Институт динамики систем и теории управления имени В.М. Матросова
Сибирского отделения Российской академии наук
Иркутск, Россия

Центр научных исследований и высшего образования
(CICESE Research Center)
Энсенада, Мексика

Материалы V Международного семинара по информационным, вычислительным и управляющим системам для распределенных сред (ICCS-DE 2023)

3 – 7 июля 2023 г., Иркутск, Россия

Иркутск: Издательство ИДСТУ СО РАН
2023

УДК: 004.7/.9

ББК: 32.96 + 32.97

Материалы V Международного семинара по информационным, вычислительным и управляющим системам для распределенные сред (ICCS-DE 2023), 3-7 июля, 2023, Иркутск, Россия. Иркутск: Изд-во ИДСТУ СО РАН, 2023. 158 с.

Proceedings of the 5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 3-7, 2023, Irkutsk, Russia. Irkutsk: ISDCT Publisher, 2023. 158 p.

ISBN 978-5-6041814-4-7

Материалы научного сборника включают избранные статьи и тезисы V Международного семинара по информационным, вычислительным и управляющим системам для распределенные сред (ICCS-DE 2023), проведенного Институтом динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (Иркутск, Россия) совместно с Центром научных исследований и высшего образования (CICESE Research Center, Энсенада, Мексика) 3-7 июля, 2023 г. в г. Иркутск, Россия. Избранные научные работы посвящены развитию теории и практики разработки информационных, вычислительных и управляющих систем, относящихся к различным предметным областям информационной, технической, экономической и других сфер человеческой деятельности. Особое внимание в них уделяется системам, функционирующим в различных распределенных средах.

Официальный сайт семинара ICCS-DE 2023: https://iccs-de.icc.ru/

ISBN 978-5-6041814-4-7

ОРГАНИЗАТОРЫ

Международный программный комитет семинара

Председатель программного комитета

Бычков И.В., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, г. Иркутск, Россия

Сопредседатель программного комитета

Andrei Tchernykh, CICESE Research Center, Ensenada, Mexico

Члены программного комитета

Jorge Mario Cortés-Mendoza, Polytechnic University of Amozoc (UPAM), San Andrés las Vegas, Amozoc, Puebla, Mexico

Dalibor Dobrilovic, University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia

Xiwang Guo, Northeastern University, Shenyang, China

Ljubica Kazi, University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia

Zeljko Stojanov, University of Novi Sad, Technical faculty “Mihajlo Pupin”, Zrenjanin, Serbia

Бабенко М.Г., Северо-Кавказский федеральный университет, Ставрополь, Россия

Еделев А.В., Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук, Иркутск, Россия

Ковтуненко А.С., Уфимский Государственный Авиационный Технический Университет, Уфа, Россия

Костромин Р.О., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Легалов А.И., Национальный исследовательский университет "Высшая школа экономики", Москва, Россия

Сидоров И.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Смагин С.И., Вычислительный Центр Дальневосточного отделения Российской академии наук, Хабаровск, Россия

Ульянов С.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Федоров Р.Е., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Феоктистов А.Г., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Черкашин Е.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Шигаров А.О., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Шокин Ю.И., Федеральный исследовательский центр информационных и вычислительных технологий, Новосибирск, Россия

Юрин А.Ю., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Организационный комитет семинара

Председатель организационного комитета

Феоктистов А.Г., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, г. Иркутск, Россия.

Члены организационного комитета

Бабанин Д.Г., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Башарина О.Ю., Уральский государственный экономический университет, г. Екатеринбург, Россия

Горский С.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Игнатьев И.Н., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Калашников М.Ю., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Кензин М.Ю., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Кононенко Г.Б., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Костромин Р.О., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Кумачев А.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Маджара Т.И., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Максимкин Н.Н., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Попова А.К., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

Столбова Г.Н., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия
Толстихин А.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия
Фереферов Е.С., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия
Чекан М.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия
Черкашин Е.А., Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук, Иркутск, Россия

СОДЕРЖАНИЕ

ПЛЕНАРНАЯ СЕССИЯ

СЕКЦИОННЫЕ ДОКЛАДЫ

СПЕЦИАЛЬНАЯ СЕКЦИЯ ДЛЯ МОЛОДЫХ УЧЕНЫХ

СПЕЦИАЛЬНАЯ СЕКЦИЯ ДЛЯ ШКОЛЬНИКОВ

Adopting Qualitative Social Science Methods for Inquiring Software Engineering Industrial Practice

Zeljko Stojanov[1]

[1]University of Novi Sad, Technical faculty "Mihajlo Pupin" Zrenjanin, Djure Djakovica BB, Zrenjanin, Serbia

Abstract

Software engineering is a young engineering discipline with a very dynamic and complex industrial practice that tries to align with the ever-changing needs of business and human society. Understanding and improvement of industrial practice is in the constant focus of the research community for over 50 years. The initial research focus has recently moved from strictly technical issues to issues related to human factors, organization issues, economics, and management. This shift in the research required using new methods for inquiring about non-technical aspects of the practice, leading to the adoption of qualitative social science methods in fieldwork and data analysis. Qualitative research methods are suitable for a deeper understanding of human behaviour and experiences, which is essential for practice observation and improvement. In this paper, personal experiences with qualitative research methods in three studies conducted in the software industry are presented. Based on the presented experience, and insight into the scientific literature, recommendations for using qualitative research methods are outlined.

Keywords

software engineering, industrial practice, field research, qualitative methods, social science methods

1. Introduction

Software engineering professional practice has been recognized as an important knowledge area for research and it was included in *Guide to the Software Engineering Body of Knowledge (SWEBOK)* [1]. Professional practice relates to implementing processes in the practice resulting in quality products and services, requiring software engineers with proper knowledge, skills, and experiences. According to SWEBOK [1], professional practice includes areas such as professionalism, group dynamics and psychology, and communication skills, all of them putting the main focus on humans in the practice. It has been recognized that the majority of costs in the software industry relate to costs of skilled labour (especially software developers), which strongly point out the importance of the human workforce [2]. Therefore, interdisciplinary research projects considering social, cultural and human factors are necessary for better understanding of the practice [3].

One of the most important aspects of software engineering industrial practice is productivity, which is generally defined as the ratio between the outputs and inputs for the given context to

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

zeljko.stojanov@uns.ac.rs (Z. Stojanov)

0000-0001-6930-5337 (Z. Stojanov)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

achieve ideal functionality and quality of products or services as well as the optimal effort of workers [4]. Software productivity includes technical, social, and economic aspects of practice, which highly depends on the organization and performance of processes, human factor, and complexity and the nature of software products [5].

Software engineering like any other technical discipline is based on mathematics and general technical and engineering principles. Therefore, in research on software engineering practice dominate quantitative methods based on the deductive approach. Dittrich et al. [6] argued that social practice theory can be useful for inquiring and explaining software engineering practice, which indicates that the use of qualitative social science methods for practice observation and understanding is preferable. Compared to quantitative methods, qualitative methods are more flexible in design, subjective, speculative, and grounded in the real context in which the phenomenon is inquired [7]. The reasons for avoiding qualitative research methods in software engineering are [8]: the lack of expertise in using qualitative methods, limited access to software engineers in industry, the problem with generalizability of research findings, time consuming, very high and rigour ethical standards, and different epistemological stance.

Effective use of qualitative social science methods for inquiring software engineering practice assumes understanding methods in order to understand obtained findings and knowledge [9]. It is important to note that qualitative methods are based on the inductive model of inquiry (bottom-up approach) that starts by observing a phenomenon, collecting data about it, and then deriving theoretical explanations of the observed phenomenon. The main aim of social science is to understand elements and patterns of human experiences and behaviour [10]. Since software engineering has been perceived as highly human dependant practice, use of qualitative social science methods can contribute to get more comprehensive and deeper understanding of the practice which is based on human characteristics, knowledge and actions.

Qualitative research methodologies have been regularly used in social sciences such as education, sociology, psychology, health sciences, social work, community development, and management [11]. There exist a variety of qualitative research genres and approaches, which are based on multiple methods and procedures for conducting data collection, analysis, interpretation, and creation of findings. The selection of the most suitable methods depends on the study objective, complexity of the social context and interactions. The main characteristics of qualitative research are: (1) it takes place in the natural world, (2) it is based on using multiple methods that respect the humanity of the participants in the study, (3) it is focused on the real context, (4) it is evolving and not rigid in structure, rather than tightly prefigured , and (5) it is fundamentally interpretive (the findings are interpreted by researchers).

Empirical studies, based on extensive fieldwork, may occur in several forms, and require different research methods. Selection of the appropriate approaches and methods is a challenging and difficult task for researchers because there is a need to get acquainted with the selected methods. Wohlin and Aurum [12] suggested selection of research methods in a process with three phases: (1) *Strategic phase* - based on the research objective research logic, purpose, and approach are selected, (2) *Tactical phase* – selection of research process (qualitative, quantitative, mixed) and methodology, and (3) *Operational phase* – selection of data collection and analysis methods. The use of qualitative methods is based on inductive research logic and interpretive approach. Actually, the use of qualitative methods assumes interpretation, interaction, and reality construction based on raw data collected as unstructured text in natural language [13].

Qualitative methods have been used in empirical software engineering studies in the last 30 years, but they are used very shyly in the beginning. Seaman [14] outlined qualitative methods mostly used in the fieldwork, and argued that they are a good choice for inquiring about human factors in software engineering. Qualitative methodologies and methods by their nature are suitable for inquiring software engineers in their real settings through intensive fieldwork. However, researchers have challenge to select the most suitable methods for their objectives. Lethbridge et al. [15] listed and classified data collection techniques based on the involvement of software engineers in their implementation, which is very important for organizing fieldwork that will not disturb the everyday work of engineers (especially when using qualitative methods that assume spending significant time in the workplace).

Based on the previous considerations, the objective of this work is to present the author's personal experiences in using various qualitative methods for researching software engineering industrial practice, and based on the experiences to propose recommendations for using qualitative methods for fieldwork in software engineering.

2. Experiences

In this section the experience with using qualitative research methods for inquiring software engineering practice will be presented. The first subsection presents the first experience with qualitative methods during PhD research, while the next two subsections present the experience with the software process improvement project in a local micro software company in Serbia, and research on software requirements practice with experts from various software companies in Serbia.

2.1. Beginnings

The first contact with qualitative research methods was during the empirical evaluation of the method and service for specifying software change requests in data-driven software as a part of the author's PhD research [16]. Qualitative data were collected in three focus groups organized after the experimental use of software in the laboratory at the faculty where the software was installed. During the focus groups used for discussion of students, the students provided answers to open-ended questions in a distributed questionnaire. The main objective of using open-ended questions was to collect students' opinions about the characteristics of the installed service without any suggestions to them. Collected answers to open-ended questions were analyzed by using the constructivist grounded theory coding technique proposed by Charmaz [17], leading to the identification of the advantages and disadvantages of the service.

In addition to a qualitative evaluation of the developed service for software change request specification, PhD research included a qualitative study on software change request process in small software companies in Banat region in Romania and Serbia [18][19]. Data were collected by using two focus groups organized in software incubators in Zrenjanin (Serbia) and Timisoara (Romania), and semi-structured interviews with software experts from software companies. Both focus groups and interviews were recorded, transcribed to raw text and prepared for data analysis.

The analysis of raw and unstructured text was based on the constructivist grounded theory coding technique [17] and writing memos for describing developed codes and concepts that represent process characteristics [20, 21]. Coding was implemented in two phases: (1) *initial coding* - identification of segments of text that relate to research objective and labelling them with *initial codes* (textual tags), and (2) *focused coding* - analysis of identified initial codes and derivation of more focused codes that are promoted to the concepts that represent the findings of data analysis. This type of analysis is classified as *inductive* and *interpretative* because the researcher interprets the initial data and construct the research findings. Interpretation and researcher thoughts about the inquired phenomenon are written in *theoretical memos* (description of initial codes, construction of focused codes, identification of concepts), while methodological dilemmas on research process are written in *methodological memos* (notes on the method of coding, notes on the process of construction of findings, notes on the way of organizing focus groups, notes on the structure of interviews). Memo writing was realized throughout the entire research process. An example of qualitative data analysis from industrial setting study based on constructivist grounded theory coding and writing reflexive memos is presented in Figure 1.

Although it is not represented in the Figure 1, the data analysis procedure is iterative, which means, for example, that after the construction of the concepts it can be returned to the refinement of the initial or focused codes.

Overview of used qualitative methods during PhD research is presented in Figure 2, while the personal experience with qualitative research methods gained during the PhD research was distilled in the chapter "Qualitative research on practice in small software companies" published in "Encyclopedia of Information Science and Technology" [22].

A gentle introduction to qualitative techniques (with recommended literature) such as focus groups, semi-structured interviews, the use of open-ended questions, qualitative coding, writing memos, and grounded theory can be found in the book "The SAGE Encyclopedia of Qualitative Research Methods" [20].

2.2. SPI project in a local micro software company

The next experience in using qualitative methods after getting PhD degree was during the implementation of Software Process Improvement (SPI) project in a local micro software company. The project started in 2011 after joint preparation of the author and the company manager. The focus in the SPI project was on improving software maintenance process in the company because company manager stated that majority of tasks in the company relate to software maintenance, which was confirmed in later statistical analysis of the company tasks' records (over 84%) [23].

The objective of the SPI project was to identify and improve software maintenance process in the company based on detailed inquiry of everyday practice. Inductive bottom-up approach to SPI was adopted since the aim was to assess and improve maintenance practice without affecting everyday work of programmers. Lightweight inductive assessment method is based on frequent feedback to company management and relevant employees in order to get the best results [24]. The main objective of the assessment method is to assess the practice and find out which segments should be improved. Practice assessment during SPI project fostered

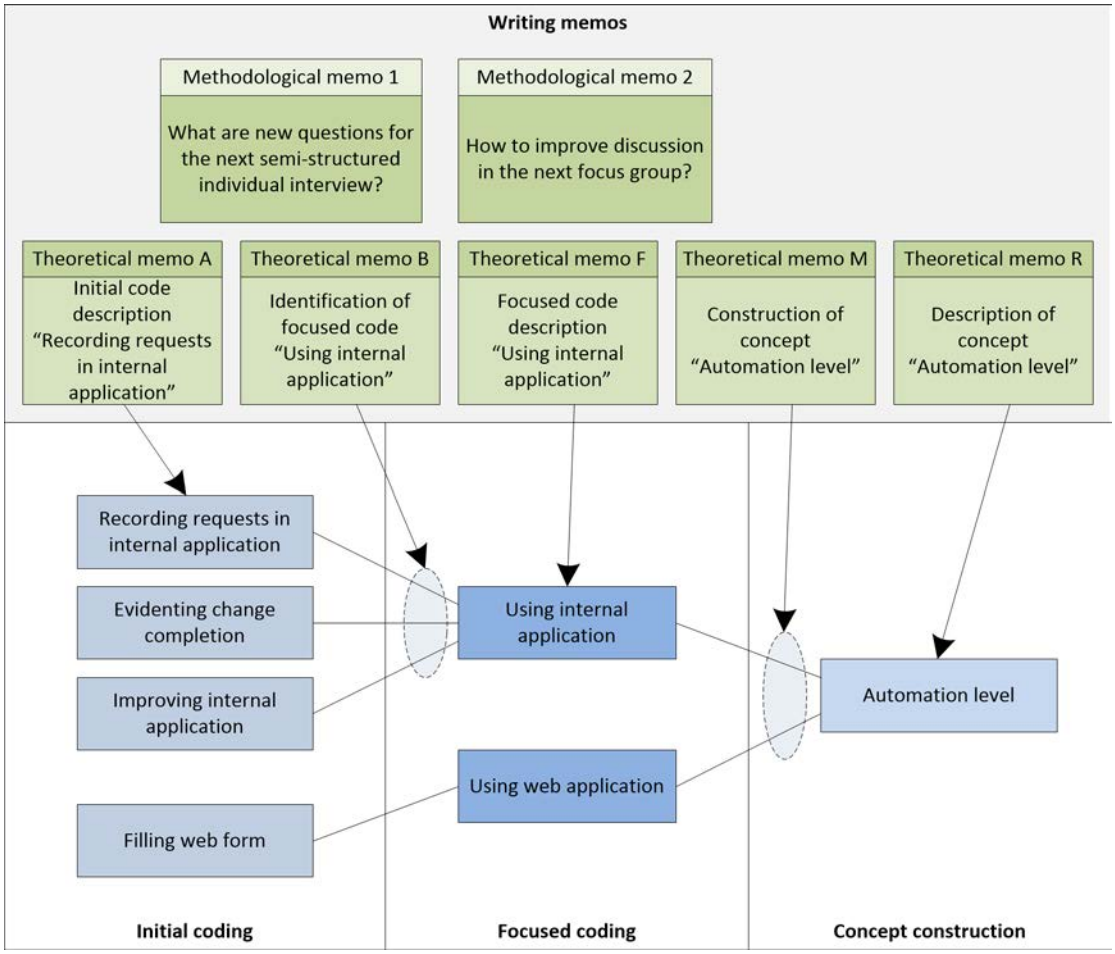


Figure 1: Elements of qualitative data analysis used in PhD research

organizational learning in the company, which resulted with systematized knowledge about maintenance practice rendered as a thematic knowledge framework [25][26].

Practice assessment was based on using qualitative social science methods for gaining deeper insight into the practice and proposing improvements to existing maintenance process. Analysis of qualitative data was accompanied with quantitative data, which confirmed the findings of qualitative data analysis. Qualitative research methods used in SPI project are presented in Figure 3.

Qualitative data were analysed with inductive thematic analysis [27], with the steps presented in Figure 3, while details on analysis were written in memos [21]. During the familiarization with the data, all textual data were prepared and imported in software MAXQDA and the reading and coding strategy were proposed. Coding was performed in software, and after that defined coding schema was exported to MS Word document which is used for other phases and identification of themes. MS Word document was printed and data analysis leading to

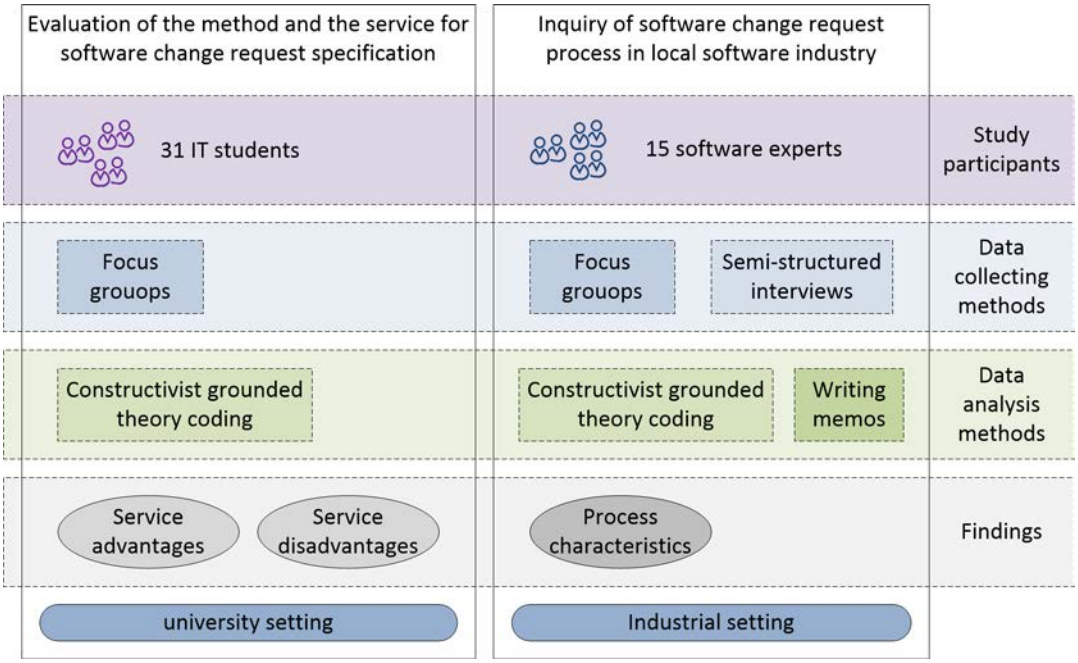


Figure 2: Overview of qualitative methods used during the author's PhD research

theme identification (process improvement proposals, and knowledge thematic framework) was performed during feedback sessions in the company. The analysis was done using colored felt-tip pens in order to emphasize the merging of codes into themes [26].

2.3. Study on software requirements practice with experts from industry

A study with experts from software industry was conducted in order to explore their experience with software requirements engineering practice, especially how they solve problems during requirements elicitation and specification [28]. In addition, software experts perceptions of soft skills in software requirements engineering (SRE) were also explored [29]. The study was created as purely qualitative, aimed at increasing understanding of problem solving skills in SRE.

Qualitative data analysis was based on inductive thematic analysis [27], while memos were written for analysing items emerged during the thematic analysis of collected data and for explaining methodological dilemmas and problems during the study [21]. Elements of qualitative research on SRE problem solving skills are presented in Figure 4.

Data collection was based on semi-structured interviews with open-ended questions. In the study participated 14 software experts from software industry. Interviews were conducted in Serbian and recorded, while the core elements of the research and findings were translated to English. Field notes were also used to prepare annotations about the interviewees, the research process and the data [30].

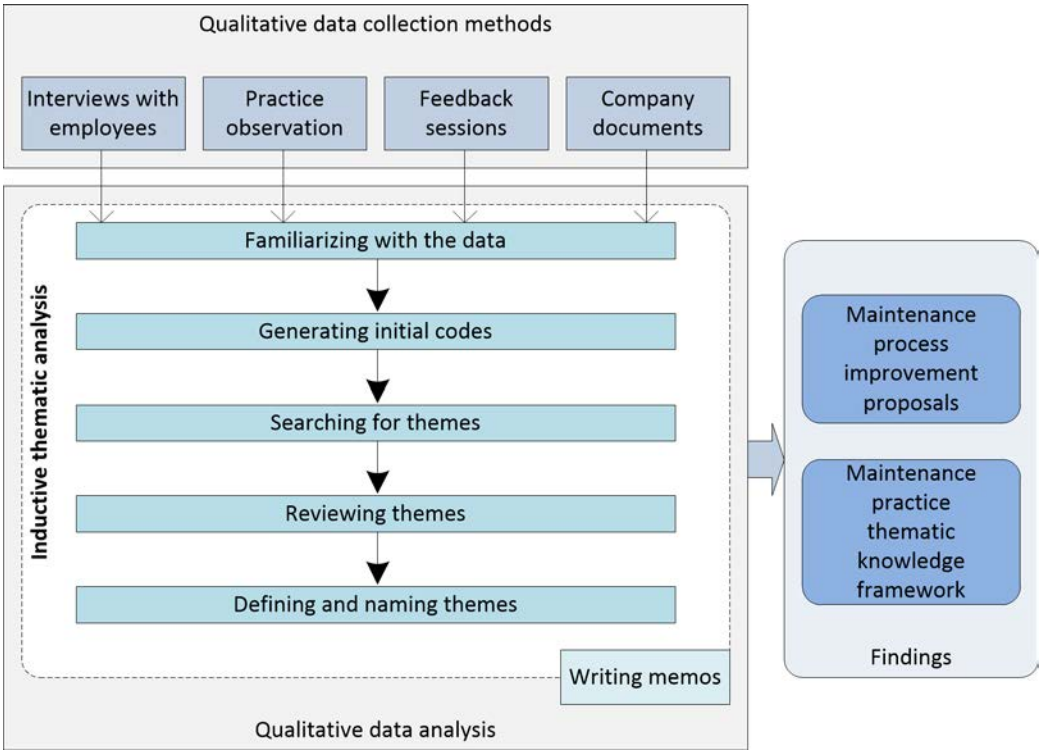


Figure 3: Qualitative methods used in the SPI project implemented in a local micro software comany

3. Recommendations

Based on the extensive insight into literature on using qualitative research methods in software engineering and presented experience in three studies with different qualitative methods, the following recommendations can be proposed:

- *Use qualitative methods to get deeper insight into specific phenomenon in the practice.* Qualitative data collection methods enable getting the real experiences and opinions through open-ended questions, which is not anticipated by the researchers. In addition, semi-structured and flexible implementation of data collecting methods such as interviews or focus groups enables setting of sub-questions and additional questions that deepen the understanding of the observed practice. In that case, it is possible to inquire something that was missed or overlooked in the preparation of the research.
- *Mixing qualitative and quantitative methods.* Mixing different types of methods and approaches in research enables getting insights from different perspectives on the same phenomenon [31]. In that case, triangulation of data sources and data analysis methods increases the validity of the research and reliability of the findings (trustworthiness in qualitative terminology) [32].
- *Work in a multidisciplinary team.* Software engineering as a human intensive practice situated in specific organizational context requires inquiry that assumes not only technical

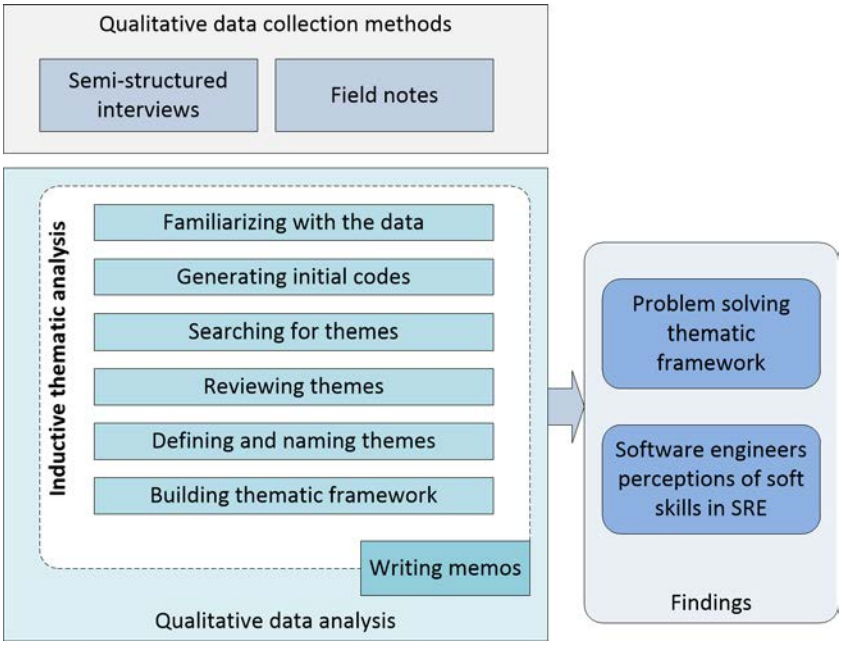


Figure 4: Qualitative methods used in a study on software requirements engineering with experts from software industry

and technological view points, but also considering people and organizational aspects. More comprehensive research on such a complex practice is possible if a research team is composed of researchers with different background, covering both technical and social aspects (sociologists, psychologists, linguists, economists, etc.). Multidisciplinary teams are necessary to research human and organizational aspects of software engineering industrial practice [33].

- *Include experts from industry in research.* Since qualitative studies on industrial practice are based on fieldwork in the real context with real experts, it is advisable to include them in data analysis and to provide them with relevant feedback to check the current research state and the relevance of the findings. This will increase the validity of the research study and the reliability of the findings.

4. Conclusion

Software engineering as one of the youngest engineering disciplines is based on mathematics and technical principles developed through a long history of engineering. However, recently it has been recognized that human and organizational factors should be understood to improve industrial practice and education of software engineers. Qualitative research methods, which originated in social science, are valuable tools for inquiring about these aspects of the practice and deserve more attention from the research community.

This work through the presentation of experience provides evidence of using qualitative

methods in different industrial contexts. The author argues that the use of qualitative methods in the research of software engineering practice can lead to better practice understanding and improvements.

References

[1] P. Bourque, R. E. D. Fairley (Eds.), Guide to the Software Engineering Body of Knowledge (SWEBOK), 3 ed., IEEE Press, Piscataway, NJ, USA, 2014.

[2] S. A. Slaughter, A Profile of the Software Industry: Emergence, Ascendance, Risks, and Rewards, 1st ed., Business Expert Press, New York, NY, USA, 2014.

[3] D. Méndez Fernández, J.-H. Passoth, Empirical software engineering: From discipline to interdiscipline, Journal of Systems and Software 148 (2019) 170–179. doi:10.1016/j.jss.2018.11.019.

[4] S. Wagner, F. Deissenboeck, Defining productivity in software engineering, in: C. Sadowski, T. Zimmermann (Eds.), Rethinking Productivity in Software Engineering, Apress, Berkeley, CA, USA, 2019, pp. 29–8. doi:10.1007/978-1-4842-4221-6_4.

[5] C. H. C. Duarte, Software productivity in practice: A systematic mapping study, Software 1 (2022) 164–214. doi:10.3390/software1020008.

[6] Y. Dittrich, C. B. Michelsen, P. Tell, P. Lous, A. Ebdrup, Exploring the evolution of software practices, in: Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ESEC/FSE 2020, Virtual Event, USA, 2020, pp. 493–504. doi:10.1145/3368089.3409766.

[7] D. Silverman, Qualitative research: meanings or practices?, Information Systems Journal 8 (1998) 3–20. doi:10.1046/j.1365-2575.1998.00002.x.

[8] M.-A. Storey, N. A. Ernst, C. Williams, E. Kalliamvakou, The who, what, how of software engineering research: a socio-technical framework, Empirical Software Engineering 25 (2020) 4097–4129. doi:10.1007/s10664-020-09858-z.

[9] M. Coccia, An introduction to the methods of inquiry in social sciences, Journal of Social and Administrative Sciences 5 (2018) 116–126. doi:10.1453/jsas.v5i2.1651.

[10] P. J. Ethington, E. L. McDonagh, The common space of social science inquiry, Polity 28 (1995) 85–90. doi:10.2307/3235187.

[11] C. Marshall, G. B. Rossman, Designing Qualitative Research, 6th ed., SAGE Publications, Thousand Oaks, US, 2016.

[12] C. Wohlin, A. Aurum, Towards a decision-making structure for selecting a research design in empirical software engineering, Empirical Software Engineering 20 (2015) 1427–1455. doi:10.1007/s10664-014-9319-7.

[13] K. Rönkkö, Interpretation, interaction and reality construction in software engineering: An explanatory model, Information and Software Technology 49 (2007) 682–693. doi:10.1016/j.infsof.2007.02.014.

[14] C. B. Seaman, Qualitative methods in empirical studies of software engineering, IEEE Transactions on Software Engineering 25 (1999) 557–572. doi:10.1109/32.799955.

[15] T. C. Lethbridge, S. E. Sim, J. Singer, Studying software engineers: Data collection techniques for software field studies, Empirical Software Engineering 10 (2005) 311–341.

doi:10.1007/s10664-005-1290-x.
[16] Z. Stojanov, D. Dobrilovic, B. Perisic, Integrating software change request services into virtual laboratory environment: Empirical evaluation, Computer Applications in Engineering Education 22 (2014) 63–71. doi:10.1002/cae.20529.
[17] K. Charmaz, Constructing Grounded Theory: A Practical Guide through Qualitative Analysis, 1st ed., Sage Publications, London, UK, 2006.
[18] Z. Stojanov, D. Dobrilovic, V. Jevtic, Identifying properties of software change request process: Qualitative investigation in very small software companies, in: Proceedings of the 9th IEEE International Symposium on Intelligent Systems and Informatics (SYSI 2011), Subotica, Serbia, 2011, pp. 47–52. doi:10.1109/SISY.2011.6034369.
[19] Z. Stojanov, Using qualitative research to explore automation level of software change request process: A study on very small software companies, Scientific Bulletin of The "Politehnica" University of Timişoara: Transactions on Automatic Control and Computer Science 57 (2012) 31–40.
[20] L. M. Given (Ed.), The SAGE Encyclopedia of Qualitative Research Methods, SAGE Publicationsi, Thousand Oaks, CA, USA, 2008.
[21] M. Birks, Y. Chapman, K. Francis, Memoing in qualitative research: Probing data and processes, Journal of Research in Nursing 13 (2008) 68–75. doi:10.1177/1744987107081254.
[22] Z. Stojanov, Qualitative research on practice in small software companies, in: M. Khosrow-Pour (Ed.), Encyclopedia of Information Science and Technology, 3rd ed., IGI Global, Hershey, PA, USA, 2015, pp. 650–658. doi:10.4018/978-1-4666-5888-2.ch062, dOI: 10.4018/978-1-4666-5888-2.ch062.
[23] Z. Stojanov, Software maintenance improvement in small software companies: Reflections on experiences, in: I. Bychkov, A. Tchernykh, A. Feoktistov (Eds.), Proceedings of the 3rd International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2021), volume 2913 of *CEUR Workshop Proceedings*, Irkutsk, Russia, 2021, pp. 182–197.
[24] Z. Stojanov, J. Stojanov, D. Dobrilovic, A lightweight inductive method for process assessment based on frequent feedback: A study in a micro software company, Journal of Engineering Management and Competitiveness 9 (2019) 134–147. doi:10.5937/jemc1902134S.
[25] Z. Stojanov, J. Stojanov, D. Dobrilovic, Knowledge discovery and systematization through thematic analysis in software process assessment project, in: Proceedings of the IEEE 13th International Symposium on Intelligent Systems and Informatics, SISY 2015, Subotica, Serbia, 2015, pp. 25–30. doi:10.1109/SISY.2015.7325405.
[26] Z. Stojanov, Thematic knowledge framework on human factor in software maintenance practice: A study in a micro software company, Journal of Software Engineering & Intelligent Systems 4 (2019) 41–57.
[27] V. Braun, V. Clarke, Using thematic analysis in psychology, Qualitative Research in Psychology 3 (2006) 77–101. doi:10.1191/1478088706qp063oa.
[28] T. Zorić, Z. Stojanov, Problem solving in software requirements elicitation and specification: Preliminary results from a qualitative study, in: Proceedings of the 7th International conference on Applied Internet and Information Technologies (AIIT2017), Zrenjanin,

Serbia, 2017, pp. 258–266.
[29] T. Zoric, Z. Stojanov, Software developers' perceptions of soft skills in software requirements engineering, Journal of Engineering Management and Competitiveness 8 (2018) 54–64. doi:10.5937/jemc1801054Z.
[30] N. H. Wolfinger, On writing fieldnotes: collection strategies and background expectancies, Qualitative Research 2 (2002) 85–93. doi:10.1177/1468794102002001640.
[31] S. N. Hesse-Biber, R. B. Johnson (Eds.), The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry, Oxford University Press, New York, NY, USA, 2015.
[32] J. A. Maxwell, Understanding and validity in qualitative research, Harvard Educational Review 62 (1992) 279–300. doi:10.17763/haer.62.3.8323320856251826.
[33] M. L. Disis, J. T. Slattery, The road we must take: Multidisciplinary team science, Science Translational Medicine 2 (2010) 22cm9. doi:10.1126/scitranslmed.3000421.

Мультиагентное управление потоками вычислительных заданий в гетерогенной распределенной вычислительной среде

Игорь Бычков [1], Александр Феоктистов [1], Роман Костромин [1]

[1] *Институт динамики систем и теории управления им. В.М. Матросова СО РАН, ул. Лермонтова, д. 134, 664033, Иркутск, Россия*

Аннотация

Рассматривается проблема управления вычислениями в гетерогенной распределенной вычислительной среде. Предложен новый подход к планированию вычислений и распределению ресурсов с использованием мультиагентных технологий. В рамках данного подхода предложены три модели согласования пользовательских критериев качества решения задач в вычислительной среде и предпочтений провайдеров ресурсов. Разработанная мультиагентная система показала преимущества в управлении вычислениями по совокупности критериев пользователей и предпочтений провайдеров ресурсов по сравнению с известными метапланировщиками GridWay и Condor DAGMan.

Ключевые слова

Управление вычислениями, мультиагентная система, гетерогенная среда, потоки заданий, мониторинг

1. Введение

Анализ современных тенденций развития в области параллельных и распределенных вычислений показывает стремительное увеличение масштаба и общей производительности создаваемых гетерогенных распределенных вычислительных сред, широкий спектр больших научных и прикладных задач, требующих их использования, необходимость обеспечения требуемого уровня обслуживания заданий в процессе решения таких задач [1]. В связи с этим возникают проблемы учета специфики предметных областей решаемых задач, разрешения противоречий между пользовательскими критериями решения задач и предпочтениями провайдеров ресурсов, смягчения разного рода неопределенностей при планировании вычислений и распределении ресурсов, обеспечения гибкого обслуживания вычислительных заданий и поддержки масштабируемости распределенных вычислений в гетерогенной среде.

Для решения подобных проблем с целью улучшения качества выполнения заданий (сокращения времени и стоимости решения задач, повышения эффективности использования ресурсов, балансировки их нагрузки и др.) целесообразно применять методы и средства, базирующиеся на интеллектуализации управления вычислениями с помощью мультиагентных технологий [2], а также экономические механизмы регулирования спроса и предложения ресурсов [3].

В этой связи именно такие методы, средства и механизмы применяются в рамках предложенного подхода к планированию вычислений и распределению ресурсов. В данной работе представлены основные аспекты мультиагентного управления потоками вычислительных заданий в гетерогенной распределенной вычислительной среде.

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: idstu@icc.ru (A. 1); agf65@yandex.ru (A. 2); roman@kostromin.net (A. 3)

ORCID: 0000-0002-1765-0769 (A. 1); 0000-0002-9127-6162 (A. 2); 0000-0001-8406-8106 (A. 3)

ICCS-DE 2023 Workshop Proceedings

2. Мультиагентная система управления вычислениями

В рамках предложенного подхода управление вычислениями на уровне среды реализуется мультиагентной системой (МАС) [4], относящейся к классу метапланировщиков. МАС обладает иерархической структурой, в рамках которой агенты строят отношения между собой на основе принципов самоорганизации путем локальных взаимодействий. На каждом уровне системы агенты могут играть разные роли, обладать специфическими правами и обязанностями, делегируемыми им пользователями и провайдерами вычислительных ресурсов среды.

Структура МАС представлена на рисунке 1. Система включает агентов пользователей, планирования вычислений, классификации и конкретизации заданий, распределения ресурсов и их мониторинга. В процессе распределения вычислительной нагрузки по выполнению задания агенты, представляющие ресурсы объединяются в виртуальное сообщество и выбирают координатора. Распределение ресурсов агентами виртуального сообщества осуществляется с помощью экономических механизмов регулирования их спроса и предложения в рамках тендера вычислительных работ.

Тендер реализуется на основе аукциона Викри [5]. В рамках тендера может быть выбрана одна из трех моделей согласования пользовательских критериев и предпочтений провайдеров. Две модели ориентируются соответственно на пользователей и провайдеров. В третьей модели используются неупорядоченные критерии и предпочтения.

