of the VLDB Endowment. vol. 8, 2015. P.938–949. DOI: 10.14778/2777598.2777603.
- [9] **Li Y. et al.** On the discovery of evolving truth. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD-2015 (Sydney, NSW, Australia, 2015, August 10–13). 2015. P.675–684. DOI: 10.1145/2783258.2783277.
- [10] **Yao L. et al.** Online truth discovery on time series data. In: Proceedings of the 2018 SIAM international Conference on Data Mining SDM-2018 (San Diego, USA, 2018, October 6–13). 2018. Society for Industrial and Applied Mathematics. P.162–170. DOI: 10.1137/1.9781611975321.19.
- [11] **Zubiaga A., Liakata M., Procter R., Wong Sak Hoi G., Tolmie P.** Analysing how people orient to and spread rumours in social media by looking at conversational threads. PLoS ONE 2016. 11(3): e0150989. DOI: 10.1371/journal.pone.0150989.
- [12] **Kochkina E., Liakata M., Zubiaga A.** All-in-one: Multi-task learning for rumour verification. In: Proceedings of 27th International Conference on Computational Linguistics COLING-2018 (Santa Fe, New-Mexico, USA, 2018, August 20–26). Association for Computational Linguistics (ACL). 2018. P.3402–3413. DOI: 10.48550/arXiv.1806.03713.
- [13] PHEME dataset for Rumour Detection and Veracity Classification. <https://www.kaggle.com/datasets/usharengaraju/pHEME-dataset>.
- [14] **Chen X., Yuan Y., Lu L., Yang J.** A multidimensional trust evaluation framework for online social networks based on machine learning. IEEE Access. vol. 7, 2019. P.175499–175513. DOI: 10.1109/ACCESS.2019.2957779.
- [15] **Vyas P., El-Gayar O.** Credibility analysis of news on twitter using LSTM: An exploratory study. In: Proceedings of 26th Americas Conference on Information Systems AMCIS 2020 (Virtual conference, 2020, August 10–14). Association for Information Systems. <https://scholar.dsu.edu/cgi/viewcontent.cgi?article=1150&context=bispapers>.
- [16] **Hirlekar V.V., Kumar A.** Tweet Credibility Detection for COVID-19 Tweets using Text and User Content Features. International Journal of Advanced Computer Science and Applications, 13(4), 2022. P.430–439. DOI: 10.14569/IJACSA.2022.0130451.

Сведения об авторе

Серый Алексей Сергеевич, 1987 г. рождения. Окончил Новосибирский государственный университет в 2010 г. Младший научный сотрудник лаборатории искусственного интеллекта Института систем информатики им. А.П. Ершова (Новосибирск). В списке научных трудов более двух десятков работ в области представления знаний и компьютерной лингвистики. Author ID (RSCI): 714554; ORCID: 0000-0001-8275-4700; Author ID (Scopus): 56403204900; Researcher ID (WoS): K-1557-2018. alexey.seryj@iis.nsk.su. ✉.



Поступила в редакцию 10.01.2023, после рецензирования 31.01.2023. Принята к публикации 11.02.2023.



Data credibility when populating ontologies and knowledge graphs

© 2023, A.S. Sery

A.P. Ershov Institute of Informatics Systems of Siberian Branch of RAS, Novosibirsk, Russia

Abstract

The problem of assessing trust in the information extracted from textual sources to populate ontologies or knowledge graphs is considered. For a unit of information or a fact, the minimum knowledge about an instance of the subject area, expressed by a single RDF triplet, is taken. The paper provides a description of a probabilistic trust evaluation model based on Markov random processes. When assessing, the model is built on the basis of available information about sources, taking into account previously extracted data. A method for assessing the credibility of information with parallel weighting of sources is also provided. The proposed approach is in demand when the quality of the data sources is unknown or unavailable. As part of testing the model, sets of numerical data of various sizes were automatically generated, experiments were carried out to weigh the sources and assess trust in the information extracted from them. It was shown that in most cases the weights of the sources calculated on the basis of the proposed model are the greater, the smaller the average deviation of the information they provide from the true one, and the confidence in facts increases with decreasing distance to the true data. Comparison with data aggregation models is made. In most cases, the aggregation based on the trust score showed the smallest average deviation from the true data among the considered models. The obtained results show that the proposed model is effective in comparison with other similar models and can be used in problems of assessing trust in facts represented by real numbers.

Key words: ontology, knowledge graph, data extraction, information trustworthiness, Markov Process.

For citation: Sery AS. Data credibility when populating ontologies and knowledge graphs [In Russian]. *Ontology of designing*. 2023; 13(1): 113-124. DOI:10.18287/2223-9537-2023-13-1-113-124.

Conflict of interest: The author declares no conflict of interest.

List of figures and tables

Figure 1 - Errors of sources depending on their actual weight

Figure 2 - Errors of sources depending on their calculated weight

Figure 3 - Assessing confidence in facts depending on the absolute error

Table 1 - Designations

Table 2 - Synthetic datasets

Table 3 - Comparison of results of MarkTE, DynaTD, Mean and Median models

References

- [1] **Baclawski K, Bennett M, Berg-Cross G, Schneider T, Sharma R, Singer J, Sriram, R.D.** Ontology summit 2020 communiqué: Knowledge graphs. *Applied Ontology*. 2021; 16(2): 229–247. DOI: 10.18287/2223-9537-2020-10-4-540-555.
- [2] **Simsek U, Umbrich J, Fensel D.** Towards a Knowledge Graph Lifecycle: A pipeline for the population of a commercial Knowledge Graph. In: A. Paschke, C. Neudecker, G. Rehm, J.A. Qundus and L. Pintscher (eds): *Proceedings of the Conference on Digital Curation Technologies Qurator-2020* (Berlin, Germany, 2020, January 20-21). CEUR Workshop Proceedings, vol. 2535, CEUR-WS.org, https://ceur-ws.org/Vol-2535/paper_10.pdf.
- [3] **Fernández-Cañellas D. et al.** Enhancing Online Knowledge Graph Population with Semantic Knowledge. In: *The Semantic Web ISWC 2020. Lecture Notes in Computer Science*, vol 12506. Springer, Cham. 2020. 183–200. DOI: 10.1007/978-3-030-62419-4_11.
- [4] **Cimmino A, García-Castro R.** Helio: a framework for implementing the life cycle of knowledge graphs. *Semantic Web*. Preprint 2022. 1–27. DOI: 10.3233/SW-233224.

- [5] **Galland A, Abiteboul S, Marian A, Senellart P.** Corroborating information from disagreeing views. In: Proceedings of the third ACM international conference on Web search and data mining WSDM-2010. (New York, USA, 2010, February 4–6). 2010. 131–140. DOI: 10.1145/1718487.1718504.
 - [6] **Li X, Dong XL, Lyons KB, Meng W, Srivastava D.** Truth finding on the deep web: Is the problem solved? In: Proceedings of the VLDB Endowment. 2012; 6(2): 97–108. DOI: 10.14778/2535568.2448943.
 - [7] **Pochampally R. et al.** Fusing data with correlations. In: Proceedings of the 2014 ACM SIGMOD international conference on Management of data SIGMOD-2014 (Snowbird, Utah, USA, 2014, June 22–27). 2014. 433–444. DOI: 10.1145/2588555.2593674.
 - [8] **Dong XL, Gabrilovich E, Murphy K, Dang V, Horn W, Lugaresi C, Sun S, Zhang W.** Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources. In: Proceedings of the VLDB Endowment. 2015; 8: 938–949. DOI: 10.14778/2777598.2777603.
 - [9] **Li Y. et al.** On the discovery of evolving truth. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD-2015 (Sydney, NSW, Australia, 2015, August 10–13). 2015. 675–684. DOI: 10.1145/2783258.2783277.
 - [10] **Yao L. et al.** Online truth discovery on time series data. In: Proceedings of the 2018 SIAM international Conference on Data Mining SDM-2018 (San Diego, USA, 2018, October 6–13). 2018. Society for Industrial and Applied Mathematics. 162–170. DOI: 10.1137/1.9781611975321.19.
 - [11] **Zubiaga A, Liakata M, Procter R, Wong Sak Hoi G, Tolmie P.** Analysing how people orient to and spread rumours in social media by looking at conversational threads. PLoS ONE 2016. 11(3): e0150989. DOI: 10.1371/journal.pone.0150989.
 - [12] **Kochkina E, Liakata M, Zubiaga A.** All-in-one: Multi-task learning for rumour verification. In: Proceedings of 27th International Conference on Computational Linguistics COLING-2018 (Santa Fe, New-Mexico, USA, 2018, August 20–26). Association for Computational Linguistics (ACL). 2018. 3402–3413. DOI: 10.48550/arXiv.1806.03713.
 - [13] PHEME dataset for Rumour Detection and Veracity Classification. <https://www.kaggle.com/datasets/usharengaraju/pHEME-dataset>.
 - [14] **Chen X, Yuan Y, Lu L, Yang J.** A multidimensional trust evaluation framework for online social networks based on machine learning. IEEE Access. 2019; 7: 175499–175513. DOI: 10.1109/ACCESS.2019.2957779.
 - [15] **Vyas P, El-Gayar O.** Credibility analysis of news on twitter using LSTM: An exploratory study. In: Proceedings of 26th Americas Conference on Information Systems AMCIS 2020 (Virtual conference, 2020, August 10–14). Association for Information Systems. <https://scholar.dsu.edu/cgi/viewcontent.cgi?article=1150&context=bispapers>.
 - [16] **Hirlekar VV, Kumar A.** Tweet Credibility Detection for COVID-19 Tweets using Text and User Content Features. International Journal of Advanced Computer Science and Applications, 2022; 13(4): 430–439. DOI: 10.14569/IJACSA.2022.0130451.
-

About the author

Alexey Sergeevich Sery (b.1987) holds a master's degree in mathematics from Novosibirsk State University (2010) and the position of Junior Researcher at the A.P. Ershov Institute of Informatics Systems of Siberian Branch of RAS. He is the author of more than 20 papers in the fields of NLP systems and Knowledge Representation. Author ID (RSCI): 714554; ORCID: 0000-0001-8275-4700; Author ID (Scopus): 56403204900; Researcher ID (WoS): K-1557-2018. alexey.seryj@iis.nsk.su. ✉.

Received January 10, 2023. Revised January 31, 2023. Accepted February 11, 2023.

МЕТОДЫ И ТЕХНОЛОГИИ ПРИНЯТИЯ РЕШЕНИЙ

УДК 004.3

Научная статья

DOI: 10.18287/2223-9537-2023-13-1-125-138

**Многоаспектное моделирование ситуаций
в задачах обеспечения функциональной безопасности
аппаратно-программных комплексов**

© 2023, В.Е. Гвоздев, О.Я. Бежаева ✉, Г.Р. Сафина

Уфимский университет науки и технологий, Уфа, Россия

Аннотация

Функциональная безопасность является латентной характеристикой аппаратно-программных комплексов (АПК) и объективно характеризуется удовлетворённостью потребителей поведением АПК. В работе рассматривается подход к решению задач, связанных с управлением проектом, на основе анализа динамических характеристик функциональной безопасности конфликтных ситуаций в системе управления проектом. Для анализа ситуаций, возникающих при управлении проектами создания АПК возможно использование моделей, известных как системные архетипы. Рассмотрены информационная сущность ситуаций и основы многоаспектного моделирования. Методическую основу исследований составляет сочетание динамических моделей параметров, характеризующих функциональную безопасность, и структурных моделей, соответствующих конфликтным ситуациям, возникающим при обеспечении требуемого уровня функциональной безопасности. Рассмотрены примеры многоаспектного моделирования ситуаций, где в качестве событий выступают проявления латентных дефектов. В результате исследований: определены информационные сущности ситуаций, предложены концептуальные основы многоаспектного моделирования ситуаций, возникающих при управлении функциональной безопасностью АПК, выделены базовые этапы построения системы структурных и динамических моделей ситуаций на разных стадиях жизненного цикла АПК. Полученные результаты могут быть использованы для принятия решений о целесообразности внесения изменений в структуру системы обеспечения функциональной безопасности АПК.

Ключевые слова: функциональная безопасность, аппаратно-программный комплекс, сетцентрическое управление, системный архетип, структурные модели.

Цитирование: Гвоздев В.Е., Бежаева О.Я., Сафина Г.Р. Многоаспектное моделирование ситуаций в задачах обеспечения функциональной безопасности аппаратно-программных комплексов // Онтология проектирования. 2023. Т.13, №1(47). С.125-138. DOI: 10.18287/2223-9537-2023-13-1-125-138.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

Возрастание роли систем обработки данных в управлении распределёнными техническими системами выдвигает на первый план обеспечение функциональной безопасности (ФБ) аппаратно-программных комплексов (АПК). Цифровая экосреда является системообразующим фактором сетцентрического управления территориально распределёнными сложными техническими системами (СТС). Базовым требованием к их информационной инфраструктуре является предоставление своевременной полной и достоверной информации, необходимой для урегулирования ситуаций, возникающих в разных частях распределённой

СТС. Потребительские свойства АПК определяется качеством управления проектами создания АПК [1], относящихся к классу субъектоцентрических систем [2].

Безопасность является латентной характеристикой АПК и объективно характеризуется удовлетворённостью потребителей поведением комплекса. В настоящей работе рассматривается подход к решению задачи, связанной с управлением проектом АПК и состоящей в выявлении конфликтных ситуаций, имеющих место в системе управления проектом.

1 Подходы к обеспечению ФБ АПК

При создании АПК обеспечение безопасного их функционирования является важнейшей задачей [3, 4]. В [1, 3] отмечается необходимость развития методологических и теоретических основ дефектологии АПК как одного из направлений исследований в области системной инженерии. В [4] выделяются классы задач, связанные с обеспечением технологической и эксплуатационной безопасности. В [5, 6] отмечается необходимость совершенствования систем управления проектами создания компонент распределённых коммуникационно-вычислительных систем. В [3, 7, 8] выделяются задачи, связанные с обеспечением конфиденциальности, целостности и доступности информации в случае, когда пользователи принимают активное участие в управлении СТС. Задачи, связанные с обеспечением ФБ АПК при решении задач управления в реальном времени, обсуждаются в [9, 10].

Под ФБ АПК понимается свойство сохранять работоспособность в соответствии с целевым назначением при случайных дестабилизирующих воздействиях и отсутствии злоумышленного влияния на программную, аппаратную составляющие и базы данных [3]. Из анализа работ [11-13] можно сделать заключение о том, что ФБ определяется не выявленными (латентными) дефектами в конструкции АПК. Дефекты являются следствием ошибок, допускаемых разработчиками АПК на разных стадиях их жизненного цикла (ЖЦ). Существуют разные подходы к управлению непреднамеренно допускаемыми дефектами на разных стадиях ЖЦ, включая предпроектную стадию. Важное значение имеет природа возникновения дефектов. На начальных стадиях ЖЦ преобладают дефекты, обусловленные неопределённостью среды использования и нечёткостью целей управления. По мере перехода от начальных стадий ЖЦ к последующим акценты перемещаются в области: неверного использования руководств по разработке; технологий использования инструментальных средств; нарушения условий применимости моделей ЖЦ АПК. Используемые методы управления дефектами можно соотнести с проактивным, активным и реактивным подходами к управлению СТС.

К настоящему времени развитие получили методы, соотносимые с активным и реактивным подходами, т.е. ориентированные на выявление ошибок по результатам специально организованных испытаний и изучения исторических данных об опыте эксплуатации АПК с целью установления закономерностей в проявлениях симптомов дефектов, а также причин возникновения дефектов. Эту группу методов, можно объединить понятием - анализ коренных причин отказов [14]. Они являются адаптацией положений теории решения изобретательских задач [15] в область управления ФБ АПК, а именно методов, относящихся к диверсионному анализу [16]. К другому развиваемому в настоящее время направлению активного подхода к обеспечению ФБ АПК относятся методы, которые можно объединить понятием анализ распространения ошибок [17-19]. Суть этих методов состоит в раннем обнаружении проявлений дефектов и сбоев и парировании этих явлений. К методам, основу которых составляет проактивный подход к управлению СТС, следует отнести методы, ориентированные на сравнительный анализ альтернатив проектных решений по критериям ФБ (анализ видов и последствий отказов; анализ дерева отказов) [20-22].

В работах [23, 24] к числу критических факторов, негативно влияющих на успех реализации проектов создания АПК, относятся функциональные¹ и нефункциональные требования² к потребительским свойствам систем. Это обусловлено сложностью формирования консолидированного мнения различных целевых групп пользователей в условиях неопределённости среды использования и размытости целей функционирования СТС [2, 25-27]. В работах [26, 27] обосновывается необходимость развития подходов к выработке консолидированных решений на основе методов конвергентного управления и развития сетевых технологий согласования решений заинтересованных сторон.

2 Информационная сущность ситуаций

Информационная сущность ситуаций, связанных с обеспечением ФБ АПК, формируется на основе следующих понятий.

События - проявление взаимодействий процессов, протекающих как внутри исследуемого объекта, так и вне его. Событие объективно. Ситуация определяется местом субъекта внутри события, его озабоченностью событием. Степень озабоченности влияет на восприятие события субъектом, составляет основу для выделения симптомов ситуаций, формирования описания ситуации [2, 25]. Описание ситуации служит основанием получения ответа на вопросы: в чём содержание события и как наилучшим образом реагировать на событие сейчас? Режим понимания ситуаций – реактивный.

Шаблоны событий предназначены для выявления наиболее вероятных условий возникновения событий, которые неоднократно имели место ранее. Режим понимания – адаптивный, что определяет выбор способов наиболее эффективного реагирования на события. Составляющая, связанная с управлением ситуацией, т.е. стимулированием возникновения /предотвращением возникновения событий, в шаблонах отсутствует.

Систематическая структура предназначена для выявления «генераторов», «виновников» возникновения событий. При формировании систематических структур необходимо определить системообразующие факторы (в том числе гипотетические, ранее не наблюдавшиеся). Установление причин возникновения ситуаций есть производная от ментальных моделей исследователей, основу которых составляют представления о ценностях. Режим понимания – креативный, определяемый ментальными моделями субъектов - конструкторов систематических структур. Выявление условий возникновения ситуаций создаёт базу для влияния на них, т.е. выбора «точек приложения рычага» [28] изменения ситуаций в нужную сторону.

Расширенное видение - понимание того, что хочется получить в результате урегулирования ситуации, какие альтернативы систематических структур позволят достичь желаемого. Новые знания изменяют положение дел как за счёт инженерных решений, так и за счёт изменения ментальных моделей с учётом понимания сложности событий и ограниченности ресурсов для урегулирования ситуаций. Это соответствует генеративному режиму понимания событий.

Информационная сущность делает возможным построение разных знаковых моделей характеристик ФБ, которые создают основу для выработки обоснованных решений по урегули-

¹ Функциональные требования описывают сервисы, предоставляемые АПК или программной системой, её поведение в определённых ситуациях, реакцию на входные данные и действия, которые система позволит выполнить пользователям.

² Нефункциональные требования фиксируют условия, которые непосредственно не связаны с поведением или функциональностью решения, а скорее описывают условия окружающей среды, при которых решение должно оставаться эффективным, или качества, которыми система должны обладать. Они также известны как атрибуты (показатели) качества или дополнительные требования. Они могут включать требования, связанные с пропускной способностью, надёжностью, масштабируемостью, скоростью, безопасностью, доступностью, информационной архитектурой и др.

рованию ситуаций. Одной из разновидностей структурных моделей ситуаций являются так называемые «системные архетипы» (СА) [29].

СА составляют основу решения двух классов задач: диагностических и прогностических. Основой решения диагностических задач является выбор из доступного множества СА того, который соответствует наблюдаемым симптомам ситуации. Это позволяет на качественном уровне определить основные причины сложившейся ситуации и подходы к её урегулированию. Рекомендации по урегулированию ситуаций, соответствующих разным СА, приведены в [29]. Одной из разновидностей прогностических задач является выявление по результатам анализа динамических характеристик поведения СТС конфликтных ситуаций в системе управления. Результатом решения задачи является параметрическая настройка существующей структуры системы управления либо внесение изменения в структуру. Получаемые оценки зависят от динамических характеристик поведения, учитываемых при выявлении ситуаций. Полученные результаты создают основу планирования системы мер по целенаправленному изменению структуры и параметров системы управления («выбора точки приложения рычага» [30]).

3 Концептуальная основа многоаспектного моделирования

Концептуальную основу моделирования СТС составляют следующие положения.

- АПК есть разновидность СТС. Это даёт основание научной адаптации моделей ситуаций, возникающих при управлении СТС иной природы, в область управления ФБ АПК.
- Признаком наличия ситуаций являются симптомы событий, которые характеризуют отклонение поведения системы от ожидаемого. Различные субъекты (аналитики, постановщики задач, проектировщики, испытатели, системные администраторы, пользователи и потребители информационных продуктов и услуг), вовлечённые в урегулирование ситуаций, в силу различия их ментальных моделей по-разному воспринимают одно и то же событие и по-разному выделяют ситуации. Последнее обстоятельство служит причиной одновременного использования разных моделей ситуаций для характеристики одного и того же события.
- Соответствие свойств системы запросам потребителей определяется качеством управления на всех стадиях ЖЦ АПК: от обследования информационных потребностей, желаний и представлений о ценностях потребителей, до эксплуатации и модернизации АПК. Свойства АПК определяются их устройством и проявляются как внешнее поведение. Ошибки в организации управления проектом представляют наибольшую опасность для результатов проекта.
- Особенностью СТС является наличие множества взаимодействующих контуров причинно-следственных связей параметров, определяющих поведение системы, а также задержек времени между оказываемыми на систему воздействиями и последствиями от их проявления. Особенности взаимодействия контуров изменяются во времени.

Известные рекомендации по урегулированию ситуаций на основе выделения СА носят качественный характер [29]. Недостаточное развитие получили структурные и математические модели, посредством которых можно обеспечить информационную поддержку выбора обоснованного подхода к урегулированию ситуаций. В настоящей статье методическую основу исследований составляет системное сочетание динамических моделей параметров, характеризующих ФБ, и структурных моделей, соответствующих конфликтным ситуациям, возникающим при обеспечении требуемого уровня ФБ.

4 Этапы построения моделей ситуаций

Основными этапами построения системы структурных и динамических моделей ситуаций в рамках архитектурного подхода, содержание которого представлено в стандарте *IEEE 1471*³, являются следующие.

- *Определение подхода к исследованиям.* Множественность точек зрения на события, влияющих на поведение показателей ФБ АПК, является проявлением свойства полиморфизма моделей. Точка зрения на событие предопределяет выбор существенных факторов, т.е. подходы к построению моделей и границы их применимости.
- *Временной горизонт.* Наличие адекватных структурных моделей ситуаций и динамических моделей изменения поведения ситуации во времени делает возможным оценить временные горизонты, в которых целесообразно обсуждать подходы к урегулированию ситуации.
- *Определение границ описания ситуации.* Этап предопределяет выделение минимально достаточного множества характеристических параметров ситуации.
- *Определение уровня детальности описания ситуации* (уровня абстракции): определяется уровнем понимания содержания ситуации, также временным горизонтом, в пределах которого предполагается урегулирование ситуации. Целью построения системы моделей является обобщение частных динамических характеристик к виду СА.
- *Значимые задержки.* Построение на основе анализа частных динамических характеристик СА создаёт основу для выделения значимых задержек в системе обеспечения ФБ. Задержки являются причиной нарастающего дисбаланса между свойствами разных частей системы. Своевременно выявленные временные задержки могут сыграть роль предохранительного клапана, препятствующего преобразованию дисбаланса в негативные последствия. Недостаточный учёт роли временных задержек может явиться причиной взрывного изменения свойств системы.

5 Примеры многоаспектного анализа ситуаций

В качестве событий в примерах многоаспектного анализа ситуации, связанных с обеспечением ФБ программной компоненты АПК, выступают проявления латентных дефектов, негативно влияющих на возможность получения ценных для потребителя информационных продуктов и услуг.

Пример 1. Влияние устранения латентных дефектов на показатели ФБ.

Описание задачи. Показателем ФБ является вероятность успешного завершения/отказа программы в серии прогонов. Считается, что при реализации серии сведения о проявлениях дефектов, их местоположении и характере лишь собираются. После завершения серии прогонов выполняется анализ коренных причин отказов, и в ПС вносятся изменения, улучшающие её свойства.

Требуется: построить совокупность моделей, способствующих повышению обоснованности принятия решения относительно объёмов ресурсов, необходимых для обеспечения требуемого уровня показателя ФБ.

Описанной задаче можно сопоставить структурную модель, представленную на рисунке 1.

Здесь:

³ Стандарт IEEE 1471-2000 - рекомендуемая практика описания архитектуры систем, интенсивно использующих программное обеспечение (*Recommended Practice for Architectural Description of Software — Intensive Systems*). <https://standards.ieee.org/ieee/1471/2187/>.

S_1 – этап подготовки исходных данных, соответствующих запросам пользователей;

S_2 – этап оформления информационных продуктов и сервисов и передачи их потребителям;

S_3 – состояние системы, при котором с вероятностью P дефекты проявляются;

S_4 – состояние системы, при котором дефекты не проявляются с вероятностью $Q = 1 - P$;

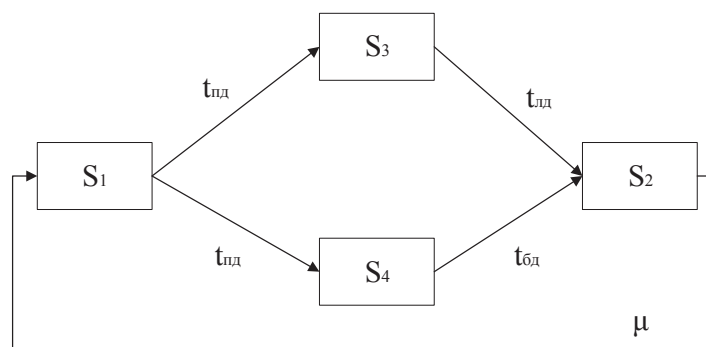


Рисунок 1 - Структурная модель точки зрения на ситуацию

$t_{нд}$ – время подготовки данных, соответствующих запросам потребителей;

$t_{лд}$ – время устранения проявившихся дефектов и последующей обработки данных;

$t_{бд}$ – время обработки данных в случае, когда латентные дефекты не проявляются;

μ – интенсивность запросов пользователей.

Имитационный эксперимент выполнен по следующему укрупнённому алгоритму.

Шаг 1. Задаётся ограничение на время обработки запроса пользователей $T_{ср}$. Если время обработки запроса превышает это значение, считается, что произошел отказ ПС [31].

Шаг 2. Посредством датчиков случайных чисел генерируются значения времени подготовки данных $t_{нд}^{(i)}$; время устранения проявившихся дефектов и последующей обработки данных $t_{лд}^{(i)}$, время обработки данных в случае, когда латентные дефекты не проявляются $t_{бд}^{(i)}$. Время между поступлениями заданий (на вход блока S_1) генерировалось посредством датчика случайных чисел на основании μ . В качестве распределения случайных величин принимался показательный закон. Его выбор обоснован тем, что он является характеристикой простейшего потока событий [32]. При этом параметры закона принимались обратными математическим ожиданиям $M[t_{нд}]$, $M[t_{лд}]$ и $M[t_{бд}]$. В ходе проведения исследований принималось, что $M[t_{нд}] < M[t_{лд}]$. Чтобы обеспечить устойчивость статистических оценок, число реализаций в серии принималось равным 1000.