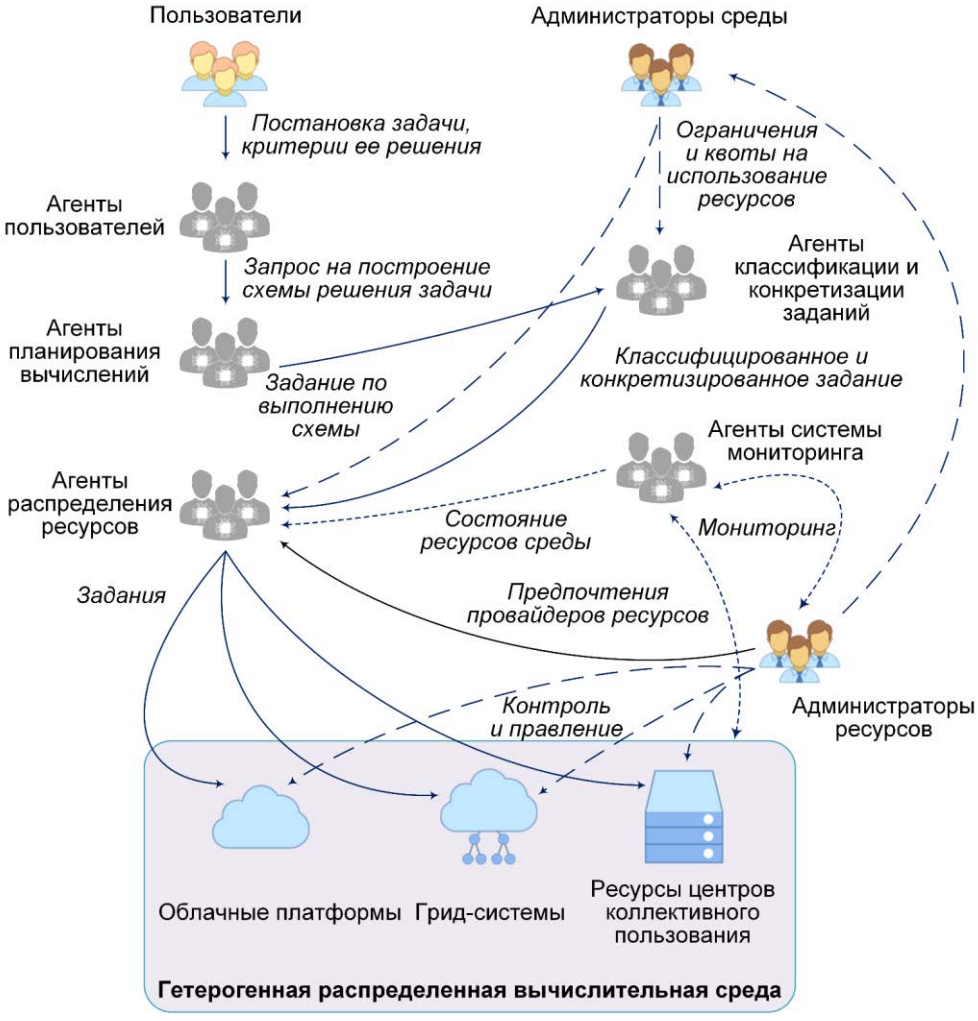


Рисунок 1: Структура МАС

Система мониторинга включает следующие уровни: пользовательский уровень, предоставляющий доступ администраторам к системным компонентам других уровней в интерактивном или пакетном режимах; уровень доступа, реализующий проверку прав доступа и серверную часть пользовательского интерфейса; уровень управления системой мониторинга в целом; промежуточный уровень распределения потоков инструкций и данных системы мониторинга в специализированных выделенных узлах; уровень, предназначенный для сбора и обработки первичных данных с вычислительных узлов, которые поступают от контрольно-измерительного оборудования. Система мониторинга использует распределенную циклическую базу данных для хранения ретроспективной и текущей информации о состоянии вычислительного и инфраструктурного оборудования, а также вычислительной истории выполнения заданий.

Система мониторинга предназначена для поддержки процессов администрирования, конфигурирования и настройки программно-аппаратного обеспечения гетерогенной вычислительной среды, а также контроля и управления вычислительными ресурсами. В частности, в системе мониторинга реализованы средства тестирования вычислительных узлов, выделяемых для выполнения заданий, и освобождения их системных ресурсов после завершения выполнения этих заданий с целью повышения эффективности использования узлов. Схема тестирования узлов и освобождения их ресурсов [6] приведена на рисунке 2.

Рисунок 2: Схема тестирования узлов и освобождения их ресурсов

Для нахождения значений пользовательских критериев и предпочтений провайдеров разработаны модели оценки времени, надежности и стоимости выполнения заданий, а также выбраны методы определения ускорения вычислений, эффективности использования ресурсов и балансировки их загрузки [7]. Базовые методы и средства машинного обучения агентов различного назначения представлены в [8]. Дополнительно разработана вероятностная нейронная сеть для агентов классификации и конкретизации заданий [9]. Вся необходимая информация о вычислительных ресурсах и процессах для применения вышеупомянутых моделей, методов и средств предоставляется системой мониторинга.

3. Вычислительные эксперименты

Оценка работы MAC при распределении потоков заданий для трех предложенных моделей тендера проведена в экспериментальной среде с двумя пулами ресурсов (таблица 1). Потоки заданий сформированы на основе типовых рабочих процессов Montage, CyberShake, Epigenomics, LIGO и SIPHT. Сравнение с GridWay и Condor DAGMan проводилось по стоимости и времени выполнения потоков, ускорению вычислений, эффективности использования ресурсов, средней загрузке процессора и среднеквадратическому отклонению от нее, характеризующему степень балансировки загрузки ресурсов.

Таблица 1

Ресурсы вычислительной среды

Ресурс	Описание
Пул 1	8 виртуальных машин со следующими характеристиками: 1 процессор Intel Xeon E5506 (1 core, 2.13 GHz, 4 MB L3 cache, 2 GB RAM DDR3-800, 4 FLOP/cycle)
Пул 2	10 виртуальных машин со следующими характеристиками: 1 процессор AMD Opteron 6276 (8 core, 2.3 GHz, 16 MB L3 cache, 32 GB RAM DDR3-1600, 4 FLOP/cycle)

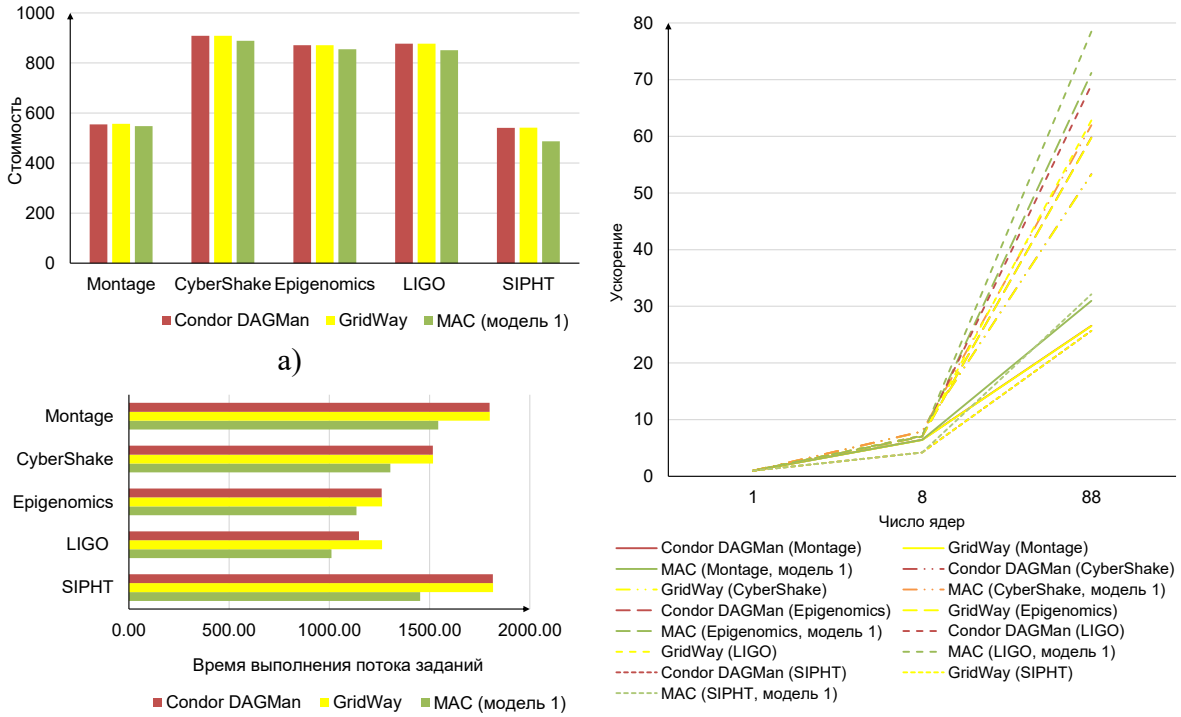


Рисунок 3: Показатели обработки потоков заданий метапланировщиками и MAC (модель 1, ориентированная на пользователей): стоимость (а) и время (б) выполнения заданий; ускорение вычислений (в)

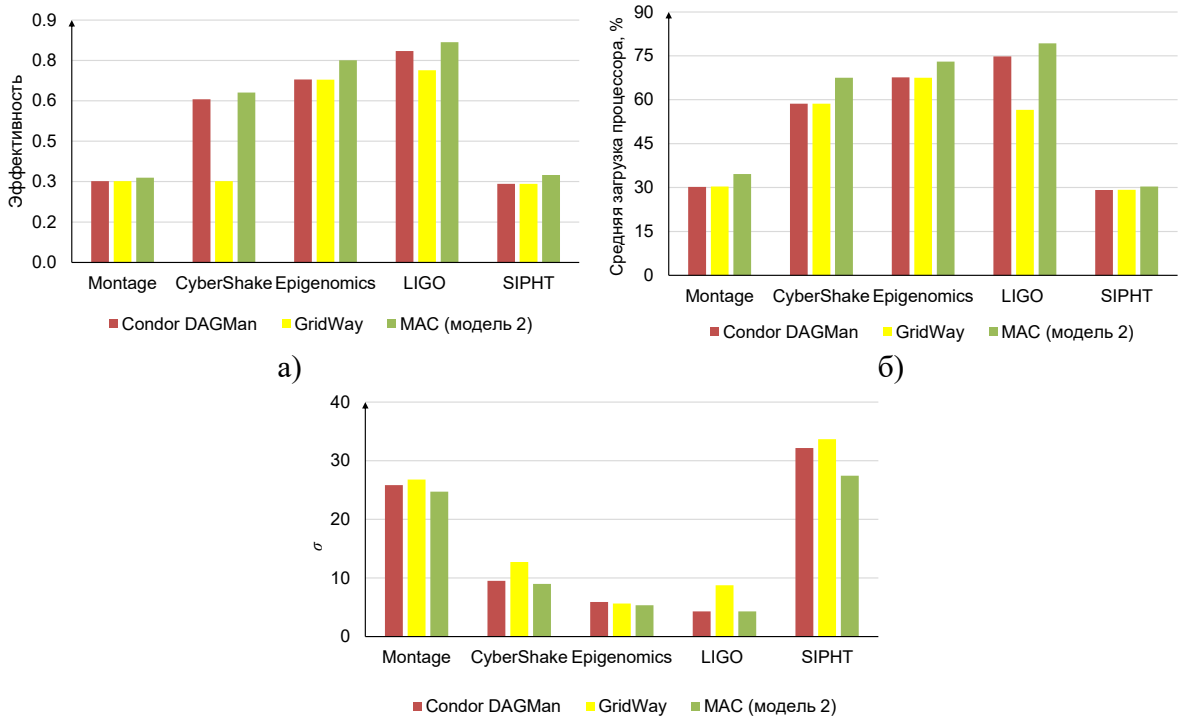


Рисунок 4: Показатели использования ресурсов метапланировщиками и MAC (модель 2, ориентированная на провайдеров): эффективность (а), средняя загрузка процессора (б), среднеквадратическое отклонение σ от средней загрузки процессора (в)

MAC с моделью 1 продемонстрировала преимущество по всем критериям пользователя для каждого потока заданий в сравнении с Condor DAGMan и GridWay (рисунок 3). В тоже время MAC с моделью 2 показала превосходство по всем предпочтениям провайдеров для каждого потока в сравнении с Condor DAGMan и GridWay. При неупорядоченном множестве критериев MAC обеспечивает сопоставимые результаты в сравнении с Condor DAGMan и GridWay.

4. Практическое применение

Разработанные модели, методы и средства мультиагентного управления вычислениями успешно применены на практике в процессе решения ряда прикладных задач. В их числе оптимизация складской логистики [10], выявление критических элементов газотранспортной сети России [11] и исследование направлений развития топливно-энергетического комплекса Вьетнама с учетом требований энергетической безопасности [12].

Применение перечисленных разработок позволило существенно ускорить процесс вычислений, сократить время решения задач и повысить эффективность использования ресурсов гетерогенной распределенной вычислительной среды.

5. Заключение

В рамках решения проблемы управления вычислениями в гетерогенной распределенной вычислительной среде рассмотрен новый подход к планированию вычислений и распределению ресурсов, основанные на использовании агентов. Разработана мультиагентная система, базирующаяся на применении трех моделей управления вычислениями: двух моделей, ориентированных соответственно на пользователей и провайдеров ресурсов, а также модели с неупорядоченными критериями и предпочтениями.

Модельные эксперименты показали очевидное преимущество MAC по сравнению с метапланировщиками GridWay и Condor DAGMan в рамках двух моделей, ориентированных

на пользователей и провайдеров. В рамках третьей модели с неупорядоченными критериями и предпочтениями МАС показала сопоставимые результаты с вышеупомянутыми метапланировщиками. Полученные преимущества в управлении вычислениями подтверждены результатами решения ряда практических задач.

Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации, проект № FWEW-2021-0005 «Технологии разработки и анализа предметно-ориентированных интеллектуальных систем группового управления в недетерминированных распределенных средах».

6. Список литературы

[1] A. Feoktistov, S. Gorsky, I. Sidorov, R. Kostromin, A. Edelev, and L. Massel ”Orlando Tools: Energy Research Application Development through Convergence of Grid and Cloud Computing,” Communications in Computer and Information Science, 2019, vol. 965, pp. 289–300.

[2] И.А. Каляев, А.И. Каляев “Метод и алгоритмы адаптивного мультиагентного диспетчирования ресурсов в гетерогенных распределенных вычислительных средах,” Автоматика и телемеханика, 2022, № 8, с. 100–122.

[3] В.В. Топорков, Д.М. Емельянов, А.С. Топоркова “Метапланирование и управление ресурсами в ГРИД,” ИТНОУ: информационные технологии в науке, образовании и управлении, 2017, № 3 (3), с. 72–80.

[4] А.Г. Феоктистов, Р.О. Костромин, Ю.А. Дядькин “Управление заданиями в гетерогенной распределенной вычислительной среде на основе знаний,” Вестник компьютерных и информационных технологий, 2018, № 2, с. 10–17.

[5] A. Feoktistov “Tender of computational works in heterogeneous distributed environment,” Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments, CEUR-WS Proceedings, 2020, vol. 2638, pp. 99–108.

[6] И.А. Сидоров, А.Г. Феоктистов “Средство тестирования выделяемых узлов кластера и осво-бождения их ресурсов в процессе обработки потока заданий,” Суперкомпьютерные дни в России: Труды международной конференции, М.: МАКС Пресс, 2019, с. 229–230.

[7] А.Г. Феоктистов “Организация предметно-ориентированных распределенных вычислений в гетерогенной среде на основе мультиагентного управления заданиями,” Автореф. дис. докт. техн. наук: 05.13.11, Иркутск: Изд–во ИДСТУ СО РАН, 2022, 38 с.

[8] А.Г. Феоктистов, Р.О. Костромин “Система машинного обучения агентов управления распределенными вычислениями,” XII мультиконференция по проблемам управления: Материалы XII мультиконференции, Ростов-на-Дону, Таганрог: Изд-во Южного федерального университета, 2019, с. 216–218.

[9] А.Г. Феоктистов ”Вероятностная нейронная сеть классификации вычислительных заданий,” Прогрессивные научные исследования – основа современной инновационной доктрины: сборник статей Международной научно–практической конференции (г. Киров, РФ, 25 ноября 2022 г.), Уфа: Аэтерна, 2022, с.131–135.

[10] A. Tchernykh, A. Feoktistov, S. Gorsky, I. Sidorov, R. Kostromin, I. Bychkov, O. Basharina, A. Alexandrov, and R. Rivera-Rodriguez “Orlando Tools: Development, Training, and Use of Scalable Applications in Heterogeneous Distributed Computing Environments,” Communications in Computer and Information Science, 2019, vol. 979, pp. 265–279.

[11] И.В. Бычков, С.А. Горский, А.В. Еделев, Р.О. Костромин, И.А. Сидоров, А.Г. Феоктистов, Е.С. Фереферов, Р.К. Федоров ”Поддержка управления живучестью систем энергетики на основе комбинаторного подхода,” Известия РАН. Теория и системы управления, 2021, № 6б с. 122–135.

[12] A. Edelev, V. Zorkaltsev, S. Gorsky, V.B. Doan, H.N. Nguyen ”The Combinatorial Modelling Approach to Study Sustainable Energy Development of Vietnam,” Communications in Computer and Information Science, 2017, vol. 793, pp. 207–218.

Численное моделирование и экспериментальные методы в анализе смежных волновых полей и геофизических сред Байкальской рифтовой зоны

Марат Хайретдинов [1,2], Александр Михайлов[1], Дарья Пинигина[2]

[1] *Институт вычислительной математики и математической геофизики СО РАН, Новосибирск, пр.акад. Лаврентьева 6, 630090, Россия*
[2] *Новосибирский государственный технический университет, пр. Карла Маркса, 20, Новосибирск, 630092, Россия*

Аннотация

На основе численного метода решения прямой и обратной задач и данных экспериментов анализируются процессы одновременного распространения волновых полей в смежных средах- земля, вода, атмосфера, лед- сложнопостроенной геофизической структуры Байкальской рифтовой зоны (БРЗ) на участке пос. Бабушкин (юго-восточная часть Байкала) – пос. Бугульдейка (северо-западная часть Байкала). Применяемый в работах вибрационный источник ЦВ-100 одновременно излучает сейсмические колебания в землю и акустические в атмосферу. Численный алгоритм решения прямой задачи восстановления сейсмического волнового основан на применении интегрального преобразования Лагерра по времени и конечно-разностной аппроксимации по пространственным координатам. Численная модель среды, используемая при расчётах распространения сейсмических волн, задавалась с учетом априорных данных о скоростном разрезе БРЗ. В качестве подхода к решению обратной задачи восстановления скоростных характеристик неоднородной среды предложен и апробирован вычислительный сеточный алгоритм, основанный на вычислении средневзвешенных скоростей в участках сетки, накладываемой на поверхность земли. Рассмотрены вопросы многолучевого распространения гидроакустических волн в водной толще оз. Байкал, образующихся на границе перехода «суша-вода» вследствие трансформации сейсмических волн. Дальнее распространение акустических волн в атмосфере от источника ЦВ-100 обязано эффектам приповерхностного волноводного распространения вдоль дневной поверхности земли, а также возможного высотного лучевого распространения с отражениями от градиентных границ в неоднородной атмосфере.

Keywords

Байкальская рифтовая зона, численное моделирование, вибрационный источник, сейсмо-гидро-акустическое волновое распространение, мгновенный снимок поля, синтетические сейсмограммы, сеточный алгоритм, скоростной разрез, сопоставительный анализ

1. Введение

В проблеме мониторинга сейсмоактивной Байкальской рифтовой зоны (БРЗ) все большую роль играют экологически чистые методы активного мониторинга сложно построенных структур как земной части БРЗ, так и водной толщи оз. Байкал. Здесь перспективным

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: marat@opg.sscc.ru (A. 1)

ICCS-DE 2023 Workshop Proceedings

высокоразрешающим и экологически чистым является метод вибрационного просвечивания Земли, позволяющий получать с высокой повторяемостью отклики среды в ответ на акты зондирования мощным вибрационным источником ЦВ-100 с амплитудой возмущающей силы 100 тс. [1,2]. Источник установлен на берегу Байкала в районе пос. Бабушкин (юго-восточная часть оз.Байкал. Кроме сейсмических волн, источник излучает в атмосферу акустические волны [3].

С использованием вибратора силами Института физики Земли РАН, Института вычислительной математики и математической геофизики СО РАН, Института геологии СО РАН в 2021 г. были выполнены эксперименты по зондированию района БРЗ вдоль профиля п.Бабушкин-п. Бугульдейка (северо-западная часть оз.Байкал). Схема зондирования представлена на рис.1. Здесь изображены типы излучаемых волн, используемые датчики и места их расстановки.

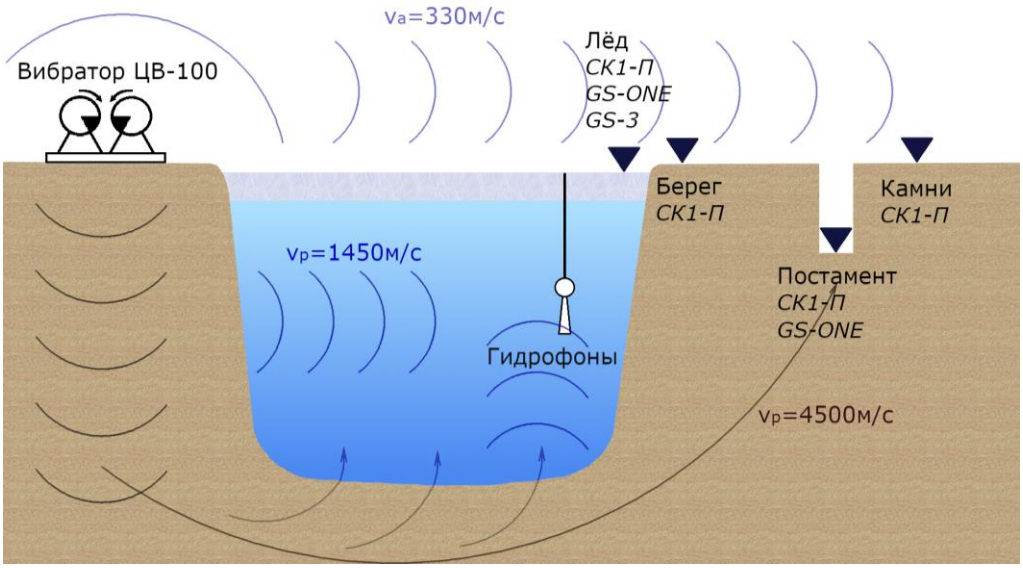


Рисунок 1: Схема проведения экспериментов на Байкале по вибрационному просвечиванию сопряженных сред «земля-вода-лед-атмосфера. Представлены : сейсмический вибратор ЦВ-100 с возмущающей силой 100 тс в диапазоне частот 6.25-10.05 Гц; СК-1П, GS-3, GS-ONE--сейсмические трехкомпонентные датчики, гидрофоны-датчики гидроакустических колебаний, «постамент»-сейсмическая обсерватория на глубине 3 м; «лед», «берег», «камни», «гидрофоны»- наименование мест установки датчиков.

Экспериментально получаемая пространственно - временная картина сейсмического волнового поля в земле и гидроакустического в воде как результат вибрационного зондирования имеют сложный характер для интерпретации. В этой ситуации для прогнозирования и повышения возможностей интерпретации картины волновых полей большую роль играет численное моделирование процессов распространения волн, рассчитываемых с учетом априорных данных по глубинным скоростным разрезам БРЗ. На сегодня такие разрезы получены рядом отечественных и зарубежных исследователей с применением широкого спектра исследований – глубинного сейсмического зондирования Земли (ГСЗ), метода преломленных волн (МПВ, КМПВ) и др. Обзор многолетних исследований в этой области представлен в работе [4].

В настоящей работе авторы опираются на современные данные по скоростному разрезу БРЗ, полученные в работе датских ученых C.Nielsen, H. Thybo [5]. Ими проанализированы свойства осадочных отложений под оз. Байкал, включая скоростной разрез среды по бортам озера на базе протяженностью 350 км. С учетом приведенных результатов в настоящей статье представлены численные методы моделирования волновых полей и восстановления скоростных характеристик неоднородных упругих сред Байкальской рифтовой зоны, основанные на решении прямых и обратных задач сейсмики. Рассматриваются вопросы

многолучевого распространения гидроакустических волн в водной толще оз. Байкал, образующихся на границе перехода «суша-вода» как результата трансформации сейсмических волн в гидроакустические. Распространение последних происходит в волноводе «лед-дно» озера с многочисленными переотражениями от границ, характерными для градиентных сред, образующихся по глубине озера- слои с выраженными температурными градиентами, подводный звуковой канал (ПЗК), границы разноплотности сред, обусловленные перепадами давлений. Экспериментально полученные записи волн в воде отражают сложность структуры волнового поля в воде.

При вибрационном зондировании смежных сред рассматриваются феномены дальнего распространения акустических волн в атмосфере от источника ЦВ-100, связанные с процессами приповерхностного волноводного распространения вдоль дневной поверхности земли, а также дополнительно высотного лучевого распространения с отражениями от градиентных границ в неоднородной атмосфере. Показана согласованность выводов по результатам численного моделирования и экспериментов.

Решение рассматриваемых задач тесно связано с созданием вычислительной технологии изучения пространственно-временной динамики волновых полей в связи с повышенной сейсмоактивностью района БРЗ.

2. Численная модель среды, алгоритм и результаты решения динамической задачи расчета волновых полей

Численная модель среды, используемая при расчётах распространения сейсмических волн, задавалась с учетом априорных данных о скоростном разрезе БРЗ [5]. Соответствующее графическое отображение численной модели среды приводится на рис.2. На представленном рисунке изображены границы слоёв и подписаны значения скоростей продольных волн V_P в этих слоях. Значения поперечных волн задавались по формуле $V_S = V_P / \sqrt{3}$. Плотность рассчитывалась по известной формуле Гарднера $\rho = 1.74 * V_p^{0.25}$. Физические характеристики слоя воды – скорость продольной волны $V_p = 1480$ м/сек, плотность $\rho = 1.0$ г/см^3. При расчётах задавалась ограниченная по пространству область размерностью $(x, z) = (90\,км, 45\,км)$. Для подавления отражения волн на границах, ограничивающих заданную область, был использован способ ограничения расчетной области идеально поглощающими PML слоями (PML аббревиатура английского Perfectly Matched Layers). Этот подход предложен для численных расчетов упругих волновых полей в работах [6,7].

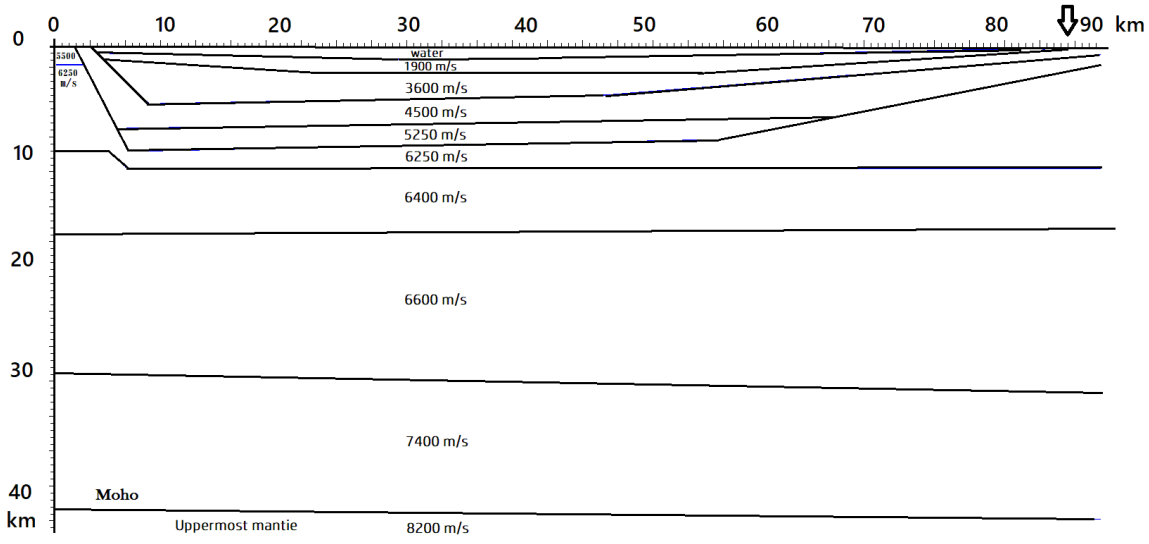


Рисунок 2: Численная модель скоростного разреза БРЗ.

2.1. Постановка задачи

Связь между компонентами напряжений и скоростями смещений в декартовой системе координат (x,z) для задачи распространения сейсмоакустических колебаний в упругой изотропной среде записывается как:

$$\begin{cases} \dfrac{\partial u_x}{\partial t} = \dfrac{1}{\rho}\left(\dfrac{\partial \sigma_{xx}}{\partial x} + \dfrac{\partial \sigma_{xz}}{\partial z} \right) + F_x(x,z) f(t) \\ \dfrac{\partial u_z}{\partial t} = \dfrac{1}{\rho}\left(\dfrac{\partial \sigma_{xz}}{\partial x} + \dfrac{\partial \sigma_{zz}}{\partial z} \right) + F_z(x,z) f(t) \\ \dfrac{\partial \sigma_{zz}}{\partial t} = (\lambda + 2\mu)\dfrac{\partial u_z}{\partial z} + \lambda \dfrac{\partial u_x}{\partial x} \\ \dfrac{\partial \sigma_{xx}}{\partial t} = (\lambda + 2\mu)\dfrac{\partial u_x}{\partial x} + \lambda \dfrac{\partial u_z}{\partial z} \\ \dfrac{\partial \sigma_{xz}}{\partial t} = \mu\left(\dfrac{\partial u_x}{\partial z} + \dfrac{\partial u_z}{\partial x} \right) \end{cases} \quad (1)$$

В этих уравнениях (u_x, u_z) - компоненты скорости смещения, $(\sigma_{xx}, \sigma_{zz}, \sigma_{xz})$ - компоненты тензора напряжений, $\rho(x,z)$ - плотность среды, $\lambda(x,z)$ и $\mu(x,z)$ - коэффициенты Ламе. F_x, F_z -составляющие силы $\vec{F}(x,z) = F_x\vec{e}_x + F_z\vec{e}_z$, описывающей распределение локализованного в пространстве источника. В качестве моделируемого источника F_x, F_z выбирается источник вертикальной силы: $F_x = 0,\ F_z = \delta(x - x_0)\delta(z - z_0)$; $f(t)$ - моделируемый временной сигнал в источнике с координатами (x_0, z_0).

Задача решается при нулевых начальных данных:

$$u_x\big|_{t=0} = u_z\big|_{t=0} = \sigma_{xx}\big|_{t=0} = \sigma_{zz}\big|_{t=0} = \sigma_{xz}\big|_{t=0} = 0 \ . \quad (2)$$

Решение рассматривается на полупространстве $z \geq 0$, с граничными условиями на свободной поверхности:

$$\sigma_{xz}(x,z,t)\big|_{z=0} = \sigma_{zz}(x,z,t)\big|_{z=0} = 0 \, . \quad (3)$$

Полагаем параметры среды $\rho(x,z), \lambda(x,z), \mu(x,z)$ - кусочно-непрерывными функциями переменных x, z.

2.2. Алгоритм решения

Алгоритм решения задачи основан на применении интегрального преобразования Лагерра по времени и конечно-разностной аппроксимации по пространственным координатам. На первом этапе, к задаче (1)-(3) применим интегральное преобразование Лагерра по времени. Для некоторой функции $F(t)$ интегральное преобразование Лагерра определяется как:

$F^m = \int_0^\infty F(t)(ht)^{-\frac{\alpha}{2}} l_m^\alpha(ht) d(ht)$ с формулами обращения:

$F(t) = (ht)^{\frac{\alpha}{2}} \sum_{p=0}^{\infty} \dfrac{m!}{(m+\alpha)!} F^p l_m^\alpha(ht)$, где $l_m^\alpha(ht)$ - ортонормированные функции Лагерра:

$$\int_0^{\infty} l_m^{\alpha}(ht) l_n^{\alpha}(ht) d(ht) = \delta_{mn} \frac{(m+\alpha)!}{m!} \ .$$

Функции Лагерра $l_m^{\alpha}(ht)$ выражаются через классические стандартизованные многочлены Лагерра $L_m^{\alpha}(ht)$. Выбираем параметр α целым и положительным, тогда:

$$l_m^{\alpha}(ht) = (ht)^{\frac{\alpha}{2}} e^{-\frac{ht}{2}} L_m^{\alpha}(ht) .$$

В результате данного преобразования исходная задача (1)-(3) сводится к двумерной пространственной дифференциальной задаче в спектральной области. Для решения ее применяется метод конечно-разностной аппроксимации производных по пространственным координатам на сдвинутых сетках с 4-ым порядком точности [8]. Для этого в расчетной области вводятся в направлении координаты z сетки ωz_1 и $\omega z_{1/2}$ с шагом дискретизации Δz, сдвинутые относительно друг друга на $\Delta z/2$:

$$\omega z_1 = (x, j\Delta z, t), \quad \omega z_{1/2} = (x, j\Delta z + \frac{\Delta z}{2}, t), \ j = 0,...,M \ .$$

Аналогично вводятся в направлении координаты x сетки ωx_1 и $\omega x_{1/2}$ с шагом дискретизации Δx, сдвинутые относительно друг друга на $\Delta x/2$:

$$\omega x_1 = (i\Delta x, z, t), \quad \omega x_{1/2} = (i\Delta x + \frac{\Delta x}{2}, z, t), \ i = 0,...,N \ .$$

На данных сетках вводятся операторы дифференцирования D_x и D_z, аппроксимирующие производные $\frac{\partial}{\partial x}$ и $\frac{\partial}{\partial z}$ с четвертым порядком точности по координатам $z = x_1$ и $x = x_2$:

$$D_x u(x,z) = \frac{9}{8\Delta x}\left[u(x+\frac{\Delta x}{2}, z) - u(x-\frac{\Delta x}{2}, z) \right] - \frac{1}{24\Delta x}\left[u(x+\frac{3\Delta x}{2}, z) - u(x-\frac{3\Delta x}{2}, z) \right],$$

$$D_z u(x,z) = \frac{9}{8\Delta x}\left[u(x, z+\frac{\Delta z}{2}) - u(x, z-\frac{\Delta z}{2}) \right] - \frac{1}{24\Delta x}\left[u(x, z+\frac{3\Delta z}{2}) - u(x, z-\frac{3\Delta z}{2}) \right].$$

Определим искомые компоненты вектора решения в следующих узлах сеток:

$u_x^m(x,z) \in \omega x_1 \times \omega z_1$, $u_z^m(x,z) \in \omega x_{1/2} \times \omega z_{1/2}$, $\sigma_{xx}^m(x,z)$, $\sigma_{zz}^m(x,z) \in \omega x_{1/2} \times \omega z_1$,

$\sigma_{xz}^m(x,z) \in \omega x_1 \times \omega z_{1/2}$.

В результате применения метода конечно-разностной аппроксимации в конечном счете получим систему линейных алгебраических уравнений. Представим искомый вектор решения $\vec{W}$ в следующем виде:

$$\vec{W}(m) = (\vec{V}_0(m), \vec{V}_1(m), ..., \vec{V}_{M+N}(m))^T ,$$

$$\vec{V}_{i+j} = (u_x^{i,j}, u_z^{i+1/2, j+1/2}, \sigma_{xx}^{i+1/2, j}, \sigma_{zz}^{i+1/2, j}, \sigma_{xz}^{i, j+1/2})^T .$$

Тогда данная система линейных алгебраических уравнений в векторной форме может быть записана как: $(A_{\Delta} + \frac{h}{2} E) \ \vec{W}(m) = \vec{F}_{\Delta}(m-1)$. В результате матрица системы сведённой задачи имеет хорошую обусловленность, что позволяет использовать быстрые методы решения систем линейных алгебраических уравнений на основе итерационных методов, типа сопряжённых градиентов, сходящиеся к решению с требуемой точностью всего за несколько итераций. На этом этапе проведения вычислений была реализована распараллеленная версия

метода сопряженных градиентов. На уровне входных данных при задании модели среды это равносильно декомпозиции исходной области на множество подобластей, равных количеству процессоров. Это дает возможность распределения памяти как при задании входных параметров модели, так и при дальнейшей численной реализации алгоритма в подобластях.

2.3. Результаты численного моделирования

Для моделирования по Байкальскому профилю использовался двухмерный случай постановки задачи в плоскости (x,z). Моделировалось волновое поле от локального источника типа вертикальная сила. В этом случае компоненты распределения источника в пространстве для системы уравнений (1) задавались как; $F_x = 0$, $F_z = \delta(x - x_0)\delta(z - z_0)$. Временной сигнал в источнике задавался в виде импульса Пузырёва:

$$f(t) = exp\left(-\frac{2\pi\, f_o (t - t_0)^2}{\gamma^2}\right) sin(2\pi\, f_0 (t - t_0)), \qquad (4)$$

где $\gamma = 4$, $f_0 = 8$ Гц, $t_0 = 0.125$ сек.

При этом моделируется волновое поле от точечного источника типа вертикальная сила, расположенного на поверхности с координатой $x_0 = 88\ км$ по оси **X**.(местоположение показано вертикальной стрелкой на рис.2). По отношению к заданной модели среды (рис.2) изложенная методика расчета позволяет отслеживать развитие во времени картины распространения волнового поля от источника по горизонтале и глубине. В качестве примера на рис.3 приведен мгновенный снимок волнового поля $\mathbf{U_Z}$ – компоненты для момента времени T=15 сек. Рисунок наглядно показывает распространение волновых фронтов колебаний в выбранных координатах.

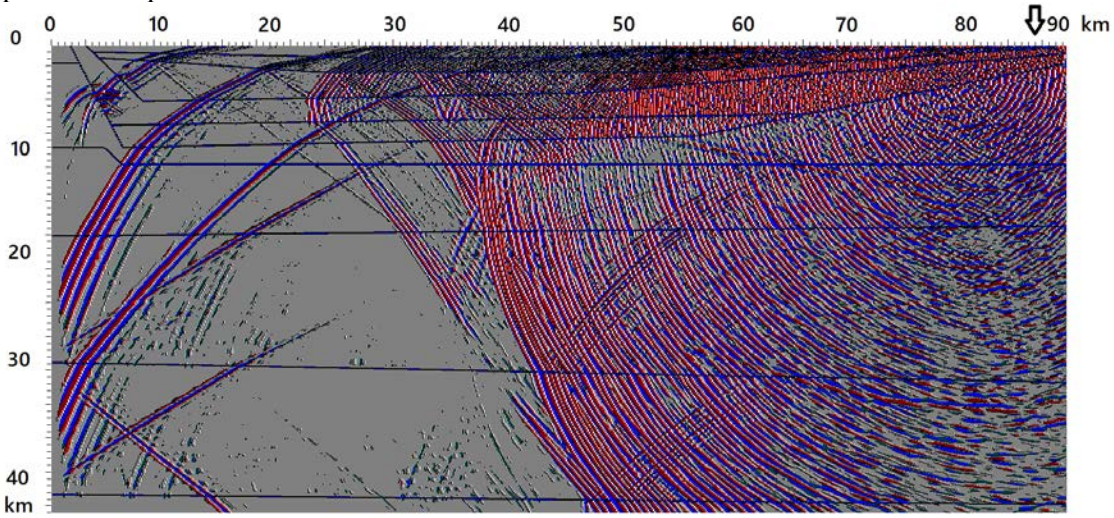


Рисунок 3: Мгновенный снимок волнового поля $\mathbf{U_Z}$ – компоненты для T=15 секунд.

По результатам расчетов получены синтетические сейсмограммы, характеризующие расположение волновых откликов в координатах «время-расстояние», где расстояние отсчитывается от источника по горизонтали. При этом получены синтетические сейсмотрассы для **P** и $\mathbf{U_z}$ – компонент волнового поля в диапазоне дальностей 76.5- 86 км. Приёмники для записи сейсмотрасс **P** компоненты давления располагались в воде на глубине $z = 20$ метров, а для $\mathbf{U_Z}$ – компоненты на поверхности земли (рис.1). Значения для **P** компоненты вычислялись по значениям нормальных компонент тензора напряжения по формуле $P = \sqrt{\sigma_{xx}^2 + \sigma_{zz}^2}$. Координата первого приёмника $x_1 = 2\ км$. Интервал между приёмниками по оси **X** составляет $\Delta x = 500$ метров. Первые три приёмника расположены на суше, а остальные в воде.

Приведенные параметры расстановки для численного эксперимента соответствуют тем, что были реализованы в реальном эксперименте в юго-восточной части Байкала на трассе пос. Бабушкин- пос. Бугульдейка.

На основе полученных волновых откликов оцениваются времена вступления волн, которые в дальнейшем используются для расчета скоростных характеристик среды по глубине. Алгоритм расчета рассматривается ниже.

3. Алгоритм и результаты оценивания скоростных характеристик сложнопостроенной среды

Решение задачи восстановления скоростной характеристики среды, по сути своей, является обратной, заключающейся в построении скоростного разреза на основе измеренных времен вступлений волн $\vec{t}$, образующих скоростной годограф в виде некоторой нелинейной интегральной функции [9]. Часто из-за большой размерности $\vec{t}$ применение строгих методов решения обратной задачи осложнено. Для этого случая предложен и апробирован подход к решению обратной задачи восстановления скоростных характеристик неоднородной среды в виде вычислительного сеточного алгоритма с адаптивным выбором шага сетки. Алгоритм основан на вычислении средневзвешенных скоростей в участках сетки, накладываемой на поверхность земли. При этом подразумевается уточнение расчетов скоростей в областях с наиболее выраженной гетерогенностью. Идея подхода состоит в том, что исследуемая область разбивается на участки [10], в которых вычисляются локальные скорости в соответствии с:

$$\vec{v_k} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\vec{V_{ij}}\cdot\frac{L_{ijk}}{L_{ij}}}{\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{L_{ijk}}{L_{ij}}}, \quad (i=1,\dots N), (j=1,\dots M) \qquad (5)$$

где $\vec{V_{ij}}$ — средняя скорость сейсмической волны от источника i до датчика j, N- число источников, M – число регистрирующих датчиков на линейном профиле, L_{ij} — расстояние, пройденное сейсмической волной от источника i до датчика j, L_{ijk} — расстояние, пройденное сейсмической волной от источника i до датчика j в k-ом участке. Таким образом, на всю исследуемую область накладывается сетка, узлами которой являются k-ые участки. Минимальный шаг сетки ограничивается верхней граничной частотой колебаний источника. Такой подход использован к восстановлению скоростного строения сложнопостроенной среды на выбранном для проведения экспериментов профиле пос. Бабушкин- пос. Бугульдейка..