Шаг 3. Выбор пути обработки данных (по ветке P либо ветке Q) для каждой заявки, поступающей на вход блока S_1 , осуществлялся случайным образом. Путь, на котором дефект проявлялся, выбирался с вероятностью P , а путь, на котором дефект не проявлялся, – с вероятностью Q . Для каждой i -й заявки генерировались $t_{нд}^{(i)}$ (когда выбиралась ветка Q), либо $t_{нд}^{(i)}$ (когда выбиралась ветка P). Общее время обработки заявки $T_{общ}^{(i)}$ определялось как сумма времён подготовки и обработки данных. При этом в качестве времени обработки данных выбиралось $t_{лд}^{(i)}$ (когда выбиралась ветка P), либо $t_{нд}^{(i)}$ (когда выбиралась ветка Q). Если соблюдалось условие $T_{общ}^{(i)} < T_{ср}$, считалось, что заявка обслужена успешно. В противном случае считалось, что произошел отказ ПС.

Шаг 4. На основе массива данных о результатах испытаний серии оценивалась вероятность отказа /безотказной работы ПС.

Шаг 5. Выполнялась оценка влияния усилий на выявление и устранение латентных дефектов на показатели ФБ. Для этого вносились изменения в значения вероятности проявления латентных дефектов при прогоне ПС по правилу: $P_l = P_{l-k} / l$. Здесь l – номер серии испытаний; k – задержка времени (выражаемая количеством серий) между внесением улучшений и проявлением последствий этого.

На рисунке 2 представлены динамические характеристики интегрального показателя ФБ – вероятности отказа $P_{отк}$ АПК при разных значениях временной задержки k .

Полученные зависимости создают основу планирования ресурсов (затрат времени на уменьшение вероятности проявления латентных дефектов) для обеспечения требуемых значений показателя ФБ. Выделенной ситуации ставится в соответствие СА «Уравновешивание с задержкой» (рисунок 3):

Контурная модель, представленная на рисунке 3, представляет собой стабилизирующий цикл. Характеристикой состояния ПС является вероятность безотказной работы/отказа.

На рисунке 4 представлен демонстрационный пример изменения вероятности отказа ПС в случае, когда усилия, направленные на уменьшение вероятности проявления латентных дефектов, не принимаются.

Пример 2. Эффективность использования ресурсов на устранение дефектов.

Описание задачи. Обеспечение ФБ реализуется на основе реактивного и проактивного подходов. В реактивном подходе осуществляется выявление и устранение проявившихся дефектов; проактивного – создание барьеров, препятствующих возникновению дефектов. При планировании проектов следует принимать обоснованные решения по выделению ресурсов на реализацию разных видов деятельности, как в рамках реактивного, так и проактивного подходов. Одним из критериев принятия решений является анализ динамики эффективности устранения дефектов.

Объективными косвенными показателями ФБ являются латентные дефекты [33]. Объективным показателем изменения ФБ является динамика количества выявленных дефектов в результате целенаправленной деятельности по выявлению и устранению латентных дефектов [1]. В силу свойства уникальности проектов [34] и субъектоцентрического характера программных проектов, модели динамики количества выявленных дефектов целесообразно строить в классе эмпирических моделей [35]. Модели, позволяющие оценить количество невыявленных дефектов, описаны, например, в [33].

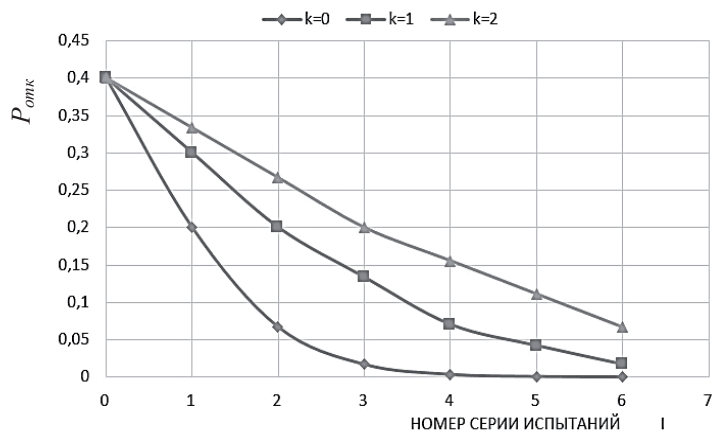


Рисунок 2 - Динамические характеристики интегрального показателя ФБ



Рисунок 3 - СА «Уравновешивание с задержкой»

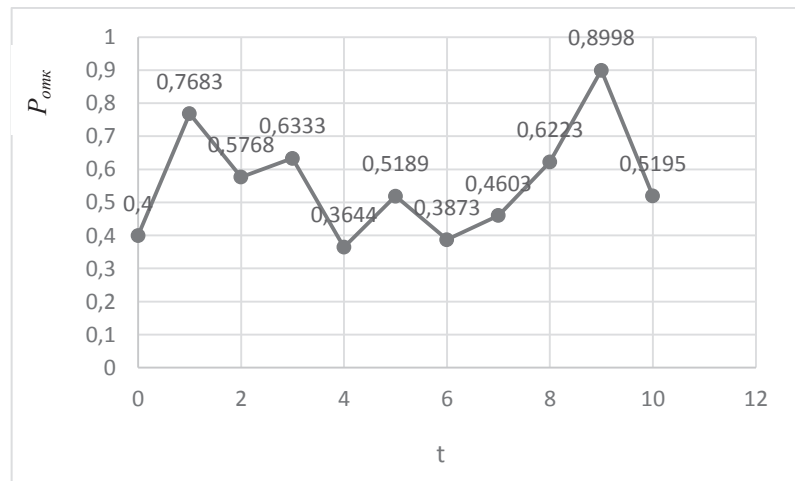


Рисунок 4 - Демонстрационный пример изменения вероятности отказа АПК при отсутствии подхода к уменьшению вероятности проявления латентных дефектов

Постановка задачи. Дано: динамика относительного показателя $D(t)$, характеризующего выявленные дефекты (например, в виде количества дефектов на 1000 строк кода); динамика затрат ресурсов на выявление дефектов $S(t)$; критерий эффективности устранения дефектов вида $E(t)=D(t)/S(t)$; нижняя граница критерия эффективности E^* , меньше которой дальнейшие затраты ресурсов на устранение дефектов становятся нецелесообразными.

Требуется: построить зависимость $E(t)$ и определить границу ресурсов (времени) t^* , соответствующую E^* .

Допущения: постулируется, что зависимость $D(t)$ имеет вид:

$$D(t) = K \cdot \exp(-\lambda t), \quad (1)$$

где K - исходное значение относительного показателя засоренности дефектами. Возможные подходы к оцениванию K описаны, например, в [33].

Зависимость затрат на устранение дефектов имеет вид:

$$S(t) = \exp(\mu t) - 1 \quad (2)$$

Пример решения задачи выполнен при значении λ и μ равных единице; $K=0.4$. На рисунке 5 представлены зависимости $D(t)$ и $S(t)$.

Временные отметки τ_i ($i=1, 2, 3, \dots$) соответствуют значениям

$D(t) = D(0)/(2)^{(i)}$, т.е. это отметки времени, в которых количество дефектов по сравнению с предыдущей отметкой уменьшилось вдвое.

В таблице 1 приведены значения $D(t)$, $S(t)$, $E(t)$, соответствующие разным отметкам времени τ_i . На рисунке 6 приведена графическая модель, характеризующая поведение $E(t)$ и построенная на основе данных, представленных в таблице 1.

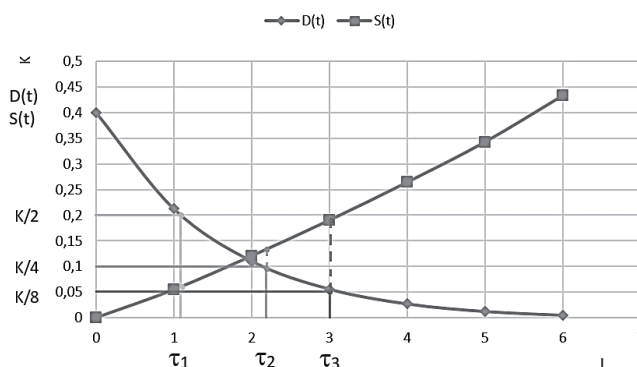


Рисунок 5 - Динамические модели $D(t)$ и $S(t)$

Таблица 1 – Значения $D(\tau_i)$, $S(\tau_i)$, $E(\tau_i)$ соответствующие разным отметкам времени τ_i .

i	τ_i	$D(\tau_i)$	$S(\tau_i)$	$E(\tau_i)$
1	1,1	0,20	0,07	3,43
2	2,2	0,10	0,14	0,87
3	3,0	0,05	2,00	0,28
4	4,0	0,02	2,26	0,10

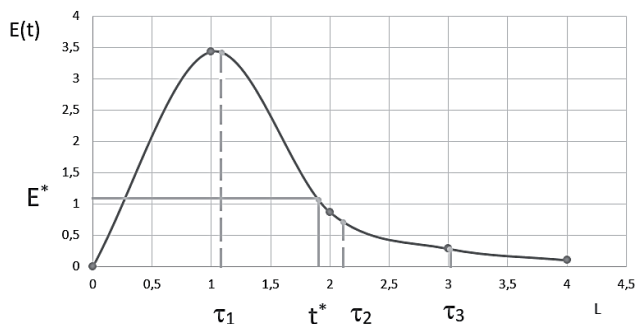


Рисунок 6 - Характеристика эффективности устранения дефектов

Зависимость $E(t)$ построена при условии, что $E(0)=0$, так как если усилий по устранению дефектов не прилагается, эффект отсутствует.

На основе зависимости $E(t)$ (рисунок 6) при заданном E^* можно оценить границы периода (ресурс времени) $T \in [0; t^*]$, в течение которого имеет смысл совершенствовать ФБ за счёт выявления и устранения дефектов.

Выделенному видению задач обеспечения ФБ АПК можно поставить в соответствие СА «Пределы роста» (рисунок 7).

С течение времени действия по устранению дефектов приводят к уменьшению их количества (усиливающий контур ФБ). Вместе с тем, с течением времени стоимость устранения дефектов возрастает. Это препятствует выделению ресурсов на другие виды деятельности, связанные с обеспечением ФБ, например, за счёт создания и совершенствования барьеров, препятствующих возникновению дефектов на разных стадиях ЖЦ АПК.

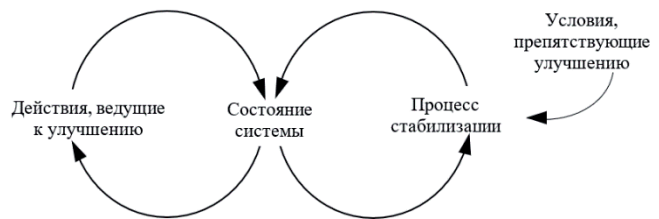


Рисунок 7 - Структурная модель, соответствующая СА «Пределы роста»

Пример 3. Анализ сбалансированности параметров проекта

Описание задачи. Уровень потребительских свойств программных продуктов, включая ФБ, определяется подходами к решению комплекса задач, связанных с реализацией программных проектов [36]. Основа успешности программного проекта – сбалансированность требований к потребительским свойствам предполагаемого продукта и объёмов ресурсов, выделяемых на реализацию проекта [31]. В [37] показано, что результатом несбалансированности ресурсов в первую очередь является недостаточная глубина испытаний.

Постановка задачи. Дано: динамика относительного показателя $D(t)$, характеризующего выявленные дефекты (например, в виде количества дефектов на 1000 строк кода); динамика затрат ресурсов на выявление дефектов $S(t)$; ограничение на бюджет проекта S^* ; ограничение на допустимое число латентных дефектов D^* .

Требуется: оценить сбалансированность показателей S^* и D^* .

Основу решения задачи составляют зависимости $D(t)$, $S(t)$ (рисунок 8).

В примере постулированы зависимости $D(t)$, $S(t)$, описанные в примере 2. Заданному значению S^* соответствует t^* , представленное на рисунке 8. Этому значению t^* соответствует показатель $D(t^*)$, указанный на оси ординат.

Различие $\omega = |D^* - D(t^*)|$ есть характеристика несбалансированности параметров проекта по показателю количества латентных дефектов.

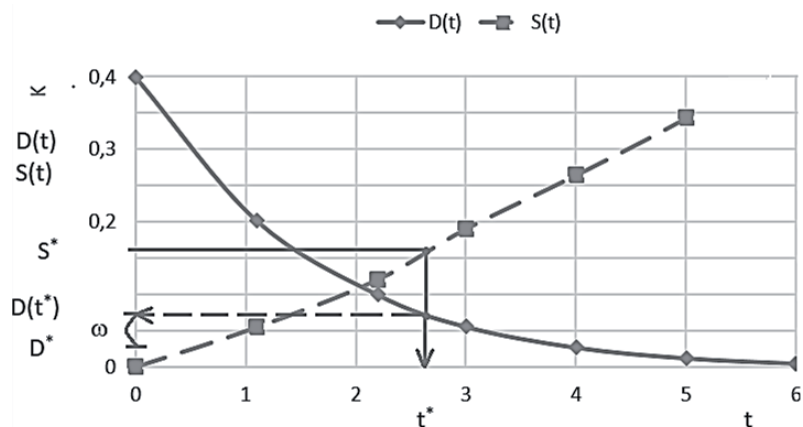


Рисунок 8 - Несбалансированность параметров проекта по показателю количества дефектов

Аналогичным образом можно оценить величину $S(t^*)$, нужную для обеспечения требуемого значения D^* (рисунок 9).

Следствие несбалансированности нефункциональных требований и объёма выделяемых ресурсов является превышение ограничений на бюджет и время реализации проекта либо несоответствие фактических значений потребительских характеристик программных продуктов желаемым (см. также [24]).