Результат оценивания скоростной характеристики средыв зависимости от дальности зондирования методом (5) представлен на рис.4 в виде гистограмм средневзвешенных скоростей V_p в локализованных диапазонах дальностей «источник-приемник» 76-90 км, указанных на оси абцисс.

4. Дальнее распространение акустических волн в атмосфере

Многофакторная модель интегрального давления как результата распространения акустических волн в нижней атмосфере от вибратора ЦВ-100 может быть описана уравнением энергетического баланса:

$$P_{\Sigma}(t,f,r) = P_{V}(f) + P_{1}(r) + P_{2}(e,\tau,\omega,\varphi) + P_{3}(1/r^{2}) \qquad (3)$$

Здесь $P_{\Sigma}(t,f,r)$ – давление в точке регистрации на удалении r от источника; $P_{V}(f)$ – частотно зависимое акустическое давление, развиваемое источником; $P_{1}(\mathrm{r})$ – поглощение инфразвука по расстоянию, определяемое неоднородностью атмосферы и покровом дневной поверхности

Земли (лес, трава и т.д.); $P_2(e,\tau,w_0,\varphi)$ – давление в пункте регистрации как функция метеопараметров: относительной влажности, температуры, скорости и направления ветра, угла φ между направлением ветра и волновым фронтом от источника; $P_3\,(1/r^2)$ – давление как результат сферической расходимости волнового фронта.

Один из факторов, способствующих дальнему распространению инфразвука (низкочастотное акустическое колебание), связан с явлением пространственной фокусировки волн в направлении ветра [11]. Другой фактор связан с явлением отражения акустических волн от верхних слоев неоднородной атмосферы. В случае благоприятных атмосферных параметров-высотного (в км) профиля температуры, компоненты горизонтального ветра, высоты отражения от верхних слоев атмосферы – дальность распространения может достигать сотен км [12]. Экспериментальным подтверждением этого вывода являются результаты одновременной регистрации вибрационных сейсмических и акустических волн, представленных на рис.5. Здесь распространение волн в Земле отражено в виде первичной сейсмической волны со временем первого вступления 16.9 с, в атмосфере акустической волны со временем вступления 262 с.

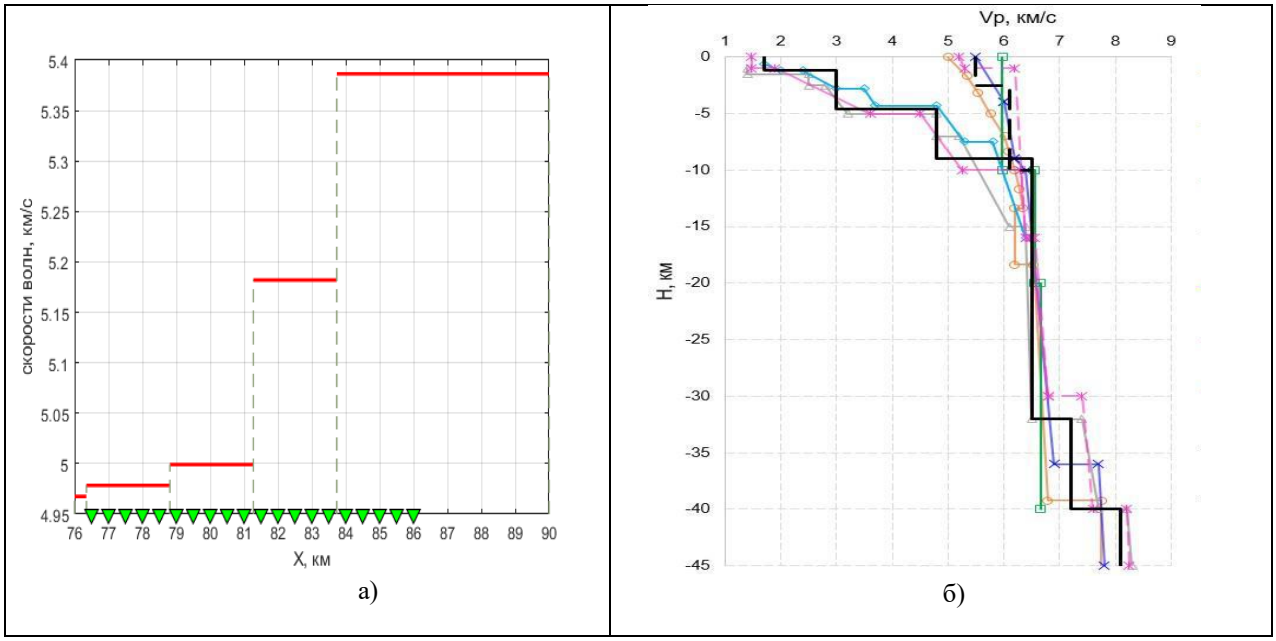


Рисунок 4: а)-гистограмма восстановленных скоростей волн V_p как функция дальности; б)-обобщенные данные по скоростям, полученные в натурных экспериментах по глубинному сейсмическому зондирования БРЗ (А.В. Беляшев, Ц.А. Тубанов [4])

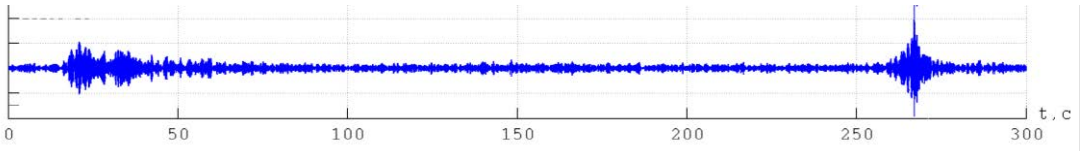


Рисунок 5: структура вибрационных волн как результат зондирования зоны оз.Байкал на удалении от вибратора 85.5 км: представлены первичные сейсмические волны со временем вступления 16.9 с и вторичные акустические волны со временем вступления 262 с.

5. Вибрационные волны в оз. Байкал

Картина гидроакустических волн в водной толще оз. Байкал имеет сложный характер вследствие выраженности эффектов переотражений, обусловленных рядом факторов. Один из основных вызван распространением волн в воде с границами «дно-лед». Наличие других

переотражений обусловлено распространением волн в подводном звуковом канале. Приповерхностный звуковой канал, впервые открытый в океане Л.М. Бреховских, получает наиболее полное развитие при образовании ледяного покрова. При наличии льда выше и ниже оси звукового канала (у нижней поверхности льда) создаются благоприятные условия для фокусировки звуковых лучей, с одной стороны, за счет льда, с другой – за счет 200-метрового слоя воды, что намного лучше, чем в летний период. В условиях зимней стратификации звук распространяется, рефрагируя по лучевым траекториям, описывающим дуги разной длины, в зависимости от угла выхода луча. Для модели сосредоточенного источника звука лучевая картина для периода времени январь-апрель, рассчитанная теоретически, представлена на рис.6а [13]. На рис.6б представлен результат регистрации волн в воде на удалении от источника 83,96 км и на глубине 20 м с помощью спускаемого гидрофона.

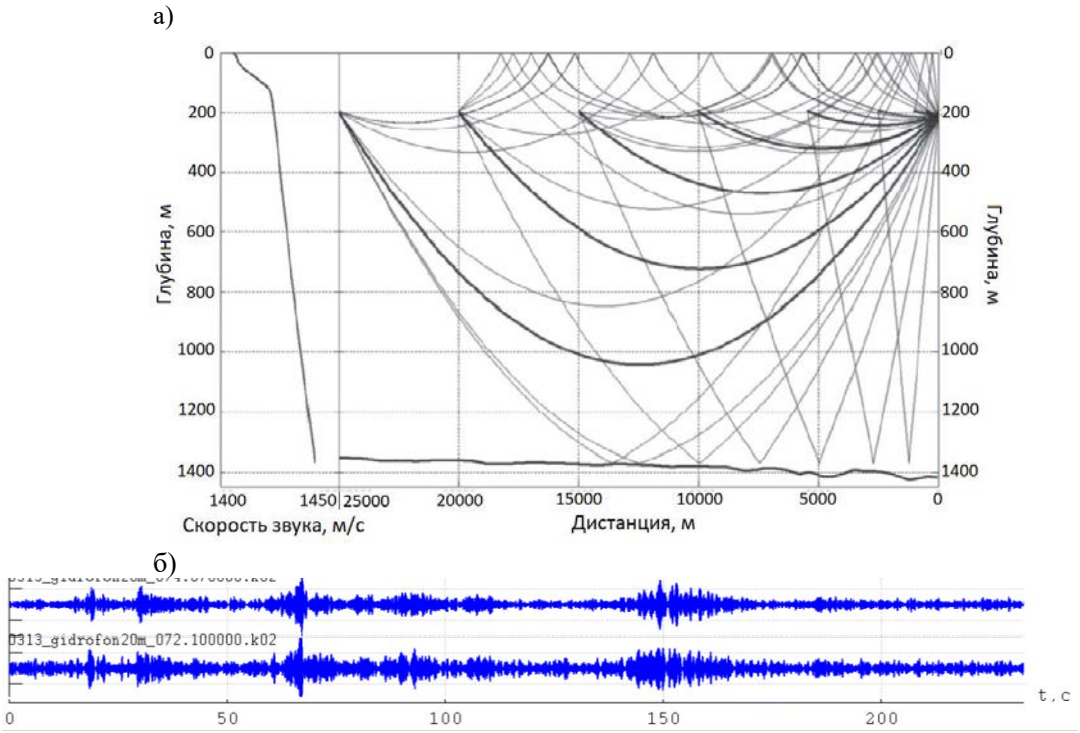


Рисунок 6: а)- лучевая картина волн в оз. Байкал; б)-результат регистрации волн в воде на глубине 20 м и удалении от источника на расстояние 83.96 км.

Из рис.6а следует, что лучи, выходящие под малыми углами скольжения, распространяясь от источника к приёмнику, отражаются от льда. Чем больше угол скольжения луча в источнике, тем глубже будет находится точка заворота луча и тем большую интегральную скорость он будет иметь. В левой части рисунка представлен вертикальный профиль скорости звука по глубине, отражающий нарастание скорости с глубиной в пределах 1402-1440 м/с.

Результаты экспериментов по регистрации волн в воде представлены в виде откликов на повторяющиеся сеансы вибрационного зондирования – рис.6б. Из анализа результатов экспериментальной регистрации в воде следует вывод о наличии многоволновой структуры отклика среды в ответ на вибрационное зондирование солитонно подобным волновым импульсом типа (4). Наряду с присутствием здесь первичных скоростных (продольных) волн на временах прихода 15.5 сек со скоростями распространения сейсмических волн в земле (около 5.35 км/с) здесь наблюдается присутствие группы волн как результата проявления процессов переотражений, что согласуется с картиной поля на рис.6а.

Из анализа результатов экспериментов по регистрации волн во льду следует, что из-за влияния трещиноватой структуры льда происходит сильное затухание волн в такой среде, вследствие чего положительные результаты регистрации являются неустойчивыми.

6. Заключение

В рамках проблемы геофизического мониторинга Байкальской рифтовой зоны на основе численного метода решения прямой и обратной задач и данных экспериментов анализируются процессы одновременного дальнего распространения волновых полей в смежных телеиле пос. Бабушкин (юго-восточная часть Байкала) – пос. Бугульдейка (северо-западная часть Байкала). На основе применения интегрального преобразования Лагерра по времени и конечно-разностной аппроксимации по пространственным координатам численным моделированием рассчитаны мгновенные снимки во времени сейсмического волнового поля, а также синтетические сейсмограммы для заданной трассы зондирования. Решение обратной задачи восстановления скоростных характеристик среды достигается на основе применения вычислительного сеточного алгоритма, показывающего хорошее согласие результатов численного моделирования и глубинного сейсмического зондирования, выполненного ранее рядом авторов в районе оз. Байкал.

Благодарности

Работа выполнена при поддержке проекта РФФИ № 20-07-00861а, госзадания FWNM–2022–0004

Литература

[1] Алексеев А.С., Глинский Б.М., Ковалевский В.В., Хайретдинов М.С. и др. Активная сейсмология с мощными вибрационными источниками / Отв. ред. Г.М. Цибульчик. Новосибирск: ИВМиМГ СО РАН, Филиал "Гео" Издательства СО РАН, 2004. 387с.

[2] Kovalevsky V. V., Fatyanov A. G., Karavaev D. A., Braginskaya L. P., Grigoryuk A. P., Mordvinova V. V., Tubano v T. A., Bazarov A. D. Research and verification of the Earth crust velocity models by mathematical simulation and active seismology methods. // Geodynamics & Tectonophysics, 2019, V.10, Iss. 3, P.569-583.

[3] Ю.М. Заславский. Излучение сейсмических волн вибрационными источниками. Институт прикладной физики РАН ,Н.-Новгород, 2007, 198 с. verification

[4] Беляшев А.В., Тубанов Ц.А. Подбор скоростных моделей для локализации сейсмических событий в пределах Байкальской рифтовой зоны. Геофизические технологии, №1, 2021, с.38-51

[5] Nielsen Christoffer, Thybo H.. Lower crustal intrusions beneath the southern Baikal Rift Zone: Evidens from full-waveform modeling of wide-angle siesmic data// Tectonophysics.-2009b.-vol.470 (3-4).-P. 298-318, doi:10.1016/j.t cto.2009.01.023.

[6] Collino F. 1996. Perfectly matched absorbing layers for the paraxial equations. J.Comput. Phys. 131(1), 164 - 180.

[7] Collino F., Tsogka C. 2001. Application of PML absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. Geophysics. 66(1), 294 - 307.

[8] Mikhailenko B.G. Spectral Laguerre method for the approximate solution of time dependent problems // Applied Mathematics Letters, 1999, № 12, P. 105–110.

[9] С.В. Гольдин. Интерпретация данных сейсмического метода отраженных волн. М.: Недра, 1979. С.343.

[10] Тагиров Х. Ю., Асланов Т. Г., Магомедов Х. Д. Определение средневзвешенной скорости сейсмической волны на участках Земли по пути ее распространения // Известия Дагестанского государственного педагогического университета. Естественные и точные науки. 2017. Т. 11. № 3. С. 108-114.

[11] Khairetdinov M.S., Kovalevsky V.V., Voskoboinikova G.M., Sedukhina G.F. Vibroseismoacoustic method in studying of geophysical fields interaction in ground atmosphere. // Proceeding of 14th International Multidisciplinary Scientific Geoconference SGEM-2014, ISBN 978-619-7105—10-0 ISSN 1314-2704. 2014, p. 925- 931(set Scopus 2015)

[12] В.Т. Гуляев, В.В.Кузнецов, В.В. Плоткин, С.Ю. Хомутов. Генерация и распространение инфразвука в атмосфере при работе мощных сейсмических сейсмовибраторов. Известия АН. Физива атмосферы и океана, 2001, т.37, №3, с. 303-312.
[13] Макаров М.М., Кучер К.М., Попов О.Е., Асламов И.А., Гранин Н.Г. Экспериментальные исследования распространения тонально-импульсных сигналов в воде оз. Байкал. // Международный журнал прикладных и фундаментальных ссследований. – 2018. – № 7. – С. 54-60; URL: https://applied-research.ru/ru/article/view.id=12329).

Методы кластеризации при анализе надёжности ЭЭС

Денис Бояркин [1], Дмитрий Крупенёв [1] и Дмитрий Якубовский [1]

[1] *Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук, Лермонтова 130, Иркутск, 664033, Россия*

Аннотация

В статье рассматривается вопрос кластеризации электроэнергетических систем (ЭЭС) на зоны надёжности, которая решается для формирования энергетических расчетных моделей ЭЭС. Традиционно данная задача решается преимущественно экспертным образом с привлечением специалистов отрасли. В статье предлагается использование методов поиска сообществ в графах для решения задачи кластеризации ЭЭС, а именно метод Лейдена. В экспериментальной части представлен результат его применения, включая метод обнаружения сообществ в графах для формирования ЭРМ объединённой энергосистемы Сибири.

Ключевые слова

Электроэнергетическая система, балансовая надёжность, зона надёжности, методы обнаружения сообществ в графах, метод Лейдена.

1. Введение

Электроэнергетические системы (ЭЭС) представляют собой структуры взаимосвязанных элементов, число которых может достигать сотен тысяч [1]. Для анализа балансовой надёжности ЭЭС [2-4], традиционно, используют методику, основанную на методе Монте-Карло [5], которая является наиболее вычислительно эффективной по сравнению с другими подходами [6]. При оценке балансовой надёжности возникают вычислительные трудности, связанные с высокой размерностью анализируемых ЭЭС. Для снижения сложности задачи проводится кластеризация ЭЭС на зоны надёжности и формирование энергетических расчётных моделей (ЭРМ). Различного рода подходы к кластеризации уже разрабатывались другими исследователями [7, 8], но ранее не применялись в задачах балансовой надёжности. При кластеризации внутри получаемых зон надёжности ЭЭС не учитываются сетевые ограничения, слабо влияющие на показатели балансовой надёжности. Формирование ЭРМ, как правило, проводится до оценки балансовой надёжности, и сформированная ЭРМ принимается неизменной для всего периода оценки балансовой надёжности.

Сейчас в практике оценки балансовой надёжности существуют различные методы кластеризации ЭЭС. В основном эти методы основаны либо на знаниях экспертов, либо на упрощенном анализе балансовой ситуации в ЭЭС. Целью этой статьи является представление методики кластеризации ЭЭС на зоны надёжности, имеющей математическую формализацию.

2. Кластеризация электроэнергетической системы на зоны надёжности

Задача кластеризации ЭЭС на зоны надёжности формулируется следующим образом: для известной структуры ЭЭС, характеристик генерирующих мощностей и графиков потребления мощности, ограничений пропускной способности линий электропередачи и сечений

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: boyarkin_denis@mail.ru (A. 1); krupenev@isem.irk.ru (A. 2); yakubovskii.dmit@mail.ru (A. 3)
ORCID: 0000-0002-7048-2848 (A. 1); 0000-0002-3093-4483 (A. 2); 0000-0001-8331-6200 (A. 3)

ICCS-DE 2023 Workshop Proceedings

электрической сети, организационно-экономических условий функционирования ЭЭС необходимо определить границы ЗН для формирования ЭРМ и дальнейшей оценки БН.

Математически задачу кластеризации ЭЭС на ЗН можно сформулировать следующим образом: задано $X = \{x_1, x_2, \dots, x_N\}$ – множество узлов расчетной схемы ЭЭС, заданных в k – мерном пространстве признаков:

$$x^k = \left\{x_1^k, x_2^k, \dots, x_N^k\right\}, k = 1, \dots, K. \quad (1)$$

Между узлами расчетной схемы ЭЭС существует функция расстояния (метрика) $\rho(x)$. Следует понимать, что функции расстояния не являются прямым расстоянием между узлами расчетной схемы. Функция расстояния формируется из комплекса рабочих характеристик энергосистемы, учитываемых при кластеризации ЭЭС.

Требуется разбить конечную выборку узлов расчетной схемы X на Z не пересекающихся подмножеств:

$$S_z, z = 1, \dots, Z;\ X = \bigcup_{z=1}^{Z} S_z, \quad (2)$$

которые образуют ЗН Z, таким образом, что каждая ЗН включает узлы, у которых метрика ρ меньше порогового значения, в обратном случае узлы сортируются в разные ЗН.

При этом каждому узлу $x_i \in X$ присваивается номер зоны надёжности $z_j \in Z$.

3. Формирование ЭРМ с использованием методов кластеризации в графах

Для максимально эффективного (быстрого и точного) решения задачи кластеризации ЭЭС на зоны надёжности и формирования ЭРМ необходимо разработать подход, который будет максимально полно отражать влияющие критерии и ограничения на уровень БН ЭЭС и позволит достаточно быстро и адаптивно проводить кластеризацию ЭЭС в процессе проведения исследований. При формировании алгоритма кластеризации ЭЭС на зоны надёжности необходимо получить функцию $a: X \to Z$, которая каждый узел $x \in X$ включает в зону надёжности $z \in Z$. В рассматриваемом случае множество Z, т.е. количество зон надёжности, неопределенно. Оптимальное число зон надёжности необходимо определить в процессе вычислений. Стоит отметить, что, как и в любом примере кластеризации, для решаемой задачи справедлива теорема невозможности Клейнберга, в которой указывается, что не существует оптимального алгоритма кластеризации. Каждый применяемый алгоритм будет вносить свою специфику в конечное решение. Это будет зависеть от выбранного критерия и метрики кластеризации.

В данном случае метрика кластеризации должна включать как технические, так и экономические признаки. Метрика должна максимально отражать специфику работы ЭЭС и влияние учитываемых признаков на кластеризацию. В качестве меры расстояния (степени похожести) может быть использована одна из следующих метрик – евклидово расстояние или его квадрат, манхэттенское расстояние, расстояние Чебышева, степенное расстояние и др.

4. Обзор алгоритмов кластеризации для решаемой задачи

Наличие взаимосвязей между кластеризуемыми объектами значительно влияет на получаемый результат и накладывает определенные ограничения. Существует множество методов кластеризации на графах [9-11], которые могут быть использованы в рамках сетевого анализа.

Методы кластеризации в графах можно разделить на несколько групп, в зависимости от используемых алгоритмов и подходов:

1. Методы на основе сходства: Эти методы определяют кластеры на основе степени сходства между вершинами графа. В таких методах могут использоваться различные метрики сходства, например, косинусное сходство или Евклидово расстояние.
2. Методы на основе расстояния: Эти методы определяют кластеры на основе расстояния между вершинами графа. В таких методах могут использоваться различные метрики расстояния, например, евклидово расстояние или расстояние Манхэттена.

3. Методы на основе модели: Эти методы определяют кластеры на основе вероятностной модели графа. Такие методы могут использоваться для кластеризации графов с большим количеством вершин и/или ребер.

4. Методы на основе графовой структуры: Эти методы определяют кластеры на основе структуры графа, такой как плотность графа, центральность вершин и т.д.

Каждая группа методов имеет свои преимущества и недостатки, и выбор конкретного метода зависит от задачи и свойств графа, который необходимо кластеризовать.

Некоторые из наиболее распространенных методов кластеризации на графах включают в себя:

Класс методов разбиения:

1. Метод k-средних [10] – это один из наиболее популярных методов кластеризации. Он основан на разделении вершин графа на заданное число кластеров, так чтобы минимизировать сумму квадратов расстояний между вершинами и центроидами кластеров. Метод k-средних можно применять, когда свойства вершин являются числовыми или когда можно преобразовать их в числовые.

2. Метод кратчайшего пути [11] - в этом методе кластеры строятся на основе расстояний между вершинами графа.

Методы иерархической кластеризации [12] – это класс методов, которые позволяют создавать иерархическую структуру кластеров. Представляет собой наиболее широкий класс методов. Эти методы могут быть использованы для кластеризации вершин графа на основе их свойств, иерархически объединяя близкие вершины. Общая идея методов данной группы заключается в последовательной иерархической декомпозиции множества объектов. В качестве условия остановки можно использовать пороговое число кластеров, которое необходимо получить, однако обычно используется пороговое значение расстояния между кластерами. Рассмотрим некоторые из иерархических методов.

1. Метод Лувена [13] – в этом методе вершины графа постепенно объединяются в кластеры, чтобы максимизировать внутрикластерную связность и минимизировать межкластерную связность.

2. Метод Spectral Clustering [14] – этот метод основан на анализе собственных значений матрицы смежности графа и позволяет разбивать граф на кластеры на основе его структурной информации.

3. Метод Edge Betweenness Clustering [15] – этот метод использует меру межкластерной связности между центральными вершинами для разделения графа на кластеры.

Методы кластеризации на основе плотности – это методы, которые основаны на анализе плотности распределения вершин графа в пространстве свойств. Они могут быть особенно полезны, если вершины графа не являются числовыми, и не могут быть преобразованы в числа.

Также можно использовать сетевые методы анализа [16-19] для кластеризации на графах. Сетевые методы – это подход к анализу данных, основанный на представлении данных в виде сетей или графов, что имеет идентичную структуру. Общая идея методов заключается в том, что пространство объектов разбивается на конечное число ячеек, образующих сетевую структуру, в рамках которой выполняются все операции кластеризации. Главное достоинство методов этой группы в малом времени выполнения.

Одним из самых используемых алгоритмов обнаружения сообществ в графах является алгоритм Лувена [13]. Анализ литературных источников показал, что для решаемой задачи наиболее подходящим является алгоритм Лейдена (Leiden) [20].

5. Кластеризация ЭЭС на зоны надёжности с использованием алгоритма Лейдена

Алгоритм Лейдена — это алгоритм обнаружения сообществ в больших сетях. Метод Лейдена является итеративным алгоритмом и разделяет узлы на непересекающиеся кластеры, чтобы максимизировать показатель модулярности для каждого кластера. Модулярность количественно определяет качество назначения узлов кластерам, то есть насколько плотно связаны узлы в

кластере по сравнению с тем, насколько они были бы связаны в случайной сети. Алгоритм Лейдена состоит из трех основных этапов:

1) Локальное перемещение узлов в кластеры, основанное на значении модулярности (метрики), которая определяется, как:

$$Q = \frac{1}{2m}\sum_{ij}(A_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j), \quad (3)$$

где: A_{ij} – вес ребра, образованный в соответствии с используемыми признаками; k_i и k_j – сумма весов ребер, присоединённых к узлам i и j; m - сумма всех весов ребер в графе; c_i и c_j – кластеры; δ - дельта-функция Кронекера.

$$\delta(c_i, c_j) = \begin{cases} 1, c_i = c_j \\ 0, c_i \neq c_j \end{cases}, \quad (4)$$

Для определения модулярности кластера c применяется следующее выражение:

$$Q_c = \frac{\sum in}{2m} - (\frac{\sum tot}{2m})^2, \quad (5)$$

где: $\sum in$ – это сумма весов ребер между узлами внутри кластера c (каждое ребро учитывается дважды); $\sum tot$ – сумма всех весов рёбер для узлов внутри кластера.

В итерационном процессе происходит перемещения узлов из одного кластера в другой. Для каждого варианта перемещения рассчитывается разница модулярностей ΔQ:

$$\Delta Q = (\frac{\sum in + 2k_{i,in}}{2m} - (\frac{\sum tot + k_i}{2m})^2) - (\frac{\sum in}{2m} - (\frac{\sum tot}{2m})^2 - (\frac{k_i}{2m})^2) \quad (6)$$

На основании наилучшего результата соответствующий узел закрепляется за соответствующим кластером.

2) Уточнение кластеров, а именно идентификация кластеров, предложенных на первом этапе. Кластеры, предложенные на первом этапе, могут разделяться на несколько кластеров. Фаза уточнения не следует жадному подходу и может объединить узел со случайно выбранным кластером, что увеличивает функцию качества (модулярности). Эта случайность позволяет более широко раскрыть пространство кластера.

3) Агрегация и повторение этапов 1 и 2 до тех пор, пока качество перемещения и объединения узлов нельзя будет повысить.

6. Экспериментальные исследования

Для оценки работы алгоритма Лейдена предлагается выполнить экспериментальные расчёты на тестовой схеме ОЭС Сибири, представленной в виде графа, в котором каждая из вершин характеризуется максимальной и минимальной генерацией и нагрузкой, линии характеризуются длиной и напряжением. Результаты кластеризации ОЭС Сибири на зоны надёжности по узлам представлен на рисунке 1.

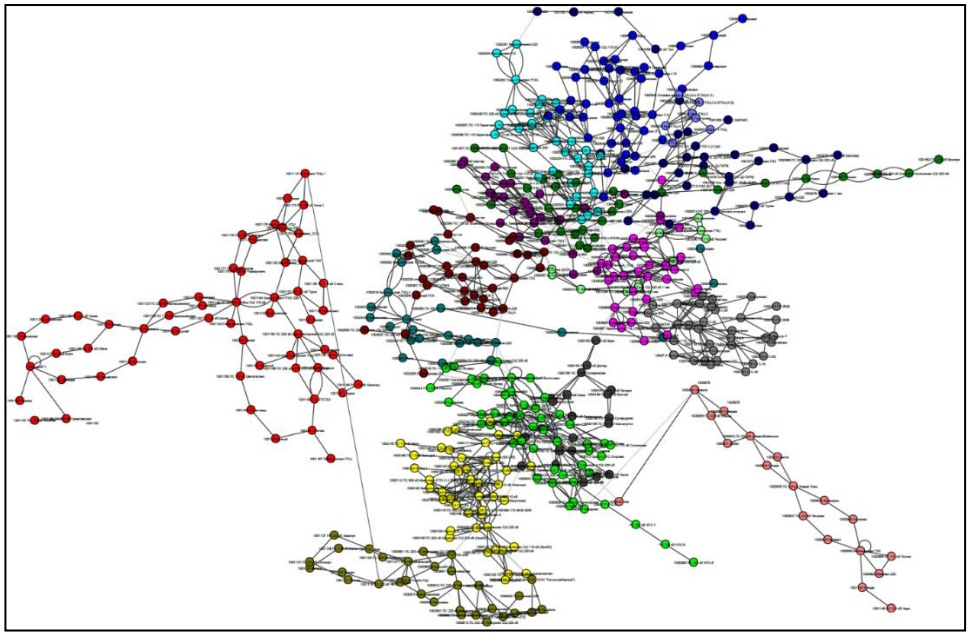


Рисунок 1: Кластеризация ОЭС Сибири на ЗН на основании применения алгоритма Лейдена (по узлам)

Результат, представленный в границах географических регионов указан на рисунке 2.

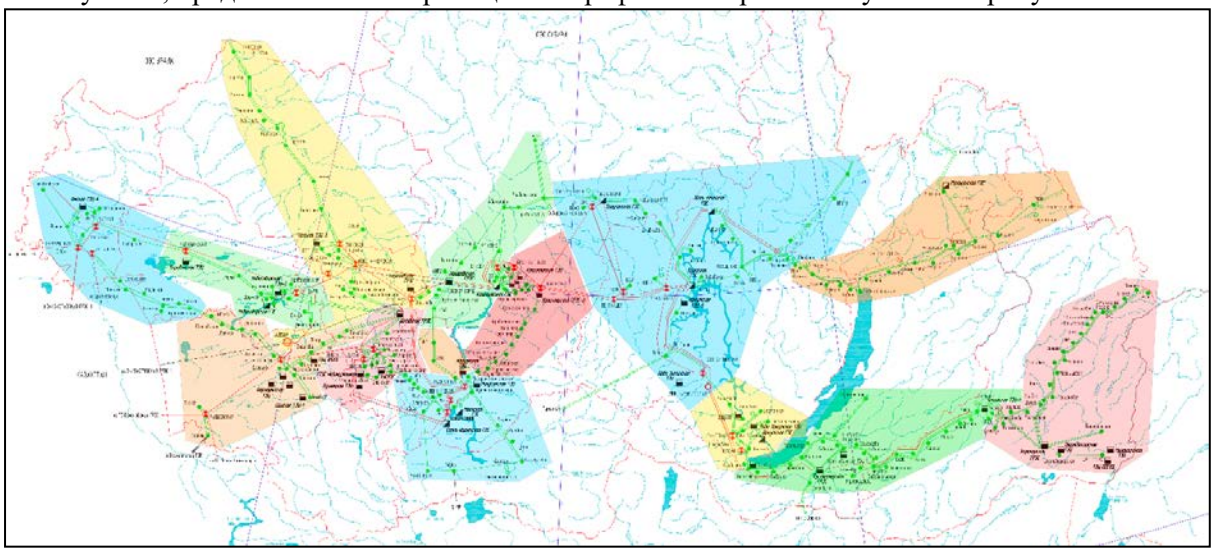


Рисунок 2: Кластеризация ОЭС Сибири на ЗН на основании применения алгоритма Лейдена (с географической привязкой)

7. Заключение

В представленном исследовании рассмотрен задача формирования энергетических расчетных моделей, которые используются при оценке балансовой надёжности электроэнергетических систем. В рамках настоящего исследования предложен алгоритм Лейдена, как один из наиболее подходящих способов кластеризации ЭЭС на зоны надёжности. По результату выполненных экспериментальных расчётов было показано, что предложенный вычислительный подход позволяет выполнять определение зон надёжности оптимальным образом, высвобождая значительные временные ресурсы. Ранее определение зон надёжности выполнялось экспертным способом предшествуя выполнению оценки балансовой надёжности. С помощью алгоритма Лейдена становится возможным определение зон надёжности во время

выполнения соответствующих расчётов, что значительно удобнее и быстрее и не требует привлечения экспертов в области.

8. Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 23-71-01090.

9. References

[1] Руденко Ю.Н., Чельцов М.Б. Надёжность и резервирование в электроэнергетических системах. Изд.-во «Наука» Сибирское отделение, Новосибирск, 1974, 262 с.
[2] Ковалёв Г.Ф., Крупенёв Д.С., Лебедева Л.М. Системная надёжность ЕЭС России на уровне 2030 г. // Электрические станции. — 2011. — № 2. — С. 44-47.
[3] Probabilistic Adequacy and Measures. Technical Reference Report Final, NERC, July, 2018.
[4] 2021 Long-Term Reliability Assessment. NERC, 2021, P. 126.
[5] Ковалев Г.Ф., Лебедева Л.М. Надёжность систем электроэнергетики. Новосибирск.: Наука, 2015.
[6] Krupenev D, Boyarkin D, Iakubovskii D. Improvement in the computational efficiency of a technique for assessing the reliability of electric power systems based on the Monte Carlo method // Reliability Engineering & System Safety. — 2020. — volume 204. doi:10.1016/j.ress.2020.107171
[7] M. Rouhani, M. Mohammadi, M. Aiello. Soft clustering based probabilistic power flow with correlated inter temporal events // Electric Power Systems Research,Volume 204,2022,107677,https://doi.org/10.1016/j.epsr.2021.107677
[8] Gheorghe, Grigoras & Scarlatache, Florina & Neagu, Bogdan. (2016). Clustering in Power Systems. Applications
[9] Мандель И. Д. Кластерный анализ. — М.: Финансы и статистика, 1988. 176 с.
[10] MacQueen, J. Some methods for classification and analysis of multivariate observations/ J. MacQueen // In Proc. 5th Berkeley Symp. On Math. Statistics and Probability, 1967. -C.281-297
[11] Graph Clustering: A Review" (2020) by Arindam Sarkar and Srikumar Ramalingam
[12] Kaufman, L.; Rousseeuw, P.J. (1990). Finding Groups in Data: An Introduction to Cluster Analysis (1 ed.). New York: John Wiley. ISBN 0-471-87876-6
[13] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 10008, 6, https://doi.org/10.1088/1742-5468/2008/10/P10008 (2008)
[14] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. Phys. Rev. E 69, 026113 – 2004
[15] Ester, M., H. P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”. In: Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, AAAI Press, pp. 226-231. 1996
[16] A Survey on Clustering Techniques for Graph Structured Data" (2019) by Foteini Beligianni, George Tsatsaronis, and Themis Palpanas
[17] Palla, G., Derényi, I., Farkas, I. et al. Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814–818 (2005). https://doi.org/10.1038/nature03607
[18] Maps of information flow reveal community structure in complex networks/Martin Rosvall and Carl T. BergstromPNAS 105, 1118 (2008)
[19] Lancichinetti, A. & Fortunato, S. Community detection algorithms: A comparative analysis. Phys. Rev. E 80, 056117, https://doi.org/10.1103/PhysRevE.80.056117 (2009)
[20] Traag, V.A., Waltman, L. & van Eck, N.J. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 9, 5233 (2019). https://doi.org/10.1038/s41598-019-41695-z

Процесс автоматизированного сопровождения профессионального развития учителей на основе индивидуальных образовательных маршрутов

Aleksandr Bykov [1,2]

[1] *Irkutsk State University, 1 Karl Marx str., Irkutsk, 664003, Russia*
[2] *Institute for Educational Development of the Irkutsk region, 1 Krasnokazach'ya str., Irkutsk, 664007, Russia*

Abstract

В данной статье приведено описание процесса сопровождения развития компетенций учителей на основе индивидуальных образовательных маршрутов, реализованного в автоматизированной информационной системе. Представлены технологические особенности реализации данного процесса, а также обозначены проблемные области для дальнейшего его исследования.

Keywords

Развитие компетенций, рекомендательная система, автоматизированная информационная система, процесс сопровождения

1. Введение

Повышение квалификации учителей на протяжении последних лет претерпевает значительные изменения, меняются подходы, участники и повышается его технологизация. Это связано с рядом факторов, прежде всего, с повышением динамики изменения содержания и технологий образования, активным развитием «сквозных» направлений развития, зафиксированных различными стратегическими документами в сфере общего образования. Наряду с положительными эффектами наблюдаются и проблемы, возникающие в результате динамичного повышения информационной насыщенности пространства профессионального педагогического образования. Среди них снижение уровня сформированности внутренней мотивации к собственному профессиональному развитию у работников системы образования, недостаточной степени поддержки педагогов со стороны методических служб и работодателя в условиях отсутствия эффективных средств навигирования в данном пространстве.

Данные проблемы эффективно решаются «в ручном режиме» при малом количестве учителей и коллективов, профессиональный уровень которых необходимо повысить. В случае необходимости организации профессионального развития больших групп таких работников требуются технологии, позволяющие решать данные проблемы в значительной степени автоматизации за счет современных технологий [4].

Представленный далее процесс внедрен и апробируется в региональной автоматизированной информационной системе iom.iro38.ru. Данная система введена в опытную эксплуатацию осенью 2022 года в государственном автономном учреждении дополнительного профессионального образования Иркутской области «Институт развития образования Иркутской области». За время опытной эксплуатации системы ее возможностями воспользовалось более 2000 учителей региона. Опыт эксплуатации демонстрирует ее высокую значимость для решения актуальных задач государственной политики в сфере развития общего образования [1], а также удовлетворения частных образовательных запросов отдельных учителей.

EMAIL: thekindbull@gmail.com (A. 1)
ORCID: 0000-0002-9509-3451 (A. 1)

ICCS-DE 2023 Workshop Proceedings

2. Условия профессионального развития учителей

Как было отмечено выше, успешность развития компетенций учителя зависит от его мотивационного состояния, степени вовлеченности работодателя и методической поддержки.

Самой сложной и слабоформализуемой проблемой в данном случае является работа с мотивацией. В отличие от коммерческой сферы, где основой работы с мотивацией являестя выявление клюевых показателей эффективности, в деятельности учителей эти показатели часто не обозначены конкретным образом, не носят системный характер.

С точки зрения предмета рассмотрения данной публикации важно отметить, что он прямо не влияет на внутреннюю мотивацию учителей к эффективному профессиональному развитию, однако, способен повысить удобство данного процесса, предоставлять педагогу ресурсы для такого развития в нужный момент времени.

Включение в данный процесс ресурсов и запросов работодателя, руководителя образовательной организации, чей сотрудник нуждается в профессиональном развитии – необходимое условие повышения эффективности профессионального развития учителей. Работодатель в данном процессе является в значительной степени заинтересованным лицом. Именно работодатель, осуществляя общее управление основными направлениями деятельности образовательной организации, уполномочен определять политику в области профессионального развития педагогов, ставить индивидуальные и командные образовательные цели перед коллективом. Обладая всем доступным набором информации о деятельности образовательной организации, только руководитель имеет возможность принимать объективные решения о необходимости повышения квалификации работников по различным направлениям, оперируя всей этой информацией в совокупности. В свою очередь, подразумевается, что данное утверждение справедливо при условии достаточного уровня сформированности у руководителя навыков управления развитием профессиональных компетенций сотрудников или наличия квалифицированной hr-службы. Очевидно, что и первое и, тем более, второе условие в недостаточной степени представлены в системе общего образования по объективным причинам.

Оказание методической поддержки в вопросах повышения профессиональной компетентности – не менее важный фактор повышения эффективности данного процесса, который усложняется постоянным изменением знаниево-технологической составляющей профессиональной деятельности учителей. Обеспечение методически обоснованного навигирования работника во множестве направлений и аспектов развития образовательной политики, изменяемом предметном содержании позволяет корректно и вовремя выявлять и нивелировать его профессиональные затруднения. В то же время важно отметить, что в настоящее время в нашей стране система методической работы с педагогами в значительной степени утратила свои фундаментальные возможности, часто не способна оперативно и компетентно дать обратную связь на любой педагогический запрос.