Величина $\gamma = |S^* - S(t^*)|$ есть характеристика несбалансированности по показателю бюджета (рисунок 9). Интегральной характеристикой несбалансированности может служить ин-

декс $I = \sum \alpha_i I_i$, где α_i – весовые коэффициенты, характеризующие значимость i -го показателя ФБ. В качестве индивидуальных индексов приняты $I_1 = \omega / D^*$; $I_2 = \gamma / S^*$.

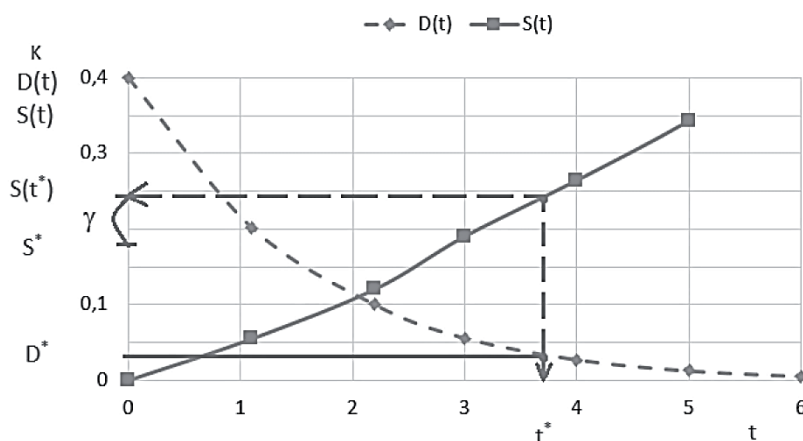


Рисунок 9 - Несбалансированность проекта по показателю бюджета



Рисунок 10 – Структурная модель, соответствующая СА «Эрозия целей»

Описанной задаче может быть поставлен в соответствие СА «Эрозия целей» (рисунок 10). В рамках этого СА несбалансированность желаемых и фактических свойств устраняется за счёт снижения требований к свойствам продукта либо за счёт увеличения объёмов ресурсов.

Заключение

В статье обоснована возможность использования СА как методической основы построения моделей ситуаций, возникающих при решении задач обеспечения ФБ АПК в условиях ограничений на ресурсы.

Дано определение информационной сущности ситуаций, позволяющее создать основу для принятия решений о целесообразности внесения изменений в структуру системы управления СТС.

Предложены концептуальные основы многоаспектного моделирования ситуаций, возникающих при обеспечении ФБ АПК, сочетающие динамические модели параметров, характеризующих ФБ, и структурные модели ситуаций.

Выделены базовые этапы построения системы структурных и динамических моделей ситуаций, возникающих при обеспечении ФБ АПК на разных стадиях их жизненного цикла.

Список источников

- [1] *ESA PSS-05-10. Guide to software verification and validation* Prepared by: ESA Board for Software Standardization and Control (BSSC). ESA PSS-05-10 Issue 1 Revision 1. March 1995. 117 p.
- [2] **Виттих В.А.** Введение в теорию интерсубъективного управления. Самара: Самарский научный центр РАН, 2013. 64 с.
- [3] **Лушаев В.В.** Функциональная безопасность программных средств. Москва: СИНТЕГ, 2004. 348 с.
- [4] **Нагибин С.Я., Пальчун Б.П., Ухлинов Л.М.** Технологическая безопасность программирования – новая проблема в области создания информационных систем // Информационное общество. 1995. Т.6. С.45-49.
- [5] *Сопряженное проектирование встраиваемых систем (Hardware/Software Co-Design)* / С.В. Быковский [и др.]. Санкт-Петербург: Университет ИТМО, 2016. 108 с.
- [6] *A Driving Assistance System with Hardware Acceleration.* – University of Gothenburg, Sweden, 2016.

- [7] **Бородакий Ю.В., Юсупов Р.М., Пальчун Б.П.** Проблема имитационного моделирования дефектоскопических свойств компьютерной инфосферы // Труды третьей Всероссийской научно-практической конференции «Имитационное моделирование. Теория и практика». Санкт-Петербург, 2007. С.87-32.
- [8] **Рот А.** Внедрение и развитие Индустрии 4.0. Основы, моделирование и примеры из практики. Москва: Техносфера, 2017. 294 с.
- [9] **Кириллов Н.П.** Концептуальная модель объекта ситуационного управления функциональным состоянием технических систем // Искусственный интеллект и принятие решений. 2012. Т.4. С.61-75.
- [10] **Мостовой А.Я.** Управление сложными техническими системами: конструирование программного обеспечения спутников ДЗЗ. Москва: Техносфера, 2016. 352 с.
- [11] **Куликов С.С.** Тестирование программного обеспечения. Базовый курс. Минск: Четыре четверти, 2017. 312 с.
- [12] **Марков А.С.** Модели оценки и планирования испытаний программных средств по требованиям безопасности информации // Вестник МГТУ им. Н.Э. Баумана. Сер. «Приборостроение». Специальный выпуск «Технические средства и системы защиты информации». 2011. С.30-103.
- [13] **Cortellessa V., Grassi V.** A modeling approach to analyze the impact of error propagation on reliability of component-based systems // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2007. P.140-156.
- [14] **Ritu Soni, Ashpinder Preet.** Cognitive approach to root cause analysis for improving quality of life: a case study for IT Industry. // International journal of informative and futuristic research (Online). Vol. 1. Issue 1, August - September 2013, 8 p.
- [15] **Sunday E.** Extension and Modification of Anticipatory Failure Determination Approach Based on I-TRIZ. University of Stavanger, Department of Mechanical and Structural Engineering, June 2014, 106 p.
- [16] **Renan Favarão Da Silva, Marco Aurélio De Carvalho.** Anticipatory Failure Determination (AFD) for product reliability analysis: A comparison between AFD and Failure Mode and Effects Analysis (FMEA) for identifying potential failure modes, Federal Technological University of Paraná (UTFPR), Curitiba, Brazil, January 2019, 24p. (DOI: 10.1007/978-3-319-78075-7_12)
- [17] Error Propagation Analysis of Software Architecture Specifications / D. Nassar [et al.] // Communication. 2006. Vol.1. P.496-501.
- [18] **Lee P.A., Anderson T.** Fault tolerance, principles and practice // Springer Sci. Bus. Media. 2012. Vol.3. P.320.
- [19] **Verzola I., Lagny A.E., Biswas J.** A Predictive Approach to Failure Estimation and Identification for Space Systems Operations // Proc. 13th international conference on space operations. – Pasadena, California, USA, 2014.
- [20] **Khater H.A., Mohamed A.B., Kamel S.M.** A Proposed Technique for Software Development Risks Identification by using FTA Model // World Acad. Sci. Eng. Technol. Int. J. Comput. Inf. Eng. 2013. Vol.7. P.105-111.
- [21] **Pentti H., Atte H.** Failure Mode and Effects Analysis of Software-Based Automation Systems // STUK-YTO-TR 190. 2002. P.36.
- [22] **Zhu Y.M.** Failure-modes-based software reading. 2017. P.7-15.
- [23] **Тимофеев А.Н.** Почему падают ИТ-проекты? // Практика проектирования систем, 2017. С.2-11.
- [24] CHAOS Report. The Standish Group International, Inc., 2018, 68 p. <https://www.standishgroup.com/news/37>.
- [25] **Виттих В.А.** Неоднородный актор и повседневность как ключевые понятия эвергетики. Самара: Самарский научный центр РАН, 2014. 12 с.
- [26] **Райков А.Н.** Конвергентное управление и поддержка решений. Москва: Издательство ИКАР, 2009. 245 с.
- [27] **Райков А.Н.** Сетевая экспертная поддержка решений // Управление большими системами. 2010. Т.30.1. С.758-772.
- [28] **Сенге П.М.** Пятая дисциплина. М.: Олимп-Бизнес; 2003. 408 с.
- [29] **Braun W.** The System Archetypes by William Braun. 2002. 25 p.
- [30] **Meadows D.H.** Thinking in Systems. Chelsea Green Publishing. 2008, 240 p.
- [31] **Hastie S., Wojewoda S.** Standish Group 2015 Chaos Report - Q&A with Jennifer Lynch. OCT 04, 2015. <https://www.infoq.com/articles/standish-chaos-2015/>
- [32] **Вентцель Е.С.** Теория вероятностей. Учебник. -12-е издание. М.: Юстиция; 2018. 658 с.
- [33] **Майерс Г.Дж.** Надежность программного обеспечения. М: Издательство Мир, 1980. 359 с.
- [34] **Бэзьюли Ф.** Управление проектом. М: Гранд-Фаир; 2002. 208 с.
- [35] **Липаев В.В.** Надежность и функциональная безопасность комплексов программ реального времени. М: Институт системного программирования РАН; 2013. 207 с.
- [36] **ESA PSS 05-11.** Guide to software quality assurance.
- [37] **Макконнелл С.** Сколько стоит программный проект. Санкт-Петербург, 2007. 297 с.

Сведения об авторах



Гвоздев Владимир Ефимович, 1956 г. рождения. Окончил Уфимский авиационный институт им. Орджоникидзе в 1978 г., д.т.н. (2000). Профессор кафедры технической кибернетики Уфимского университета науки и технологий. В списке научных трудов более 370 работ в области прикладного статистического анализа, информационной поддержки управления программными системами, информационной поддержки управления состоянием территориальных систем. AuthorID (РИНЦ): 174520. ORCID 0000-0002-1481-0982. Author ID (Scopus): 7101700484. wega55@mail.ru.



Бежаева Оксана Яковлевна, 1977 г. рождения. Окончила Уфимский государственный авиационный технический университет в 2000 г., к.т.н. (2004). Заведующая кафедрой технической кибернетики Уфимского университета науки и технологий. В списке научных трудов более 100 работ в области разработки моделей и программного обеспечения сложных систем, информационной поддержки управления программными проектами и системами. AuthorID (РИНЦ): 271220. ORCIDID 0000-0002-3373-7266. Author ID (Scopus): 57216845244. obezhava@gmail.com.



Сафина Гульнур Радиковна, 1997 г. рождения. Окончила Уфимский государственный авиационный технический университет в 2022 г., направление «Информатика и вычислительная техника». В списке научных трудов 10 работ в области программно-аппаратных комплексов технических систем. lafleur300997@gmail.com.

Поступила в редакцию 17.11.2022, после рецензирования 9.02.2023. Принята к публикации 1.03.2023.



Scientific article

DOI: 10.18287/2223-9537-2023-13-1-125-138

Multi-aspect modeling of situation in the functional safety control tasks of hardware and software complexes

© 2023, V.E. Gvozdev, O.Ya. Bezhaeva ✉, G.R. Safina

Ufa University of Science and Technology, Ufa, Russia

Abstract

Functional safety is a latent characteristic of hardware and software complexes (HSC) and is objectively characterized by consumer satisfaction with the HSC behavior. The paper considers an approach to solving project management problems based on the dynamic characteristics analysis of functional safety of conflict situations in the project management system. To analyze situations that arise when managing HSC projects, it is possible to use models known as system archetypes. The information essence of situations and the foundations of multi-aspect modeling are considered. The methodological basis of the research is a combination of dynamic models of parameters that characterize functional safety and structural models, corresponding to conflict situations that arise when ensuring the required level of functional safety. Examples of multi-aspect modeling of situations where manifestations of latent defects act as events are considered. As a result of the research, the information essences of situations are determined, the conceptual foundations of multi-aspect modeling of situations that arise when managing the HSC functional safety are proposed, and the basic stages of building a system of structural and dynamic models of situations at different stages of the HSC life cycle are identified. The results obtained can be used to make decisions about the advisability of making changes to the system structure for ensuring the functional safety of hardware and software complexes..

Key words: functional safety, hardware and software complex, network-centric management, system archetype, structural models.

Citation: Gvozdev VE, Bezhaeva OYa, Safina GR. Multi-aspect modeling of situation in the functional safety control tasks of hardware and software complexes [In Russian]. *Ontology of designing*. 2023; 13(1): 125-138. DOI: 10.18287/2223-9537-2023-13-1-125-138.

Conflict of interest: The authors declare no conflict of interest.

List of figures and tables

- Figure 1 - Structural model of the point of view on the situation
 Figure 2 - Dynamic characteristics of the integral indicator of functional safety
 Figure 3 - The system archetype of "Balancing with delay"
 Figure 4 - The probability of failure of hardware and software complexes in the absence of a systematic approach to reducing the probability of latent defects
 Figure 5 - Dynamic models $D(t)$ and $S(t)$
 Figure 6 - Defect elimination efficiency characteristic
 Figure 7 - Structural model corresponding to the Limits of Growth archetype
 Figure 8 - Imbalance of project parameters in terms of the defects amount
 Figure 9 - Imbalance of the project in terms of budget
 Figure 10 - Structural model corresponding to the Eroding goals system archetype
 Table 1 - The values of $D(\tau_i)$, $S(\tau_i)$, $E(\tau_i)$ corresponding to different timestamps of τ_i .

References

- [1] *ESA PSS-05-10*. Guide to software verification and validation Prepared by: ESA Board for Software Standardization and Control (BSSC). ESA PSS-05-10 Issue 1 Revision 1. March 1995. 117 p.
- [2] *Vittich VA*. Introduction to the theory of intersubjective management. [In Russian]. Samara: Samara Scientific Center of the Russian Academy of Sciences, 2013; 64 p.
- [3] *Lipaev VV*. Functional safety of software tools. [In Russian]. Moscow: SINTEG, 2004; 348 p.
- [4] *Nagibin SYa, Palchun BP, Ukhlinov LM*. Technological security of programming - a new problem in the field of information systems. [In Russian]. *Information society*, 1995; 6: 45-49.
- [5] *Bykovsky SV*. Conjugated design of embedded systems (Hardware/Software Co-Design) [In Russian]. Saint Petersburg: ITMO Research Institute, 2011; 125 p.
- [6] A Driving Assistance System with Hardware Acceleration. – University of Gothenburg, Sweden, 2016.
- [7] *Borodaki YuV, Yusupov RM, Palchun BP*. The problem of simulation modeling of flaw detection properties of computer infosphere. [In Russian]. Proceedings of the third All-Russian scientific and practical conference "Simulation. Theory and practice". St. Petersburg, 2007; P.87-32.
- [8] *Roth A*. Implementation and development of Industry 4.0. Fundamentals, modeling and examples from practice. Moscow: Technosphere, 2017; 294 p.
- [9] *Kirillov NP*. Conceptual model of the object of situational control of the functional state of technical systems. [In Russian]. Artificial intelligence and decision making. 2012; 4: 61-75.
- [10] *Mostovoy AYa*. Complex technical systems management: design of remote sensing satellite software. [In Russian]. Moscow: Technosphere. 2016; 352 p.
- [11] *Kulikov SS*. Software testing. Basic course. [In Russian]. Minsk: Four quarters, 2017; 312 p.
- [12] *Markov AS*. Models of evaluation and planning of software testing according to information security requirements. [In Russian]. Bulletin Bulletin of Bauman Moscow State Technical University. Special issue "Technical means and information security systems". 2011; P.30-103.
- [13] *Cortellessa V, Grassi V*. A modeling approach to analyze the impact of error propagation on reliability of component-based systems // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2007. P.140-156.
- [14] *Soni R, Preet A*. Cognitive approach to root cause analysis for improving quality of life: a case study for IT Industry. // International journal of informative and futuristic research (Online). 2013; 1(1): 8.
- [15] *Sunday E*. Extension and Modification of Anticipatory Failure Determination Approach Based on I-TRIZ. University of Stavanger, Department of Mechanical and Structural Engineering, June 2014, 106 p.
- [16] *Renan Favarão Da Silva, Marco Aurélio De Carvalho*. Anticipatory Failure Determination (AFD) for product reliability analysis: A comparison between AFD and Failure Mode and Effects Analysis (FMEA) for identifying potential failure modes, Federal Technological University of Paraná (UTFPR), Curitiba, Brazil, January 2019, 24p. DOI: 10.1007/978-3-319-78075-7_12.
- [17] *Nassar D*. Error Propagation Analysis of Software Architecture Specifications. *Communication*. 2006; 1: 496-501.
- [18] *Lee PA*. Fault tolerance, principles and practice / P.A. Lee, T. Anderson // Springer Sci. Bus. Media. 2012; Vol. 3, 320 p.

- [19] **Verzola IA** Predictive Approach to Failure Estimation and Identification for Space Systems Operations / I. Verzola, A.E. Lagny, J. Biswas // Proc. 13th international conference on space operations. – Pasadena, California, USA, 2014.
 - [20] **Khater HA**. A Proposed Technique for Software Development Risks Identification by using FTA Model / H.A. Khater, A.B. Mohamed, S.M. Kamel // World Acad. Sci. Eng. Technol. Int. J. Comput. Inf. Eng. 2013; Vol. 7, P. 105-111.
 - [21] **Pentti H, Atte H**. Failure Mode and Effects Analysis of Software-Based Automation Systems. STUK-YTO-TR 190. 2002; 36p.
 - [22] **Zhu YM**. Failure-modes-based software reading. 2017; P.7-15.
 - [23] **Timofeev AN**. Why IT projects are falling? [In Russian]. Practice of system design, 2017; P. 2-11.
 - [24] CHAOS Report. The Standish Group International, Inc., 2018, 68 p. - Available: <https://www.standishgroup.com/news/37>
 - [25] **Vittich VA**. Heterogeneous actor and everyday life as key concepts of evergetics. [In Russian]. Samara: Samara Scientific Center of the Russian Academy of Sciences, 2014; 12p.
 - [26] **Raikov AN**. Convergent management and decision support. [In Russian]. Moscow: ICAR Publishing House, 2009; 245 p.
 - [27] **Raikov AN**. Network expert support of solutions [In Russian]. Management of large systems. 2010; Vol. 30.1, P.758-772p.
 - [28] **Senge PM**. Fifth discipline. [In Russian]. Moscow: Olympus-Business. 2003; 408 p.
 - [29] **Braun W**. The System Archetypes by Wiiliam Braun. 2002; 25p.
 - [30] **Meadows DH**. Thinking in Systems. Chelsea Green Publishing. 2008; 240 p.
 - [31] **Hastie S, Wojewoda S**. Standish Group 2015 Chaos Report - Q&A with Jennifer Lynch. OCT 04, 2015. <https://www.infoq.com/articles/standish-chaos-2015/>
 - [32] **Ventzel ES**. Theory of Probability. [In Russian]. Textbook. -12th edition. Moscow: Justice, 2018; 658 p.
 - [33] **Myers GJ**. Software reliability. [In Russian]. Moscow: Mir Publishing House, 1980; 359 p.
 - [34] **Baguli F**. Project Management. [In Russian]. Moscow: Grand-Fair, 2002; 208 p.
 - [35] **Lipaev VV**. Reliability and functional safety of real-time software complexes. [In Russian]. Moscow: Institute of System Programming of the Russian Academy of Sciences, 2013; 207 p.
 - [36] **ESA PSS 05-11**. Guide to software quality assurance.
 - [37] **McConnell S**. How much does a software project cost? [In Russian]. Peter Publisher; 2007. 296 p.
-

About the authors

Vladimir Efimovich Gvozdev (b. 1956) graduated from the Ufa Aviation Institute in 1978, D. Sc. Eng. (2000). Professor of the Department of Technical Cybernetics at the Ufa University of Science and Technology (UUST). He is a co-author of about 370 scientific articles and abstracts in the field of applied statistical analysis, information support for managing software systems, information support for managing the state of territorial systems. ORCID 0000-0003-4874-0895. AuthorID (RSCI): 174520. Author ID (Scopus): 7101700484. wega55@mail.ru.

Oksana Yakovlevna Bezhaeva (b. 1977) graduated from the Ufa State Aviation Technical University in 2000, PhD. (2004). Head of the Department of Technical Cybernetics at the Ufa University of Science and Technology (UUST). She is a co-author of about 100 scientific articles and abstracts in the field of development of models and software for complex systems, information support for managing software projects and systems. ORCID 0000-0002-3373-7266. AuthorID (RSCI): 271220. Author ID (Scopus): 57216845244. obezhaeva@gmail.com. ✉

Gulnur Radikovna Safina (b. 1997) graduated from the Ufa State Aviation Technical University in 2022 in the direction of Informatics and Computer Engineering. She is a co-author of 10 scientific articles and abstracts in the field of software and hardware complexes of technical systems. lafleur300997@gmail.com.

Received November 17, 2022. Revised February 9, 2023. Accepted March 1, 2023.



Онтологический нечётко-возможностный подход к созданию модели управления печью пиролиза

© 2023, И.Т. Кимяев

ООО «Норникель Спутник», Москва, Россия

Аннотация

Для объектов хозяйственной деятельности рассматриваются разработка и применение прикладного подхода к созданию подсистемы принятия управленческих решений, устраняющих возможные семантические разрывы между смежными слоями производственно-технологической иерархии при описании разноуровневых информационных потоков. Созданная с помощью предложенного подхода подсистема принятия решения способна обрабатывать разноуровневые потоки данных в комплексной структуре поддержания указанных объектов в жизнеспособном состоянии. Предложенный подход базируется на вертикально-интегрированных онтологических моделях, созданных с использованием стандартных программных средств. Выявленные семантические связи между концептами внутри формализованных разноуровневых онтологических моделей восстанавливаются с помощью явных и неявных знаний экспертов о ведении производственных операций. В качестве методологической основы извлечения и формализации знаний экспертов в виде нечёткого аналитического полинома выбран нечётко-возможностный подход. Применение предложенного подхода показано на примере управления процессом пиролиза высокомолекулярного углеводородного сырья в специализированных печах.

Ключевые слова: принятие решений, интегрированная система управления, онтологическое моделирование, нечётко-возможностный подход, печь пиролиза, административно-управленческая группа, семантика отношений.

Цитирование: Кимяев И.Т. Онтологический нечётко-возможностный подход к созданию модели управления печью пиролиза // *Онтология проектирования*. 2023. Т.13, №1(47). С.139-149. DOI:10.18287/2223-9537-2023-13-1-139-149.

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Введение

Технологическая компания является сложным объектом хозяйственной деятельности (ОХД), включающим в себя множество связанных вещественными и энергетическими потоками объектов управления (ОУ) типа технологических линий и агрегатов, сгруппированных в физические, территориальные или функционально-логические домены [1].

Эффективность работы и жизнеспособность каждого отдельного ОУ и ОХД в целом невозможна без наличия многоуровневой интегрированной информационно-управляющей системы (ИУС), позволяющей полностью или частично автоматизировать функции принятия управленческих решений.

Согласно принципу необходимого разнообразия, при создании интегрированной ИУС уровни её функциональной и структурной сложности следует привести в соответствие с уровнями функциональной и структурной сложности ОХД [2].

В современных комплексах ОХД ↔ ИУС существует виртуальная «стена» [3] между различными уровнями управленческой ОТ/ИТ-иерархии (где ОТ - операционные и ИТ - информационные технологии в нотации *Purdue/ISA-95* [4], см. рисунок 1). Возникновение «стены» обусловлено традициями и подходами к созданию программно-аппаратных ком-

плексов рассматриваемых предметных областей (ПрО), проектированию контрольно-управленческих процедур и сетевых протоколов, обеспечению мер информационной и физической безопасности и т.д.

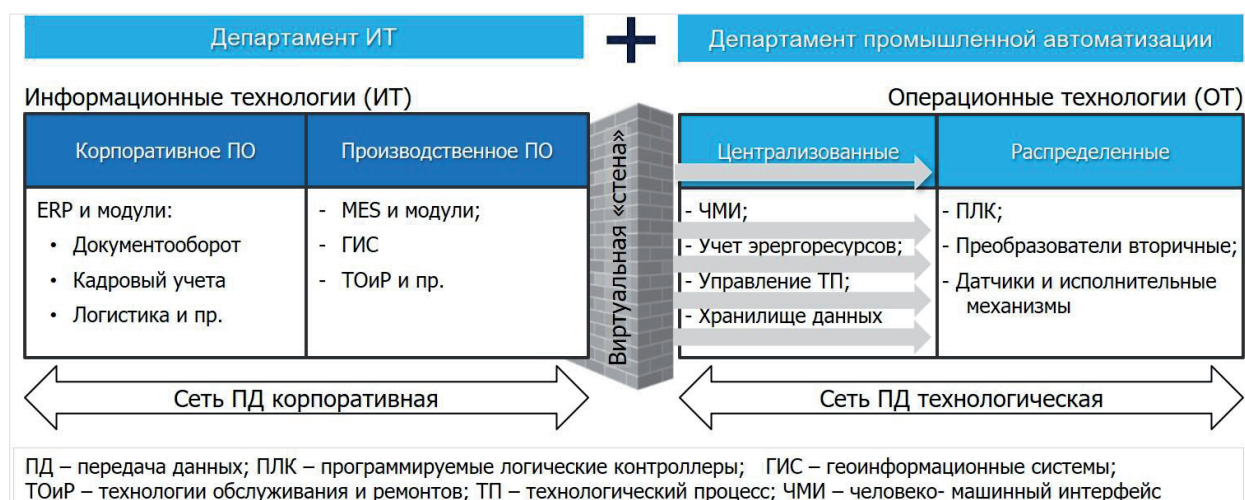


Рисунок 1 - Виртуальная функционально-технологическая «стена» между уровнями управленческой ОТ/ИТ-иерархии

В настоящее время происходит взаимопроникновение («диффузия») ИТ ↔ ОТ, которая проявляется, с одной стороны, в применении ИТ-решений для проектирования архитектуры систем контроля и управления технологическим оборудованием, с другой – востребованностью в информационных системах производственно–корпоративного уровня данных, полученных от средств измерения в «сыром» или минимально обработанном виде.

Отмеченная «диффузия» практически не коснулась актуальных для каждого управленческого уровня подходов и методик синтеза управляющих процедур и алгоритмов для их последующей автоматизации, т.к. они базируются на различных онтологических, математических, терминологических, логико-лингвистических и пр. секторах знаний.

Исторически сформировавшаяся «серая зона» в области обработки информационных потоков на стыке разнородных слоев управленческой иерархий затрудняет выработку высокоэффективных управленческих решений, гармонизированных со всеми участниками производственной деятельности.

Сложившаяся ситуация обуславливает широкое вовлечение в обработку разноуровневой информации и принятие на её базе управленческих решений специалистов экспертного уровня: лиц, вырабатывающих решение (ЛВР); лиц, обосновывающих решение (ЛОР); лиц, принимающих решения (ЛПР). В совокупности ЛВР, ЛОР, ЛПР представляют собой административно-управленческую группу (АУГ) из специалистов различного профиля: технологи, энергетики, механики, экономисты и др., целенаправленная деятельность которой обеспечивает эффективность работы отдельных производственных компонентов и в целом ОХД.

Особую сложность для АУГ при выработке единого управленческого решения, наиболее эффективного в текущей ситуации и в актуальный момент времени, являются различные для каждого участника группы представления об эффективности [2], а также граничные условия их применимости. В частности, для каждого операционного, производственного и/или корпоративного ЛВР, ЛОР и ЛПР периодичность опроса и состав контролируемых параметров (факторов) и возмущений, частота формирования и доступ к управляющим воздействиям, а также формализованные или неявные критерии эффективности, существенно различны.

Таким образом, формирование подходов и методологической базы для разработки эффективных моделей, способных обеспечивать обработку разноуровневых информационных

потоков и решение задач управления производством (с учётом множества критериев и показателей), является актуальной научно-прикладной проблемой.