Таким образом, справедливо заключить, что эффективность профессионального развития учителей выражается в сбалансированности внешних и внутренних факторов, которые позволяют своевременно выявлять профессиональные затруднения, формировать на их основе актуальный образовательный запрос и выполнять его. Процесс профессионального развития должен определяться не только внутренними мотивами отдельно взятого работника, но и подлежать управлению извне работодателем и быть обеспечен методическим сопровождением.

Учитывая представленные выше условия профессионального развития учителей и современные тенденции развития компетентностного подхода допустимо предположить, что в общем виде профессиональное развитие – непрерывный процесс на определенном отрезке времени, который можно представить в виде функции

$$y = f(t, p_1, p_2, \ldots, p_n), \quad (1)$$

где $p_{i=\overline{1,n}}$ – оценочные характеристики актуального состояния i-й составляющей профессиональной компетентности учителя в момент времени t.

3. Содержание профессионального развития как система

В данных условиях можно заключить, что в качестве аргументов функции (1) можно приводить оценочные характеристики уровня сформированности отдельных элементов системы знаний, умений, личностных качеств учителя, которые необходимы для качественного исполнения его должностных обязанностей.

При определении состава данной системы сообразно обратить внимание на возрастающую практическую значимость профессиональных стандартов. Согласно Трудовому кодексу Российской Федерации, профессиональный стандарт представляет собой характеристику квалификации, необходимой работнику для осуществления определенного вида профессиональной деятельности. В свою очередь, профессиональные стандарты разрабатываются и применяются в соответствии с правилами, утвержденными Постановлением Правительства Российской Федерации от 22.01.2013. №23 «О правилах разработки, утверждения и применения профессиональных стандартов» [3].

Профессиональные стандарты имеют утвержденную структуру детализации обозначенных выше характеристик, оперируя понятиями трудовой функции, трудового действия, знания, умения. Для обеспечения согласованности данной системы компетенций с нормативно-правовым обеспечением профессионального развития учителей важно сформировать первичный иерархический справочник-кодификатор профессиональных компетенций именно на основе профстандарта «Педагог (педагогическая деятельность в сфере дошкольного, начального общего, основного общего, среднего общего образования) (воспитатель, учитель)». В практической деятельности этот справочник позволит учителям использовать его в качестве основы для формулирования профессиональных образовательных запросов.

Однако, имеющиеся формулировки знаний и умений в профессиональных стандартах часто не позволяют однозначно идентифицировать конкретный запрос, например, по работе с отдельным педагогическим инструментом или технологией. Определение образовательных запросов должно сопровождаться осознанием его содержания и принятием его как компетентностного дефицита самим работником. В рамках предмета рассмотрения данной статьи осознание содержания дефицита выражается его фиксацией в виде свободного описания по определенному правилу в контексте выбранной компетенции из справочника. Заметим, что это же справедливо для знаний и умений в конкретной предметной области, не отраженных в профессиональном стандарте.

4. Выявление профессиональных дефицитов

Для формирования образовательного запроса учитель сперва должен получить информацию об имеющихся профессиональных дефицитах. Данная задача может быть решена различными способами, в т.ч. в ходе рефлексии имеющегося опыта самим человеком, участия в тестированиях, экзаменах и т.п. Однако, и другие формы, например, результаты участия в конкурсах, демонстрационных (открытых) учебных занятиях, соревнованиях, олимпиадах, семинарах, курсах повышения квалификации также могут являться эффективными инструментами решения данной задачи. Разнообразие представленных способов позволяет обеспечить диверсификацию и наибольшую полноту выявленных дефицитов. В автоматизированной среде данный подход должен предполагать значительное разнообразие организаций, способных влиять на своего рода профиль дефицитов отдельного учителя. Важно, чтобы организации, имеющие такой доступ к профилям дефицитов учителей понимали общие подходы, правила и принципы, проходили своего рода аккредитацию, носили статус доверенных.

Профиль дефицитов должен описывать актуальное состояние профессиональных затруднений работника, в т.ч. позволяющее выявить наиболее актуальные точки дальнейшего профессионального развития.

На Рисунке 1 представлено взаимодействие учителя со своим профилем дефицитов. Предполагается, что профиль дефицитов состоит из таких обязательных атрибутов, как наименование компетенции, способ и дата выявления, уровень.

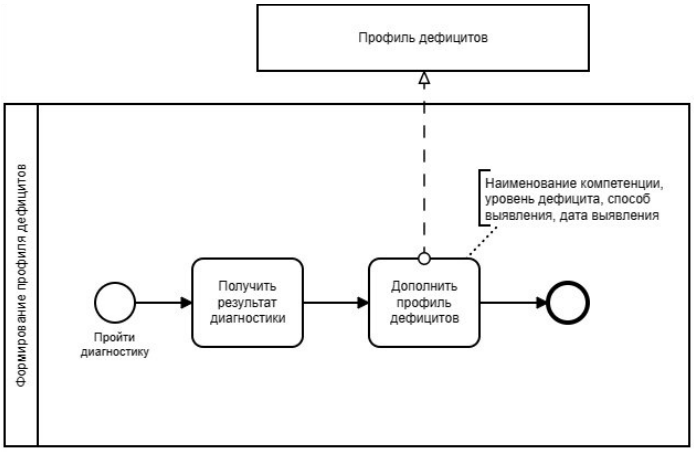


Рисунок 1: BPMN-диаграмма формирования профиля дефицитов

Важно заметить, что в ходе изложения мы сделали переход к дефицитарному принципу работы с профессиональным развитием учителя и ввели понятие «профиль дефицитов». Данный принцип не ставит необходимостью получение актуального состояния всех составляющих профессиональной компетентности учителя, но определяет необходимость выявления профессиональных дефицитов – формирования профиля дефицитов – а уже на его основе выявлять актуальные в конкретный момент времени образовательные запросы. Однако, по мнению автора, по мере развития и становления процесса будет правильным внедрить гибридный подход управления профессиональным развитием учителей, объединив возможности дефициатарного и ресурсного подходов. Последний в данном случае направлен на выявление личностно-профессиональных особенностей каждого учителя [3] и их правильного направления.

Таким образом, функция (1), отражающая совокупность всех возможный компетенций, становится неактуальной. К тому же она не может быть вычислена в приемлемые сроки с применением адекватных человеческих и технических ресурсов. Базой для дальнейшей автоматизации становится профиль дефицитов, который также можно представить в виде некоторой функции

$$y' = f'(t', d_1, d_2, \ldots, d_k), \quad (2)$$

где $d_{j=\overline{1,k}}$ – оценочная характеристика уровня дефицита k-й компетенции. В идеальном случае функция (2) является разницей между актуальным значением функции (y) и ее предельным верхним значением при максимальных значениях всех p_i. Однако, и этот случай не рассматриваем по той же причине невозможности достижения в адекватных условиях.

Можно допустить, что ждя функции (2) отдельно взятый аргумент d_j можно вычислить на определенной шкале по конкретному измерителю (например, тестированию). В данном способе предлагается использовать 100-балльную шкалу, а оценочные характеристики $d_j \in R_{\geq 0}$.

Возьмем во внимание еще одно очевидное утверждение, что формирование профиля дефицита – это не непрерывный процесс. Учитывая специфику трудовой деятельности учителей для функции (2) можно сделать допущение считать длиной шага времени t' учебный год. Поэтому его значением в данном случае можно пренебречь на промежутках времени внутри учебного года.

5. Рекомендательная система

На основе построенного профиля дефицитов необходимо обеспечить корректную навигацию учителя в пространстве образовательных ресурсов, которые в наибольшей степени будут способствовать ликвидации выявленных профессиональных дефицитов. Для решения этой задачи важно обеспечить общее хранилище метаданных о различных образовательных ресурсах. Под образовательным ресурсом будем понимать совокупность информации, потенциально представляющая образовательный и научный интерес с точки зрения профессионального

развития учителя, определенная различными формами: книги, онлайн-курсы, семинары, программы повышения квалификации и т.п.

При этом рекомендательная система включает в себя два уровня функционирования: простой и «интеллектуальный». В первом случае – уровне простого функционирования – важно обеспечить привязку каждого ресурса к одному или нескольким элементам справочника компетенций. Это позволит сразу информировать пользователя о новых ресурсах, которые могут представлять для него интерес.

Для более эффективного формирования списка рекомендаций и качественного повышения их пертинентности необходимо привлечение различных механизмов формирования персонализированных рекомендаций: коллаборативной и контентной фильтрации, образующих интеллектуальный уровень функционирования рекомендательной системы.

В общем виде, рекомендательной системой предусмотрено построение матрицы $P = |N| \times |M|$, где N – количество учителей (пользователей), M – количество образовательных ресурсов, а каждый элемент матрицы – это действительная или прогнозируемая заинтересованность отдельного учителя в конкретном ресурсе.

Коллаборативная фильтрация – метод прогнозирования интересов пользователя относительно образовательного ресурса на основании существующего пользовательского опыта других пользователей со сходным пользовательским опытом взаимодействия со всем множеством ресурсов.

Для обеспечения функционирования данного метода необходимо иметь базу пользовательского опыта пользователей (транзакций) с имеющимися образовательными ресурсами, с помощью которой опыт взаимодействия i-го пользователя с j-м ресурсом можно представить функцией $\tilde{y}_{ij} = f(u_i, r_j, ...)$, значения которой и заполнят матрицу P. В общем случае задачей метода коллаборативной фильтрации является заполнение всей матрицы, даже в тех случаях, когда у пользователя отсутствует опыт взаимодействия с каким-либо ресурсом. В данной методике для коллаборативной фильтрации используется классический вариант реализации, основанный на поиске ближайших соседей, чьи индивидуальные образовательные маршруты наиболее коррелируют с маршрутом выбранного пользователя, корреляция в данном случае рассчитывается косинусным расстоянием. Однако, поиск наилучших методов коллаборативной фильтрации в данном случае является открытым для будущих исследований.

В разработанной автоматизированной методике пользовательский опыт отражается выбором образовательного ресурса в качестве составляющей индивидуального образовательного маршрута данного пользователя. В будущем, при более широком внедрении данной методики поиск видов транзакций, безусловно, будет расширяться.

Коллаборативная фильтрация, как видно, стремится к получению некоторых усредненных значений интересов пользователей и потенциально ограничивает перечень рекомендованных ресурсов. Это обуславливается несколькими факторами, первый из которых – технический, заключающийся в возрастающей вычислительной ресурсоемкости метода по мере увеличения количества ближайших соседей. Их малое количество, очевидно, будет в значительной степени ограничивать перечень рекомендаций, а слишком большое – будет создавать «шумные» рекомендации, не соответствующие действительности. Принимая во внимание обозначенный ранее тезис о низкой внутренней мотивации широкой части педагогических работников системы образования к профессиональному развитию, существует риск, что отдельные ресурсы не попадут в рекомендации пользователей или станут неактуальными до этого момента ввиду того, что немногие добавят его себе в индивидуальный маршрут. Поэтому важно дополнить функционирование персональных рекомендаций для пользователя посредством другого метода – контентной фильтрации, опирающегося на неформализованную информацию в профилях дефицитов пользователей и метаданных образовательных ресурсов.

Контентная фильтрация позволяет предоставлять клиентам рекомендации на основе текстовых формулировок профессиональных затруднений. Как было отмечено ранее, методикой предусмотрена необходимость “детализации” каждого затруднения в свободной форме. На основе данной информации проводится сравнение на подобие детализации образовательного запроса, описанной самим пользователем и атрибута «Аннотация» различных образовательных ресурсов по ключевым словам. В более совершенном виде требуется реализация анализа на

подобие посредством семантического анализа текстов и внедрению методов машинного обучения.

6. Индивидуальный образовательный маршрут

На основе представленных рекомендаций учителю необходимо сформировать индивидуальный образовательный маршрут, который представляет собой совокупность данных об образовательных ресурсах, которые ему необходимо изучить или посетить, чтобы достичь поставленной цели. В рамках данной методики индивидуальный маршрут помимо совокупности ресурсов содержит в обязательном порядке цель и данные о предполагаемом продукте реализации. Вся эта информация определяет содержательные рамки отдельного индивидуального образовательного маршрута, способствует систематизации и осознанию клиентом своих дефицитов и правильному планированию деятельности по их ликвидации. Вместе с тем, возвращаясь к ранее обозначенным усложняющим данный процесс факторам – вовлеченности работодателей и методической поддержке – важно включить их в данный процесс. Инициация формирования индивидуального образовательного маршрута происходит только с разрешения работодателя, а уже на этапе формирования маршрута подразумевается возможность включения консультационной поддержки со стороны назначенного конкретному педагогу методиста в режиме электронной переписки.

Важной составляющей процесса работы с индивидуальным образовательным маршрутом является его завершение и оценка состоятельности. Учитывая, что педагогическую деятельность школьного учителя со всеми ее составляющими трудно формализовать на сколь-нибудь приемлемом для целей системного анализа уровне отметим, что знания и умения в конкретной предметной области возможно проверить с помощью тестирований, которые при правильной разработке могут предоставить достоверную информацию об уровне владения такими знаниями и умениями. В то же время общепедагогические компетенции в большинстве случаев могут быть адекватно оценены лишь субъективно, в прямом наблюдении со стороны компетентного методиста. В данном случае учитель должен на стадии формирования маршрута точно представлять структуру и критерии оценки продукта, через который в конечном итоге будет оцениваться достижение им цели образовательного маршрута. Таким образом, для эффективности реализации данного процесса в автоматизированном режиме на стадии проектирования маршрута должны быть адекватно сформулированы цель и предполагаемый образовательный продукт деятельности по его реализации. С учетом предыдущих параграфов на Рисунке 2 приведена BPMN-диаграмма процесса работы учителя с индивидуальным образовательным маршрутом от формирования профиля дефицитов до представления образовательного продукта.

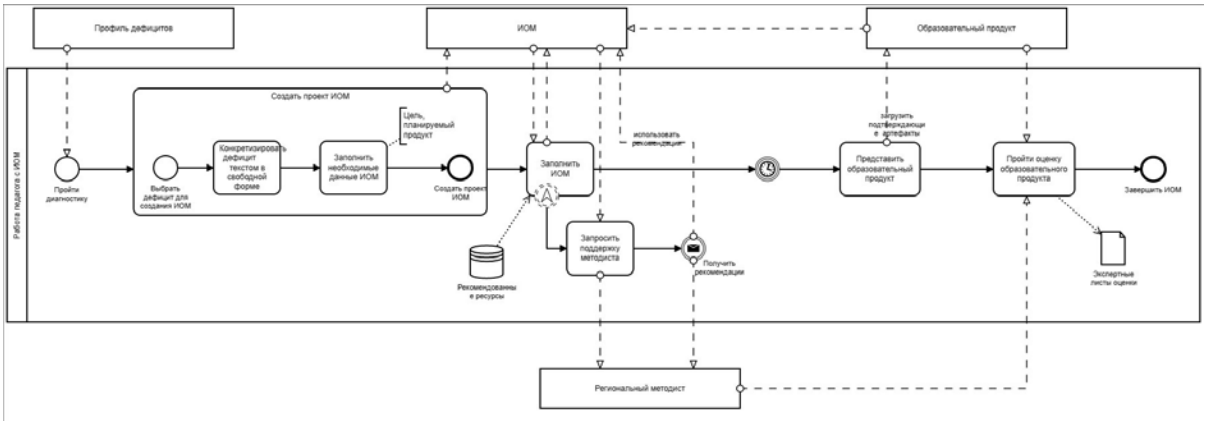


Рисунок 2: BPMN-диаграмма процесса работы учителя с индивидуальным образовательным маршрутом

Заметим, что собственно индивидуальный маршрут хоть и является лишь частью всего процесса, в то же время представляет его технологическую и содержательную основу.

7. Заключение

Изучение литературы и обзор существующих практик по представленному вопросу говорит о том, что в различных государствах вопросу развития профессиональных компетениц учителей уделяется значительное внимание. Однако, ряд экономических, социальных и технологических вызовов способствуют формированию кризисной ситуации в данном вопросе. Вместе с тем возможности современных информационнх технологий в решении стоящих проблем не используются в достаточной мере [4]. В России в настоящее время отсутствует какой-либо продукт федерального уровня, в функционировании которого были бы задействованы все заинтересованные субъекты. Вместе с тем, отдельные регионы и организации разрабатывают собственные решения с учетом региональной или институциональной специфики.

Из успешных практик, направленных на поддержку развития профессиональных компетенций, можно отметить закрытые системы университетов и крупных коммерческих компаний, которые формируют собственный контур содержания обучения в контексте необходимых образовательных задач.

В то же время представленный процесс и его реализация в существующей автоматизированной системе ставят своей целью создать избыточную совокупность образовательных ресурсов для направленного использования в индивидуальных образовательных маршрутов учителей. Данная цель предоставляет ряд возможностей, главной из которых является качественная деверсификация возможностей профессионального развития учителей. Вместе с тем и оставляет ряд открытых вопрос, технологическое решение которых в рамках системного анализа еще не представлено в российской и зарубежной практике. Среди таких вопросов остаются указанные выше измерители практических умений, оценка оптимальности и достаточности ресурсов индивидуального образовательного маршрута в условиях широкого масштаба применения и динамики изменения содержания и технологий образования.

8. Список литературы

[1] Быков, А. С. Концепция региональной автоматизированной информационной системы построения индивидуальных маршрутов повышения профессионального мастерства работников системы общего и дополнительного образования / А. С. Быков // Динамические системы и компьютерные науки: теория и приложения (DYSC 2021) : Материалы 3-й Международной конференции, Иркутск, 13–17 сентября 2021 года. – Иркутск: Иркутский государственный университет, 2021. – С. 205-209. – EDN HMFNLP.

[2] Гриненко Т.Г. Оценка как инструмент развития персонала организации // УПИРР. 2016. №1. URL: https://cyberleninka.ru/article/n/otsenka-kak-instrument-razvitiya-personala-organizatsii (дата обращения: 17.06.2023).

[3] Сезонова, О. Н. Модель управления развитием профессиональных компетенций кадров организации // Среднерусский вестник общественных наук. 2015. №2. URL: https://cyberleninka.ru/article/n/model-upravleniya-razvitiem-professionalnyh-kompetentsiy-kadrov-organizatsii (дата обращения: 15.06.2023)

[4] Даггэн, С. Искусственный интеллект в образовании: Изменение темпов обучения, Аналитическая записк ПИТ ЮНЕСКО, ред. С. Ю. Князева, пер. с англ. А. В. Паршакова, Институ ЮНЕСК по информационны технология в образовании, 2020

Logic Programming Approach to Observability Testing

Artem V. Davydov[1], Aleksandr A. Larionov[1] and Nadezhda V. Nagul[1]

[1]*Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences, 134 Lermontov St, Irkutsk, 664033, Russian Federation*

Abstract

The paper presents a logic-based way of checking observability of partially-observed discrete-event systems. The automated theorem proving technique in the calculus of PCFs is employed. A known algorithm is realized with the help of the inference search for a specially designed positively-constructed formula. The main advantage of the presented PCF-based approach and the employment of the ATP technique is the declarative description of the algorithms considered.

Keywords

discrete event systems, positively constructed formulas, automated theorem proving, supervisory control, partial observation

1. Introduction

The class of logical discrete-event systems (DES) is a wildly used modelling formalism for man-made complex objects [1, 2, 3]. A logical DES is often represented by a finite-state automaton which transitions from state to state are labeled with the letters of some finite alphabet. Sequences of such transitions may be considered as words of a regular language. Since each transition is associated with an event occurring in the system thus the language describes behaviour of the system from the high-level, or symbolic, point of view. That is why studying DES is naturally embraced by the paradigm of automated theorem proving (ATP).

A small overview of the state-of-the-art on ATP in robotics is presented in [4], while a detailed review on planning in robotics, including the use of ATP, is presented in [5]. The work [6] lies at the intersection of ATP and machine learning, presenting a reinforcement learning toolkit for experiments on guiding ATP in the calculus of connections. The core of the toolkit is a Prolog-based proves i.e. a program for automated reasoning in Prolog language.

We suggest a new way of study and design logical DES based on the the ATP in the calculus of positively constructed formulas (PCFs). The origins of the PCF calculus as a complete method for ATP may be found in [7, 8, 9] while a detailed discussion of its characteristics and capabilities is presented in [10]. In the series of the previous works it was shown how the basic problems of the supervisory control theory (SCT) [11] for logical DES may be solved using the technique

5th International Workshop on Advanced Information and Computation Technologies and Systems, July 03–07, 2023, Irkutsk, Russia

artem@icc.ru (A. V. Davydov); bootfrost@zoho.com (A. A. Larionov); sapling@icc.ru (N. V. Nagul)

0000-0002-8703-3096 (A. V. Davydov); 0000-0001-7116-9338 (A. A. Larionov); 0000-0003-1439-3274 (N. V. Nagul)

presented. The issues considered so far include controllability checking to determine if a formal language restricting DES functioning may be guaranteed by a supervisor [12], supremal controllable sublanguage of a given specification language construction [13] and a monolithic supervisor realization [14].

This paper deals with partially observed DES in which occurrence of some events are unavailable for observation. In this case the property called observability of a formal language determines those specifications on DES functioning that may be ensured by the supervisory control. Some effective tests for observability have been already suggested, e.g. in [15] and [16], where the latter is based on the algebraic operations on processes that represent DES. After necessary preliminaries provided in the next section, in section 3 it will be shown how the polynomial algorithm from [17] may be realized by the usage of PCFs. The main advantage of the presented PCF-based approach and the employment of the ATP technique is the declarative description of the used algorithms. In this case, the programmer only describes (declares) the properties of the required result, and the solution (method) is provided by the logical programming system, as a result of searching for the logical inference of some goal statement. This is a step up from programming in imperative languages (e.g. C/C++, Java, etc.) as the programmer does not have to worry about the low-level details of the program.

2. The PCF-calculus

The calculus of positively-constructed formulas (PCFs) is based on the refutation of the negation of an original statement which is to be proved. The main idea is the following: if the negation of the statement has been proved to be false then the statement itself is valid. The language of PCFs is a restricted variant of the language of the first-order logic (FOL), which consists of first–order formulas (FOFs) built out of atomic formulas, or atoms, with the help of operators $\&$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$, quantifier symbols $\forall$ and $\exists$, and constants $True$ and $False$. Note that any FOF can be represented as PCF since the PCF-language is a special way of writing classical FOFs, as well as the conjunctive normal form, the disjunctive normal form, etc. The converting algorithm is presented in [8].

For easier reading, a PCF may be represented graphically as a tree structure. For example, consider a PCF representation of a FOF

$$\mathcal{F} = \neg\big(\forall x \exists y P(x,y) \rightarrow \exists z P(z,z)\big).$$

An image $\mathcal{F}'$ of $\mathcal{F}$ in the PCF language is $\mathcal{F}' = \forall\colon \varnothing\{\exists\colon \varnothing\{\forall x\colon \varnothing\{\exists y\colon P(x,y)\}, \forall z\colon P(z,z)\{\exists\colon False\}\}\}$. The tree–like form of the latter is

$$\forall\colon \varnothing \cdot \exists\colon \varnothing \left\langle \begin{array}{l} \forall x\colon \varnothing \text{ —— } \exists y\colon P(x,y) \\ \forall z\colon P(z,z) \cdot \exists\colon False. \end{array} \right.$$

The root $\forall\varnothing$ of a PCF tree is called a PCF $root$. Each PCF root's child $\exists_X A$ is called a PCF $base$, the conjunct A is called a $base\ of\ facts$, and a PCF rooted from the base is called a $base\ subformula$. The PCF base children $\forall_Y B$ are called $questions$ to the parent base. The subtrees

of the questions are called $consequents$. If a question has no consequent then the question is referred to as $goal\ question$, and it is identical to $False$.

The only axiom of the PCF calculus is $\forall\varnothing\colon \varnothing$, i.e., $False$. The inference rule ω in the PCF calculus is based on the search for so-called $answering\ substitutions$, i.e., such substitutions of variables in terms that satisfy certain conditions. Any finite sequence of PCFs $\mathcal{F}, \omega\mathcal{F}, \omega^2\mathcal{F}, \ldots, \omega^n\mathcal{F}$, where $\omega^s\mathcal{F} = \omega(\omega^{s-1}\mathcal{F}), \omega^1 = \omega, \omega^n\mathcal{F} = \forall$, is called an $inference$ of $\mathcal{F}$ in PCF calculus (with the axiom $\forall$). An answer to the goal question refutes the corresponding base. When all bases of the PCF are refuted then the PCF $\mathcal{F}$ reduces to $\forall$, i.e. $False$. This means that $\mathcal{F}$ as the negation of the statement under consideration is unsatisfiable; therefore, the statement itself is true. The details on the PCF-calculus may be found in [7, 8, 9, 10].

3. Checking observability of a regular language

Consider a discrete event system (DES) in the form of a generator $\mathcal{G} = (Q, \Sigma, \delta, q_0, Q_m)$ of a formal language [11]. Here Q is the set of states q; Σ the set of events; $\delta\colon \Sigma \times Q \to Q$ the transition function; $q_0 \in Q$ the initial state; $Q_m \subset Q$ the set of marker states. Let $L(\mathcal{G})$ be a language generated by $\mathcal{G}$, and $L_m(\mathcal{G})$ be a language marked by $\mathcal{G}$. Let $\mathcal{G}$ be partially observable and the observation function is defined as the natural projection $P : \Sigma^* \to \Sigma_o^*$ which erases unobservable events belonging to a set $\Sigma_{uo} = \Sigma \setminus \Sigma_o$. Let $\overline{L}$ be the set of all strings that are prefixes of words of L, i.e. $\overline{L} = \{s | s \in \Sigma^* \text{ and } \exists t \in \Sigma^* : s \cdot t \in L\}$.

In the SCT framework, a supervisor existence criterion for partially-observed DES requires, among other, a specification language to be observable. A language K is *observable* (with respect to $L(\mathcal{G})$ and P) if $\forall s, t \in \Sigma^*\,(P(s) = P(t) \to (\forall\sigma \in \Sigma)(s\sigma \in \overline{K}\,\&\,t\sigma \in L(\mathcal{G})\,\&\,t \in \overline{K} \to t\sigma \in \overline{K}))$.

There are several algorithms exist to check if the regular language K is observable. For example, an algorithm was presented in [1] that is polynomial with respect to the size of the automata generating K. The main idea of the algorithm is constructing an automata for tracking two words s_1, s_2 of K which have the same projection $P(s_1) = P(s_2)$ but $s_1 \in \overline{K}$ while $s_2 \notin \overline{K}$. Let the regular language K is recognized by the finite-state automaton H. The algorithm from [1] suggests to consider two copies of the automaton H and one copy of the automaton $\mathcal{G}$ and to design an automaton T with the states of the form (h_1, h_2, q) where $h_1 \in Q_H$, $h_2 \in Q_H$, $q \in Q_{\mathcal{G}}$, and the single state $dead$. The existence of the state $dead$ denotes unobservability of the language K considered because by construction for $s_i \in \overline{K}$, $i = 1, 2, 3$, and some event a, $s_2 = s_3$, $P(s_1) = P(s_2)$. $s_1 a \in \overline{K}$ while $s_2 a \notin \overline{K}$, $s_3 a \in L(\mathcal{G})$.

We are going to demonstrate how this algorithm may be realized with the help of ATP in the PCF-calculus. To do this, some preprocessing is necessary. The following list of predicates will be exploited next: a predicate $Q(q^i)$ which corresponds to all states of the automaton $\mathcal{G}$, a predicate $E(\sigma^j)$ that defines all events of automata, terms $\delta_G(q_1^i, \sigma^i, q_2^i)$ and $\delta_H(q_1^i, \sigma^i, q_2^i)$ that determine transitions from the state q_1^i of automaton G or H to the state q_2^i labeled with an event σ^i. Let term $F_X(q^i, \sigma^j)$ mean that there is a transition from the state q^i of automaton X labeled by the event σ^j. Let $NoF_G(q^i, \sigma^j)$ mean the opposite, i.e. that there is no such transition. Consider the PCF $\mathcal{F}_{Prep}$ (1) with the base $B_{Prep} = \{Q(q^i), \delta_G(q_1^i, \sigma^i, q_2 i), \delta_H(q_1^i, \sigma^i, q_2^i), E(\sigma^j)\}$. The inference of $\mathcal{F}_{Prep}$ adds proper atoms $F_X(q^i, \sigma^j)$ and $NoF_X(q^i, \sigma^j)$ into the base.

$$\mathcal{F}_{Prep} = \exists B_{Prep} \begin{cases} \forall \sigma, q\ E(\sigma), Q(q) \text{ ——————— } \exists\ NoF_G(q,\sigma), NoF_H(q,\sigma) \\ \forall \sigma, q, q_1\ \delta_G(q,\sigma,q_1), NoF^*_G(q,\sigma) - \exists\ F_G(q,\sigma) \\ \forall \sigma, q, q_1\ \delta_H(q,\sigma,q_1), NoF^*_H(q,\sigma) - \exists\ F_H(q,\sigma) \end{cases} \quad (1)$$

Note the operator $*$ used for the deletion of the unnecessary atoms which were added by the first question. One of the essential features of the calculus of PCFs which will be used for analysis of DES is that we can build a *non-monotonic* inference by adjusting the definition of the inference rule. For this, an operator * is exploited. In particular, after applying the inference rule, the atoms in the base that participated in the matching search with the marked atoms in question should be removed from the base. In general, the operator * affects the property of completeness of the PCF calculus but for the problem considered in this paper the inference using * is always correct.

Denote the resulting base B'_{Prep}. Let $B_{Obs} = B'_{Prep} \cup \{Q^G_0(q^G_0), Q^H_0(q^H_0), E_c(\sigma^j), E_o(\sigma^j), E_{uo}(\sigma^j)\}$, where $Q^X_0(\_)$ determines the initial state of automata $X, E_c(\sigma^j), E_o(\sigma^j), E_{uo}(\sigma^j)$ are intuitively clear. The predicate $T(\_,\_,\_)$ will be used to define states of the testing automaton T. The predicate $\delta_T(q^1_H, q^2_H, q_G, \sigma_1, \sigma_2, \sigma_3, t^1_H, t^2_H, t_G)$ is equal to the phrase"there is a transition labeled $\sigma_1, \sigma_2, \sigma_3$ from the state (q^1_H, q^2_H, q_G) to the state (t^1_H, t^2_H, t_G) of T. The PCF $\mathcal{F}_{Obs}$ in Fig. 2 constructs the testing automaton T. The questions of $\mathcal{F}_{Obs}$ are listed as formulas $R_1 - R_6$.

$$\mathcal{F}_{Obs} = \exists B_{Obs} \begin{cases} R_1 \\ R_2 \\ \dots \\ R_6 \end{cases} \quad (2)$$

$$R_1 : \forall q_H, q_G\ Q^G_0(q_G), Q^H_0(q_H) - \exists\ Q^T_0(q_H, q_H, q_G), T(q_H, q_H, q_G)$$

$$R_2 :\ \forall \sigma, q^1_H, q^2_H, q_G\ T(t^1_H, q^2_H, q_G), E_c(\sigma), F_H(q^1_H, \sigma), NoF_H(q^2_H, \sigma), F_G(q_G, \sigma) - \exists\ dead(q^1_H, q^2_H, q_G, \sigma)$$

$$R_3 : \forall x, y, z, s\ dead(x, y, z, s)$$

$$R_4 : \forall \sigma, q^1_H, q^2_H, q_G, t^1_H, t^2_H, t_G\ T(q^1_H, q^2_H, q_G), E_o(\sigma), \delta_H(q^1_H, \sigma, t^1_H), \delta_H(q^2_H, \sigma, t^2_H), \delta_G(q_G, \sigma, t_G) - \exists\ T(t^1_H, t^2_H, t_G), \delta_T(q^1_H, q^2_H, q_G, \sigma, t^1_H, t^2_H, t_G)$$

$$R_5 : \forall \sigma, q^1_H, q^2_H, q_G, t^1_H\ T(q^1_H, q^2_H, q_G), E_o(\sigma), \delta_H(q^1_H, \sigma, t^1_H) - \exists\ T(t^1_H, q^2_H, q_G), \delta_T(q^1_H, q^2_H, q_G, \sigma, \epsilon, \epsilon, t^1_H, q^2_H, q_G)$$

$$R_6 : \forall \sigma, q^1_H, q^2_H, q_G, t^2_H, t_G\ T(q^1_H, q^2_H, q_G), E_o(\sigma), \delta_H(q^2_H, \sigma, t^2_H), \delta_G(q_G, \sigma, t_G) - \exists\ T(q^1_H, t^2_H, t_G), \delta_T(q^1_H, q^2_H, q_G, \epsilon, \sigma, \sigma, q^1_H, t^2_H, t_G)$$

Example 1. Consider the DES presented by the generator $\mathcal{G}$ in Fig. 1 and specification language K generated by the automaton $\mathcal{H}_{unobs}$ in Fig. 2. Let $\Sigma_{uo} = \{u, v\}$. It may be noted that the

strings $s = uv$ and $t = \epsilon$ are those that cause the conflict in the system. Indeed, the occurrence of event a leads to the situation when $P(s) = P(t)$, $sa \in \overline{K}$ but $ta \notin \overline{K}$ what violates the observability condition.

For this example the prover Bootfrost developed for ATP in the PCF-calculus (https://github.com/snigavik/bootfrost) finds the conflict in 8 steps of the inference. Below the clipped listing of the prover log is provided. The listing illustrates the process of adding new atoms into the base that define the states and transitions of the automaton T responsible for checking observability. We have reduced the technical information where possible as well as a few inference steps which play no role for the inference. As one can see, the minimum inference for PCF $\mathcal{F}_{Obs}$ with the base corresponding to the automata in Figures 1 and 2 consists of 5 steps.

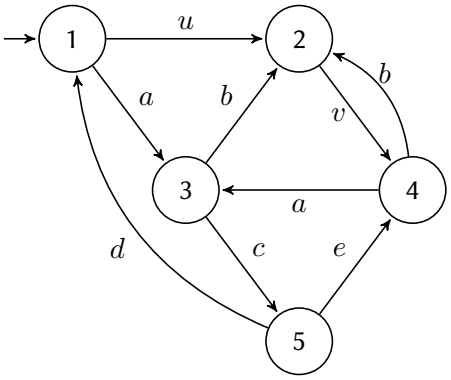


Figure 1: Automaton $\mathcal{G}$.

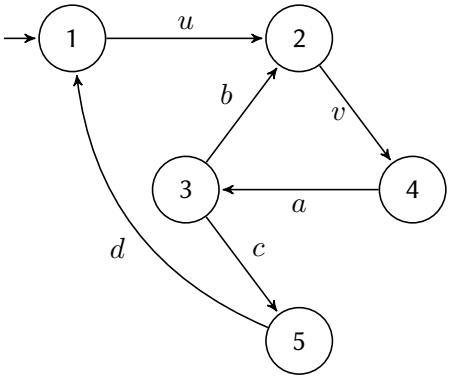


Figure 2: Automaton $\mathcal{H}_{unobs}$.

```
Arguments { formula: "wsUnObsTest.pcf", strategy: "general", limit: 10000 }
Current formula:
Base: Q(1), Q(2), Q(3), Q(4), Q(5), Q0G(1), Q0H(1),
  E("a"), E("b"), E("c"), E("u"), E("v"), E("d"), Ec("a"), Ec("b"),
  Euo("u"), Euo("v"), Eo("a"), Eo("b"), Eo("c"), Eo("d"),
  dG(1, "u", 2), dG(2, "v", 4), dG(4, "b", 2), dG(1, "a", 3), dG(3, "b", 2),
  dG(4, "a", 3), dG(3, "c", 5), dG(5, "d", 1), dG(5, "e", 4),
  FG(1, "a"), FG(1, "u"), FG(2, "v"), FG(4, "b"), FG(3, "b"), FG(4, "a"),
```

```
  FG(3, "c"), FG(5, "d"), FG(5, "c"),
  NoFG(2, "a"), NoFG(3, "a"), NoFG(5, "a"), NoFG(1, "b"), NoFG(2, "b"),
  NoFG(5, "b"), NoFG(1, "c"), NoFG(2, "c"), NoFG(4, "c"), NoFG(2, "u"),
  NoFG(3, "u"), NoFG(4, "u"), NoFG(5, "u"), NoFG(1, "v"), NoFG(3, "v"),
  NoFG(4, "v"), NoFG(5, "v"), NoFG(1, "d"), NoFG(2, "d"), NoFG(3, "d"), NoFG(4, "d"),
  dH(1, "u", 2), dH(2, "v", 4), dH(3, "b", 2), dH(4, "a", 3),
  dH(3, "c", 5), dH(5, "d", 1), FH(4, "a"), FH(3, "b"), FH(1, "u"),
  FH(2, "v"), FH(3, "c"), FH(5, "d"), NoFH(1, "a"), NoFH(2, "a"),
  NoFH(3, "a"), NoFH(5, "a"), NoFH(1, "b"), NoFH(2, "b"), NoFH(4, "b"),
  NoFH(5, "b"), NoFH(1, "c"), NoFH(2, "c"), NoFH(4, "c"), NoFH(5, "c"),
  NoFH(2, "u"), NoFH(3, "u"), NoFH(4, "u"), NoFH(5, "u"), NoFH(1, "v"),
  NoFH(3, "v"), NoFH(4, "v"), NoFH(5, "v"), NoFH(1, "d"), NoFH(2, "d"),
  NoFH(3, "d"), NoFH(4, "d")
Questions:
(0) !qH.39,qG.40 Q0H(qH), Q0G(qG)
  ? T(qH, qH, qG)
(1) !qH1.42,qH2.43,qG.44,e.45 Ec(e), FH(qH1, e), NoFH(qH2, e), FG(qG, e),
           T(qH1, qH2, qG)
  ? dead(qH1, qH2, qG, e)
(2) !x.47,y.48,z.49,s.50 dead(x, y, z, s)
(3) !qH1.51,qH2.52,qG.53,e.54,qtH1.55,qtH2.56,qtG.57 T(qH1, qH2, qG), Eo(e),
      dH(qH1, e, qtH1), dH(qH2, e, qtH2), dG(qG, e, qtG)
  ? T(qtH1, qtH2, qtG), dT(qH1, qH2, qG, e, e, e, qtH1, qtH2, qtG)
(4) !qH1.59,qH2.60,qG.61,e.62,qtH1.63 T(qH1, qH2, qG), Euo(e), dH(qH1, e, qtH1)
  ? T(qtH1, qH2, qG), dT(qH1, qH2, qG, e, eps, eps, qtH1, qH2, qG)
(5) !qH1.65,qH2.66,qG.67,e.68,qtH2.69,qtG.70 T(qH1, qH2, qG), Euo(e),
      dH(qH2, e, qtH2), dG(qG, e, qtG)
  ? T(qH1, qtH2, qtG), dT(qH1, qH2, qG, eps, e, e, qH1, qtH2, qtG)
================================ Step 0 ================================
Try question 0
New answer has been found: ...
0: {qG -> 1, qH -> 1}+134
Terms added to the base: T(1, 1, 1)
Terms used in the base: Q0H(1), Q0G(1),
Current formula:
Base: ..., T(1, 1, 1)
Questions: ...
================================ Step 1 ================================
Try question 4
New answer has been found: ...
4: {qH1 -> 1, e -> "u", qtH1 -> 2, qG -> 1, qH2 -> 1}+348
Terms added to the base: T(2, 1, 1), dT(1, 1, 1, "u", eps, eps, 2, 1, 1)
Terms used in the base: T(1, 1, 1), Euo("u"), dH(1, "u", 2),
Current formula:
Base: ..., T(1, 1, 1), T(2, 1, 1), dT(1, 1, 1, "u", eps, eps, 2, 1, 1)
================================ Step 6 ================================
Try question 4
New answer has been found: ...
4: {qH1 -> 2, qH2 -> 1, qtH1 -> 4, qG -> 1, e -> "v"}+1220
Terms added to the base: T(4, 1, 1), dT(2, 1, 1, "v", eps, eps, 4, 1, 1)
Terms used in the base: T(2, 1, 1), Euo("v"), dH(2, "v", 4),
Current formula:
Base: ..., T(1, 1, 1), T(2, 1, 1), dT(1, 1, 1, "u", eps, eps, 2, 1, 1),
  T(4, 1, 1), dT(2, 1, 1, "v", eps, eps, 4, 1, 1)
================================ Step 7 ================================
Try question 1
New answer has been found: ...
1: {e -> "a", qH1 -> 4, qG -> 1, qH2 -> 1}+584
Terms added to the base: dead(4, 1, 1, "a")
Terms used in the base: Ec("a"), FH(4, "a"), NoFH(1, "a"), FG(1, "a"), T(4, 1, 1),
Current formula:
Base: ..., T(1, 1, 1), T(2, 1, 1), dT(1, 1, 1, "u", eps, eps, 2, 1, 1),
```

```
  T(4, 1, 1), dT(2, 1, 1, "v", eps, eps, 4, 1, 1),
  dead(4, 1, 1, "a")
================================ Step 8 ================================
Try question 2
New answer has been found: ...
2: {z -> 1, x -> 4, s -> "a", y -> 1}+222
Branch has been removed
Result: Refuted
```

4. Conclusion

This paper continues the work on developing a new way of formalizing and solving control problems for the important class of dynamic systems known as DES. The PCF calculus provides powerful tools for dealing with sophisticated control problems, and above it was shown how the PCFs inference helps checking observability of regular languages.

In the future works the PCF-based approach to constructing supervisors for partially-observed DES will be suggested. Moreover, if the specification language is either not controllable or not observable, one may be interested in finding the less restricting controllable and observable sublanguage of the specification. While there are effective algorithms to construct the supremal controllable sublanguage of the given language, the supremal observable sublanguage does not exist. Only a maximal observable sublanguage may be found which is not unique in general. Finding these languages and implementing them in control systems using the PCF-approach is the line of our future research. Results obtained will be embedded at the different levels of the hierarchical control system for mobile robots.

Acknowledgments

The research is supported by the Ministry of Science and Higher Education of the Russian Federation, project no. 121032400051-9.

References

[1] S. Lafortune, Discrete event systems: Modeling, observation, and control, Annual Review of Control, Robotics, and Autonomous Systems 2 (2019). doi:10.1146/annurev-control-053018-023659.

[2] C. Seatzu, M. Silva, J. H. van Schuppen (Eds.), Control of discrete-event systems, Springer London, 2013. doi:10.1007/978-1-4471-4276-8.

[3] W. M. Wonham, K. Cai, Supervisory Control of Discrete-Event Systems, Springer International Publishing, 2019.

[4] A. Davydov, A. Larionov, Action planning for robots using the first order logic calculus of positively constructed formulas, in: 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT), volume 1, 2020, pp. 727–732. doi:10.1109/CoDIT49905.2020.9263814.

[5] E. Karpas, D. Magazzeni, Automated planning for robotics, Annual Review of Control, Robotics, and Autonomous Systems 3 (2020) 417–439.

[6] Z. Zombori, J. Urban, C. E. Brown, Prolog technology reinforcement learning prover, in: International Joint Conference on Automated Reasoning, Springer, 2020, pp. 489–507.
[7] S. N. Vassilyev, Machine synthesis of mathematical theorems, The Journal of Logic Programming 9 (1990) 235–266. doi:10.1016/0743-1066(90)90042-4.
[8] A. K. Zherlov, S. N. Vassilyev, E. A. Fedosov, B. E. Fedunov, Intelligent control of dynamic systems, Fizmatlit, Moscow, 2000. In Russian.
[9] A. Davydov, A. Larionov, E. Cherkashin, On the calculus of positively constructed formulas for automated theorem proving, Automatic Control and Computer Sciences 45 (2011) 402–407. doi:10.3103/s0146411611070054.
[10] A. Larionov, A. Davydov, E. Cherkashin, The calculus of positively constructed formulas, its features, strategies and implementation, International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija 2013 (2013) 1023–1028.
[11] P. J. Ramadge, W. M. Wonham, Supervisory control of a class of discrete event processes, SIAM Journal on Control and Optimization 25 (1987) 206–230. doi:10.1137/0325013.
[12] A. Davydov, A. Larionov, N. Nagul, On checking controllability of specification languages for des, in: 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020, pp. 1151–1156. doi:10.23919/MIPRO48935.2020.9245154.
[13] A. Davydov, A. Larionov, N. V. Nagul, The construction of controllable sublanguage of specification for des via pcfs based inference, in: I. Bychkov, A. Tchernykh (Eds.), Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments, ICCS-DE 2020, Irkutsk, Russia, July 6-7, 2020, volume 2638 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020, pp. 68–78.
[14] A. Davydov, A. Larionov, N. Nagul, Application of the PCF calculus for solving the problem of nonblocking supervisory control of discrete event systems, Journal of Physics: Conference Series 1864 (2021) 012048. URL: https://doi.org/10.1088/1742-6596/1864/1/012048. doi:10.1088/1742-6596/1864/1/012048.
[15] H. Cho, S. I. Marcus, Supremal and maximal sublanguages arising in supervisor synthesis problems with partial observations, Mathematical systems theory 22 (1989) 177–211. URL: https://doi.org/10.1007/BF02088297. doi:10.1007/BF02088297.
[16] K. Inan, An algebraic approach to supervisory control, Math. Control. Signals Syst. 5 (1992) 151–164. URL: https://doi.org/10.1007/BF01215843. doi:10.1007/BF01215843.
[17] C. G. Cassandras, S. Lafortune, Introduction to Discrete Event Systems, Springer US, 2008.

A Conceptual Model of Agile Meetings' Problems and Their Relationships with Project Issues in IT Industry

Maja Gaborov[1], Zeljko Stojanov[1], Srđan Popov[2], Dragana Kovač[1]

[1]*University of Novi Sad, Technical faculty "Mihajlo Pupin", Djure Djakovica bb, Zrenjanin, Serbia*
[2]*University of Novi Sad, Faculty of Technical Science, Trg Dositeja Obradovica 6, Novi Sad, Serbia*

Abstract

Agile project management methodologies are applied in the IT industry. Agile meetings are the core of agile methodologies and deserve the attention of both researchers and industry practitioners. The aim of this paper is to present a conceptual model of problems in agile meetings and their relationship with project issues in the IT industry. Problems are identified through literature analysis, and the model is based on the interpretation of identified problems and their relationships. The presented model can be used as a starting point for a deeper and more comprehensive examination of agile meetings, as well as for the development of guidelines and software tools for managing potential problems in the IT industry. A conceptual model can help practitioners understand how important it is to organize a meeting well. We believe that in the future it would be good to create software that instead of holding daily standup meetings, employees could write their impressions. We used Visio for modeling. Using this conceptual model, one can see the problems in organizing meetings and expand the conceptual model with more problems that can potentially arise.

Keywords

Agile methods, agile meeting problems, conceptual model, project, software industry

1. Introduction

A project is a temporary undertaking that has a beginning and an end. It is undertaken to create a unique product, service or result. A project is an enterprise that is started to meet market demand, to take advantage of business needs, to satisfy customer requirements, etc. Meeting the project deliverables without compromising its quality, scope and time frame within a limited budget is the key to project success. Project management is the most important tool that every organization needs to fulfill to achieve project results [1]. Today, agile methodologies are mainly used for project management. Agile software development (ASD) is an umbrella term for a set of incremental and iterative development methods [2]. In recent decades, the adoption of ASD has been extremely fast in the software industry worldwide [2, 3, 4]. ASD has a clear emphasis on people and a rapid response to change [2, 5]. The main goal of the methodology is to deliver the product to the customer on time [6, 7]. Some other specifics of agile methodologies can be highlighted, such as: monitoring project progress through the organization of meetings, client feedback is discussed at meetings, frequent meetings take a lot of time for project implementation, etc. [8] Meetings are necessary during project implementation. Meetings are events where everything should be discussed and potential problems should be prevented and problems solved. In industry surveys [2, 9], Scrum was identified as the most commonly used agile method. Henriksen et al concluded that publishing an agile manifesto increased the success rate of agile software development projects [10]. Given that agile methodologies are able to provide innovation and competitiveness, further research is encouraged to find new ways to reduce failure rates [10,11].

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: *maja.gaborov@tfzr.rs, zeljko.stojanov@uns.ac.rs, srdjanpopov@uns.ac.r*s, dragana.milosavljev@tfzr.rs
ORCID: 0000-0002-3810-6156; 0000-0001-6930-5337; 0000-0003-1215-3111; 0000-0002-5421-4376


ICCS-DE 2023 Workshop Proceedings

Since in recent years there has been a boom in research in the field of agile methodologies, in the software industry there is a need for clear guidelines on agile implementation and new issues, where there is a great focus on organizing meetings, all in order to increase the success of projects.

Based on the above, the goal of this paper is to present a conceptual model of problems in agile "meetings", their relationships, as well as the relationship with project problems in IT companies. The rest of the paper is structured as follows. The second section outlines related work in the area of conceptual models in agile methodologies. The third part presents a conceptual model of problems with agile meetings and their relationship with project issues in IT companies. The fourth section contains a discussion of the model, research implications, and validity of the research. The final part contains concluding remarks and directions for further research.

2. Related work

"Conceptual modeling is the activity of deciding what to model and what not to model – 'model abstraction'. A conceptual model is a non-software specific description of the computer simulation model describing the objectives, inputs, outputs, content, assumptions, and simplifications of the model" [12,13]. Shakya et al. consider the size of the project as a critical success factor, which contributes to the success of the project through the successful implementation of agile methodology in a software company in Nepal. It is predicted that "project size" will have the greatest impact on "ability to respond to change"[1].

A study conducted by Silvius et al. [14], plans to investigate the relationship between sustainability considerations and the perception of project success. Based on a literature review of the two main variables, sustainability and project success, a conceptual model for the study was developed that showed that the relationship between sustainability and project success is not a simple one. In the model, nine dimensions of sustainability are identified and the measures for project success are clustered into six criteria. With this model, a more detailed understanding of how considering different dimensions of sustainability may affect the individual criteria of project success.

3. Methods and development of conceptual model

Conceptual modeling is the abstraction of a model from a real or proposed system [15,16]. This process of abstraction involves some level of simplification of reality [15]. In this paper, a conceptual model was created based on the literature analysis. The development of the conceptual model is based on the analysis of published works that were collected and analyzed by using a method for systematic review of literature [17], which was done for the purposes of the doctoral dissertation. Selected studies are used for the development of a conceptual model that present relationships between problems with agile meetings and project issues.

Meeting issues and project issues are represented in the conceptual model in Figure 1. The relationships between meeting issues and project issues are highlighted by dashed lines. Meeting issues are marked with blue ellipses and project issues are marked with gray rectangles. Problems from one group are also related to each other. Problems in meetings are connected with blue lines, while problems of a project nature are marked in black. All the problems of individual meetings are explained in the following subsections, as well as the relationship between these problems and project issues.

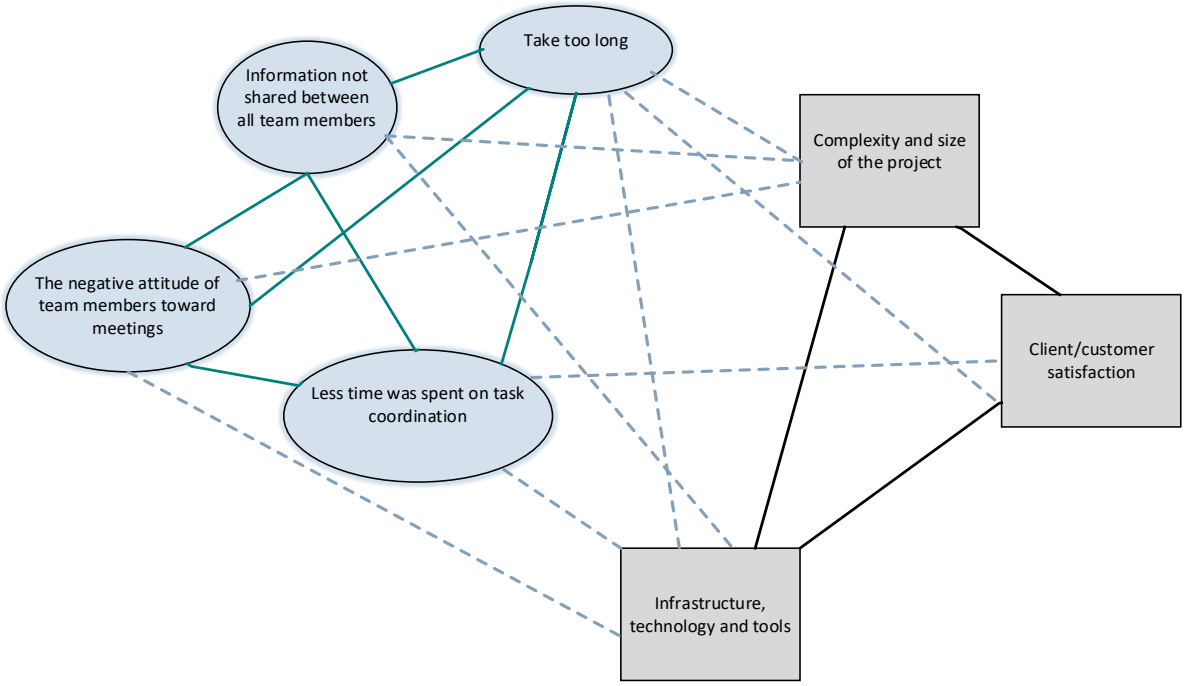


Figure 1. Conceptual model of agile meeting problems and relationships with project issues

The model shows the problems from the meetings identified in references [6,17,18,19,18,20,21,22]: meetings take a long time; information is not shared between all team members; negative attitude of team members towards meetings (reduces job satisfaction, trust, well-being); less time was spent on task coordination (more was spent on working out some other problematic issues). Also, the model shows project issues: project complexity and project size [1,23], client/customer satisfaction [1,24,25], and Technology, Infrastructure and tools [26,23,27].

3.1. Agile meetings problems

Based on the literature search, there were several problems related to meetings.

3.1.1. Take too long

In the searched literature, the meetings lasted a long time. It may happen that the meetings are longer than planned, but there are benefits from such meetings. Empirical evidence shows that many respondents complained that meetings took too long [6].

3.1.2. Information not shared between all team members

In a study [22] that describes and interprets the lived experiences of agile developers with daily stand-up meetings based on the experiences of 19 professional agile developers, developers who have experience with such meetings say that they were too short to enable clear identification and resolution of problems or lacked meaningful outcome. With experienced developers, meetings were too short to facilitate clear problem identification and resolution, or lacked a meaningful outcome. Older developers experienced these meetings differently than younger developers in terms of sharing information, taking an interest in others' work, monitoring progress, and facilitating decision making.

3.1.3. The negative attitude of team members toward meetings

Team members believed that meetings are held frequently, attitudes towards meetings became negative [6]. A subsequent study [28] included 60 members from 15 teams in five countries. Many team members have had a negative experience leading a meeting, which reduces job satisfaction, trust, and co-workers' well-being. A study [19], involving three companies in Malaysia, Norway, Poland and the United Kingdom, was conducted to obtain personal attitudes about meetings. The factors that most contributed to a positive attitude towards the daily stand-up meeting were the sharing of information with the team and the opportunity for discussion and problem solving. The factors that most contributed to the negative attitude are reporting the status to the manager and too high frequency of meetings and too long duration compared to other work activities.

3.1.4. Less time was spent on task coordination

Stray et al. [21] analyzed eight daily meetings of two software development teams. On average, only a small percentage of each meeting was focused on development tasks. They found that the majority of meetings were spent on working out problematic issues and discussing possible solutions. Very little time was spent on task coordination.

3.2. Project issues and problems

Based on the literature search, there were several problems related to project problems.

3.2.1. Complexity and size of the project

Shakia et al. [1] believe that when organizing a meeting, it is important to take into account the complexity and size of the project. If it is a bigger project, it takes more time to complete the whole project and generally more people are involved in the project.

3.2.2. Client/customer satisfaction

In the study by Uikei et al. [24], where the data was collected from a sample population of project managers from different software organizations, who apply agile methods in the development of their projects, the emphasis is on the customer or product owner, who represents an important subject in the development of agile software. It is very important to organize meetings so that the product owner agrees with the customer. It provides product requirements and information and ultimately approves the working software. According to the author's research, most of the time the customer is actively involved [24].

3.2.3. Infrastructure, technology and tools

In the study by Stadler et al. [26], more than half of the respondents stated that poor infrastructure and Internet availability were one of their biggest problems. It is very important to hold quality meetings through technology, especially when it comes to geographically distant teams. Communication and collaboration in distributed teams build on reliable internet connectivity technology like communication hardware and software. A regularly reported issue is the lack of those requirements like an unsteady network connection which massively impedes synchronous and frequent communication, in turn having a negative impact on coordination and control. Their results show that applying agile methods in distributed teams with low geographical distance poses no problem but rather brings forth several benefits. Common agile practices were mastered successfully by replacing face-to-face communication with a variety of digital communication channels. These practices not only improve the software engineering process but furthermore pose additional communication channels and information radiators which in turn improve collaboration. Daily meetings and general meetings of short duration (with durations typically not planned much longer than one hour) are done smoothly with audio or video conferencing. When it comes to informal, complex communication situations on the other hand, all

interviewees reported that they prefer to bring their teams together in one physical location if possible. Interacting with remote colleagues is done with the excessive use of collaboration and management tools which function as information radiators to keep distant team members informed. Although no substantial barrier, there are aspects that require increased attention like the scheduling of meetings or extra effort to uphold communication between sites [26,12]. Inconsistency of tools among teams is one of the problems in the study by Beecham et al. [27]. There can be an unfortunate interaction between geographic distance, temporal distance, and technology [27].

3.3. Relationships between agile meetings' problems and project issues

The fact is that meetings often last a long time. This can happen due to the complexity and size of the project. Usually, large projects involve more people and therefore it is more difficult to organize meetings, and when they are organized they can last longer because there are more people and more ideas and the like. Because of all this, people can develop a negative attitude towards meetings. As a result, it may happen that not all information is shared at meetings that would be important for some employees and that less time is spent working on specific tasks. Also, meetings can last longer because it can happen that the teams are not in the same location, so it is more difficult to agree on everything because they are not in daily live contact, and communication can be difficult, for example, if the Internet connection is not good. If communication is difficult, it might be forgotten to share all the information that some team members would need, then some team members would waste a lot of time in meetings instead of solving their tasks, and thus would create stress in people and a negative attitude towards meetings. In addition to all this, it is important that the client/customer is involved in the project and that their requirements are respected because the goal of the project is to deliver the product to the client/customer on time. If meetings last longer and slow down the entire process of project implementation, where employees spend less time on project implementation, i.e. concretely solving tasks, it can affect the client/customer who may become dissatisfied.

4. Discussion

Based on the literature review, the conceptual model with relationships between the mentioned problems was created and described. From the conceptual model it can be concluded that it is necessary that the organizer of the meeting (project) influences the duration of the project, so that only important issues are resolved at the meetings. It would not be desirable for meetings to last a long time because employees will spend a lot of time in meetings, and they will have less time to solve their tasks within their working hours, and this leads to employee dissatisfaction and possibly to worse results. In addition, it is necessary to plan the meetings well, considering that some projects can be large where more employees are engaged, so there may not be a need for all those employees to be at those meetings and the like. Also, it is advisable to ensure that all information reach every person for whom it is important. In addition to all this, the requirements of clients/customers should not be neglected because the ultimate goal is to deliver a quality product to the customer/client on time. The customer/client should be satisfied with the product.

4.1. Research implications

In this section, the research implications for practitioners from industry and researchers from academia will be discussed.

This conceptual model can be used by industry practitioners. Although this conceptual model contains only specific project problems and problems that arise in meetings, it can help practitioners understand how important it is to organize a meeting well, because it is very important that the employees of the company do not feel bad, as well as because the clients feel satisfied. Also, this could help practitioners to understand that they should try to automate the meeting process and thus make it easier for employees since they have a lot of meetings and need a lot of time to do personal tasks. It would be desirable to create software that instead of holding meetings, would allow each of the

employees to write their impressions and thus reduce attendance at meetings, but would still provide important information for all members.

Researchers can learn a lesson from this work on how to create a conceptual model based on literature analysis of specific problems. This study by Davies et al [29], found that Visio is clearly the best tool for modelling business systems. We used that modelling tool. Using this conceptual model, researchers can see some of the problems in organizing meetings and expand this conceptual model with more problems that can potentially arise. Furthermore, the presented conceptual model can serve as a starting point for a more detailed examination of the problems in organizing meetings.

4.2. Validity of the research

Although stated implications indicate several benefits of this work, the authors are aware that there exist some constraints that affect the validity of the research. Therefore, reliability of the research findings can be increased by discussing constraints that affect validity of the study [30], which includes considering internal and external validity.

Internal validity relates to all activities performed during the research that leads to construction of the research findings. These activities relate to finding relevant studies, extracting data, construction of the conceptual model and its discussion. For the selection of studies, we followed guidelines for performing systematic literature reviews [31], and clearly describe the whole process. We applied inclusion/exclusion criteria and quality assessment criteria to minimize this type of threat to select the most relevant literature. It is possible that we missed some case studies that were published in digital libraries we did not search. However, our aim was to search the most influential libraries in which papers are published after rigour review process. Data extraction from available papers was difficult because many studies did not explicitly mention and explain each of the problems we observed, requiring interpretation of the data, which includes personal bias. This threat was minimized by having some of the issues appear in multiple papers (by different authors). In addition, all authors of this paper participated in the discussion and development of the findings, and finally agreed on the defined problems in the conceptual model.

External validity relates to the generalizability of the findings presented in this study. The findings relate to problems in meeting organized in projects that apply agile methodologies, and therefore, applicability in projects that implement other methodologies (in general, waterfall-based methodologies) is questionable. However, majority of recent projects and research in IT industry are based on agile methodologies, which increases generalizability of the presented conceptual model. In addition, the detailed description of the process that leads to construction of the conceptual model can be used for developing similar conceptual models for other aspects of managing IT projects, which additionally increases the usability and generalizability of the presented study and findings.

5. Conclusion

This paper discusses the importance of agile meetings and the problems that may arise during their implementation. This is shown with the conceptual model. The model was created based on the literature analysis. The advantage of this model is that managers and employees in IT industry can see specific problems that occur in projects, as well as problems related to meetings, where they can more easily recognize which of these problems are potentially occurring in their companies. Well-conducted meetings are of great importance for projects and companies in general. At well-organized and conducted meetings, people will come to common solutions that are important for the project, none of the employees will feel bad, they will not think that they are wasting time because the meeting facilitates the understanding of all tasks. Also, it is important to have the right technology and tools for the job. If the teams are remote, the communication needs to be really good and the teams have the right technology to communicate. All these issues influence the fact that all tasks are done in a good way and that the clients are satisfied with the final outcome.

In the future, the authors will develop a more comprehensive model that will also represent other issues related to the meeting, such as human resources issues or external factors. The research findings of this study build a strong case for the need to rethink how meetings are conducted in the IT industry

and how to mitigate or prevent problems in practice. Future work will also be directed toward developing guidelines and software tools to assist IT professionals in avoiding and solving problems related to agile meetings.

6. References

[1] Shakya, P., & Shakya, S. (2020). Critical success factor of agile methodology in software industry of nepal. Journal of Information Technology, 2(03), 135-143.

[2] Shastri, Y., Hoda, R., & Amor, R. (2021). The role of the project manager in agile software development projects. Journal of Systems and Software, 173, 110871.

[3] T. Dingsøyr, S. Nerur, V. Balijepally, and N. B. Moe, "A decade of agile methodologies: Towards explaining agile software development," Journal of Systems and Software, vol. 85, no.6, pp. 1213-1221. (2012).

[4] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic 14 review," Inform. and Software Technology, vol. 50, no. 9-10, pp. 833-859, 2008.

[5] V. Stray, N.B. Moe., and D.I. Sjoberg, Daily stand-up meetings: start breaking the rules. IEEE Software, 37(3), 70-77, 2018

[6] V.G. Stray, Y. Lindsjørn, and D. I. Sjøberg, Obstacles to efficient daily meetings in agile development projects: A case study. In 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 95-102). IEEE. (2013).

[7] N. B. Moe, T. Dingsøyr, and K. Rolland. To schedule or not to schedule? An investigation of meetings as an inter-team coordination mechanism in large-scale agile software development. (2018).

[8] S. Milovanović, Primena agilnih metoda razvoja softvera u mobilnim tehnologijama, INFOTEH-Jahorina, Vol 13, (2017). Bosnia and Hercegovinia

[9] Digital.ai Software Inc. (2020). VersionOne 14th Annual State of Agile Report. Digital.ai 2 Software Incorporated. [Online]. Available: https://stateofagile.com/#ufh-i-615706098-14th3 annual-state-of-agile-report/7027494

[10] C. Tam, E.J.da Costa Moura, T. Oliveira, and Varajão, J. The factors influencing the success of on-going agile software development projects. International Journal of Project Management, 38(3), 165-176. (2020)

[11] Conforto, E. C., Amaral, D. C., da Silva, S. L., Di Felippo, A., & Kamikawachi, D. S. L. (2016). The agility construct on project management theory. International Journal of Project Management, 34(4), 660-674

[12] M. Gaborov, Z. Stojanov, M. Kavalić, I. Vecštejn and S. Popov, "A conceptual model of agile meetings' problems and their relationships with organizational issues in IT industry," 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina, 2023, pp. 1-6, doi: 10.1109/INFOTEH57020.2023.10094204.

[13] Robinson, S., Arbez, G., Birta, L. G., Tolk, A., and Wagner, G. Conceptual modeling: definition, purpose and benefits. In 2015 winter simulation conference (wsc) (pp. 2812-2826). IEEE.

[14] Silvius, A. G., & Schipper, R. (2015). A conceptual model for exploring the relationship between sustainability and project success. Procedia Computer Science, 64, 334-342.

[15] Robinson, S. (2006, December). Conceptual modeling for simulation: issues and research requirements. In *Proceedings of the 2006 winter simulation conference* (pp. 792-800). IEEE.

[16] Robinson, S. (2008). Conceptual modelling for simulation Part I: definition and requirements. *Journal of the operational research society*, *59*, 278-290.

[17] B. Kitchenham, O. P Brereton,.D. Budgen,, M. Turner, J. Bailey, and S. Linkman, Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1), 7-15. (2009)

[18] V. Stray, N.B. Moe., and D.I. Sjoberg, Daily stand-up meetings: start breaking the rules. IEEE Software, 37(3), 70-77, 2018.

[19] V. Stray., D. I. Sjøberg, and T. Dybå, The daily stand-up meeting: A grounded theory study. Journal of Systems and Software, 114, 101-124. (2016).

[20] V. Stray., N. B. Moe, and G. R Bergersen. Are daily stand-up meetings valuable? A survey of developers in software teams. In International Conference on Agile Software Development (pp. 274-281). Springer, Cham. (2017).
[21] V. G.Stray, N. B Moe, and Aurum, A. Investigating daily team meetings in agile software projects. In 2012 38th Euromicro Conference on Software Engineering and Advanced Applications (pp. 274-281). IEEE. (2012).
[22] K. Singh,and J. Strobel Exploring lived experiences of agile developers with daily stand-up meetings: a phenomenological study. Behaviour& Information Technology, 1-21. (2022).
[23] Badampudi, D., Fricker, S. A., & Moreno, A. M. (2013). Perspectives on productivity and delays in large-scale agile projects. In International Conference on Agile Software Development (pp. 180-194). Springer, Berlin, Heidelberg.
[24] Uikey, N., & Suman, U. (2012, September). An empirical study to design an effective agile project management framework. In Proceedings of the CUBE International Information Technology Conference (pp. 385-390).
[25] Wiesche, M. (2021). Interruptions in agile software development teams. Project Management Journal, 52(2), 210-222.
[26] Stadler, M., Vallon, R., Pazderka, M., & Grechenig, T. (2019). Agile distributed software development in nine central European teams: Challenges, benefits, and recommendations. International Journal of Computer Science & Information Technology (IJCSIT) Vol, 11.
[27] Beecham, S., Noll, J., & Richardson, I. (2014, August). Using agile practices to solve global software development problems--a case study. In 2014 IEEE International Conference on Global Software Engineeering Workshops (pp. 5-10). IEEE.
[28] V. Stray, N. B. Moe, and D.I. Sjoberg, Daily stand-up meetings: start breaking the rules. IEEE Software, 37(3), 70-77. (2018).
[29] Davies, I., Green, P., Rosemann, M., Indulska, M., & Gallo, S. (2006). How do practitioners use conceptual modeling in practice?. *Data & Knowledge Engineering*, *58*(3), 358-380.
[30] Wright, H. K., Kim, M., & Perry, D. E. (2010, November). Validity concerns in software engineering research. In *Proceedings of the FSE/SDP workshop on Future of software engineering research* (pp. 411-414).
[31] Kitchenham, B. A., Budgen, D., & Brereton, P. (2015). *Evidence-based software engineering and systematic reviews* (Vol. 4). CRC press.

The multi-depot vehicle routing problem for group monitoring operations on discrete-continuous space

Maksim Kenzin, Igor Bychkov and Nikolai Maksimkin

Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian Academy of Sciences, 134 Lermontova St, Irkutsk, 664033, Russian Federation

Abstract

The paper presents the periodic monitoring problem for a group of mobile agents. The problem is to construct reliable group route for agents providing the required quality of the operational area coverage and ensuring well-timed return of each agent to its point of primary dislocation. Besides the fact that such points (depots) are required to place individually as a part of the of problem solving, an additional feature of the task is probabilistic nature of the coverage procedure itself. A mathematical model of the described problem is formulated in terms of cyclic vehicle routing problems. An evolutionary algorithm together with a specialized set of genetic operators and heuristics have been developed and software-implemented for its effective solution.

Keywords

vehicle routing problem, cyclic routing, variable multi-depot, probabilistic area coverage, hybrid evolutionary algorithm

1. Introduction

Operational coverage challenges in different formulations regularly arise in a variety of fields, ranging from smart farms and environmental monitoring to search and rescue operations and even air defence missions. In general, the coverage problem assumes two implementation options: using a static surveillance system covering the entire monitoring area, and using a group of mobile vehicles. Due to the high resource intensity of the first approach, it is the use of mobile, mostly autonomous vehicles that are becoming an increasingly popular option [1].

Since the radius of action and the number of such vehicles (agents) are usually insufficient to simultaneously cover the entire area under surveillance, their collective travel route has to be constructed so as to cover it sequentially (step-by-step) in a minimum number of movements [2]. At the same time, the large-scale nature of coverage tasks creates a number of difficulties in solving the problem of constructing effective group trajectories. The most complicated among them are constraints imposed by the limited resources of the operating agents.

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

gorthauers@gmail.com (M. Kenzin); idstu@icc.ru (I. Bychkov); mnn@icc.ru (N. Maksimkin)

0000-0001-7050-3154 (M. Kenzin); 0000-0002-1765-0769 (I. Bychkov)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

2. Problem formulation

Let O be a continuous operational area, which is required to be monitored on a regular basis. For the sake of clarity, in the following examples we will consider a rectangular operational area $O = \{(x, y)|x \in [0, X], y \in [0, Y]\}$.

Inside the area O there is a network of m control (dislocation) points $V = \{v_1, v_2, ..., v_m\}$, $v_i \in O$, $i = \overline{1, m}$ connected by a network of roads E. Thus, $G = (V, E)$ is a connected undirected graph of m vertices fully located within the operational area.

The monitoring task is carried out by the group of n identical mobile vehicles (agents) $a_1, a_2, ..., a_n$ that are able to detect unauthorized objects. Being located on any control point $v \in V$, an agent provides area coverage within the range of the detection equipment r. The probability of the object detection is inversely proportional to the distance and is defined by a function $p(v, (x, y)) \in [0, 1]$. The most common example of such a function is a linear decreasing function $p = 1 - \min(d, r)/r$, where d stands for the distance between the agent and the object.

Each agent can travel on the graph G on a discrete-time $T = \{T_0, T_1, T_2, ...\}$. Between two time steps an agent can either remain on its position or traverse exactly one edge (and only if the edge length does not exceed the given value L). At the same time, the first option (idle) is undesirable since the agent's static position makes it easier for tracing, which may eventually threaten the safety of agent and of the entire operation.

Since the travel resource of each agent is bounded by t time steps, several maintenance stations (so called depots) must be deployed at some dislocation points in order to support uninterrupted monitoring of the area. Thus, each agent must always start its movement from the depot and return there by the end of t time steps. In this paper, we do not consider the resource costs of deploying the depots and assume that any number of them can be arranged.

Hence, the group route for the problem is in the following form:

$$R = \Big((v_{11}, v_{12}, ..., v_{1t}), (v_{21}, v_{22}, ..., v_{2t}), ..., (v_{n1}, v_{n2}, ..., v_{nt})\Big), \tag{1}$$

where $v_{ij} \in V$, $i = \overline{1, n}$, $j = \overline{1, t}$. The coverage efficiency is defined as the average probability of detecting a random object within the area O at each time step:

$$F(R) = \sum_{j=1}^{t} \{P_j(x, y)|(x, y) \in O\}/tS_O \in [0, 1], \tag{2}$$

where S_O is the area of the operational area, and $P_j(x, y)$ estimates the probability of detecting an object in (x, y) by all agents from their positions according to the step j of the group route R. Here the probability of detection by the whole group is calculated as the probability that a situation will not occur where no agent detects an object:

$$P_j(x, y) = 1 - \prod_{i=1}^{n} (1 - p(v_{ij}, (x, y))).$$

This leads us to the following vehicle routing problem: to form a set of n cyclic routes of length t (1) that would provide the most efficient coverage (2).

3. Evolutionary algorithm

To solve the suggested problem we propose using a problem-based modification of genetic algorithm. When implemented correctly, such population-based approaches allow for efficient construction and reconstruction of feasible high-quality solutions at low time costs. In addition, it provides high scalability and flexibility along with wide hybridization capabilities to address the essential features of the problem [3].

The designed variation of the genetic algorithm is equipped with a set of problem-oriented constructive and improvement heuristics and uses a Monte Carlo-based fitness function with a penalty component for feasibility violation. The algorithm treats group routes (1) as chromosomes, changing and combining them gradually to find the nearest to the optimal one.

3.1. Fitness Function

Since it seems inefficient to directly evaluate the coverage quality (2) for the entire monitoring area in terms of computational and time costs, two variants of the operational area discretization can be proposed. The first traditional option is using a sampling grid with a given cell size c. The second approach is to use the Monte Carlo method, randomly rearranging s sampling spots each time a calculation is required. In this case, the number of spots can vary from generation to generation, depending on the required accuracy and calculation speed. The spot number also defines whether the focus on the exploration or exploitation of the search space will prevail. In addition, we can manipulate the distribution function for those cases in which the whole area coverage importance is not uniform.

Thus, we propose to estimate the coverage effectiveness of the group route R through the average probability of non-detection of a random set of s objects located at points $\{(x_1, y_1), ..., (x_s, y_s)\}$ correspondingly:

$$f_1(R) = \sum_{j=1}^{t} \sum_{k=1}^{s} P_j(x_k, y_k)/ts \in [0, 1]. \tag{3}$$

When genetic operators rearrange vertices, the feasibility of the routes themselves may be violated, causing agents to travel along non-existent roads or to fail in returning to the depot at the end of the cycle. One could simply forbid such damaging changes, or conduct a specially written restoration procedure, but in our experience, it is preferable to keep such solutions and charge them with a penalty fee as follows:

$$f_2(R) = \sum_{i=1}^{n} \Big(|1 - l(v_{i1}, v_{it})| + \sum_{j=2}^{t} |1 - l(v_{ij}, v_{ij-1})| \Big), \tag{4}$$

where $l(v_i, vj)$ returns the travel distance between two dislocation points (minimal number of time steps required to get from one to the other). Using a penalty function (4), we penalize agents both for traveling on invalid graph edges, for not returning to the depot and for idling.

Thus, the final fitness function is a weighted sum of two functions (3) and (4):

$$f(R) = w \cdot (1 - f_1(R)) + f_2(R) \longrightarrow \min .$$

3.2. Constructive and Improvement Heuristics

Most intelligent meta-heuristic algorithms use custom-designed constructive heuristics to generate an initial population of reasonable quality. We propose using two different heuristics that consistently generate routes for one agent after another following the corresponding procedure:

- The first heuristic builds a random path on a graph G for each agent i starting from a random dislocation point v_{i1} and may end up being unfeasible;
- The second "feasible" heuristic firstly defines a random depot station v_{i1} for each agent i, then defines a random transitional point v_{ie} such that $l(v_{i1}, v_{ie}) \leq t/2$, and then connects these two dislocation points by two different shortest paths, if possible.

A standard three-step scheme of selection, crossover, and mutation are used to reproduce new solutions. Firstly, a 2-way tournament gives a chance to different-quality solutions to undergo procreation. Then the uniform crossover generates offspring solutions by choosing n random agent's routes from two parents. Finally, the multi-mode mutation is applied to slightly change the resulting solution. The most efficient mode is replacing, which destroys the whole route of a random agent and then recreates it using one of the constructive heuristics described above.

4. Conclusion

Being software-implemented, the proposed approach has demonstrated high efficiency along with good reliability and scalability through several conducted series of simulation studies. It allows for the fast generation of feasible high-quality solutions and their gradual improvement towards local and global optima. High customizability makes it easy to adapt the computation process to the specific problem statement. Still, both the proposed mathematical model and the designed genetic algorithm are open to various possible extensions and developments.

Acknowledgments

This research was funded by the Russian Science Foundation, project no. 22-29-00819.

References

[1] H. S. Munawar, A. W. Hammad, S. T. Waller, Disaster region coverage using drones: Maximum area coverage and minimum resource utilisation, Drones 6 (2022). doi:10.3390/drones6040096.

[2] N. Drucker, H.-M. Ho, J. Ouaknine, M. Penn, O. Strichman, Cyclic-routing of unmanned aerial vehicles, Journal of Computer and System Sciences 103 (2019) 18–45. doi:10.1016/j.jcss.2019.02.002.

[3] Çağrı Koç, T. Bektaş, O. Jabali, G. Laporte, Thirty years of heterogeneous vehicle routing, European Journal of Operational Research 249 (2016) 1–21. doi:10.1016/j.ejor.2015.07.020.

Сервис мониторинга сферы туризма территории на основе анализа информационных веб-ресурсов

Ольга А. Николайчук [1], Юлия В. Пестова [1], Александр И. Павлов [1] and Дмитрий Е. Косогоров [1]

[1] *Институт динамики систем и теории управления им. В.М. Матросова СО РАН, ул. Лермонтова, д. 134, Иркутск, 664033, Россия*

Аннотация

В статье выполнена постановка задачи разработки информационной технологии мониторинга сферы территориального туризма, описана концепция сервиса, обоснованы методы анализа и визуализации данных, сформирована онтологическая модель, определено содержание информационных панелей (дашбордов), реализован прототип сервиса, представлены результаты сбора, обработки и представления полученных в результате исследования данных. В настоящее время в рамках предложенной концепции сервиса реализован прототип функции «формирование туристского профиля территории» на примере Байкальской территории. В дальнейшем планируется развить данный сервис и наполнить его данными согласно предложенной концепции.

Ключевые слова

Туризм, открытые источники данных, мониторинг, онтология, интеграция геоданных, анализ текстов, веб-сервис

1. Введение

Туризм является одной из наиболее быстро растущих отраслей мировой экономики, важным источником валютных поступлений и рабочих мест, тесно связан с социальным, экономическим и экологическим благополучием страны. Всемирная туристская организация определила понятие устойчивого туризма: «туризм, в полной мере обеспечивающий учет его нынешних и будущих экономических, социальных и экологических последствий при удовлетворении потребностей туристов, индустрии туризма, окружающей среды и принимающих общин». Достижение устойчивого туризма — это непрерывный процесс, требующий постоянного мониторинга воздействия и принятия необходимых превентивных и/или корректирующих мер, когда это необходимо.

Существуют региональные системы мониторинга (обсерватории), например, shapetourism, которая в виде интерактивных карт представляет инструмент интерпретации динамики туризма на основе четырех измерений: репутации, привлекательности, конкурентоспособности и устойчивости, и охватывает 52 страны средиземноморского региона (https://www.quantitas.it/data/shapetourism/build/index.php). Другой вид систем – это системы мониторинга отдельных достопримечательностей или территорий, например, в Австралии (https://www.destinationnsw.com.au/about-us), Португалии [1], Китае [2], Франции [3].