1 Процесс пиролиза как объект многофакторного и многоуровневого управления

В качестве базового ОХД рассмотрен один из важных технологических этапов в нефтехимической отрасли, а именно, производство олефинов, которое является главной технологической цепочкой и источником основных «строительных блоков» - этилена, пропилена, бутена, бутадиена и бензола [5]. Наиболее распространенной технологией для олефиновых цепочек является высокотемпературный крекинг высокомолекулярного углеводородного сырья - пиролиз, который происходит в специализированных технологических агрегатах - печах пиролиза (ПП) с последующим выделением путём очистки (фракционированием) из пирогаза (ПГ) товарного этилена и сопутствующих фракций. На современных олефиновых производствах количество ПП может достигать восьми и более одновременно работающих единиц с общей производительностью в миллионы тонн в год.

Устойчивость и качество работы пиролизного комплекса определяют устойчивость, качество и экономические показатели работы всех последующих производственных переделов. При этом каждая отдельная ПП может работать на своём, отличном от соседнего агрегата, сырье (широкая фракция лёгких углеводородов, бензин, газойль, нефтя и пр.) с вытекающими отсюда существенными различиями в выборе критериев и режимов управления.

На рисунке 2 представлена типовая ПП, внутри которой находятся четыре змеевика А, В, С, D. Для ведения полнофункционального, иерархически и горизонтально взаимоувязанного, управления пиролизным комплексом используют наборы контролируемых (физически измеряемых) и управляемых (реализуемых) технологических переменных¹: CV и MV (Таблица 1).

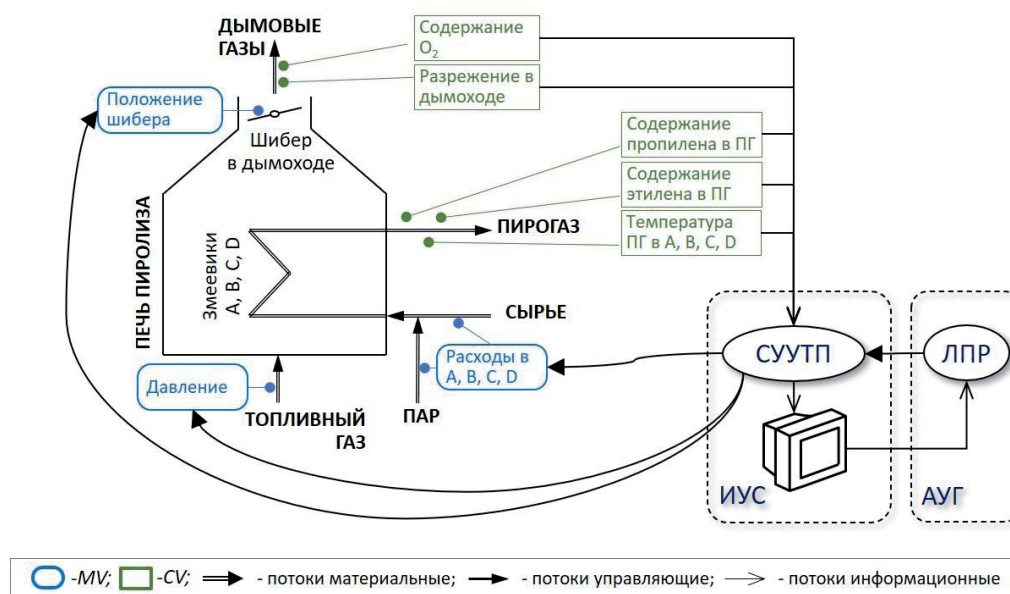


Рисунок 2 - ПП как ОУ со змеевиками А, В, С, D и типовым набором физических измеряемых CV и реализуемых MV

¹ Предложенный набор переменных (CV - controlled variables, MV - manipulated variables), приведённый для общего представления о сложности и многофакторности процесса, является типовым практически для любой ПП, при этом не является исчерпывающим.

Таблица 1 - Контролируемые и управляемые технологические переменные для управления ПП

Описание	Переменная ПП	Диапазон измерений		Целевое значение	Ед. изм.
		min	max		
Режим работы ПП	CV1	2	3		ед
Содержание углеродной фракции C ₄ , C ₅ в сырье	CV2.45	35	45		%
Содержание углеродной фракции C ₆₊ в сырье	CV2.6	11	30		%
Температура ПГ целевая	CV3	400	900	851...854	°C
Содержание углеродной фракции C ₂ в ПГ целевое	CV4.2	0	100	>33	%масс
Содержание углеродной фракции C ₃ в ПГ целевое	CV4.3	0	100	<15	%масс
Содержание углеродной фракции C ₂ в ПГ текущее	CV5.2	0	100	>33	%масс
Содержание углеродной фракции C ₃ в ПГ текущее	CV5.3	0	100	<15	%масс
Температура ПГ в А, В, С, D	CV6.A...D	400	900	851...854	°C
Отклонение температуры ПГ в А, В, С, D от средней	CV7.A...D	-4	4	2	°C
Отклонение температуры ПГ в А, В, С, D от целевой	CV8.A...D	-4	4	2	°C
Градиент температуры ПГ в А, В, С, D	CV9.A...D	-4	4	2	°C/мин
Температура ПГ средняя	CV10	400	900	851...855	°C
Отклонение средней температуры ПГ от целевой	CV11	-4	4	2	°C
Градиент температуры ПГ средней текущий	CV12	-4	4	2	°C/мин
Продолжительность эксплуатации ПП	CV13	0,2	1		год
Качество готовой продукции ²	CV14	0,2	1		ед отн
Общая операционная прибыль ²	CV15	0,2	1		ед отн
Энергозатраты ²	CV16	0,2	1		ед отн
Производительность ²	CV17	0,2	1		ед отн
Расход сырья в А, В, С, D	MV1.A...D	0	5000	4600	кг/ч
Давление ТГ перед горелками	MV2	0	5	0,8...1,2	кг/см ²

На рисунке 3 представлена онтологическая модель (ОМ) многоуровневого процесса принятия управленческих решений для ПП. За оперативное (уровня *L2*) управление столь сложными ОУ отвечает, как правило, специфический класс автоматических систем управления технологическими процессами, называемых «усовершенствованными» (СУУТП) [6]. Системы данного класса в составе интегрированной ИУС решают задачу рутинного многофакторного поддержания основных режимных показателей ПП в установленных ЛПР целевых диапазонах.

Достижение перечисленных выше целей возможно путём последовательного / циклического с синхронизацией по времени решения следующих иерархических задач управления:

- поддержание максимальной (или минимальной, в зависимости от установленного к выработке целевого продукта) допустимой температуры ПГ с регламентированными отклонениями;
- многопараметрическое управление подачей сырья и ТГ:
 - а. общая («грубая») стабилизация средней температуры ПГ в заданном (целевом) диапазоне;

² Для данных CV диапазоны измерений приведены в относительных единицах, поскольку для реального олефинового производственного комплекса могут быть представлены различными метриками и колебаться в широких диапазонах.

- б. формирование равномерного температурного профиля ПГ по змеевикам с минимальным отклонением от целевого значения;
 - с. корректирующая («тонкая») стабилизация температуры ПГ по потокам (змеевикам) с минимальным отклонением от целевого значения;
- подача пара разбавления по заданному режимному соотношению «пар – сырье»;
- минимизация содержания кислорода O_2 в дымовых газах (в пределах нормированных значений по разрежению в камере).

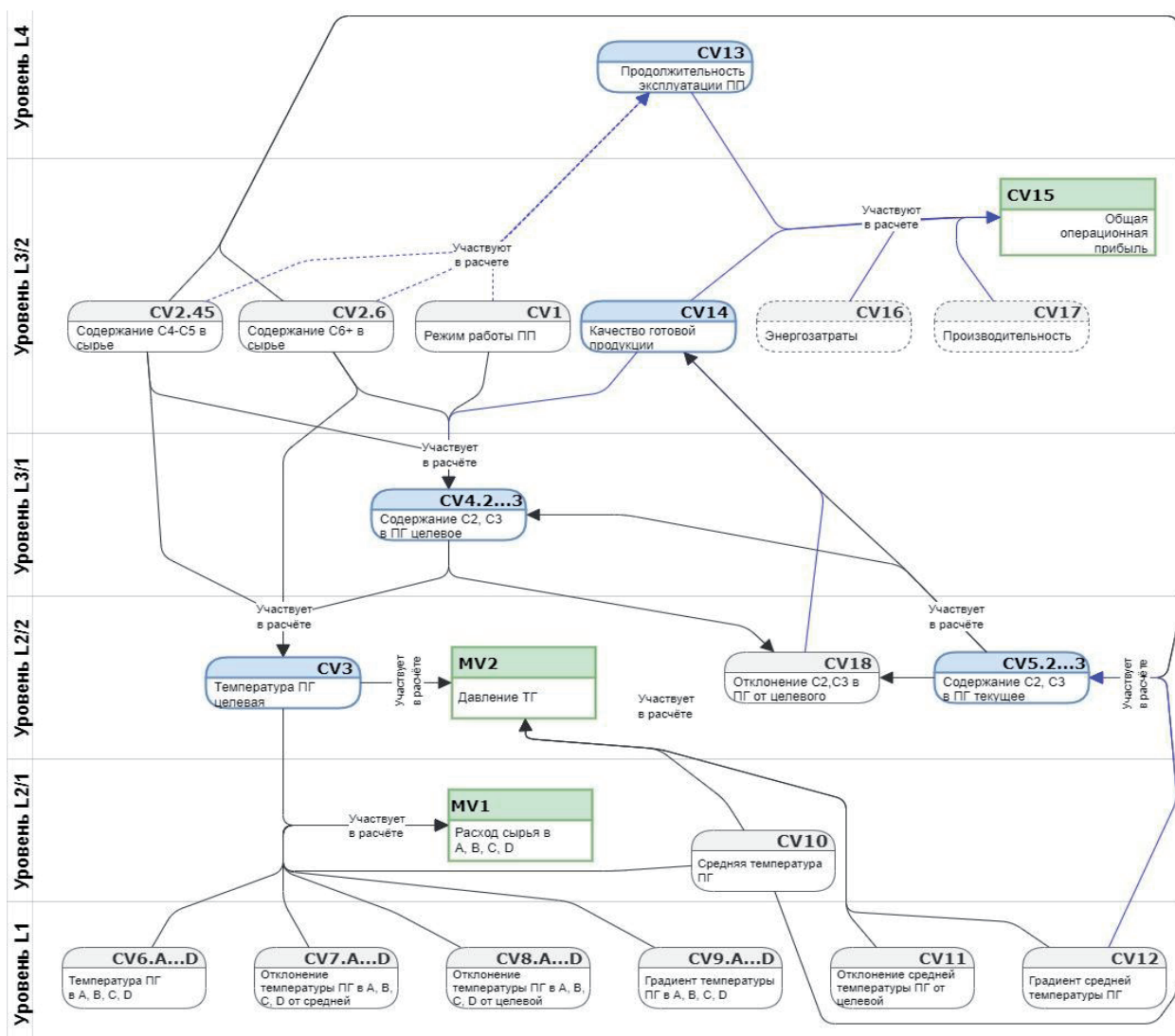


Рисунок 3 - Онтологическая модель многоуровневого процесса принятия управленческих решений для ПП

Создание многофакторных и многокритериальных подсистем управления на производственном уровне $L3$ (например, системы глобальной динамической стабилизации – СГДО) и выше для олефинового производства сопряжено с существенными капитальными, операционными и временными издержками, с множеством организационно-технологических ограничений, совокупность которых для большинства реальных производственных компаний является препятствием к их внедрению. В этой связи поддержание сложного многокомпонентного ОХД, как комплекса по производству олефинов на базе ПП, в эффективном и жизнеспособном состоянии путём решения производственно-экономических задач, может обеспе-

чить только специализированная АУГ в составе различных предметно-ориентированных экспертов: химики-технологи, экономисты, диспетчеры, логисты и др.

Типовыми задачами управленческой иерархии производственного уровня являются³: управление (прогнозирование) процессом закоксовывания змеевиков ПП [7], например, с целью единовременного вывода всех ПП в плановый ремонт; обеспечение максимальной операционной прибыли олефинового производства с учётом количества и стоимости потребляемых материальных и энергетических ресурсов и др.

Достижение представленных производственно-экономических целей совокупными усилиями экспертов в составе АУГ требует от них принятия решений в своём секторе знаний (онтологий) и ответственности при безусловном учёте информации, поступающей от смежных производственно-технологических и др. служб о состоянии ПП как ОУ. Например, решение о суточном повышении производительности ПП не может быть принято без учёта информации о работоспособности механического оборудования, количестве доступного сырья, уровне закоксованности змеевиков и пр.

Для многокритериального выбора управленческих решений в приведённых задачах и ограничениях необходимо обработать множество разноуровневых количественных и качественных информационных потоков с различной дискретизацией, со слабо формализованными внутренними связями и применить соответствующего уровня сложности методики и подходы к моделированию.

Для решения этих задач предложен подход к логико-математическому синтезу многофакторных моделей, способных отобразить разноуровневые логико-понятийные подходы, используемые каждым членом АУГ на базе реальных технологических данных и неформализованных знаний.

2 Синтез многоуровневой ИУС на базе онтологий и нечётко-возможностного подхода

Одним из эффективных способов идентификации знаний об устройстве сложной ПрО и механизмах функционирования компонентов в её составе является онтологическое моделирование [8, 9] на основе формальных методик дескриптивной логики (ДЛ). Методика [9] нацелена на представление, в данном случае - реальных производственных сущностей (классов сущностей) - в иерархическом виде как абстрактных функциональных эквивалентов (концептов). Для комплекса ИУС ↔ ОХД методика ОМ/ДЛ позволяет средствами формальной семантики эффективно описать явные и неявные, конструктивно и технологически обусловленные, взаимосвязи (отношения, предикаты) между компонентами ОХД, вещественными, энергетическими и информационными потоками и др. В рамках описания ПрО ОМ даёт возможность определить семантику и способы воздействия на компоненты ОХД членов АУГ и/или существующих систем класса ОТ/ИТ, что является идентификацией соответствующих информационных потоков.