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: nikoly67@mail.ru (A. 1); yupest@gmail.com (A. 2); Asd@icc.ru (A. 3); dmitriy.kosogorov@yandex.ru (A. 4)
ORCID: 0000-0002-5186-0073 (A. 1); 0000-0001-5874-7965 (A. 2); 0000-0002-7753-7514 (A. 3)

ICCS-DE 2023 Workshop Proceedings

Таким образом, в мире определена цель развития устойчивого туризма, отмечено, что реализация данной цели возможна только при осуществлении постоянного мониторинга данной сферы на основе всесторонних данных.

Необходимо отметить еще один фактор – это значительное увеличение объема данных, связанных со сферой туризма. Данные создаются в среде Интернет туристическими агентствами, туристическими объектами и потребителями туристических услуг. Имеющиеся данные характеризуются массовостью, разнородностью, иногда слабой структурированностью, неполнотой и противоречивостью. Взрывное увеличение количества данных создает потребность в разработке методов и программных инструментов для интеграции источников данных, поиска релевантных данных, их извлечения и интерпретации [3].

Целью данной работы является описание концепции сервиса мониторинга туризма Иркутской области, как части Байкальской территории, определение методов и средств технологии реализации данного сервиса и представление отдельных результатов его прототипирования.

2. Туризм Иркутской области

Иркутская область является одним из крупнейших регионов России. Юго-восточная граница Иркутской области проходит по озеру Байкал [4], которое в 1996 году было внесено в Список объектов Всемирного наследия ЮНЕСКО [5].

Безусловно озеро Байкал является объектом притяжения туристов. В Национальном туристическом рейтинге-2021 Иркутская область заняла 15 место и вошла в золотую двадцатку регионов, наиболее привлекательных для отечественных и иностранных туристов [6].

В течение последних 20 лет численность размещенных лиц в коллективных средствах размещения области неуклонно росла и в 2019 году превысила 1 млн. человек, а их доходы в 2021 г. составили 6877621,3 тыс. руб. Данный факт свидетельствует, с одной стороны, об активном развитии сферы туризма в Иркутской области, а с другой – о повышении различных экологических рисков на территории уникального природного объекта. Мониторинг и управление данными рисками является насущной задачей.

3. Концепция сервиса мониторинга сферы туризма

Для решения задачи мониторинга сферы туризма на заданной территории на основе открытых данных предлагается разработать сервис, обеспечивающий выполнение следующих функций:

1. Оперативное получение информации о социально-эколого-экономических показателях регионального туризма (объем показателей шире набора показателей официальной статистики и может собираться оперативно в режиме онлайн в отличие от возможностей сбора данных в результате административного управления):

- формирование туристского профиля территории (региона, района, местности):
  - виды туризма (рекреационный, событийный, экологический, этнографический, активный, деловой, сельский, детский, водный и круизный, социальный туризм, самодеятельный, лечебно-оздоровительный, экстремальный, паломнический, культурно-познавательный, гастрономический и др.),
  - «бренды» территории, достопримечательности (музеи, религиозные объекты, объекты природно-заповедного фонда, охотничье-рыболовные объекты, объекты сельского, промышленного, делового, военно-патриотического туризма, горнолыжные объекты),
  - площадки для наблюдения за природой,
  - туристические события,
  - туристические маршруты,
  - экологические туристские тропы,
  - экскурсии,

    - туристическо-информационные центры,
    - объекты размещения,
    - транспортные услуги, транспортная доступность,
    - объекты общественного питания,
    - климат,
- оперативный мониторинг (в том числе неорганизованного туризма)
    - видов туристических услуг (продуктов),
    - туристических потоков, их направления и плотность,
    - туристических маршрутов,
    - отзывов для выявления их тональности и проблем в сфере туризма (экологических, транспортных, качества оказываемых услуг, их безопасности),
    - рекреационной (антропогенной) нагрузки территории, определение экологического риска,
    - населения, занятого в туризме (количество рабочих мест, образование),
    - рейтинга региона (района, местности),
    - влияния экстремальных событий (пандемия, лесные пожары – наличие дыма, гарей).
    - мониторинг
    - мест неорганизованного отдыха на основе распознавания космических снимков (размещение в палатках),
    - неучтенных в официальном реестре туристических баз (мест размещения туристов) на основе распознавания космических снимков,
    - информационной популярности достопримечательностей,
- районирование (ранжирование) территорий:
    - специализация видов туризма,
    - плотность туристических потоков,
    - рекреационная нагрузка,

2. Поддержки принятия решений для малого и среднего предпринимательства:

- определение территории для бизнеса,
- выбор, обоснование и формирование туристических продуктов для бизнеса,

3. Поддержки принятия решений для клиентов туристических услуг – подбор туристических продуктов.

В настоящее время существуют различные способы реализации сервиса и организации к нему общего доступа, начиная от служб, предоставляющих различные услуги по аренде или размещению собственного оборудования в дата-центрах и заканчивая облачными платформами, предлагающими доступ к вычислительным ресурсам или сервисам по схеме «Infrastructure as a Service» (IaaS) или «Platform as a Service» (PaaS). При создании программной системы небольшой группой разработчиков (исследователей) наиболее предпочтительным является подход, основанный на использовании сервисов облачных платформ. Данный выбор обусловлен следующими факторами: отсутствие необходимости установки, настройки и сопровождения используемых при функционировании целевой программной системы средств и библиотек; отсутствие необходимости закупки оборудования и создания собственной инфраструктуры; снижение риска проекта, т.к. процесс исследования предполагает высокую вероятность существенных изменений требований и функциональности целевой программного системы; возможность автоматического масштабирования при увеличении нагрузки при использовании программной системы.

Наиболее популярными облачными платформами в настоящее время являются «Amazon Web Services» (https://aws.amazon.com), «Google Cloud Platform» (https://cloud.google.com), однако, с точки зрения удобства использования, в частности, поддержки русского языка, обеспечение требований ФЗ-152 РФ по защите персональных данных, наиболее целесообразным является использование облачных платформ, учитывающих специфику Российской Федерации, например, «VK Cloud» (https://mcs.mail.ru/cloud-platform) или «Яндекс.Облако» (https://cloud.yandex.ru). Последняя платформа наиболее предпочтительна, исходя из ее функциональности. В процессе разработки сервиса предлагается проблемно-ориентированные

функции оформить в виде независимых вычислительных блоков – элементов «бессерверной (serverless)» инфраструктуры, интеграция которых, вместе с решением стандартных задач по обеспечению хранения данных и отображения результатов их обработки, будет осуществлена с помощью существующих средств облачной платформы.

В настоящее время в рамках предложенной концепции сервиса реализуется функция «формирование туристского профиля территории» на примере Байкальской территории.

4. Источники информации

В настоящее время официальные статистические данные о состоянии региональной сферы туризма Байкальского региона не отражают полную информацию о средствах размещения, объектах питания, оказываемых услугах и их качестве и т.д. [7-10]. Один из главных недостатков этих данных – это отсутствие их точной территориальной привязки для обеспечения возможности оценки рекреационной нагрузки рассматриваемой территории.

В работе предлагается провести анализ, выявить информационную модель данных, осуществить их сбор, интеграцию и визуализацию по различным критериям следующих источников данных: сайт агентства по туризму Иркутской области, веб-ресурсы объектов размещения и общественного питания, отзывы групп социальных сетей Вконтакте и ОК.

5. Методы сбора и обработки данных

Для реализации функции сервиса формирования туристического профиля территории необходимо решить задачи сбора данных, их статистической обработки и визуализации.

5.1. Методы сбора данных

Получение данных веб-ресурсов в основном выполняется с использованием следующих методов: онлайн опросы, запросы к базам данных, API (Application Programming Interface) – это интерфейс обмена данными между приложениями, web-scraping (скрапинг). Для решения поставленных задач web-scraping является наиболее перспективным. Процесс web-scraping состоит из основных этапов [11, 12]:

1. Извлечение разметки страницы веб-источника, что осуществляется с использованием HTTP-запросов к ресурсу и сохранением полученных данных в переменную или файлы кода HTML веб-страницы.
2. Извлечение информации из структуры кода HTML осуществляется на основе поиска по заданному пути к элементам разметки кода: названию тегов, атрибутов и их значений.
3. Сохранение данных в структурируемом виде и дальнейшая их обработка.
4. Опционально: повтор действий.

В рамках текущей работы можно отметить особенности алгоритма реализации метода: его уникальность для каждого источника данных и необходимость с течением времени внесения изменений, т.к. разметка сайта может меняться его разработчиками.

5.2. Методы очистки и интеграции данных

Собранные данные из различных источников на основе разрабатываемых онтологий подвергаются обработке, в том числе опционально геокодингу (если неизвестны координаты объекта) и преобразованиям с целью получения списка уникальных объектов (Рисунок 1). Ручная проверка набора данных необходима, поскольку в разных источниках встречаются записи об одном и том же объекте с ошибками в написании значений рассматриваемых свойств. По идентифицированным значениям находится максимальная дистанция между координатами записей, которую определяем как пороговое значение. Далее осуществляется поиск ближайшей

записи по среднему значению нормированных показателей дистанций: расстояние Левенштейна используется при поиске названий объекта, населенного пункта, адреса и дистанция между координатами по формуле Хаверсина.

После идентификации объектов выполняются расчеты средних значений стоимости, пользовательских оценок (рейтинги), суммирование количества отзывов, определение категорий услуг и их популярности.

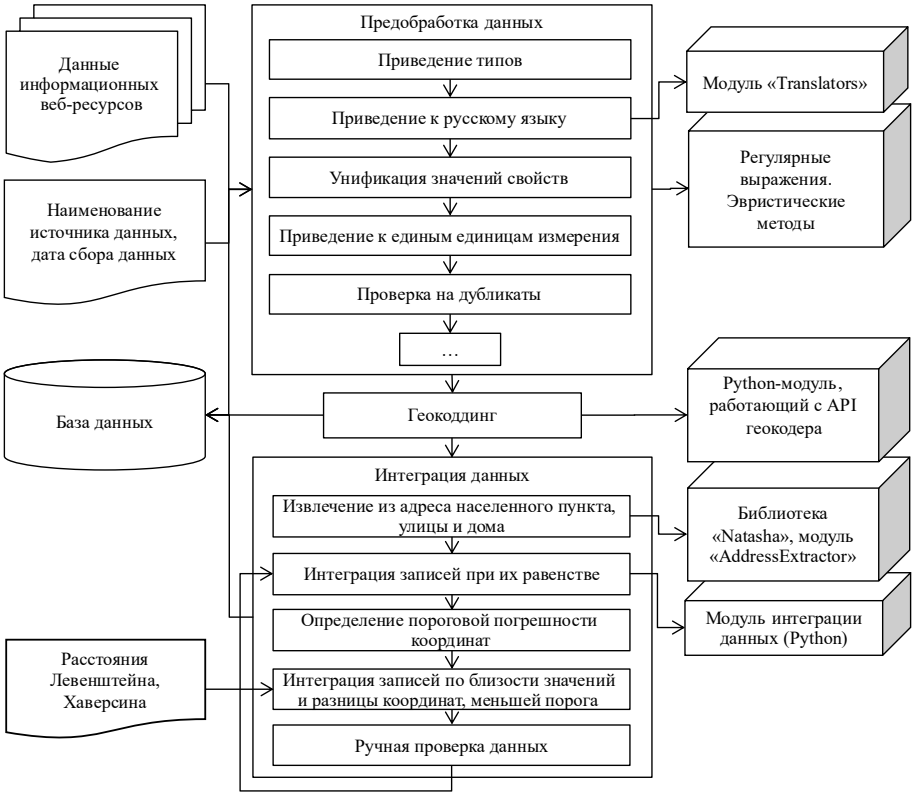


Рисунок 1: Этапы алгоритма интеграции данных из нескольких источников

5.3. Методы анализа текстов (отзывов)

Полученные данные в виде коротких текстов (отзывов) из социальных сетей подвергаются следующим этапам обработки (данный раздел подготовлен с участием Ивана Поддубного, email: poddubnyyiv@yandex.ru, ORCID: 0000-0002-2431-8692):

1. Подготовка данных к анализу включает исключение иностранных текстов или их перевод.
2. Токенизация текста – разбиение текста на отдельные слова, исключая все остальные элементы (знаки пунктуации, смайлы и другие символы).
3. Выделение именных сущностей (Named Entity Recognition, NER), в контексте решаемой задачи определение локаций популярных мест отдыха, с помощью библиотеки Stanza [13].
4. Лемматизация текста – приведение слов в начальную форму или стемминг – выделение основы слова.
5. Удаление стоп-слов (высокочастотные слова, которые не несут существенной информации).
6. Формирование словаря решаемой задачи, что позволяет провести семантический анализ текста для определения: сущностей, действий, описаний с привязкой к локациям – база знаний о флоре и фауне, объектах инфраструктуры и других характеристиках территории.

7. Сентимент-анализ текста – определение тональности на основе классификации характеристик на классы «негатив» и «позитив», что позволит дополнительно выделить проблемы и предпочтения локаций (территории).

Каждый этап анализа текста (токенизация, лемматизация, синтаксический анализ, выделение именованных сущностей) осуществляется с помощью нейронных сетей.

5.4. Методы визуализации данных

Основной задачей визуализации является обеспечение поддержки пользователя в процессе восприятия, понимания и осмысления информации и формирования новых знаний, а также обеспечение минимизации усилий по выполнению когнитивных задач в сравнении с текстовым представлением данных. Для визуализации и обработки данных все большую актуальность приобретают технологии облачных вычислений – BI-системы/платформы (Business Intelligent System), которые обеспечивают: возможность анализировать и отображать различные данные и информацию, связанные с рассматриваемой задачей. Среди существующих платформ выделим Tableau, Power BI, Yandex DataLens и др. Данные платформы обладают следующим основным функционалом: визуализация данных, автоматический поиск зависимостей и скрытых взаимосвязей, создание интерактивных отчетов, глубокая аналитика, интеграция с другими инструментами (Excel, Google Sheets), машинное обучение.

Данные системы позволяют эффективно визуализировать данные с помощью информационных панелей̆ (так называемых дашбордов, с англ. dashboard — «приборная панель»), состоящих из нескольких графиков и диаграмм с возможностью интерактивного взаимодействия с пользователем.

В случае, когда сервис создается на многофункциональной платформе, такой как Яндекс.Облако, то целесообразно использовать средство визуализации данной платформы — Yandex DataLens.

6. Реализация сервиса мониторинга

Предлагаемую технологию сбора, обработки и представления социально-эколого-экономических данных об оперативном состоянии региональной сферы туризма планируется реализовать с помощью средств облачной платформы «Яндекс.Облако» на основе подхода микросервисной архитектуры. В качестве средства реализации вычислительных блоков предлагается использовать облачные функции («Cloud Function»), для организации каналов обмена данными – очереди сообщений («Message Queue»), для общей интеграции сервис – «API Gateway», для хранения данных – PostgreSQL [14].

Текущее вид архитектуры программной системы, реализующей предлагаемую технологию отображен на Рисунке 2. Процесс сбора информации организован на основе двух очередей сообщений, где первая очередь, содержит информацию об источниках, которые требуется обработать в виде «одно сообщение — один источник», например запрос к API социальной сети «ВКонтакте» и его параметры (Рисунок 2, Задания на извлечение информации), а вторая содержит результаты обработки, которые требуется записать в базу данных PostgreSQL в виде «сообщение — одна строка целевой таблицы» (Рисунок 2, Очередь ввода данных в БД). Для обработки сообщений в очереди используется встроенный механизм триггеров, который обеспечивает передачу сообщений в облачную функцию-обработчик в виде массива с фиксированной максимальной длиной. Для извлечения информации из внешних источников используются следующие функции: get_vk() и get_ok(), чтобы сформировать данные постов сообществ и комментариев к ним из социальных сетей. Для записи данных в базу данных используется вариант компонента data-controller из платформы автоматизации создания интеллектуальных систем как сервиса (AI PaaS) [15], использующий снимок состояния структуры базы данных для минимизации требуемых прав доступа и повышения быстродействия. В штатном режиме управление процессом добавления сообщений в очередь «Задания на извлечение информации» реализовано в рамках спецификации API Gateway, однако для нужд тестирования и отладки у разработчиков существует возможность доступа к очередям

в режиме «только запись» (выделено пунктиром на Рисуноке 2) на основе встроенного HTTPS интерфейса очереди.

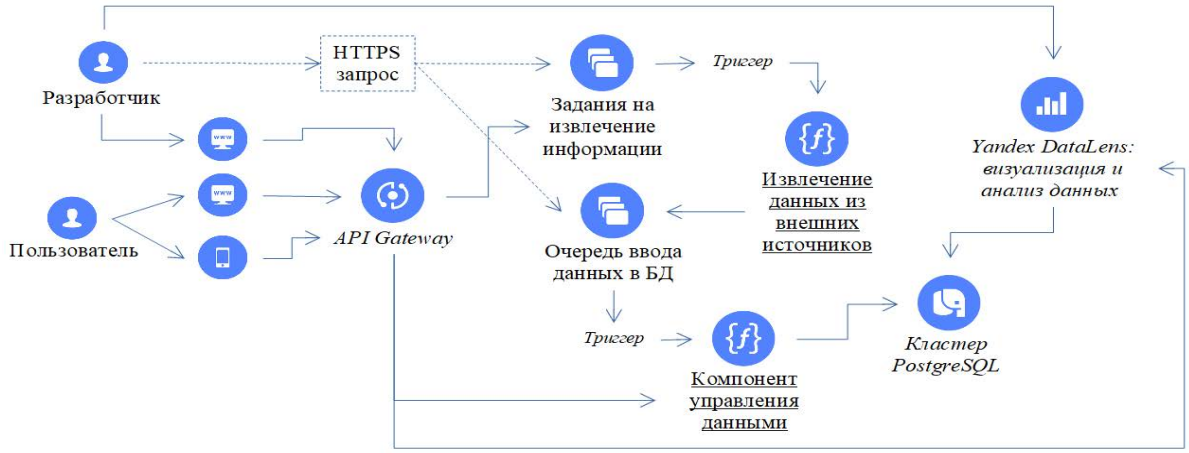


Рисунок 2: Архитектура программной системы

Интеграция отдельных блоков в программную систему выполняется на основе сервиса API Gateway путем создания спецификации в формате OpenAPI 3.0, которая позволяет в декларативном представить функции целевой программной системы, в виде набора URL адресов и указать каким образом будет осуществляться обработка поступающих на эти адреса запросов.

7. Результаты сбора и анализа данных

В настоящее время собрана и обработана информация о следующих туристических объектах: коллективных средствах размещения – 1163 (685) (количество объектов, о которых собрана информация, в скобках – количество идентифицированных объектов, включая геолокацию), услугах – 61 (16), объектах общественного питания – 809, достопримечательностях – 395 (352), туристических маршрутах и экскурсиях – 94.

Реализованы информационные панели (дашборды):

- о коллективных средствах размещения и их услугах с возможностью фильтрации по рейтингу (от «плохо» до «великолепно»), населенным пунктам/районам и категории услуг: отображение количества средств размещений и средней стоимости по населенным пунктам/районам в виде комбинированного графика, коллективных средств размещения в виде карты пропорциональных объектов по номерному фонду, количества средств размещений по категории/подкатегории услуг в виде столбчатой диаграммы и тепловой карты,
- об объектах общественного питания с возможностью фильтрации по районам и типу кухни: отображение карты пропорциональных объектов общественного питания по среднему чеку и типу кухни, агрегированных количественных показателей по населенным пунктам/районам для описания территорий по заданной мере (дорогие-дешевые, по количеству объектов общественного питания) в виде комбинированного графика, тепловой карты по количеству объектов общественного питания в разрезе районов и типа кухни.
- о достопримечательностях и популярных местах отдыха: отображение на карте и обеспечение возможности получения описательной информации о достопримечательностях, плотности точек достопримечательностей, популярных мест отдыха,
- о туристических маршрутах и экскурсиях с возможностью фильтрации по названию маршрута: отображение на карте населенных пунктов, где организованы туристические маршруты или экскурсии, и схем маршрутов и экскурсий.

8. Заключение

В статье выполнена постановка задачи разработки информационной технологии мониторинга сферы территориального туризма, описана концепция сервиса, обоснованы методы анализа и визуализации данных, сформирована онтологическая модель, определено содержание информационных панелей (дашбордов), реализован прототип сервиса, представлены результаты сбора, обработки и представления полученных в результате исследования данных. В настоящее время в рамках предложенной концепции сервиса реализован прототип функции «формирование туристского профиля территории» на примере Байкальской территории. В дальнейшем планируется развить данный сервис и наполнить его данными согласно предложенной концепции.

Работа выполняется при поддержке проекта Российского научного фонда №23-28-00844 «Мониторинг сферы регионального туризма на основе анализа данных из открытых источников».

[1] F. Lobao, M. Aparicio, M. Neto, SMART TOURISM -CITY TOURISM RADAR: A Tourism Monitoring Tool at the City of Lisbon, in 19.the Conference of the Portuguese Association of Information Systems, CAPSI ’2019, Lisbon, Portugal, 2019, pp. 1-21.
[2] H. Li, M. Hu, and G. Li, Forecasting tourism demand with multisource big data, Annals of Tourism Research 83 (2020) 102912. doi: 10.1016/j.annals.2020.102912.
[3] F. Soualah-Alila, M. Coustaty, C. Faucher, R. Wannous, TourinFlux Project: Contribution of Semantic Web Technologies for Tourism Data Management in: 6th edition of the multidisciplinary AsTRES conference, University of Western Brittany, 2016. doi: 10.1145/2810133.2810137.
[4] Иркутская область. URL: https://ru.wikipedia.org/wiki/Иркутская_область.
[5] Байкал. URL: https://ru.wikipedia.org/wiki/Байкал.
[6] Иркутская область заняла 15 место в Национальном туристическом рейтинге. URL: http://www.irk.ru/news/20220113/rating/.
[7] Д.А. Котельников, Формирование системы показателей для ведения мониторинга устойчивого развития туристских территорий, Сборник статей по материалам всероссийского научно-исследовательского конкурса научных инноваций: перспективы развития науки в современном мире, Уфа, 2020, с. 41-50.
[8] Ю.А. Лебедева, Организация мониторинга качества туристских услуг на муниципальном уровне, Среда, Чебоксары, 2020.
[9] Ю.Д. Шмидт, Н.В. Рубцова, Формирование системы мониторинга эффективности функционирования сферы туристско-рекреационных услуг региона // Материалы V Всероссийской научно-практической конференции «Интеллектуальный и ресурсный потенциалы регионов: активизация и повышение эффективности использования» Под науч. ред. А.П. Суходолова, Н.Н.Даниленко, О.Н.Баевой. 2019. С. 520-524.
[10] Туризм. Федеральная служба государственной статистики. URL: https://rosstat.gov.ru/statistics/turizm.
[11] Р. Митчелл, Современный скрапинг веб-сайтов с помощью Python, 2 издание, Питер, СПб, 2021.
[12] А.А. Москаленко, О.Р. Лапонина, В.А. Сухомлин, Система управления доступом к ресурсам веб приложений на основе анализа поведения пользователя, International Journal of Open Information Technologies 8 (9) (2020) 30-35.
[13] P. Qi, Y. Zhang, Y.J. Zhang, C.D. Bolton, Manning Stanza: A Python Natural Language Processing Toolit for Many Human Languages. URL: https://arxiv.org/pdf/2003.07082.pdf.
[14] Yandex Cloud. Документация. URL: https://cloud.yandex.ru/docs.
[15] A.I. Pavlov, A.B. Stolbov A.B., A.A. Lempert, Towards extensibility features of knowledge-based systems development platform, in: I.V. Bychkov, Dimitar Karastoyanov (Eds.), 4th Scientific-Practical Workshop Information Technologies: Algorithms, Models, Systems, volume 2984 of CEUR Workshop Proceedings, Irkutsk, 2021, pp. 87-94.

Algorithms for logic circuits synthesis in the functional-flow programming language

Darya Romanova [1,2]

[1] Siberian Federal University, Svobodny pr.79, 660041, Krasnoyarsk, Russia
[2] Krasnoyarsk State Agrarian University Mira pr. 90, 660049, Krasnoyarsk, Russia

Abstract

Ensuring the portability of designed solutions is one of the priority tasks of digital integrated circuits' development. The paper proposes a method for designing logic circuits using an architecture-independent functional parallel programming language. A subset of this language is presented for designing logic circuits of varying complexity. On the basis of this subset, algorithms for the synthesis of logic circuits, limited and unlimited in size, have been developed.

Keywords

Parallel programming, functional-flow language, logic circuit, HDL, algorithm

1. Introduction

Recently, there has been a trend towards the complication of digital integrated circuits (IC). Most of the methods for developing IC are tightly tied to a specific implementation platform, so even when making minor changes to the program, it is often necessary to completely rewrite the code. Therefore, the task of ensuring the portability of designed solutions is the highest priority in the digital circuits' development, including combinational circuits, which are the basis for the construction of integrated circuits. Portability allows the developer to optout of being tied to the target implementation platform that provides more efficient solutions for key product specifications such as speed, chip area, etc.

To design digital ICs, it is more efficient to use functional languages that have a more powerful abstraction mechanism compared to imperative languages [1]. The solutions can be found in such languages as Wired [2], Haskell (library Lava)[3], Chisel [4].

One of the approaches that ensure the portability of the designed solutions is the use of the functional-flow parallel model (FFPM) and the corresponding programming languages [5]. Therefore, if we present combinational logic design algorithms using FFPM, then the most technically effective solutions can be achieved.

2. Functional-flow model and language

The developed method for designing logical circuits is based on FFPM. The model is based on the concept of infinite resources and supports implicit control of data readiness, which makes it possible to ensure its resource independence of the designed solution.

The FPPP model is given by the triple:

$$M = (G, P, S_0),$$

where G is a dataflow graph that defines the structure of the program, P is a set of rules that determine the dynamics of the model, S_0 is the initial labeling of the graph. A dataflow graph (DG) is a collection of a set of vertices responsible for program-forming operators and a set of arcs that define links between them. The example of DG is showed in figure 1.

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: daryaooo@mail.ru
ORCID: 0000-0002-9020-4802

ICCS-DE 2023 Workshop Proceedings

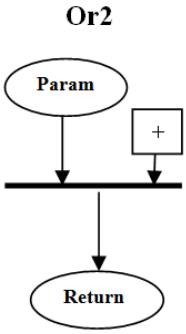


Figure 1: Dataflow graph of logic element "Or"

Functional-flow parallel programming (FFPP) language supports FFPM and parallelism at the operation level, which is important for independent, end-to-end design of digital circuits [6]. There are no cycles in the language, which avoids conflicts when using the same data. Instead of cycles recursion is used.

3. A subset of the FFPP language for describing logical circuits

To develop logic circuits in the FFPP language, a corresponding subset was identified in accordance with the specifics of designing combinational logic on an FPGA (field-programmable gate array) and the implementation features of the hardware and software on a chip [7].

In the logic design subset, there are only two data types: "int" and "bool". To develop the logic circuit programs, the logical data type "bool" is used, in which the argument can take the values true or false. To carry out auxiliary actions that are not included in the logical circuits, the integer type "int" is used, while a 32-bit representation is used. The subset also imposes restrictions on the use of program-forming statements, showed in the table 1.

Table 1

Program-forming operators using in the subset

Operator	Description
+	Logical addition (Or)
-	Logical inversion (Inv)
*	Logical multiplication (And)
|	list length operator using only with "bool" type
<, >, <=, =>,=, !=	comparison operators
(),[]	listing operators using only with "bool" and "int" types

Some operators of the FFPP language are absent in this subset, since they are not used to develop combinational circuits. This design method consists in the development of logic circuits in the FFPP language and their optimization (architecture-independent level) with subsequent translation into hardware description languages (architecture-dependent level). The logical circuits developed in the FFPP language are represented in the form of DG with the help of an appropriate compiler, and then transferred to specific solutions in the Verilog/VHDL languages. At the same time, the semantics of the original program and the logic of its functioning are preserved.

4. Algorithm for designing logic circuits in FFPP language

Features of the FFPP language make it possible to develop logic circuits of varying complexity, limited and unlimited in size. When designing digital circuits at the logical level, two types of circuits are created: combinational, which are circuits without memory, and automata with memory, the

output values of which depend not only on the input data, but also on the internal state of the system [10]. In the case of developing limited combinational circuits, initially in the FFPP language, the circuits' design algorithm is shown in Figure 2.

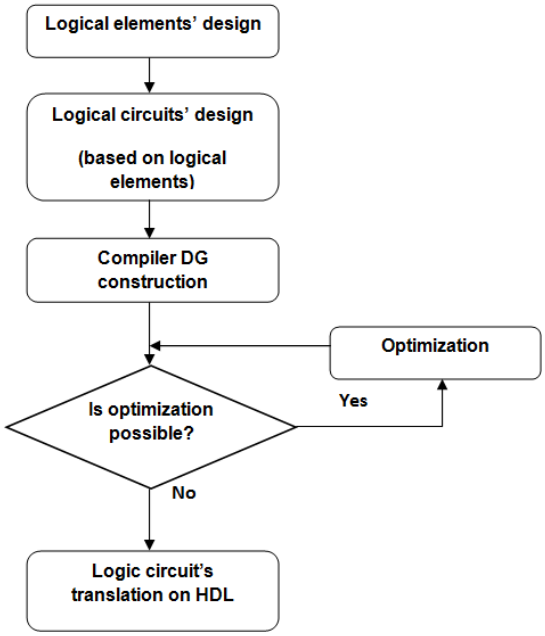


Figure 2: Algorithm for designing combinational circuits in the FFPP language with translation into a hardware description language

Initially, functions are created that model the logical elements AND, OR, NOT, AND-NOT, OR-NOT, XOR. Further, on the basis of these logical elements, combinational circuits of large sizes are designed. In this case, the description of the same circuit can be done in several ways.

Algorithm for designing logic circuits that are unlimited in size is based on use of recursion. In this case, recursion is used to set multidimensional repeating fragments [8]. The initially constructed recursive program involves working with any number of input data. And accordingly, a specific linear solution for the FPGA is formed at the output.

The algorithm for large combinational circuits in this case consists of several steps:

1. Development of a program for a logic circuit using recursion;
2. To obtain a specific solution for a circuit of a certain size in the developed program, when it is launched, the parameter N is introduced, which specifies a certain number of inputs and the amount of elements necessary to solve a specific problem on the FPGA, and the circuit is transformed into a specific circuit solution.

For example, using the MUX_N program, you can build four-channel, eight-channel multiplexers, just by setting the input parameter N equal to 4 and 8 (figure 3).

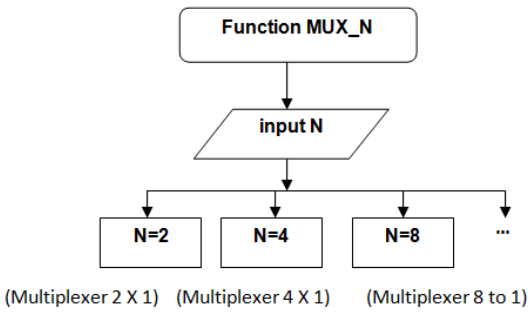


Figure 3: Recursive "unrolling" scheme for multiplexer

In the case of automaton circuits, such an implementation is impossible, since in this case their implementation is formed using cycles.

5. Conclusion

The paper singles out a subset of the FFPP model and language for logic circuits' design. In the FFPP language, functions were created that simulate the main logic elements and a set of combinational circuits of various sizes, as well as general algorithms for the synthesis of logic circuits were developed. For instrumental support of the developed method of designing logical circuits, a translator is being developed that allows transferring the functions of logical circuits created in the FFPP language to hardware description languages.

6. References

[1] J. Dongarra, G. Bosilca, A. Bouteiller, A. Danalis, M. Faverge, T. Herault. PaRSEC: A programming paradigm exploiting heterogeneity for enhancing scalability. IEEE Computing in Science and Engineering, 6(15): 36–45, 2013.

[2] E. Axelsson, K. Ciaessen, M. Sheeran. "Wired: wire-aware circuit design." Proceedings of the Conference on Correct Hardware Design and Verification Methods (CHARME '05), vol. 3725, 2005, pp. 5—19.

[3] A. Gill, T. Bull, G. Kimmell. et al. "Introducing Kansas Lava." Implementation and Application of Functional Languages, ser. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Sep. 2009, pp. 18—35.

[4] J. Bachrach at al. Chisel: Constructing hardware in a Scala embedded language, DAC Design Automation Conference (2012), 1212-1221.

[5] A.I. Legalov A functional language for creating architecture-independent parallel programs, Vychislitel'nye tehnologii, 2005, vol. 10, no. 1, pp. 71-89 (in Russian)

[6] D. S. Romanova, O.V. Nepomnyashchiy, I.N. Ryzhenko, A.I. Legalov, N.Y. Sirotinina Parallelism reduction method in the high-level VLSI synthesis implementation, Trudy ISP RAN/Proc. ISP RAS, vol. 34, issue 1, 2022. pp. 59-72.DOI: 10.15514/ISPRAS–2022–34(1)–5

[7] O.V. Nepomnyashchiy, I.N. Ryzhenko, A.I. Legalov The method of architecture-independent high-level synthesis of VLSI, Izvestiya SfedU, Technical science, 2018, no.8 (202), pp. 38–47 (in Russian)

[8] D. S. Romanova, A.I. Legalov Design of logic circuits based on the functional-flow programming model, Mathematical methods in technologies and engineering, 2022, no.10, pp. 147-150.

О проектировании потока работ в платформе создания систем, основанных на знаниях

Александр Б. Столбов [1], Александр И. Павлов [1] and Анна А. Лемперт [1]

[1]*Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (ИДСТУ СО РАН), ул. Лермонтова 134, Иркутск, Россия*

Аннотация

В работе проведён анализ возможности использования подхода потока работ в рамках разрабатываемой платформы создания систем, основанных на знаниях. К настоящему времени платформа уже содержит компоненты для обеспечения хранения данных, проектирования концептуальных моделей, создания баз знаний продукционного типа и ряд других специализированных компонентов. На основе обзора существующих языков описания и систем управления потоками работ выявлены ключевые особенности их проектирования, реализации и применения. Обсуждаются вопросы, связанные с разработкой собственного компонента платформы, использующего модели и методы потока работ; указаны функциональные требования к такому компоненту; рассмотрены особенности архитектуры и реализации прототипа компонента.

Ключевые слова

Системы, основанные на знаниях, поток работ, программная инженерия

1. Введение

В настоящее время системы, основанные на знаниях (СОЗ), являются важным элементом современных информационных технологий. Как правило, СОЗ используются при решении слабо формализованных задач. Например, в предисловии [1] отмечается, что применение СОЗ особенно актуально в ситуациях, когда объем доступной информации чрезмерно велик для непосредственной обработки лицами, принимающим решения, и в то же время существует потребность в точных, обоснованных и объяснимых решениях. С точки зрения программной инженерии разработка прикладной СОЗ может быть выполнена на языке программирования общего назначения или с помощью специализированных инструментов (Drools, G2 и т.д.). В статье рассматриваются вопросы, связанные с созданием одного из таких инструментов – платформы разработки СОЗ [2].

Платформа проектируется и реализуется как современный веб-ориентированный программный комплекс для программистов, инженеров по знаниям и экспертов-предметников и обеспечивает их необходимыми инструментами в процессе создания прикладных СОЗ. Одно из ключевых требований к платформе – возможность реконфигурации ее архитектуры без значительных усилий по переписыванию исходного кода прикладной СОЗ. В связи с этим платформа разрабатывается на основе компонентного подхода с открытой архитектурой, где для решения проблемно-ориентированной задачи формируется один или несколько программных компонентов. На текущий момент разработаны и описаны следующие основные компоненты платформы: компонент управления данными [2], компонент представления и редактирования данных [3], компонент редактирования модели предметной области [4], компонент рассуждения на основе правил [5], компонент проектирования агентных имитационных моделей [6], а также

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: stolboff@icc.ru (A. 1); asd@icc.ru (A. 2); lempert@icc.ru (A. 3)
ORCID: 0000−0001−6513−7030 (A. 1); 0000−0002−7753−7514 (A. 2); 0000-0001-9562-7903 (A. 3)

ICCS-DE 2023 Workshop Proceedings

создан прототип потока работ [7]. В данной работе обсуждаются вопросы, связанные с дальнейшим развитием одного из ключевых элементов платформы – компонента потока работ (КПР).

2. О походе к разработке на основе потока работ

Поток работ можно рассматривать как способ организации «работ» (модулей, задач, действий), связанных с некоторой бизнес-логикой. Эту организацию можно представлять с различных точек зрения, слоёв или перспектив (термин из [8]):

- Слой управления оперирует конструкциями для обеспечения выполнения параллельного и последовательного выполнения задач, в том числе их разделения и объединения; итерации и выбора по ним; синхронизации.
- В слое данных рассматриваются вопросы, связанные с передачей различных типов переменных при обмене информацией межу задачами, а также с пред и пост условиями переходов между задачами.
- В слое ресурсов внимание уделяется типам и свойствам ресурсов (программное обеспечение, оборудование, персонал), их доступности, качеству, производительности.
- В слое реализации обрабатывается информация о том, как спецификация потока работ привязана к реальным действиям и исполнителям с применением программного и пользовательского интерфейсов или, например, описания должностных инструкций.

В зависимости от уровней целей и выразительности используемых языков описания модель спецификация потока работ может содержать от одного до четырех слоёв. К настоящему времени в области исследований, посвященных формализму потоку работ, накоплен большой теоретический базис: предложены разнообразные модели и методы проектирования. Например, серию публикаций под общим названием «Workflow Patterns Initiative» можно рассматривать как подробный справочник, содержащий классифицированные по разным критериям методы проектирования потока работ [9].

Для моделирования потока работ было предложено несколько нотаций и языков, обеспечивающих описание под разные цели. Например, был предложен ряд интуитивно понятных и выразительных нотаций для графического задания моделей в виде диаграмм (BPMN, EPC, UML AD), получивших широкое распространение. Тем не менее на данный момент они не могут в полной мере быть использованными для автоматической трансляции модели в исполняемый код. Другое семейство языков (BPML, BPEL, WS-CDL, XLANG), наоборот, предназначено для решения задачи автоматического развёртывания модели потока работ в некоторой исполнимой среде. Их недостаток заключается в сложности восприятия проектировщиком, особенно, не обладающего компетенциями в области программирования или администрирования вычислительных систем. В настоящее время с целью устранения указанных недостатков ведутся разработки в области создания новых языков (BPDM, XPDL, YAWL) и модификации существующих.

Для теоретического анализа потока работ традиционно используют различные типы сетей Петри (обычная, раскрашенная, предикатная и др.), где слой управления рассматривается в терминах позиций, переходов и дуг. Например, при разработке типовых моделей потока работ (Standard Workflow Models) [10] на основе сетей Петри было проведено сравнение различных методов проектирования.

Для выполнения потока работ разрабатываются специализированные программные системы, историю появления которых можно рассматривать как результат поэтапной эволюции архитектуры программ. Изначально приложения обеспечивали свою функциональность (хранение данных, пользовательский интерфейс, бизнес-логика и т.д.) полностью самостоятельно. В дальнейшем в период с 1970-х по 1990-е годы произошло постепенное выделение в отдельные блоки систем управления базами данных, пользовательского интерфейса, а также элементов для обработки потока работ. Тем не менее, в случае потока работ существовали определённые проблемы, сдерживающие процесс их появления и обособления. Например, долгое время использование формализма потока работ не рассматривалось как действительно новая функциональность. Кроме того, отсутствие полноценных возможностей

интеграции сторонних средств в ранних продуктах отпугивал многих потенциальных пользователей. Важным стимулирующим фактором появления и быстрого распространения нового поколения систем обработки потока работ было появление веб- и облачных средств, а также последующий быстрый рост приложений с сервис-ориентированной архитектурой.

Традиционно, основной областью применения потока работ были приложения, связанные с экономической деятельностью. Поэтому достаточно быстро выделился специальный подраздел – системы управления бизнес-процессами (business process management system, BPMS), что в последствии привело к путанице в терминологии. В ходе анализа соответствующих публикаций [11-13] можно сделать вывод, что, в настоящее время следует разделять эти технологии: 1) BPM – это процессно-ориентированная технология управления, с применением возможностей ИТ; 2) системы управления на основе потока работ – это технология, которую можно внедрять в различные проблемные и предметные приложения. Тем не менее с точки зрения программной архитектуры в соответствии с этапами жизненного цикла продукта системы управления на основе потока работ являются подмножеством BPM: проектирование процесса, конфигурация системы, внедрение процесса, диагностика (только для BPM). В настоящее время разрабатывается и активно используется множество систем обоих типов. При этом можно констатировать, что в современных условиях компонент управления потоком работ является таким же важным и востребованным элементом сложных программных комплексов как подсистемы баз данных или пользовательского интерфейса.