Широко распространённым программно-техническим инструментом решения прикладных ОМ-задач для различных секторов знаний является семейство языков *OWL (Web Ontology Language)* стандарта *W3C - World Wide Web Consortium*), которые представляет комплекс открытых платформ *Open Web Platform* [10].

Рисунок 3 представляет ОМ принятия управленческих решений при ведении пиролиза, которая разработана с использованием семантических элементов *OWL*. Для формирования данной ОМ были использованы данные об аппаратно-технологическом, информационном и

³ Приведённый перечень задач управления не являются исчерпывающим, и для конкретного технологического комплекса может быть существенно скорректирован.

Для построения достоверной модели поведения исследуемых ОХД ↔ ИУС ↔ АУГ формализация режимов и состояний ОХД в виде статической взаимосвязи разнородных множеств (концептов) не является достаточной. Например, управляющая модель для сложного ОХД, как основа ИУС, должна быть способна по формализованному тем или иным способом алгоритму рассчитать количественные управляющие воздействия на вещественные и энергетические потоки. Возникает необходимость комбинирования методик ОМ/ДЛ с другими, функционально их дополняющими математическими подходами.

Одним из наиболее эффективных походов к определению количественных отношений между разнородными компонентами ПрО является использование декларативных знаний экспертов с последующим их представлением в формализованном виде средствами нечётко-возможностного подхода (НВП) [11]. НВП предполагает возможность преобразования качественного характера знаний экспертов о причинно-следственных взаимосвязях ОХД ↔ ИУС ↔ АУГ, семантика отношений которых предварительно выявлена, в количественные оценки (например, через дефаззификацию).

Нечёткие функции отклика, для которых предполагается восстановление её количественных величин для явных входных факторов и их неявных межкомпонентных взаимосвязей по методике НВП, представляются аппроксимирующими полиномами:

$$\tilde{y}^{(k)} = \tilde{b}_0 + \sum_{i=1}^n \tilde{b}_i^{(k)} \tilde{x}_i^{\circ} + \sum_{u,j=1}^n b_{ju}^{(k)} \tilde{x}_j^{\circ} \tilde{x}_u^{\circ}, \quad j \neq u, \quad (1)$$

где k – количество нечётких функций отклика, с помощью которых реализуются управляющие воздействия на компоненты ОХД; n – количество нечётких переменных, задающих факторное пространство работы ОХД; $\tilde{b}_0 \dots \tilde{b}_n$ – коэффициенты, отражающие количественную оценку знаний, опыта и интуицию членов АУГ применительно к ПрО; p – количество нелинейных коэффициентов; $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_r)$, $r = \overline{1, m}$ – нечёткие переменные, задающие факторное пространство работы ОХД на основе декларативных знаний членов АУГ; $m = n + p$.

Для идентификации количественных значений, представленных в виде ОМ (рисунок 3) семантических связей (предикатов), с помощью модели (1) в качестве множества входных переменных x и множества «выходных» переменных y могут выступать как CV , так и MV . Из этого следует, что наборы множеств CV и MV могут выступать в качестве концептов (субъектов и объектов), отношения между которыми определены предикатом «участвует в расчёте» в зависимости от целевой функции и необходимых для её расчёта входных факторов.

Примеры успешного применения НВП для создания управляющих алгоритмов (систем) класса, например, СУУТП можно почерпнуть в работах [12-14].

Из приведённого следует, что комбинированное применение методик ОМ/ДЛ и НВП позволит создать единую, содержащую количественные оценки, модель семантических межкомпонентных отношений ОХД ↔ ИУС ↔ АУГ. Это обеспечит формирование единого семантического пространства вертикальных и горизонтальных взаимосвязей, доступа к актуальным данным для ОТ/ИТ-систем и членов АУГ, снизит риск образования разрыва в оценке текущей ситуации, который зачастую возникает при изолированном проектировании элементов ИУС для каждого слоя управленческой иерархии.

Заключение

Предложен подход к использованию слабоструктурированной информации о ПрО при создании логико-семантической модели интегрированной ИУС для сложных ОХД, который заключается в комбинации двух методик: ОМ на базе ДЛ и НВП.

Ключевой особенностью данного подхода является возможность с его помощью выявить и дать количественную оценку связанным с производственной деятельностью сложным закономерностям независимо от их принадлежности к уровню управленческой иерархии.

Идентифицированные таким образом закономерности в обработке разноуровневых информационных потоков позволяют снизить риск образования межуровневого семантического разрыва в данных потоках, который зачастую возникает при раздельном проектировании элементов ИУС для каждого уровня управленческой иерархии. Исключению подобного разрыва также способствует использование для смежных уровней общих методик и инструментов моделирования. На основе предложенного подхода возможно создать единую многокритериальную модель взаимоотношений между компонентами ОХД ↔ ИУС ↔ АУГ для решения задач производства.

Список источников

- [1] ГОСТР ИСО 15746-1-2016 Системы промышленной автоматизации и интеграция. Интеграция возможностей усовершенствованного управления технологическими процессами и оптимизации для производственных систем Часть 1 Структура и функциональная модель.
- [2] **Микони С.В.** Теория принятия управленческих решений / 2-е изд.. Санкт-Петербург: Лань, 2022. 384 с.
- [3] Gartner. Strategic Roadmap for IT/OT Alignment Published 20 February 2020. ID G00466844.
- [4] ISA95, Enterprise-Control System Integration. <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95>.
- [5] **Дональд Л. Бардик, Ульям Л. Леффлер.** Нефтехимия / Пер. с англ. З. Свитанько. М.: Изд-во «Олимп Бизнес», 2005. 496 с.
- [6] **Sandip Lahiri K.** Multivariable predictive control. Applications in industry ISBN: 9781119243519 John Wiley & Sons Limited, 2017.
- [7] **Хафизов И.Ф., Мусин Р.Р.** Современные тенденции развития процесса пиролиза / *Вестник технологического университета*. 2015. Т.18, №2. С.231-234.
- [8] **Антонов В.В., Конев К.А.** Интеллектуальный метод поддержки принятия решений в типовой ситуации // *Онтология проектирования*. 2021. Т.11, №1. С.126-136. DOI: 10.18287/2223-9537-2021-11-1-126-136.
- [9] **Цуканова Н.И.** Онтологическая модель представления и организации знаний. Москва: Гор. линия-Телеком, 2015. 272 с.
- [10] The World Wide Web Consortium. <https://www.w3.org/>.
- [11] **Кимяев И.Т., Spesivtsev A.V.** Ontological and Fuzzy-Possibility Approach to the Synthesis of the DM Functional Equivalent for Management of Hierarchical Systems. In: Silhavy, R. (eds) Artificial Intelligence Trends in Systems. CSOC 2022. Lecture Notes in Networks and Systems, vol 502. Springer, Cham. DOI: 10.1007/978-3-031-09076-9_53.
- [12] **Спесивцев А.В., Домшенко Н.Г.** Эксперт как «интеллектуальная измерительно-диагностическая система» // Сб. докладов. XIII Международная конференция по мягким вычислениям и измерениям SCM 23-25 июля 2010, Санкт-Петербург. Санкт-Петербург: Издательство СПбГЭТУ «ЛЭТИ», 2010. Т.2. С.28–34.
- [13] **Спесивцев А.В., Дайманд И.Н., Лазарев В.И., Кащук А.П.** Интеллектуальная автоматизированная система управления процессом плавки ПВ-3 Медного завода ЗФ ОАО «ГМК «Норильский никель» // *Известия вузов. Цветная металлургия*. 2014. №5. С.64–69.
- [14] **Кимяев И.Т., Спесивцев А.В.** Нечётко-возможностный подход как инструмент управления сложностью интегрированных информационно-управляющих систем. Международная конференция «Региональная информатика (РИ-2022)». Санкт-Петербург, 26-28 октября 2022 г.: Материалы конференции. СПОИСУ. СПб, 2022. 626 с. ISBN 978-5-00182-047-5. С.40-41.

Сведения об авторе

Кимяев Игорь Тимофеевич, 1972 г. рождения. Окончил Норильский индустриальный институт (НИИ) в 1996 г., аспирантуру МИСиС (2001), к.т.н. Инженер-исследователь в области автоматизации технологических процессов и производств. В списке научных трудов более 30 работ в области АСУТП и интеллектуальных систем для сложных производственных объектов. ORCID 0000-0002-5362-4445. igor95a@mail.ru.



Поступила в редакцию 18.12.2022, после рецензирования 16.02.2023. Принята к публикации 20.02.2023.



Ontological fuzzy-possibility approach to creating a pyrolysis furnace control model

© 2023, I.T. Kimyaev

LLC Nornickel Sputnik, Moscow, Russia

Abstract

For objects of economic activity, the development and implementation of an applied approach to creating a subsystem for making managerial decisions that eliminate possible semantic gaps between adjacent layers of the production and technological hierarchy when describing multi-level information flows are considered. The decision subsystem created with the help of the proposed approach is capable of processing multi-level data flows in a complex structure of maintaining these objects in a viable state. The approach proposed by the author is based on vertically integrated ontological models created using standard software tools. The identified semantic links between concepts within the formalized multi-level ontological models are restored using the explicit and implicit knowledge of experts about the conduct of production operations. A fuzzy-possibility approach was chosen as a methodological basis for extracting and formalizing expert knowledge in the form of a fuzzy analytical polynomial. The application of the proposed approach is shown by the example of controlling the process of pyrolysis of high-molecular hydrocarbon raw materials in specialized furnaces.

Key words: decision making, integrated control system, ontological modeling, fuzzy-possibility approach, pyrolysis furnace, managerial group, relationship semantics.

Citation: Kimyaev IT. Ontological fuzzy-possibility approach to creating a pyrolysis furnace control model [In Russian]. *Ontology of designing*. 2023; 13(1): 139-149. DOI: 10.18287/2223-9537-2023-13-1-139-149.

Conflict of Interest: The author declares no conflict of interest.

List of figures and tables

Figure 1 - Virtual functional and technological "wall" between the levels of the OT/IT hierarchy management

Figure 2 - Pyrolysis furnace as an op-amp with A, B, C, D coils and a typical set of physical measured CVs and implemented MVs

Figure 3 - Ontological model of a multi-level management decision-making process for pyrolysis furnace

Figure 4 - Fragment of the OM of the management decision-making process for the pyrolysis furnace, formed by means of the Protege software.

Table 1 - Controlled and managed process variables for pyrolysis furnace control

References

- [1] GOST R ISO 15746-1-2016 Industrial automation systems and integration. Integrating advanced process control and optimization capabilities for manufacturing systems. Part 1 Structure and functional model [In Russian].
- [2] **Mikoni SV**. Theory of managerial decision making: textbook for universities [In Russian]. 2nd ed., corrected. and additional - St. Petersburg: Lan, 2022. 384 p.
- [3] Gartner. Strategic Roadmap for IT/OT Alignment Published 20 February 2020. ID G00466844.
- [4] ISA95, Enterprise-Control System Integration. <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95>.
- [5] **Burdick DL, Leffler WL**. Petrochemicals in Nontechnical Language. M., Olympus Business, 2005
- [6] **Sandip Lahiri K**. Multivariable predictive control. Applications in industry ISBN: 9781119243519 John Wiley & Sons Limited, 2017.
- [7] **Khafizov IF, Musin RR**. Modern trends in the development of the pyrolysis process [In Russian]. Bulletin of the Technological University. 2015, v.18, v.2, p.231-234 (RINC, HAC)
- [8] **Antonov VV, Konev KA**. Intelligent decision support method in a typical situation [In Russian]. *Ontology of designing*. 2021; 11(1): 126-136. DOI: 10.18287/2223-9537-2021-11-1-126-136.

-
- [9] **Tsukanova NI.** Ontological model of representation and organization of knowledge [In Russian]. Moscow: Gor. line-Telecom, 2015. 272 p.
- [10] The World Wide Web Consortium. <https://www.w3.org/>.
- [11] **Kimyaev IT, Spesivtsev AV.** Ontological and Fuzzy-Possibility Approach to the Synthesis of the DM Functional Equivalent for Management of Hierarchical Systems. In: Silhavy, R. (eds) Artificial Intelligence Trends in Systems. CSOC 2022. Lecture Notes in Networks and Systems, vol 502. Springer, Cham. DOI: 10.1007/978-3-031-09076-9_53.
- [12] **Spesivtsev AV, Domshenko NG.** Expert as an "intelligent measuring and diagnostic system" [In Russian]. Sat. reports. XIII International Conference on Soft Computing and SCM Measurement July 23-25, 2010, St. Petersburg. - St. Petersburg: Publishing house of SPbGETU "LETI", 2010; 2: 28–34.
- [13] **Spesivtsev AV, Daimand IN, Lazarev VI, Kashchuk AP.** Intelligent automated control system for the smelting process PV-3 of the Copper Plant of the Polar Branch of OJSC MMC Norilsk Nickel [In Russian]. *News of universities. Non-ferrous metallurgy*. 2014; 5: 64–69.
- [14] **Kimyaev IT, Spesivtsev AV.** Fuzzy-possibility approach as a tool for managing the complexity of integrated information and control systems international conference [In Russian]. "Regional Informatics (RI-2022)". St. Petersburg, October 26-28, 2022: Conference Proceedings. SPOISU. St. Petersburg, 2022. 626 p. ISBN 978-5-00182-047-5. P.40-41.
-

About the author

Igor Timofeevich Kimyaev, (b. 1972) graduated from the Norilsk Industrial Institute (NII) in 1996, postgraduate study at MISiS (2001), Ph.D. Engineer-researcher in the automation field of technological processes and productions. The list of scientific papers includes more than 30 works in the field of process control systems and intelligent systems for complex production facilities. ORCID 0000-0002-5362-4445. igor95a@mail.ru.

Received December 18, 2022. Revised February 16, 2023. Accepted February 20, 2023.

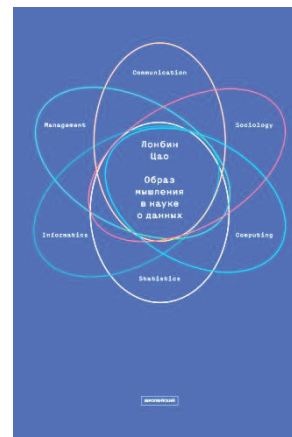
Лонбин Цао¹**Образ мышления в науке о данных:**

Наступающая научно-техническая и экономическая революция.

СПб. : Издательство Европейского университета в Санкт-Петербурге. 2022. 552 с.

Springer International Publishing AG, 2018. Перевод с английского А.В. Климонтович.

Научный редактор В.И. Городецкий



Современный быстроразвивающийся мир данных, интенсивное использование данных и новые возможности научного поиска и бизнеса на их основе породили науку о данных — новую парадигму исследований и разработок, которая использует методы и возможности экспериментальной, теоретической и вычислительной науки. Как любая новаторская область знания, наука о данных порождает много споров и вопросов: что именно она из себя представляет, как проникает в различные профессии, образование, бизнес, экономику и другие сферы нашей жизни, как конкурирует в этих областях с другими науками и как их изменяет. В этой книге представлена по возможности полная картина науки о данных как новой научной и технологической парадигмы, а также учебной дисциплины. Автор подробно рассказывает об эволюции образа мышления, основанного на данных, о развитии входящих в науку о данных дисциплин, о том, как и почему эта наука становится движущей силой новой цифровой экономики. В книге описывается, как мышление на основе данных трансформирует наши представления об управлении и стратегии развития предприятия, как оно формирует прорывные научные исследования и технологические инновации. Книга может быть полезна для менеджеров данных, бизнес-аналитиков, лиц, принимающих стратегические решения, руководителей исследований и преподавателей, ответственных за продвижение актуальной научной, инновационной и промышленной повестки и за разработку учебных курсов следующего поколения, а также для всех, кто хочет понять, как наука о данных меняет наш мир.

В книгу включен дополнительный материал: толковый англо-русский словарь терминов науки о данных, подготовленный профессором В.И. Городецким².

Фрагмент толкового англо-русского словаря терминов науки о данных³

Analytics. *Аналитика, аналитика данных:* Раздел науки о данных, включающий множество теорий, технологий, инструментальных средств и процессов, которые дают возможность глубоко понять и эффективно выявить практически полезные гипотезы, зависимости и знания, скрытые в данных. Аналитическая обработка данных включает дескриптивную аналитику, предсказательную и предписывающую аналитику.

Artificial Intelligence. *Искусственный интеллект (ИИ):* Наука о получении, представлении и использовании знаний. Её развитие опирается в основном на информатику, вычислительную технику, машинное обучение, теорию взаимодействия и теорию систем, использует вычислительную логику, планирование, обучение с подкреплением, статистическое/вероятностное представление, рассуждения на основе прецедентов, эволюционные вычисления и глубокий поиск. Эти науки покрывают большую (но не всю) часть существующих частных направлений исследований по ИИ. Их краткий список включает эмуляцию памяти, извлечение информации, её кодирование, хранение, индексацию и поиск, машинное обучение, обнаружение связей и закономерностей в данных, распознавание речи, языка, голоса, изображений и мультимедийных материалов, нейросетевые вычисления, генеративные нейросети, нечёткие и эволюционные вычисления, цифровые двойники объектов и др.

Domain data intelligence. *Интеллект предметной области:* Понятие, которое характеризует знания, интегрированные в предметную область (ПрО), к которой относятся используемые данные или системы данных. Характеризует значимые свойства ПрО и их понимание, появляется в контексте свойств ПрО. Контекст может быть представлен релевантными атрибутами описания ПрО, знаний и метазнаний о ней, а также другими ресурсами, специфичными для ПрО. Включение качественных и количественных знаний о ПрО может способствовать более глубокому пониманию её сложностей и их ключевой роли в выявлении неизвестных знаний и решений из данных о ПрО в форме рекомендуемых действий и механизмов принятия решений, например в области биоинформатики или социологических вычислений.

¹ Лонбин Цао (р. 1969) — бизнес-аналитик, профессор информационных технологий в Сиднейском технологическом университете (Австралия), главный редактор журнала *International Journal of Data Science and Analytics*.

² Продолжается дискуссия о содержании эволюционирующих терминов в области, связанной с данными, обработкой данных, ..., искусственным интеллектом. Редакция журнала поддержала профессора В.И. Городецкого, внёсшего свой вклад в терминологический спор и предложившего своё толкование важнейших терминов, и надеется на конструктивное их обсуждение читателями журнала.

³ Толковый словарь составлен профессором В.И. Городецким в основном по материалам перевода книги Longbing Cao. *Data Science Thinking. The Next Scientific, Technological and Economic Revolution*. International Publishing AG, part of Springer Nature 2018. Смысл некоторых терминов обсуждался, в том числе, и с автором этой книги Longbing Cao.

Environment data intelligence. *Интеллект окружающей среды:* Интеллект, скрытый в данных о внешнем окружении проблем и задач науки о данных. Факторы окружающей среды могут быть описаны данными в терминах ПрО, организационных, социальных, человеческих, сетевых и/или других связанных с ними понятий. Факторы окружающей среды в задаче о данных или в системе данных зависят от ПрО и конкретной задачи. Кроме предметно-зависимых контекстуальных факторов имеются общие моменты, связанные с окружающей средой, например взаимодействия и отношения внутри контекста, взаимодействия и отношения между контекстом и системой данных, влияние контекстуальных факторов и отношений на систему данных, динамика и эволюция среды и её воздействие на задачу. Эти факторы важны для понимания данных о среде и манипулирования ими.

Data science, datalogy. *Наука о данных, даталогия:* Термины являются синонимами, хотя последний вариант используется редко даже в зарубежном сообществе. Можно ожидать, что со временем для науки о данных будет принят более краткий термин даталогия. Это грамматический и отчасти смысловой аналог названий ряда других наук, например биологии.

Human intelligence (in data science). *Человеческий интеллект (в науке о данных):* Интеллект отдельной личности в контексте науки о данных. Это интеллект, который привносится человеком в процесс решения проблем, связанных с данными, или в исследование систем данных. При решении сложных задач науки о данных человеческий интеллект играет ключевую роль в понимании обрабатываемых данных и их контекста, в проектировании и реализации решений, а также в выявлении ценности данных. Человеческий интеллект в науке о данных может иметь форму явного привлечения эмпирических знаний людей, их мнений, намерений, ожиданий, форму явного привлечения человека к работе с данными в реальном времени, использования оценок отдельных экспертов или их групп при рассмотрении сложных проблем науки о данных. Он может иметь форму неявного или косвенного привлечения образного мышления человека, его рассуждений, основанных на воображении, привлечения эмоционального интеллекта человека, его вдохновения, способности к мозговому штурму, умения рассуждать и познавать с помощью конвергентного мышления посредством взаимодействия с другими членами команды при решении различных проблем науки о данных. В зависимости от уровня сложности и предъявляемых требований человеческий интеллект может выполнять разные роли в науке о данных.

Human-like machine intelligence. *Человекоподобный ИИ:* Интеллект, ориентированный на разработку новых механизмов в ИИ современного уровня. Ключевые моменты человеческого интеллекта, которые в настоящее время плохо моделируются в ИИ, но которые следовало бы включить в следующее поколение систем ИИ, а именно: интуиция, энтузиазм, любознательность, мышление на основе воображения, вдохновение, креативное мышление, иерархическая сложность, иерархический интеллект и их взаимосвязи. Наиболее трудная задача современных исследований в области человекоподобного ИИ — понять, как можно выполнить моделирование работы человеческого мозга и как можно реализовать человеко-подобный мыслительный процесс.

Network data intelligence. *Интеллект данных сети (сетевой интеллект):* Интеллект, интегрированный в сеть, которая используется для решения задач, связанных с данными или с системой данных. Сетевой интеллект возникает из веб-интеллекта или из широкомасштабных сетевых процессов и сетевых ресурсов, а также из активности пользователей, например, в социальных сетях, при работе с мобильными сервисами, сетевыми системами и при работе с другими живыми существами. Сетевой интеллект реализуется в таких аспектах, как распределение информации и ресурсов по сети, представление и учёт связей между распределёнными объектами и системами, отношения и взаимодействия между сетевыми узлами, скрытые сообщества и группы, формирующиеся в сети, информация и другие ресурсы, о которых становится известно из сетей. Сетевой интеллект содержится в веб-ресурсах, облачной инфраструктуре и облачных вычислительных ресурсах, в информации, извлекаемой средствами поиска, в способностях к структуризации данных распределённых хранилищ и внешней среды.

Knowledge. *Знание:* Представляет собой форму обработанной информации в терминах смеси информации, процедурно описанных действий или пропозициональных правил. Знание может быть субъективным или объективным, известным или неизвестным, применимым на практике или нет и разумным или нет. См. место понятия «знание» в пирамиде DIKIW.

Wisdom. *Мудрость:* Представляет собой высокоуровневые правила, которые являются продуманным результатом обработки информации, знаний или просто результатом вдохновения, интуиции, приобретённых из опыта или в процессе интеллектуальной деятельности. Мудрость указывает на высокие способности её носителя, на его метазнания, разум, умение применять знания на практике, на его способность к суждениям или принятию правильных решений в нужном месте, в нужное время и для нужных целей. Мудрость может быть нематериальной, уникальной, личной, интуитивной или может быть результатом умственного вдохновения. По сравнению со знанием, мудрость находится вне времени. Она является всесторонней, общей и эмоциональной, передаётся из поколения в поколение и между культурами.

Data DNK. *ДНК данных:* Даталогическая «молекула» данных, состоящая из фундаментальных и типовых элементов: поведение (англ. *behavior*), сущность (англ. *entity*), отношение (англ. *relationship*) и свойство (англ. *property*). В науке о данных ДНК данных приписывают роль, аналогичную той, которую биологическая ДНК играет в живых организмах. Четыре элемента в ДНК данных, а именно поведение, сущность, отношение и свойство (BERP) представляют различные фундаментальные аспекты данных. Сущность может быть объектом, примером (экземпляром), человеком, организацией, системой или частью подсистемы или окружающей среды. Свойство касается атрибутов, которые описывают сущность. Поведение описывает действия или динамику сущности или набора сущностей. Отношение описывает взаимодействие сущностей и свойств, включая взаимодействие значений свойств.

DIKIW. (от англ. *Data, Information, Knowledge, Intelligence, Wisdom*): DIKIW: Сокращённое название когнитивной пирамиды обобщения данных по уровням Данные — Информация — Знания — Интеллект — Мудрость.

Information (in data science). *Информация:* Представляет собой описание данных (объектов) в организованном виде с определённой целью или в виде, имеющем определённый смысл. Информация может быть структурированной (организованной) или функциональной (целенаправленной), субъективной или объективной (основанной на фактах) и т. д. Обычно, чтобы подчеркнуть отличие информации от данных, говорят, что информация — это семантическое восприятие данных человеком с позиций её содержания, представления и использования. См. также место понятия «информация» в пирамиде DIKIW.

Data. *Данные:* Дискретные или непрерывные факты, сигналы (например, датчиков, которые могут быть субъективными или объективными) или символы (знаки), которые относятся к объекту (к физической или виртуальной сущности или к событию). Данные лежат на нижнем уровне когнитивной системы DIKIW и могут быть субъективными или объективными, иметь или не иметь смысл и ценность.

Brain informatics. *Информатика мозга:* Научное направление в области исследований мозга человека и живых организмов, которое использует методы и модели информатики и искусственного интеллекта для изучения и исследования процессов и механизмов, используемых мозгом человека и живых организмов для решения своих задач, в частности задач принятия решений.

Computational intelligence. *Вычислительный интеллект:* Направление в ИИ, которое использует аналоги некоторых механизмов принятия решений человеком и живыми существами, в частности для решения отдельных задач науки о данных. Включает методы и технологии нейронных сетей, нечёткой логики и эволюционных вычислений. Эти технологии способны решать ряд специфических задач науки о данных лучше, чем другие того же назначения.

Data coupling. *Связи в данных:* Любые отношения между данными, например совместная встречаемость, пространственное соседство, функциональная зависимость, взаимодействие, ассоциация, корреляция, причинность, сходство, упорядоченность во времени, синонимия и т. п.

Data engineering. *Инженерия данных:* Относится к процессам извлечения, поиска, сбора, подготовки, предварительной обработки и управления данными. Выполняется инженерами по данным в интересах специалистов по анализу данных.

Data intelligence. *Интеллект данных:* Это знания, встроенные в данные (скрытые в данных). Могут быть выявлены с помощью методов машинного обучения. Критерием успешности выявления интеллекта данных является то, в каком объёме и до какой степени специалист по данным сможет понять, осмыслить и уловить свойства данных, их сложности и потенциал.

Raw data. *Сырые (необработанные) данные:* Данные, которые ещё не обработаны или не готовы для использования. Необработанные данные называют также первичными данными.

Data things. *Вещи данных:* Понятие, которое используется в экономике данных. Оно отличается от физических объектов (вещей) традиционной экономики. Вещи данных — это самоорганизующиеся объекты, которые сами обеспечивают себя необходимыми данными (англ. *self-datafied*), виртуализируются и устанавливают связи между собой, формируя новые вещи данных более высокого уровня агрегирования и фабрики вещей данных.

Data products. *Информационные продукты:* Информация, знания, программные системы, программные инструменты и т. д., которые являются результатами использования методологий, методов, алгоритмов и инструментов науки о данных. Являются результатами обработки данных или получаются с использованием данных, могут быть представлены множеством найденных свойств, предсказаний значений переменных, состояний системы или событий, происходящих в ней. Могут иметь форму сервисов, рекомендаций, решений для исполнения, извлечённых знаний, моделей, алгоритмов, новых парадигм, систем или приложений.