Типовая архитектура компонент управления потоком работ в соответствии с эталонной моделью Workflow Management Coalition включает в себя:

1. подсистему внедрения потока работ, содержащую в том числе движок исполнения потока работ;
2. подсистему проектирования спецификации потока работ;
3. клиентские приложения, т.е. точку доступа пользователей-исполнителей;
4. множество внешних сервисов, которые вызываются компонентом в процессе исполнения потока работ;
5. инструменты администрирования и мониторинга;
6. набор соответствующих программных интерфейсы между элементами 1-5.

Исполняемая в некоторой системе модель потока работ называется экземпляром процесса. При этом конкретная модель может выполняться с помощью нескольких экземпляров, запускаемых одновременно: как правило, эти экземпляры независимы и не имеют ссылок друг на друга. В настоящее время существует два основных подхода к реализации движка исполнения потока работ: основанный на токенах и основанный на экземплярах. Первый берет своё начало от сетей Петри и активно использовался в п оследние десятилетия. Второй подход получил популярностью позже и основан на объектно-ориентированной парадигме программирования, что позволяет повысить эффективность интеграции потока работ как дополнительного компонента в существующие сложные программные комплексы.

3. Особенности применения потока работ в платформе

В предыдущей работе [7] представлены модель спецификации потока работ на основе базовых и композитных операций, где базовая операция выполняется за счёт функциональности платформы или за счёт вызова внешних сервисов, а композитная – представляет собой результат комбинации других операций с использованием последовательных, разветвляющихся и циклических операторов. При этом главным требованием к компоненту потока работ является возможность явно обрабатывать элементы концептуальной модели, т.е. в дополнение к стандартному множество допустимых типов данных (строка, число, логический) предлагается рассматривать элементы концептуальной модели. Поэтому в работе [7] появилась возможность учитывать в процессе проектирования тип элемента концептуальной модели: понятие, свойство, отношение, экземпляр.

В развитие этой идеи в новой версии компонента необходимо также учитывать и значение элемента, например, «обыкновенное дифференциальное уравнение», «уравнение водного баланса» и т.п. В дополнение к этому требованию желательна также поддержка других

формализмов, используемых для описания сущностей в платформе: элементов баз знаний (например, условий правил или самих правил), элементов агентной имитационной модели (агент, объект, роль).

Для организации процесса проектирования и выполнения потока работ в новой версии компонента необходимо разработать следующие модули:

- парсер спецификаций веб-сервисов (формат OpenAPI) в базу описания базовых операций,
- парсер спецификаций онтологий (формат OWL) в базу концептуальных моделей
- редактор семантического аннотирования, где элементы потока работ (т.е. базовые и композитные операции, а также их входные и выходные параметры) связываются со значениями элементов концептуальной модели и, возможно, с другими формализмами;
- визуальный редактор спецификации, поддерживающие среди прочего возможность настраивать способы отображения и обработки узлов, соединений и портов в зависимости от связанных с ними значений элементов концептуальной модели;
- интерпретатор потока работ с возможностью вызова функций внешних систем (на примере работы с REST API веб-сервисами.

К настоящему времени разработан прототип новой версии компонента, взаимодействие модулей которого с пользователем представлено на рисунке 1.

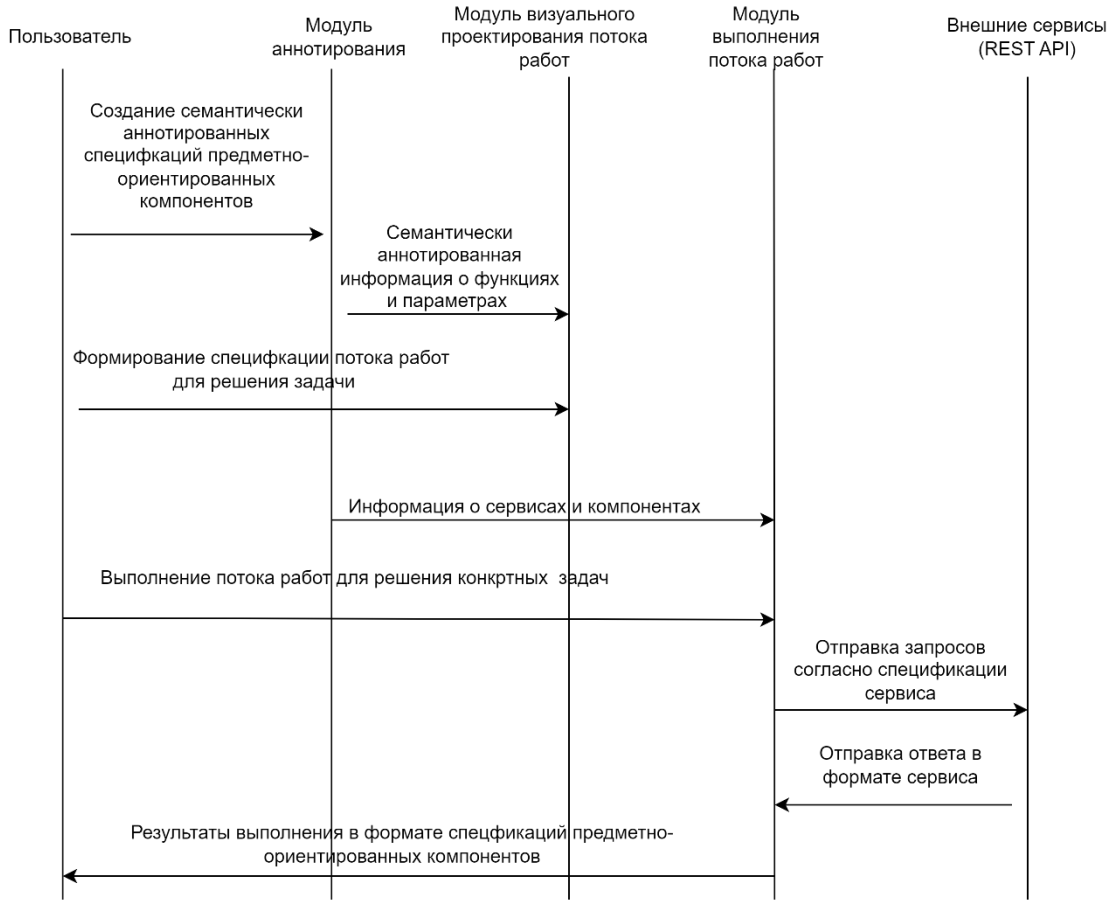


Рисунок 1: Процесс работы пользователя с компонентом

4. Заключение

Проведён краткий анализ современного состояния в области подходов к разработке систем, использующих формализм потока работ, рассмотрена типовая архитектура таких систем, сделаны выводы о характеристиках существующих языков описания потока работ. На основе

выводов по анализу предложена архитектура новой версии компонента поток работ для платформы создания систем, основанных на знаниях. Актуальность разработки вызвана необходимостью явно использовать формализованные знания об элементах потока работ в процессе проектирования. В новой версии появится возможность импорта информации из языка спецификаций и онтологий, редактор семантического аннотирования элементов проблемно-ориентированных компонентов платформы, а также возможность вызывать внешние веб-сервисы с использованием интерфейса REST API.

5. Благодарности

Результаты получены в рамках госзадания Минобрнауки России по проекту «Теоретические основы, методы и высокопроизводительные алгоритмы непрерывной и дискретной оптимизации для поддержки междисциплинарных научных исследований» (№ гос регистрации: 121041300065-9), «Методы и технологии облачной сервис-ориентированной цифровой платформы сбора, хранения и обработки больших объёмов разноформатных междисциплинарных данных и знаний, основанные на применении искусственного интеллекта, модельно-управляемого подхода и машинного обучения» (№ гос регистрации: 121030500071-2).

6. Литература

[1] Alor-Hernndez, G., Valencia-Garca, R. (eds.), Current Trends on Knowledge-Based Systems, volume 39 of Intelligent Systems Reference Library series, Springer, Cham Switzerland, 2017

[2] Nikolaychuk O.A., Pavlov A.I., Stolbov A.B.: The software platform architecture for the component-oriented development of knowledge-based systems. In: Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1234–1239. IEEE (2018). https://doi.org/10.23919/MIPRO.2018.8400194

[3] *Pavlov A.I., Stolbov A.B.* Domain-oriented specialization tools for knowledge-based systems development platform // CEUR Workshop Proceedings: 3rd Scientific-Practical Workshop Information Technologies: Algorithms, Models, Systems (ITAMS 2020; Irkutsk, 3 September 2020). 2020. Vol. 2677

[4] *Павлов А.И., Николайчук О.А., Столбов А.Б.* Редактор концептуальных моделей: Свидетельство о государственной регистрации программ для ЭВМ №2016617626 от 11.07.2016 М.: Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2016.

[5] *Павлов А.И., Николайчук О.А., Столбов А.Б.* Web-ориентированный редактор продукционных правил: Свидетельство о государственной регистрации программ для ЭВМ №2016663618 от 13.12.2016 М.: Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2016.

[6] Pavlov A.I., Stolbov A.B., Dorofeev A.S. The architecture of the software tool for the agent-based simulation models specification development // CEUR Workshop Proceedings: Proc. of 2nd Scientific-Practical Workshop Information Technologies: Algorithms, Models, Systems (ITAMS'2019). 2019. Vol. 2463. pp. 112-121.

[7] Pavlov A.I., Stolbov A.B., Dorofeev A.S. The workflow component of the knowledge-based systems development platform // Proc. of 2nd Scientific-Practical Workshop Information Technologies: Algorithms, Models, Systems (ITAMS'2019). 2019. Vol. 2463. pp. 47-58.

[8] Russell N., ter Hofstede A.H.M., Edmond D., van der Aalst, W.M.P.: Workflow Data Patterns: Identification, Representation and Tool Support. In: L. Delcambre et al., eds Proceedings of the 24th International Conference on Conceptual Modeling (ER 2005), LNCS, vol. 3716, pp. 353–368. Springer, Verlag Berlin.

[9] Workflow Patterns Initiative Homepage, http://www.workflowpatterns.com.

[10] Kiepuszewski, B., ter Hofstede, A.H.M., van der Aalst, W.M.P.: Fundamentals of Control Flow in Workflows. Acta Informatica 39(3), 143–209 (2003)

[11] van der Aalst .M.P., van Hee, K.: Workflow Management – Models, Methods and Systems. MIT Press, Cambridge MA London England (2002). https://doi.org/10.1016/S0933-3657(03)00011-3
[12] Hill, J.B., Pezzini, M. and Natis, Y.V.: Findings: confusion remains regarding BPM terminologies. Gartner Research, Stamford CT (2008)
[13] Georgakopoulos, D., Hornick, M. and Sheth, A.: An overview of workow management: from process modeling to workow automation infrastructure. Distributed and Parallel Databases 3(2), 119–153 (1995)

Биологически вдохновленная многоцелевая стратегия управления мобильными агентами при решении задачи обследования поля концентрации

Антон Толстихин[1]

[1]*Институт динамики систем и теории управления имени Матросова СО РАН, Лермонтова 134, Иркутск, 664033, Российская федерация*

Аннотация

В настоящее время задача обследования полей концентрации получила широкую популярность в сфере управления автономными мобильными агентами за счет широкого спектра практических и фундаментальных задач, описываемых с ее помощью. В данной статье рассматривается комбинированная децентрализованная стратегия управления, включающая как элементы биологически вдохновленных, так и градиентных подходов. Ее ключевой особенностью является многозадачность, заключающаяся в возможности параллельного решения нескольких задач: локализации и мониторинга нескольких источников и восстановления границ заданной линии уровня.

Keywords

Sampling problem, поле концентрации, стратегия управления, автономные агенты, роевой интеллект

1. Введение

Обследование поля концентрации или же Sampling problem [1] (в иностранных источниках) – это класс задач, подразумевающий коллективное изучение некоторой заданной области множеством скоординированных агентов. Объектом обследования в данном случае является объект, процесс или явление, которое в общем случае можно описать некоторой функцией вида:

$$f(t,q) : T \times Q \to \mathbb{R}, \qquad Q \subseteq \mathbb{R}^p, \qquad T = [0,\infty),$$

где $p \geq 2$. При этом, важным ограничением является наличие возможности агентов взаимодействовать с полем концентрации – измерять значение функции – только своих текущих координатах и с заданной периодичностью. В зависимости от природы происхождения поля принято разделять их на следующие три класса: химические [2, 3], примером которых может служить поле солености; физические [4], представляемые, например, термоклинами; биологические [5], описывающие поведение одной или нескольких популяций организмов. Благодаря такой вариативности и широкому спектру как

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

madstayler93@gmail.com (Толстихин)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

практических, так и фундаментальных задач, которые можно сформулировать на основе sampling problem, данное направление обрело высокую популярность в научной среде.

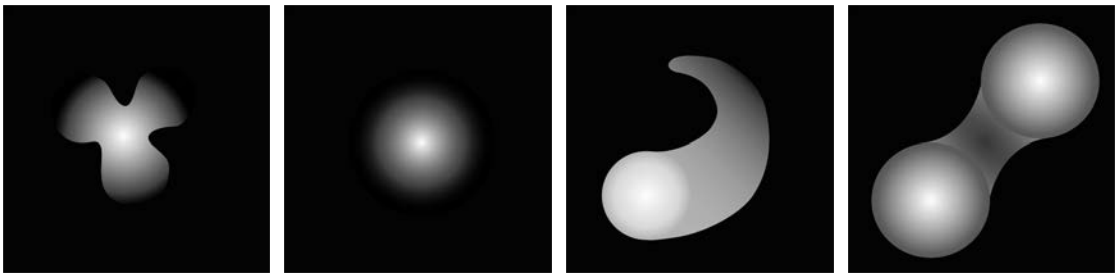
Обычно принято выделять три основных цели обследования:

- локализация одного или нескольких экстремумов функции [6] (источников), описывающей поле концентрации, в случае нестационарного поля часто сопровождается последующим отслеживанием их перемещений в пространстве (мониторинг);
- восстановление линий уровня [7], частным и наиболее интересным случаем которой является поиск нулевой линии уровня (фронта), отделяющей область с положительными значениями величины поля от области с нулевыми;
- картографирование [8] обследуемой области, зачастую носящая вспомогательный характер для рассматриваемой задачи и смежных с ней.

В данной работе предлагается децентрализованная мультиагентная стратегия управления, призванная обеспечить параллельное решение задач локализации и восстановление линии уровня. В ее основе лежат элементы биологически вдохновленных [9] и градиентных [10] подходов, а также механизмы роевой самоорганизации.

2. Постановка задачи

Помимо простейших стационарных моделей полей концентрации, часть из которых представлены на рисунке 1 и которые использовались на начальных этапах исследования, в данном исследовании рассматриваются две нестационарные модели, имеющие биологическую природу. Данный выбор обусловлен тем, что в большинстве случаев химические и физические процессы протекают значительно медленнее. В свою очередь это позволяет в рамках реальных задач пренебречь их изменчивостью, рассматривая их в качестве стационарных, что сильно упрощает задачу обследования поля концентрации.



a) b) c) d)

Рис. 1: Примеры упрощенных стационарных моделей полей концентрации.

Первая модель описывает взаимозависимое распространение двух постоянно растущих популяций в пространстве и времени, которую можно представить в виде следующей системы и являющейся частным случаем [11]:

$$f(t,q) = \left(\phi_1(t) - \frac{1}{r^2}\|q - q_0\|^2\right)^3,$$

$$g(t,q) = \left(\frac{29}{2048}\right)^3 \left(\phi_2(t) - \frac{1}{r^2}\|q - q_0\|^2\right)^3,$$

$$t \in [0,\infty), q \in \mathbb{R}^2.$$

$$\begin{cases} \dot{\phi}_1(t) = \frac{29}{6}\phi_2(t) - \frac{11}{6}\phi_1(t), \\ \dot{\phi}_2(t) = \frac{4879}{12288}\phi_1(t) - \frac{4357}{12288}\phi_2(t), \end{cases} \qquad \phi_i(0) = 1.$$

Данная модель имеет ряд особенностей:

- Каждое поле концентрации имеет так называемую «базу» или источник q^e, представляемый окружностью радиуса r и центром в q_0, концентрация внутри которого равна концентрации на его периметре. При этом, источник един для обеих популяций и не может перемещаться, но значения функций поля внутри него растет экспоненциально.
- Линии уровня представляют из себя концентрические окружности, скорость роста радиуса которых также увеличивается экспоненциально с течением времени.
- Исходя из биологического прообраза модели, значения концентрации не могут принимать отрицательные значения.

Вторая модель описывает стайное перемещение одной или нескольких независимых популяций фиксированной размерности по замкнутой территории. В данном случае поведение каждой особи q^a популяции B моделируется согласно законам Рейнольдса [12], а функция поля имеет следующий вид:

$$f(t,q) = (a * h)(t,q),$$

$$a(t,q) = \begin{cases} 1 & if \quad \exists k \in B : q = q_k^a \\ 0 \end{cases}, \quad h(t,q) = \begin{cases} 1 & if \quad \|q\| < r \\ 0 \end{cases},$$

где $(a * h)$ – операция свертки, r – радиус ядра свертки, являющийся управляемым параметром и представляющий из себя характеристики некоторого датчика, способного подсчитывать количество особей в заданной окрестности.

Ключевым отличием модели является отсутствие ярко выраженных источников поля. В данном случае под источником будем понимать локальные экремумы функции поля, иными словами:

$$q_j^e(t) = q \in Q : f(t,q) > f(t,q+\epsilon) \quad \forall\|\epsilon\| < \epsilon_0, j = 1 \ldots n_e,$$

где ϵ_0 – некоторая малая окрестность, n_e – количество отслеживаемых источников.

Для оценки текущего качества решения каждой из подзадач введем следующие критерии качества:

$$M_e(t) = \max_{j=(1,n_e)} \min_{i \in G_{base}} \|q_i(t) - q_j^e(t)\|, \tag{1}$$

$$M_f(t) = \sum_{i \in G_f ront} \frac{f(t, q_i(t))}{|G_{front}|} - s, \tag{2}$$

где $q_i(t)$ – координаты агентов, использующихся для решения задачи; G_{base} и G_{front} – соответственно подмножества агентов, занимающихся поиском источников и линий уровня; s – искомое значение линии уровня. Критерий $M_e(t)$ может принимать только положительные значения, соответственно, его минимизация демонстрирует, что в окрестности каждого источника находится как минимум один агент, производящий его мониторинг. С другой стороны, критерий $M_f(t)$ может принимать как положительные значения, что говорит об отставании агентов от желаемой линии уровня, так и отрицательные, сигнализирующие об ее опережении.

3. Стратегия управления

Пусть используемые для решения задачи агенты представляют из себя интеграторы второго порядка, а их динамика описывается следующей системой:

$$\dot{q}_i = v_i, \quad \dot{v}_i = u_i, \quad i \in G = 1, 2, ..., n, \quad \|v_i\| \leq v_{max}, \quad \|u_i\| \leq u_{max},$$

где $q_i \in \mathbb{R}^2$, $v_i \in \mathbb{R}^2$ и $u_i \in \mathbb{R}^2$ – соответственно, положение, скорость и управление i-го агента, v_{max} и u_{max} – предельно допустимые значения скорости и управления, G – множество доступных агентов мощностью n. Каждый агент способен с заданной периодичностью измерять величину поля концентрации в точке своего текущего местоположения. Помимо этого, будем считать, что каждый агент может связаться с любым другим для того, чтобы запросить его текущие координаты и значение последнего произведенного замера.

Предлагаемая стратегия задает управление каждого агента в виде взвешенной суммы нескольких сил, воздействующих на него:

$$u_i = c_1 F_i^c + c_2 F_i^g + c_3 F_i^s + c_4 F_i^f + c_5 F_i^r + c_6 F_i^b,$$

где $c_1 - c_6$ – некоторые положительные коэффициенты. При этом, для обеспечения параллельности решения подзадач локализации источников и восстановления линий уровня, множество агентов разбивается на два непересекающихся подмножества G_{base} и G_{front} такие, что $|G_{base}| \ll |G_{front}|$. В зависимости от принадлежности к этим подмножествам влияние некоторых сил на агентов исключается путем обнуления соответствующих коэффициентов. Так, для агентов, обеспечивающих поиск источников, c_4 и c_5 принимаются равными нулю, а для агентов множества G_{front} это происходит с коэффициентами $c_1 - c_3$.

Поскольку задача обследования поля концентрации подразумевает необходимость параллельного поиска и мониторинга нескольких источников, множество G_{base} разбивается на несколько поисковых кластеров таких, что

$$\tau = \{\tau_1, \tau_2, ..., \tau_m\}, \quad \forall j, k : j \neq k \rightarrow \tau_j \cap \tau_k = \emptyset, \quad |\tau_j| \approx |\tau_k| \geq 3.$$

Агенты, принадлежащие разным кластерам, находятся между собой в конкурентных отношениях. Более правильным с точки зрения биологической природы предложенной стратегии управления является термин «аменсализм», в рамках которого одни агенты

оказывают негативное эффект на других, не испытывая со стороны последних ни положительного ни негативного влияния. Это выражается в том, что группа, находящаяся ближе других к источнику, отталкивает другие кластеры, не позволяя им приближаться.

Градиентная сила F_i^g направляет робота вдоль рассчитанной оценки градиента к ожидаемому экстремальному значению поля и определяется как

$$F_i^g = \sum_{j \in \tau_k} \frac{q_{ij}(f(t, q_j(t)) - f(t, q_i(t)))}{\|q_{ij}\|}, \quad i \in \tau_k,$$

где $\|q_{ij}\|$ – евклидова норма вектора $q_{ij} = q_i - q_j$. Благодаря данной силе достигается самоорганизация движения кластера, при которой каждый робот движется в том же направлении и с той же скоростью, а также реализуется механизм эксплуатации.

С другой стороны, кооперирующая сила F_i^c обеспечивает реализацию агентами стайного поведения, в частности, избегание столкновений и центрирование стаи. Было предложено определять данную силу следующим образом:

$$F_i^c = -\sum_{j \in \tau_k} \nabla q_i U_{ij}^c(\|q_{ij}\|), \quad i \in \tau_k,$$

где $U_{ij}^c(\|q_{ij}\|)$ – искусственная потенциальная функция, которая определяет взаимодействие агентов, ∇q_i обозначает градиент относительно компонент вектора q_i В качестве потенциальной функции $U_{ij}^c : \mathbb{R}^+ \to \mathbb{R}^+$ предлагается функция

$$U_{ij}^c(\|q_{ij}\|) = \alpha \left(\frac{1}{2}(\|q_{ij}\| - d_{ij}^A)^2 + \beta \ln \|q_{ij}\| + \beta \frac{d_{ij}^A}{\|q_{ij}\|} \right),$$

где $\alpha, \beta \in \mathbb{R}^+$ – некоторые управляющие параметры; $d_{ij}^A > 0$ определяет желаемое расстояние между агентами. Функция (ссылка) построена на основе потенциальной функции, предложенной в [13], путем добавления параметра β, позволяющего варьировать размер области, на которой эта функция является строго выпуклой. Таким образом, под воздействием данной силы агенты одного кластера стремятся образовать формацию, которая в случае минимального необходимого количества агентов имеет вид правильного треугольника с гранями длинной d_{ij}^A.

Последней силой, непосредственно влияющей на процесс локализации источников, является сегрегирующая F_i^s сила, которая отвечает за вышеописанное взаимодействие между различными кластерами и задается следующим образом:

$$F_i^s = -\sum_{j \notin \tau_k} \nabla q_i U_{ij}^s(\|q_{ij}\|), \quad i \in \tau_k,$$

$$U_{ij}^s(\|q_{ij}\|) = \begin{cases} 0 & \|q_{ij}\| > d_{ij}^B \\ \alpha \left(\frac{1}{2}(\|q_{ij}\| - d_{ij}^B)^2 + \beta \ln \|q_{ij}\| + \beta \frac{d_{ij}^B}{\|q_{ij}\|} \right) & f(t, q_j(t)) > f(t, q_i(t) \end{cases},$$

где $d_{ij}^B \gg d_{ij}^A$ – минимальное желаемое расстояние между кластерами.

Каждый агент множества G_{front} при поиске заданной линии уровня ориентируется на положение одного из поисковых кластеров . На начальном этапе эта связь задается с ближайшим кластером, но агент может принять независимое решение о его смене при определенных условиях. Сила F_i^f, названная фронтальной из-за ее изначального предназначения в обнаружении фронта поля концентрации, обеспечивает притяжение агента к текущему местоположению источника или отталкивание от него в зависимости от того ниже его текущее значение замера относительно целевой линии уровня или выше. При этом, за координаты источника принимаются усредненные координаты агентов связанного кластера. Иными словами:

$$F_i^f = \sum_{j \in \tau^*} \frac{(2g_i(f(t,q_i)) - 1)q_{ij}}{\|q_i j\|},$$

$$g_i(f(t,q_i)) = \begin{cases} 0 & f(t,q_i) \le f_l \\ 3\left(\frac{f(t,q_i)-f_l}{f_u-f_l}\right)^2 - 2\left(\frac{f(t,q_i)-f_l}{f_u-f_l}\right)^3 & f_u \ge f(t,q_i) \ge f_(l,) \\ 1 & f(t,q_i) \ge f_u \end{cases},$$

где f_l и f_u – соответственно нижняя и верхняя границы целевых значений, τ^* – текущий поисковый кластер, связанный с агентом. Функция $g_i(f(t,q_i))$ представляет из себя так называемую SmoothStep функцию, обеспечивающую плавность переключения между режимами притягивания и отталкивания в окрестности искомой линии уровня, из-за чего управление не терпит разрывов.

Релаксирующая сила F_i^r применяется для равномерного распределения агентов по периметру линии уровня и имеет следующий вид:

$$F_i^r = \sum_{i,j \in \tau^*} \frac{q_{ij}}{\|q_{ij}\|}\left(e^{d_{ij}^C - \|q_{ij}\|} + a\right),$$

где d_{ij}^C – некоторое желаемое расстояние между агентами, a – положительный коэффициент, определяющий скорость распределения агентов по периметру. При решении задачи могут возникать ситуации, при которых слишком большое количество агентов занято оконтуриванием одной линии уровня, в то время как остальные страдают от их нехватки. В этом случае будет выполняться условие

$$\left(\|c_4 F_i^f\| < \left|\frac{c_4 F_i^f \cdot c_5 F_i^r}{\|c_5 F_i^r\|}\right|\right) \wedge (c_4 F_i^f \cdot c_5 F_i^r < 0),$$

где $\cdot$ – оператор скалярного произведения. При обнаружении данного события агент принимает решение сменить связанный кластер на следующий по очереди в цикличном списке. Кроме того, побочным эффектом влияния двух этих сил является выдерживание агентами определенной формации при перемещении по областям со значением функции поля ниже целевой.

Наконец, последней к агентам применяется граничная сила F_i^b, обеспечивающая их удерживание в рамках обследуемой области и не позволяющая покинуть ее пределы под воздействием других сил. Она задается следующим образом:

$$F_i^b = [f_1^b, ..., f_p^b],$$

$$f_k^b = e^{x_k^{min}+w-x_k} - e^{x_k-x_k^{max}+w},$$

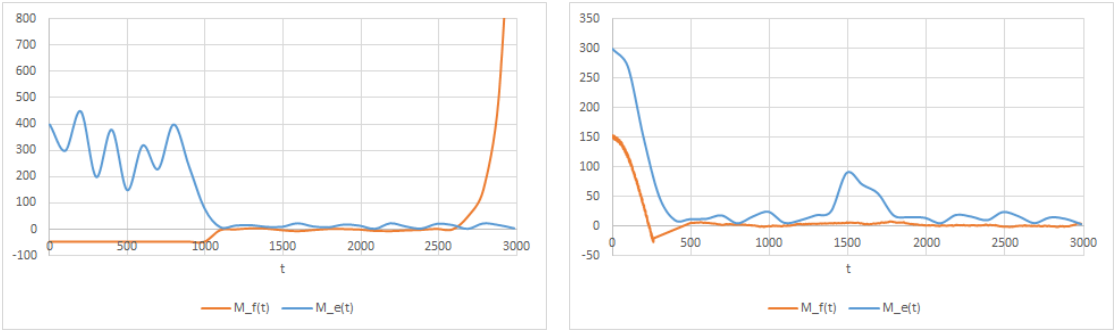
где x_k^{min} и x_k^{max} – компоненты векторов, ограничивающих обследуемую область, x_k – компоненты вектора q_i, а w – размер демпинговой зоны по периметру обследуемой области.

4. Заключение

Для оценки работы предложенного подхода была проведена серия компьютерных экспериментов с использованием вышеописанных моделей полей концентрации. На рисунке 2 представлены графики изменения критериев эффективности (1-2) для наиболее характерных вариантов поведения агентов при решении задачи. Условия проведения экспериментов приведены в таблице 1.

Таблица 1
Условия проведения экспериментов

Характеристика	Модель растущей популяции	Модель стайной популяции
Размер области	1000 × 1000 метров	1000 × 1000 метров
Время симуляции	3000 модельных секунд	3000 модельных секунд
$\|\tau\|$	1	3
$\|G_{front}\|$	50	50
v_{max}	3 $\frac{m}{s}$	3 $\frac{m}{s}$
u_{max}	0.5 $\frac{m}{s^2}$	0.5 $\frac{m}{s^2}$
Целевая линия уровня	45	45



a) b)

Рис. 2: Результаты экспериментов для модели растущей популяции (a) и модели стайной популяции (b).

На данных графиках можно условно выделить три отрезка, описывающие разные этапы решения задачи. Первый этап заключается в формировании поисковых кластеров

и поиска ими источников. Следует обратить внимание, что для первой модели данный этап занимает значительно больше времени ($t \in [0, 800]$). Это обусловлено тем, что благодаря особенностям данной модели на ранних этапах функция поля концентрации принимает нулевые значения практически во всех точках обследуемой области. Из-за этого градиентная сила не оказывает воздействия на агентов, которые осуществляют дрейф под действием остальных сил. Однако, при обнаружении хотя бы одного ненулевого замера, происходит быстрая локализация источника и переход ко второму этапу.

В его рамках наблюдается мониторинг источников и вывод агентов множества G_{front} на целевую линию уровня, а также их равномерное распределение по ней. Тесты показали, что в среднем отклонение метрики $M_e(t)$ на данном этапе не превышает 20 метров, что, учитывая размер обследуемой области и выдерживаемое между агентами расстояние (25 метров), является допустимым. Однако, из-за трудно прогнозируемого передвижения источников в стайной модели, могут наблюдаться непродолжительные прерывания мониторинга.

Наконец, третий этап характерен только для модели растущей популяции. Учитывая экспоненциальный рост скорости движения линий уровня, гарантированно случается ситуация, при которой ограничение скорости движения агентов делает невозможным дальнейшее решение задачи.

Последним важным выявленным на этапе тестирования стратегии фактом является ограничение классов линий уровня, поддающихся отслеживанию. Данный подход может применяться в тех случаях, когда ограниченная линией уровня область является выпуклой или представляет из себя звездную область [14] относительно координат источника. В случае невыпуклой области (рисунок 1c) могут наблюдаться протяженные участки периметра, которые поисковые агенты не способны обнаружить. Однако, потенциально возможны ситуации (рисунок 1d), при которых правильные границы могут быть восстановлены за счет коллективных действий нескольких подгрупп агентов.

Acknowledgments

Работа выполнена при поддержке Российского Научного Фонда, грант № 22-29-00819.

Список литературы

[1] J. Hwang, N. Bose, S. Fan, Auv adaptive sampling methods: A review, Applied Sciences 9 (2019). doi:10.3390/app9153145.

[2] Y. Zhang, R. McEwen, J. Ryan, et al., A peak-capture algorithm used on an autonomous underwater vehicle in the 2010 gulf of mexico oil spill response scientific survey, J. Field Robotics 28 (2011) 484–496. doi:10.1002/rob.20399.

[3] X.-x. Chen, J. Huang, Odor source localization algorithms on mobile robots: A review and future outlook, Robotics and Autonomous Systems 112 (2019) 123–136. doi:10.1016/j.robot.2018.11.014.

[4] H. C. Woithe, U. Kremer, Feature based adaptive energy management of sensors on autonomous underwater vehicles, Ocean Engineering 97 (2015) 21–29. doi:10.1016/j.oceaneng.2014.11.015.

[5] J. Das, K. Rajan, S. Frolov, et al., Towards marine bloom trajectory prediction for auv mission planning, Proceedings - IEEE International Conference on Robotics and Automation (2010) 4784–4790. doi:10.1109/ROBOT.2010.5509930.

[6] W. Naeem, R. Sutton, J. Chudley, Chemical plume tracing and odour source localisation by autonomous vehicles, Journal of Navigation 60 (2007) 173 – 190. doi:10.1017/S0373463307004183.

[7] S. Petillo, Autonomous adaptive oceanographic feature tracking on board autonomous underwater vehicles, Woods Hole Oceanographic Institution: Woods Hole (2015).

[8] M. Mysorewala, L. Cheded, D. Popa, A distributed multi-robot adaptive sampling scheme for the estimation of the spatial distribution in widespread fields, EURASIP Journal on Wireless Communications and Networking 2012 (2012). doi:10.1186/1687-1499-2012-223.

[9] P. Maes, M. J. Mataric, J.-A. Meyer, et al., Locating Odor Sources in Turbulence with a Lobster Inspired Robot, 1996, pp. 104–112.

[10] E. Fiorelli, P. Bhatta, N. Leonard, I. Shulman, Adaptive sampling using feedback control of an autonomous underwater glider fleet, Proceedings of the 13th International Symposium on Unmanned Untethered Submersible Technology (UUST), Durham, NH, USA, (2003) 1–16.

[11] A. Kosov, E. Semenov, New exact solutions of the diffusion equation with power nonlinearity, Sibirsk. Mat. Zh. 63 (2022) 1290–1307. doi:10.33048/smzh.2022.63.610.

[12] C. W. Reynolds, Flocks, herds and schools: A distributed behavioral model, SIGGRAPH Comput. Graph. 21 (1987) 25–34. doi:10.1145/37402.37406.

[13] V. Santos, A. Grandi Pires, R. Javanmard Alitappeh, et al., Spatial segregative behaviors in robotic swarms using differential potentials, Swarm Intelligence 14 (2020). doi:10.1007/s11721-020-00184-0.

[14] E. Schechter, Handbook of Analysis and Its Foundations, Academic Press, San Diego, 1997. doi:10.1016/B978-012622760-4/50000-5.

Извлечение иерархических пар из заголовков электронных таблиц

В.В. Парамонов[1,2], А.О. Шигаров[1,2]

[1]*Институт динамики систем и теории управления имени В.М. Матросова СО РАН, Россия, 664033, гор. Иркутск, ул. Лермонтова, 134*

[2]*Институт математики и информационных технологий Иркутского государственного университета, Россия, 664003, гор. Иркутск, бул. Гагарина, 20*

Аннотация

В электронных таблицах содержится большое количество ценной информации, агрегирование и последующий анализ которой имеют практическую значимость и могут быть использованы в различных предметных областях. Для эффективного автоматизированного извлечения данных необходимо понимание их семантической структуры (отношения). Отношения между ячейками электронной таблицы возможно получить посредством анализа заголовка таблицы. Для этого необходимо выделить пары взаимосвязанных ячеек, что даст возможность построить иерархическое дерево. В данной работе рассматриваются эвристические методы к определению взаимосвязанных ячеек в заголовках, формированию пар предок-потомок. Также предложена методика оценки производительности предложенных эвристик.

Keywords

электронные таблицы, семантика, заголовки таблицы, эвристика

1. Введение

Электронные таблицы являются весьма распространенным способом представления однотипных данных. Так они повсеместно используются для хранения и представления данных статистической и финансовой информации, различных научных наблюдений. Такой способ представления понятен и достаточно прост в использовании, поскольку не требует наличия никаких дополнительных знаний и специальных навыков для работы на базовом уровне. В сети Интернет циркулирует огромное количество электронных таблиц, представленных в формате Excel и содержащих ценную информацию из различных предметных областей [1, 2, 3]. Извлечение и последующая интеграция данных из электронных таблиц является очень важной частью современной бизнес аналитики. Последние исследования демонстрируют, что только в США более 55 миллионов пользователей работают с электронными таблицами из них более половины используют электронные таблицы для ведения бизнеса [4]. Это позволяет сказать, что Excel является одним из самых популярных форматов электронных таблиц.

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

slv@icc.ru (В.В. Парамонов); shigarov@icc.ru (А.О. Шигаров)

0000-0002-4662-3612 (В.В. Парамонов); 0000-0001-5572-5349 (А.О. Шигаров)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

Проанализировать и извлечь информацию из огромного количества таблиц можно лишь в автоматическом режиме. При этом многие проблемы интеграции информации зависят от спецификаций и схем организации данных. В то же время следует отметить, что электронные таблицы, в первую очередь, создаются людьми и для дальнейшего использования людьми. Это обуславливает различные подходы как к компоновке как самих таблиц, так и организации данных. Так на одном листе Excel может находиться несколько таблиц. Кроме самих таблиц может быть также дополнительная метаинформация, такая как названия, комментарии описания. Компоновка таблиц также отличается. Лишь в редких случаях у таблиц встречается одноуровневый заголовок [5].

Одной из особенностей, электронных таблиц не представляющих каких-либо трудностей для человека, но сказывающейся на корректности автоматического понимания данных являются иерархические заголовки. В этом случае они определяют структуру и иерархическую зависимость между табличными данным. Понимание зависимости между данными позволит более эффективно и полно извлекать табличную информацию. Пример иерархического заголовка приведен на рис. 1.

Sex, age, and selected first-listed diagnosis \1	Discharges	
	Number (1,000)\2	Per 1,000 persons \3
Both sexes		
Total \3 \4	34 667	116,2
Male		
All ages \4	13 990	95,5
Under 18 years \4	1 515	40,2
Injury	167	4,4
Pneumonia	101	2,7
Asthma	\5 99	\5 2,6
18-44 years \4	2 701	47,7

Рис. 1: Фрагмент таблицы с иерархическим заголовком.

Так, надо установить, что в заголовке, данные “Number (1,000) \2” и “Per 1,000 persons \3” являются нижними уровнями иерархии по отношению к “Discharges”. Наиболее интересный случай установления иерархий данных для боковика — самого левого столбца таблицы. На рис. 2 приведены его уровни иерархий, где 0 - самый высокий уровень. Автоматическое определение иерархий позволяет установить связи между данными, повысить качество их излечения. В связи с этим актуальна задача автоматизировать определение пар ячеек “родитель” — “потомок” для составления иерархии.

Отметим, что наличие сложных структур организации данных, в свою очередь, затрудняет процессы автоматической обработки, понимания таблиц и извлечению данных. В результате ряда исследований были сформулированы требования к формированию электронных таблиц [6, 7, 8]. Однако, в действительности эти требования остаются лишь требованиями. Разработчики электронных таблиц оформляют их таким образом, чтоб было удобнее с ними работать.

Уровень иерархии	Значения
0	Both sexes
1	Total \3 \4
1	Male
2	**All ages \4**
3	Under 18 years \4
4	Injury
4	Pneumonia
4	Asthma
3	18-44 years \4

Рис. 2: Пример иерархии вертикального заголовка

2. Обзор связанных работ

Научная проблема установления иерархий в заголовках таблиц является актуальной. Существует множество исследовательских работ, посвященных анализу функций как ячеек таблиц, так и анализу их структуры [9, 10, 11, 12, 13]. Рассмотрим некоторые из работ в данной области сконцентрированных именно на построении иерархий заголовков. Так в работах [11, 14, 12] представлен метод построения иерархий в заголовках. Особенность данного метода заключается в том, что он является независимым от предметной области. Авторы рассматривают построение иерархических отношений в боковике. Данный метод направлен на определение отношений Предок — потомок для данных и, как следствие построение иерархического дерева. Для получения иерархических пар предложены два подхода:

- основанные на эвристиках;
- использующие машинное обучение на основе метода оперных векторов.

В обоих случаях для каждой из возможных пар строится вектор свойств, учитывающий значения признаков, описывающих компоновку — взаимное расположение (соседство) ячеек, графическое форматирование — отступы, шрифты, заливку, а также текстовое содержание — заполнители, ключевые слова (“total”, “sum”), лексическое и семантическое сходство ячеек, а также семантическое сходство данных. В результате построения образуются пары родитель — ребенок, по которым и строится дерево иерархий. При этом основное внимание сфокусировано на определение иерархий именно в боковике.

В работе [5] рассматривается эвристический подход к определению семантики заголовков таблиц. В работе отмечается, что для понимания электронной таблицы необходимо проанализировать её структуру, обнаружив пары ячеек с различными типами взаимосвязей в электронной таблице, и установить иерархические индексные связи между ними.

3. Алгоритм построения иерархических пар

Предлагаемый нами алгоритм построения иерархических пар ячеек основан не идеи сопоставления векторов свойств пар ячеек, предложенных в работах Chen Z. В качестве источника данных рассматриваются таблицы в формате Excel из набора SAUS200 (“The 2010 Statistical Abstract of the United States”) [15]. Для поиска границ таблиц были использованы адаптированные подходы, переложенные в методики FrameFinder [11]. Данная методика имеет ряд ограничений — на одном листе возможно несколько таблиц, но таблицы располагаются друг под другом. Для упрощения анализа, таблицы начинаются с первого столбца. Набор SAUS200 соответствует данным ограничениями. В результате обработки, для каждой каждого Excel-документа был получен сопроводительный файл, содержащий номера строк и названия функциональных областей таблицы и метаинформации (“Title”, “Footnote”, “Header”, “Data”, “Blank”). Так как рассматривается боковик — крайний левый столбец таблицы Excel, в диапазоне “Data”.

Нами был найден единственный вариант программной реализации[1] алгоритмов, предложенных в Senbazuru [11, 12] в котором частично были реализованы некоторые из представленных в работах подходы, в основном касающиеся построения вектора свойств. Построение вектора занимает достаточно длительное время так как они строятся для всевозможных пар ячеек. И только после анализа векторов, большая часть не учитывается, т.к. образующие их пары ячеек не являются иерархией. В рассмотренных источниках не приводится полное описание эвристического алгоритма. Оговаривается лишь то, что используется вектор сопоставляющий 10 унарных и 14 бинарных свойств [11]. В адаптированной же версии мы используем вектор, состоящий из 15 свойств пары табл. 3, из которых 8 свойств (№1–8) из реализации Senbazuru и 7 (№9–15) предложены в рамках адаптации. При этом свойства 5 и 15 имеют множественные значения. Например, позиция разделителя по отношению к остальному содержимою ячейки (свойство 5); данные в ячейки соответствуют числу, строке или представляются смешанным типом (свойство 15). При этом анализ типа данных ячеек на уровне Excel не проводится, т.к. в большинстве случаев он не специфицируется.Также было сделано допущение, что если способ выравнивания данных по вертикали/горизонтали не определен, то считаем, что выравнивание по краю/левому нижнему соответственно.

Также было выявлено, что при распознавании отношений предок-потомок между ячейками таблиц коллекции SAUS200 нецелесообразно применять оценку семантической близости значений ячеек [12], поскольку упоминания сущностей в боковиках оказываются достаточно близкими семантически.

Рассмотрим работу данного алгоритма. Учитывая, что любая таблица интерпретируется сверху вниз, то пары рассматриваются только ячеек, имеющих меньший индекс — ячейка родитель, больший индекс – ячейка потомок. В соответствии с нашим предположением, ячейки в которых отсутствует какая-либо информация не несут смысловой нагрузки при построении иерархий данных поэтому они игнорируются. Следующий шаг — поиск ячеек, образующих иерархические пары. Будем считать, что ячейки образуют иерархическую пару, если отличаются их стили оформления, присутствуют

[1] https://github.com/rackingroll/Senbazuru

Таблица 1
Свойства ячеек

№	Свойство	Описание
1	BFeatureAdjacent	Ячейки являются соседними
2	BFeatureChildindentationGreater	Отступ в ячейке потомке больше, чем в родительской ячейке
3	BFeatureChildindexGreater	Индекс ячейки потомка больше чем индекс родительской ячейки
4	BFeatureChildSizeSmaller	Размер шрифта в ячейке потомке меньше чем размер шрифта родительской ячейки
5	BFeatureContainColonAndTotal	В рассматриваемо паре и между ними присутствуют такие слова как “sum”, “total”, “:”
6	BFeatureBoldDiffer	В рассматриваемой паре отличается стиль “полужирный”
7	BFeatureItalicDiffer	В рассматриваемой паре отличается стиль “наклонный”
8	BFeatureUnderlineDiffer	В рассматриваемой паре отличается стиль “подчеркнутый”
9	BFeatureBackgroundDiffer	В рассматриваемой паре отличается цвет фона
10	BFeatureParentIsEmptyCell	Родительская ячейка - пустая
11	BFeatureChildIsEmptyCell	Ячейка потомок - пустая
12	BFeatureIndentationDifferent	В рассматриваемой паре данные в ячейках имеют различный отступ
13	BFeatureHorisontalAligmentDifferent	Текст в рассматриваемой паре имеет различное выравнивание по горизонтали
14	BFeatureVerticalAligmentDifferent	Текст в рассматриваемой паре имеет различное выравнивание по вертикали
15	BFeatureDataTypeDifferent	Текст в рассматриваемой паре может относиться к различным типам данных

стоп-слова, отличаются типы данных представление строками. Пара ячеек является предположительно иерархической, если присутствует отличие в одном из свойств 1, 4, 6, 7, 8, 9, 12, 13, 14, 15 и при этом обязательно, что ячейка-потомок не начинается на стоп-слово (свойство 5). Если при этом ячейки являются смежными (расположены рядом или между ними есть только пустые ячейки), то считаем такие ячейки иерархической парой. В противном случае требуется проверить, являются ли ячейки иерархией одного уровня (иерархическая пара) или принадлежат различным уровням (не являются иерархической парой). Для этого фиксируется позиция предполагаемой ячейки предка, и рассматриваются ячейки потомки от выбранной и до ячейки-предка не включительно. В случае, если какая-то из пар ячеек имеет похожие свойства, то тогда считаем, что гипотеза об иерархической паре подтверждается. Будем считать, что вектора свойств считаются похожими, если все свойства, отвечающие за стили оформления совпадают. Типы данных в пере ячеек могут либо совпадать, либо одна из ячеек пары имеет смешанный тип. Рассмотрим работу алгоритма на примере фрагмента боковика таблицы рис. 3. Ячейки 1 и 3 являются смежными и имеют разные стили. Поэтому они являются

1	**All causes**
2	
3	Major cardiovascular diseases
4	Diseases of heart
5	Acute rheumatic fever and chronic rheumatic heart disease
6	Hypertensive heart disease
7	Acute myocardial infarction
8	Other acute ischemic heart diseases
9	Other forms of chronic ischemic heart disease
10	Atherosclerotic cardiovascular disease, so described
11	All other forms of chronic ischemic heart disease
12	Other heart diseases
13	Acute and subacute endocarditis
14	Diseases of pericardium and acute myocarditis

Рис. 3: Пример иерархии вертикального заголовка

иерархической парой (1, 3), где 1 — ячейка предок, 3 – ячейка потомок. Далее в таблице отсутствует какая-либо пара ячеек имеющих такой же вектор свойств, как и у (1, 3). Придерживаясь описанных шагов алгоритма получаем следующие пары (3, 4), (4, 5), (4, 6), (4, 12), (6, 7), (6, 8), (6, 9), (9, 10), (9, 11), (12, 13), (12, 14). Идентифицированные иерархии ячеек сохраняются в виде текстового файла. По ним возможно построить полное дерево иерархии данных, на основании вертикального заголовка (боковика) таблицы.

4. Оценка эффективности полученного решения

В работах, посвященных поиску и установлению иерархий табличных данных [11, 14, 12, 5] даются оценки производительности описанных методов, однако, нет информации о методики, которая использовалась для этого. Для оценки предложенного решения нами был подготовлен набор данных для оценки производительности предложенного метода.

Была проведена разметка эталонных данных коллекции таблиц SAUS200. Таблицы набора были сопровождены цветовой разметкой уровней вложенности иерархий заголовков. Для извлечения информации об иерархии пар на основе цветовой разметки ячеек таблицы, был разработан комплекс VBA-макросов генерации эталонных данных. Это позволило автоматически сгенерировать эталонные данные —– наборы пар ячеек, отражающих отношениям предок-потомок для иерархий заголовков в формате, предложенном в работах Chen [11, 12]. Всего сгенерировано 129 иерархий заголовков (из 200 таблиц), содержащих от 2 до 419 пар ячеек. Оставшиеся 71 таблица имели заголовки плоской структуры. В эталонных данных были обнаружены 5442 пары иерархий.

В результате тестирования разработанного подхода было извлечено 3775 иерархий заголовков в 136 таблицах. Доля корректно распознанных пар ячеек составила 43%. Количество некорректно распознанных пар колеблется от 1 до 49, присутствующих в 80 таблицах набора. При этом, следует учитывать, что количество строк в таблицах значительно отличается.

Следует отметить, что корректность установления иерархических пар очень сильно зависит от экспертной интерпретации табличных данных. В силу этого часть заголовков имеют неоднозначную интерпретацию в автоматическом и ручном режимах.

5. Заключение

В работе рассмотрен эвристический подход к поиску иерархических пар в вертикальном заголовке таблицы. Эвристики основываются на генерации вектора характеристик для каждой пары ячеек. Вектор строится на оценках взаимного расположения (соседства) ячеек, графического форматирования — отступы, табуляции, шрифты, наличие заливки, а также текстовом содержании — заполнители, ключевые слова, типы данных.

Полученные в результате тестирования показатели не являются очень высокими ввиду неоднозначности определения некоторых из иерархий. Однако они сопоставимы с результатами, полученными с использованием эвристических алгоритмов в работах [11, 12]. За счет меньшего использования характеристик удалось повысить быстродействие при построении векторов.

Благодарности

Результаты получены в рамках госзадания Минобрнауки России по проекту “Методы и технологии облачной сервис-ориентированной цифровой платформы сбора, хранения и обработки больших объёмов разноформатных междисциплинарных данных и знаний, основанные на применении искусственного интеллекта, модельно-управляемого подхода и машинного обучения” (номер гос. регистрации 121030500071-2).

Список литературы

[1] L. A. Kappelman, J. P. Thompson, E. R. McLean, Converging end-user and corporate computing, Commun. ACM 36 (1993) 79–92. URL: https://doi.org/10.1145/163298.163314. doi:10.1145/163298.163314.

[2] A. Shigarov, A. Altaev, A. Mikhailov, V. Paramonov, E. Cherkashin, Tabbypdf: Web-based system for pdf table extraction, in: R. Damaševičius, G. Vasiljevienė (Eds.), Information and Software Technologies, Springer International Publishing, Cham, 2018, pp. 257–269.

[3] R. Rastan, H.-Y. Paik, J. Shepherd, S. H. Ryu, A. Beheshti, Texus: Table extraction system for pdf documents, in: J. Wang, G. Cong, J. Chen, J. Qi (Eds.), Databases Theory and Applications, Springer International Publishing, Cham, 2018, pp. 345–349.

[4] Y. Zhang, X. Lv, H. Dong, W. Dou, S. Han, D. Zhang, J. Wei, D. Ye, Semantic table structure identification in spreadsheets, in: Proceedings of the 30th ACM SIGSOFT International Symposium on Software Testing and Analysis, ISSTA 2021, Association for Computing Machinery, New York, NY, USA, 2021, p. 283–295. URL: https://doi.org/10.1145/3460319.3464812. doi:10.1145/3460319.3464812.

[5] G. Li, R. Li, Z. Wang, C. H. Liu, M. Lu, G. Wang, Hitailor: Interactive transformation and visualization for hierarchical tabular data, IEEE Transactions on Visualization and Computer Graphics 29 (2023) 139–148. doi:10.1109/TVCG.2022.3209354.

[6] K. W. Broman, K. H. Woo, Data organization in spreadsheets, The American Statistician 72 (2018) 2–10. URL: https://doi.org/10.1080/00031305.2017.1375989. doi:10.1080/00031305.2017.1375989.

[7] Y. Wang, R. Wang, C. Jung, Y.-S. Kim, What makes web data tables accessible? insights and a tool for rendering accessible tables for people with visual impairments, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22, Association for Computing Machinery, New York, NY, USA, 2022. URL: https://doi.org/10.1145/3491102.3517469. doi:10.1145/3491102.3517469.

[8] P. Alexander, Principles of data design in spreadsheets, International Journal of Open Information Technologies 11 (2023) 102–105.

[9] E. Koci, M. Thiele, O. Romero, W. Lehner, Cell Classification for Layout Recognition in Spreadsheets: 8th International Joint Conference, IC3K 2016, Porto, Portugal, November 9–11, 2016, Revised Selected Papers, 2019, pp. 78–100. doi:10.1007/978-3-319-99701-8.

[10] M. D. Adelfio, H. Samet, Schema extraction for tabular data on the web, Proceedings of the VLDB Endowment 6 (2013) 421–432.

[11] Z. Chen, M. Cafarella, Automatic web spreadsheet data extraction, in: Proceedings of the 3rd International Workshop on Semantic Search Over the Web - SS@ '13, ACM Press, 2013. doi:10.1145/2509908.2509909.

[12] Z. Chen, S. Dadiomov, R. Wesley, G. Xiao, D. Cory, M. Cafarella, J. Mackinlay, Spreadsheet property detection with rule-assisted active learning, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management - CIKM '17, ACM Press, 2017. doi:10.1145/3132847.3132882.

[13] X. Wu, H. Chen, C. Bu, S. Ji, Z. Zhang, V. S. Sheng, Huss: A heuristic method for understanding the semantic structure of spreadsheets, in: 2022 IEEE International Conference on Knowledge Graph (ICKG), IEEE, 2022, pp. 329–336.

[14] Z. Chen, M. J. Cafarella, J. Chen, D. Prevo, J. Zhuang, Senbazuru: A prototype spreadsheet database management system, Proc. VLDB Endow. 6 (2013) 1202–1205.

[15] A. Shigarov, V. Paramonov, V. Khristyuk, Spreadsheet data extraction from real-world tables of saus (the 2010 statistical abstract of the united states): Case study, 2021. URL: https://figshare.com/articles/dataset/Spreadsheet_Data_Extraction_from_Real-World_Tables_of_SAUS_The_2010_Statistical_Abstract_of_the_United_States_Case_Study/14371055/2. doi:10.6084/m9.figshare.14371055.v2.

Web Table Extraction: Problem Statement

Alexey O. Shigarov[1]

[1]*Matrosov Institute for System Dynamics and Control Theory, SB RAS, 134 Lermontov St, Irkutsk, 664033, Russian Federation*

Abstract

Representing relational data on the Web is most naturally done through web tables. These tables contain a valuable information that can be utilized for various applications such as question answering, knowledge base population, and natural language generation. However, web tables typically lack accompanying semantics that allow for automatic interpretation of the presented information. The goal of Web Table Extraction (WTE) is to recover these missing semantics so that facts can be extracted from the tables. This process involves addressing issues such as web table discrimination and header detection. By automatically representing extracted data in relational databases, WTE provides a solution to this problem. This paper focuses on the problem statement of WTE and discusses some of the current research limitations.

Keywords

table recognition, table extraction, table mining, web tables

1. Introduction

Exploring the Web, the contemporary research [1, 2, 3, 4] collected hundreds of millions of tables containing relational data: "*the Web is the largest repository of data available, with over 150 million high-quality tables*" as reported in [5]. The abundance of relational data contained in web tables makes them an attractive source for various applications, including question answering, knowledge base population, and natural language generation. However, the lack of accompanying semantics in these tables presents a challenge for automatic interpretation of the presented information. Web Table Extraction (WTE) aims to address this issue by recovering missing semantics and extracting facts from the tables [6, 7, 8]. This process involves resolving problems such as web table discrimination and header detection, and automatically representing extracted data in relational databases. Despite the potential benefits of WTE, current research faces limitations in effectively handling the extraction and integration of tabular data for real-world applications.

In a general case, WTE involves the extraction of tables from web pages while attempting to recover both their physical and logical structure. As a result, the interrelated cells of a table become accessible to be automatically read in a logical order. WTE is breaking down into four distinct tasks: Table Detection (TD), which identifies tables on a web page and outlines

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

shigarov@icc.ru (A. O. Shigarov)

0000-0002-0877-7063 (A. O. Shigarov)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

their boundaries; Table Structure Recognition (TSR), which retrieves the relative positions and content of each cell in a given table, also known as the *physical structure* [9]; Table Functional Analysis (TFA), which assigns predefined types to cells or regions within a table to indicate their functional role, such as header or data; and Table Structural Analysis (TSA), which infers logical relationships between cells, also known as the *logical structure* of the table [9].

The remainder of this paper is organaized as follows: the first part characterizes the problem of WTE, regarding to the sources, tasks, and context (Section 2); the second part discusses the common limitations of the current research (Section 3); some directions for further research that would help overcome these limitations are considered in the conclusion.

2. Problem Statement

WTE is intended for web tables encoded in HTML. Web pages contain table borders and cell structure designated explicitly by tags (`<table>`, `<tr>`, `<td>`, and `<th>`), i. e. the physical structure is accessible. Nonetheless, both TD and TSR are relevant to web tables in general. HTML tables are often used simply as a means of layout to arrange web content in a grid [2]; this content is not relational data. TD should at least filter out such layout tables to eliminate them in further processing. TD in web pages is usually reduced to *Web Table Discrimination* (WTD), dividing HTML tables into two kinds: *genuine* (data) and *non-genuine* (non-data) [10, 11, 12, 7].

An HTML table can have an inappropriate physical structure. For example, rows and columns can be decorative or repetitive [13]; sometimes tables can be coded using listing tags [14], one can also come across nested tables (an HTML cell can contain other HTML tables) [15]. Generally, the analysis of the visual appearance of an HTML table instead of parsing its DOM tree should be preferable [15]. TSR corrects such cases to make the physical structure of the table clear for further processing [7].

TFA obligatorily differentiates *metadata cells* from *data cells* [16, 13]. This task can be reduced to *Table Type Classification* (TTC) and *Header Detection* (HD). TTC assigns one of predefined types (classes) to a table [2, 5, 3], e. g. *listing*, *attribute/value*, or *matrix* [2] (*relational*, *entity*, or *matrix* respectively in [4]). Some types can also be further differentiated by their orientation [4], e. g. *listing*/*relational* can be either *vertical* or *horizontal*. A type taxonomy can include a class of *non-genuine* (*non-data*) tables, for example, the so-called "*single-layer*" classification of [3]. In such cases, TTC takes over the function of WTD. HD separates rows/columns into two functional table regions, namely, header and data [17, 18, 9, 6]. Note that header regions can subdivided into *head*, *stub*, and *stub head*. Moreover other table regions such as *body*, *cut-in*, and *footer* can also be detected. In general, HD should be regarded as *Region Detection* (RD). HD/RD is closely related to *Cell Classification* (CC) [19, 20]. The latter is used to separate cells into two or more predefined types, e. g. *metadata* or *data* [13].

TTC and HD/RD facilitate TFA and TSA, substantially simplifying these tasks in some cases. For, example, if we define formal rules of the type-specific layout, they probably allow us to automatically infer the functional annotation. The relationships between header cells and data ones can be inferred too. More complex cases described in [21, 22] require more advanced analysis, for example, the extraction of header hierarchies.

Simple cases, e. g. vertical listings [2], are transformed to relational database tables directly,

so the CSV format is convenient to represent WTE output. However, the complex cases (e. g. header hierarchies) [21, 22] require to use a more appropriate notation such as [23, 24].

Note that WTE methods dealing with HTML tables can be usually applied to CSV files. Some of the methods even include a step of the conversion from HTML to CSV format, e. g. [18, 9]. WTE also covers wiki-tables [25, 26]. Although TD/TSR are missed when processing CSV files and wiki-tables but TFA/TSA do not lose their relevance at all.

Some sources provide not only tabular data but also a context explaining their meaning. For example, tables can be surrounded in sources by some text descriptions (titles, captions, footnotes, etc.). TME extracts metadata (topics, units, time, geography, etc.) from the tables as well as from the surrounding context [27, 28]. On the one hand, TME is an application of TE because the extraction of metadata from tables relies on TE methods. On the other hand, the metadata extracted from the surrounding context can enrich tables, making them more appropriate for the further processing.

An interesting example is presented in [29]: numerical data of Wikipedia tables are often not accompanied by units explicitly, but such omitted units can be estimated by using surrounding contexts. It is supposed that the knowledge of the units presented in tables can improve the results of TU. Another example: [30, 31] utilizes the semantic markups (RDFa/Microdata annotations) that are originally inserted within web pages to support semantic indexing in search engines. Particularly, such annotations can provide important information about entities listed in web tables.

3. Current Limitation

The current problem statments [6, 32, 7, 33] are typically limited by wide-spread table layout types. It is presumed that tables falling under the same type possess share the same physical and logical properties. While this approach facilitates table type classification and enables to propose some type-specific techniques. However, it does not encompass all possible properties that may arise in real-life tables.

Several taxonomies of table types are established by conducting extensive web crawls [2, 3, 4]. They distinguish those types which are widespread on the Web, including, “listing”, “entity”, “matrix”, and “enumeration” (Fig. 1, *a*–*e*). Accordingly, the existing solutions primarily address straightforward these prevalent table types. The current research paid none or very little attention to more sophisticated cases as reported in [5, 7, 8]. Some of them are shown in Fig. 1, *f*–*n*. There are only a few attemts to study heading hierarchy [34, 35], key-value pairs [36, 35], and the derived data [37, 38, 39].

The existing taxonomies of table types [2, 5, 3, 4] are incomplete, as they conflate layout types with functional ones. Specifically, a table may serve functionally as a cross-tabulation, while simultaneously possessing either a one-way (flat) or multi-way (matrix) layout. Moreover, extant taxonomies fail to differentiate between various types and properties of tables. For instance, merged or multivalued cells may be present in both flat and matrix layouts.

The current research is also limited by the “atomic” cells [7]. It is assumes that any cell can contain only one “atomic” data item. However, in real-life tables it can be frequently observed *structured cells* [7], the cases when a cell contains two or more distinct data items. (The obvious

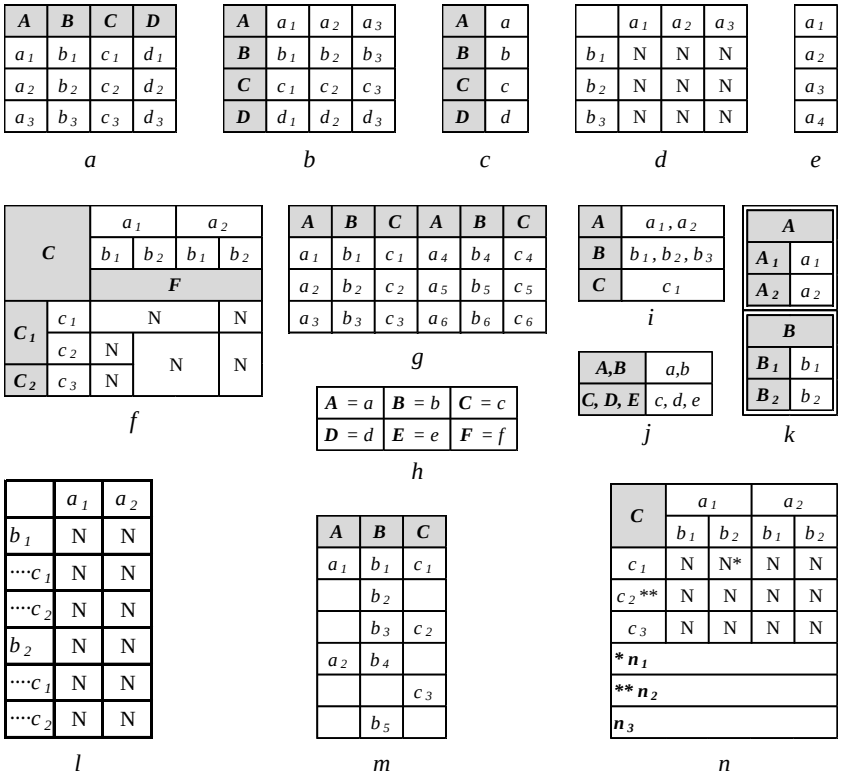


Figure 1: Simple cases: vertical/horizontal listing (*a*/*b*); entity (*c*); matrix (*d*); enumeration (*e*). Complex cases: merged cells (*f*); splitted tables (*g*); key-value pairs (*h*); simple/composed multivalued cells (*i*/*j*); nested tables (*k*); heading hierarchy (*l*); factorized cells (*m*); notes (*n*). The following notation is used: $A, B, C, \ldots$ — metadata (names of attributes or categories); $a_i, b_i, c_i, \ldots$ — data (values of categories); N — data (values of entries that are typically numbers).

example is units of measure, e.g. \$, %, or ℃, which are preceded or followed by the quantities in cells.) Surveying the literature addressing the data extraction from HTML tables, [7] indicate that only a few proposals deal with structured cells: [36] admit that a key-value pair can be placed in the same cell; [40] support structured cells containing nested tables or lists. [35] give the following formulations of the WTE tasks: TFA recognizes atomic data items in cells and assigns the predefined functional roles ("*label*" and "*entry*") to them; TSA extracts pairs of interrelated atomic data items ("*entry-label*" and "*label-label*").

4. Concusion

Two directions for further research can be indicated. The first of them is the compilation of an inventory of table types and their associated properties, encompassing both common and infrequent types, has the potential to advance the WTE. The achievement of this objective would be facilitated by the development of a formal specification for table types. To the best of our knowledge, no existing formats of such specifications currently exist. It is worth noting that

certain domain-specific languages (DSLs) have been designed to automate the transformation of tabular data and provide descriptions of table structures [41, 42, 43, 44, 45, 46, 35]. However, none of these DSLs are suitable for the aforementioned inventory goals. Consequently, it is recommended that ad hoc tools should be developed specifically for the purpose of specifying table types and creating a library of such types.

The second direction is the involvement the structural cells. There exist two distinct approaches for treating structured cells, namely the division of structured cells into atomic ones [40] or the direct handling of atomic data items within structured cells [35]. Regardless of the chosen approach, the development of advanced solutions necessitates the techniques for recognizing atomic data items within structured cells. Both direction for further research warrant greater attention.

Acknowledgments

This work was supported by the Ministry of Science and Higher Education of the Russian Federation, the Project No. 121030500071-2.

References

[1] M. J. Cafarella, A. Halevy, D. Z. Wang, E. Wu, Y. Zhang, WebTables: exploring the power of tables on the web, Proc. VLDB Endowment 1 (2008) 538–549. doi:10.14778/1453856.1453916.

[2] E. Crestan, P. Pantel, Web-scale table census and classification, in: Proc. 4th ACM Int. Conf. on Web Search and Data Mining, 2011, pp. 545–554. doi:10.1145/1935826.1935904.

[3] J. Eberius, K. Braunschweig, M. Hentsch, M. Thiele, A. Ahmadov, W. Lehner, Building the dresden web table corpus: a classification approach, in: Proc. IEEE/ACM 2nd Int. S. on Big Data Computing, 2015, pp. 41–50. doi:10.1109/BDC.2015.30.

[4] O. Lehmberg, D. Ritze, R. Meusel, C. Bizer, A large public corpus of web tables containing time and context metadata, in: Proc. 25th Int. Conf. on World Wide Web, 2016, pp. 75–76. doi:10.1145/2872518.2889386.

[5] L. R. Lautert, M. M. Scheidt, C. F. Dorneles, Web table taxonomy and formalization, ACM SIGMOD Record 42 (2013) 28–33. doi:10.1145/2536669.2536674.

[6] S. Zhang, K. Balog, Web table extraction, retrieval, and augmentation: a survey, ACM Trans. Intell. Syst. Technol. 11 (2020). doi:10.1145/3372117.

[7] J. C. Roldán, P. Jiménez, R. Corchuelo, On extracting data from tables that are encoded using html, Knowledge-Based Systems 190 (2020) 105157. doi:10.1016/j.knosys.2019.105157.

[8] A. Shigarov, Table understanding: Problem overview, WIREs Data Mining and Knowledge Discovery 13 (2023) e1482. doi:10.1002/widm.1482.

[9] D. W. Embley, M. S. Krishnamoorthy, G. Nagy, S. Seth, Converting heterogeneous statistical tables on the web to searchable databases, Int. J. Doc. Anal. Recog. 19 (2016) 119–138. doi:10.1007/s10032-016-0259-1.

[10] Y. Wang, J. Hu, Detecting tables in html documents, in: Proc. 5th Int. W. on Document Analysis Systems V, DAS'02, 2002, p. 249–260.
[11] M. J. Cafarella, A. Halevy, D. Z. Wang, U. C. Berkeley, E. Wu, Uncovering the relational web, in: Proc. 11th Int. W. on Web and Databases, 2008.
[12] J.-W. Son, S.-B. Park, Web table discrimination with composition of rich structural and content information, Appl. Soft Comput. 13 (2013) 47–57. doi:10.1016/j.asoc.2012.07.025.
[13] J. C. Roldán, P. Jiménez, P. Szekely, R. Corchuelo, TOMATE: A heuristic-based approach to extract data from HTML tables, Information Sciences 577 (2021) 49–68. doi:10.1016/j.ins.2021.04.087.
[14] K. Lerman, L. Getoor, S. Minton, C. Knoblock, Using the structure of web sites for automatic segmentation of tables, in: Proc. 2004 ACM SIGMOD Int. Conf. on Management of Data, SIGMOD '04, 2004, pp. 119–130. doi:10.1145/1007568.1007584.
[15] W. Gatterbauer, P. Bohunsky, Table extraction using spatial reasoning on the CSS2 visual box model, in: Proc. National Conference on Artificial Intelligence, volume 2, 2006, pp. 1313–1318.
[16] M. D. Adelfio, H. Samet, Schema extraction for tabular data on the web, Proc. VLDB Endowment 6 (2013) 421–432. doi:10.14778/2536336.2536343.
[17] S.-W. Jung, H.-C. Kwon, A scalable hybrid approach for extracting head components from web tables, IEEE Transactions on Knowledge and Data Engineering 18 (2006) 174–187. doi:10.1109/TKDE.2006.19.
[18] G. Nagy, S. Seth, Table headers: an entrance to the data mine, in: Proc. 23rd Int. Conf. on Pattern Recognition, 2016, pp. 4065–4070. doi:10.1109/ICPR.2016.7900270.
[19] G. Nagy, Learning the characteristics of critical cells from web tables, in: Proc. 21st Int. Conf. on Pattern Recognition, 2012, pp. 1554–1557. URL: https://ieeexplore.ieee.org/document/6460440.
[20] J. Pujara, A. Rajendran, M. Ghasemi-Gol, P. A. Szekely, A common framework for developing table understanding models, in: Proc. ISWC 2019 Satellite Tracks, 2019.
[21] S. Seth, R. Jandhyala, M. Krishnamoorthy, G. Nagy, Analysis and taxonomy of column header categories for web tables, in: Proc. 9th IAPR Int. W. on Document Analysis Systems, 2010, pp. 81–88. doi:10.1145/1815330.1815341.
[22] G. Nagy, S. Seth, D. Jin, D. W. Embley, S. Machado, M. Krishnamoorthy, Data extraction from web tables: the devil is in the details, in: Proc. Int. Conf. on Document Analysis and Recognition, 2011, pp. 242–246. doi:10.1109/ICDAR.2011.57.
[23] X. Wang, Tabular abstraction, editing, and formatting, Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada, 1996.
[24] M. Hurst, Towards a theory of tables, Int. J. Doc. Anal. Recog. 8 (2006) 123–131. doi:10.1007/s10032-006-0016-y.
[25] C. S. Bhagavatula, T. Noraset, D. Downey, Methods for exploring and mining tables on wikipedia, in: Proc. ACM SIGKDD Workshop on Interactive Data Exploration and Analytics, 2013, pp. 18–26. doi:10.1145/2501511.2501516.
[26] E. Muñoz, A. Hogan, A. Mileo, Using linked data to mine rdf from wikipedia's tables, in: Proc. 7th ACM Int. Conf. on Web Search and Data Mining, 2014, pp. 533–542. doi:10.1145/2556195.2556266.

[27] K. Bai, P. Mitra, C. L. Giles, Y. Liu, Automatic extraction of table metadata from digital documents, in: Proc. 6th ACM/IEEE-CS Joint Conf. on Digital Libraries, 2006, pp. 339–340. doi:10.1145/1141753.1141835.

[28] Y. Liu, K. Bai, P. Mitra, C. L. Giles, Tableseer: Automatic table metadata extraction and searching in digital libraries, in: Proc. 7th ACM/IEEE-CS Joint Conf. on Digital Libraries, 2007, pp. 91–100. doi:10.1145/1255175.1255193.

[29] M. Yoshida, K. Matsumoto, K. Kita, Table topic models for hidden unit estimation, in: Information Retrieval Technology, volume 9994 LNCS, 2016, pp. 302–307. doi:10.1007/978-3-319-48051-0_23.

[30] Z. Zhang, Towards Efficient and Effective Semantic Table Interpretation, in: The Semantic Web – ISWC 2014, volume 8796 LNCS, 2014, pp. 487–502. doi:10.1007/978-3-319-11964-9_31.

[31] Z. Zhang, Effective and efficient semantic table interpretation using TableMiner+, Semantic Web 8 (2017) 921–957. doi:10.3233/SW-160242.

[32] D. Burdick, M. Danilevsky, A. V. Evfimievski, Y. Katsis, N. Wang, Table extraction and understanding for scientific and enterprise applications, Proc. VLDB Endowment 13 (2020) 3433–3436. doi:10.14778/3415478.3415563.

[33] S. Bonfitto, E. Casiraghi, M. Mesiti, Table understanding approaches for extracting knowledge from heterogeneous tables, WIREs Data Mining and Knowledge Discovery 11 (2021) e1407. doi:10.1002/widm.1407.

[34] Z. Chen, M. Cafarella, Automatic web spreadsheet data extraction, in: Proc. 3rd Int. W. on Semantic Search Over the Web, 2013, pp. 1:1–1:8. doi:10.1145/2509908.2509909.

[35] A. Shigarov, A. Mikhailov, Rule-based spreadsheet data transformation from arbitrary to relational tables, Inform. Syst. 71 (2017) 123–136. doi:10.1016/j.is.2017.08.004.

[36] Y. Yang, W.-S. Luk, A framework for web table mining, in: Proc. 4th Int. W. on Web Information and Data Management, 2002, pp. 36–42. doi:10.1145/584931.584940.

[37] E. Koci, M. Thiele, O. Romero, W. Lehner, A machine learning approach for layout inference in spreadsheets, in: Proc. Int. Joint Conf. on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2016, pp. 77–88. doi:10.5220/0006052200770088.

[38] E. Koci, M. Thiele, O. Romero, W. Lehner, Table identification and reconstruction in spreadsheets, in: Advanced Information Systems Engineering, volume 10253 LNCS, 2017, pp. 527–541. doi:10.1007/978-3-319-59536-8_33.

[39] E. Koci, M. Thiele, W. Lehner, O. Romero, Table recognition in spreadsheets via a graph representation, in: Proc. 13th IAPR Int. W. on Document Analysis Systems, 2018, pp. 139–144. doi:10.1109/DAS.2018.48.

[40] W. W. Cohen, M. Hurst, L. S. Jensen, A flexible learning system for wrapping tables and lists in html documents, in: Proc. 11th Int. Conf. on World Wide Web, 2002, pp. 232–241. doi:10.1145/511446.511477.

[41] V. Hung, B. Benatallah, R. Saint-Paul, Spreadsheet-based complex data transformation, in: Proc. 20th ACM Int. Conf. on Information and Knowledge Management, 2011. doi:10.1145/2063576.2063829.

[42] S. Gulwani, W. R. Harris, R. Singh, Spreadsheet data manipulation using examples, Commun. ACM 55 (2012). doi:10.1145/2240236.2240260.

[43] D. W. Barowy, S. Gulwani, T. Hart, B. Zorn, FlashRelate: extracting relational data from

semi-structured spreadsheets using examples, in: Proc. 36th ACM SIGPLAN Conf. on Programming Language Design and Implementation, 2015, pp. 218–228. doi:10.1145/2737924.2737952.

[44] A. Shigarov, Table understanding using a rule engine, Expert Syst. Appl. 42 (2015). doi:10.1016/j.eswa.2014.08.045.

[45] R. Singh, S. Gulwani, Transforming spreadsheet data types using examples, SIGPLAN Not. 51 (2016) 343–356. doi:10.1145/2914770.2837668.

[46] Z. Jin, M. R. Anderson, M. Cafarella, H. V. Jagadish, Foofah: transforming data by example, in: Proc. ACM SIGMOD Int. Conf. on Management of Data, 2017, pp. 683–698. doi:10.1145/3035918.3064034.

Модель минимизации дефицита мощности электроэнергетических систем с учетом баланса активной и реактивной мощности

Дмитрий Якубовский [1], Дмитрий Крупенёв [1] и Денис Бояркин [1]

[1] *Федеральное государственное бюджетное учреждение науки Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук, ул. Лермонтова, д. 130., г. Иркутск, 664033, Россия*

Abstract

Работа освещает разрабатываемую модель минимизации дефицита активной мощности с учетом баланса активной и реактивной мощности. Мотивацией к формированию и тестированию данной модели является необходимость повышения адекватности моделей данного класса, а также необходимость ее внедрения в разрабатываемую методику оценки плановой надежности. В статье обозначены требования к модели и процесс ее формирования в основе которого лежат модели минимизации дефицита активной мощности и модели формата optimal power flow. При формировании модели был обозначен необходимый набор параметров, а именно - модули напряжений на шинах, углов фаз напряжений, активного и реактивного сопротивления, а также сформированы балансовые ограничения активной и реактивной мощности. В результате анализа экспериментов с использованием схем с 3-мя узлами выявлена разница в результате минимума дефицита мощности (от 3,05 до 4,46%) по сравнению с задачей на основании модели минимизации дефицита активной мощности без учета реактивной мощности в балансовых ограничениях.

Keywords

Минимизация дефицита мощности, плановая надежность, оптимальное потокораспределение, математическая модель, активная и реактивная мощность

1. Введение

В современных условиях работы электроэнергетических систем, их постоянном укрупнении и усложнении эффективное планирование их развития является ключевым фактором для обеспечения надежности и стабильности поставки электроэнергии потребителям. Одной из базовых операций необходимых для проведения качественного планирования развития является оценка балансовой надежности ЭЭС [1 - 3]. В свою очередь, методика оценки, в основе которой лежит метод статистических испытаний (Монте-Карло) включает в себя несколько этапов – формирование случайных состояний системы на конкретный период, минимизации дефицитов мощности этих состояний и расчет показателей надежности в качестве основного результата.

Важной частью методики является модель минимизации дефицита мощности, так как качество, корректность и сложность модели напрямую влияют на результаты расчетов. На сегодняшний день представлено множество различных программных комплексов [4 - 19] для оценки балансовой надежности: ANTARES, Transmission Reliability Evaluation of Large-Scale Systems (TRELSS), Siemens PTI PSS/E TPLAN, CORAL, Grid Reliability and Adequacy Risk Evaluator (GRARE), PLEXOS, Multi-Area Reliability Simulation (MARS), GridView, NARP,

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: yakubovskii.dmit@mail.ru (A. 1); krupenev@isem.irk.ru (A. 2); boyarkin_denis@mail.ru (A. 3)
ORCID: 0000-0001-8331-6200 (A. 1); 0000-0002-3093-4483 (A. 2); 0000-0002-7048-2848 (A. 3)

ICCS-DE 2023 Workshop Proceedings

MARELI; а также ОРИОН-М, ПОТОК-3, ЯНТАРЬ и «Надежность». В большей части комплексов расчеты проводятся с использованием потоковых, линейных или линеаризованных моделей. В части комплексов реализованы модели минимизации затрат, что отличается от оригинальной задачи минимизации дефицита мощности и в свою очередь ставит под сомнение полученные решения в рамках оценки балансовой надежности.

Существующие модели минимизации дефицита мощности ЭЭС с линейными потерями имеют допущения, в рамках которых не полностью учитываются потери при перетоках мощности. Данная проблема была учтена в [3] и разработана новая модель, где потери мощности зависят от квадрата передаваемой мощности и является близкой по физическому смыслу к реальной работе ЭЭС. Однако такие модели полноценно не учитывают реактивное сопротивление, разные уровни напряжений и углы фаз что позволяет проводить расчеты для укрупненных схем, но с допущениями. Тем не менее, такой подход и другие упрощения могут привести к искажению результатов работы системы и их несоответствию реальным условиям, в том числе к накоплению погрешностей и ошибок при многократных расчетах различных состояний.

Несмотря на эти допущения, существующая методика оценки балансовой надежности позволяет проводить базовые расчеты и анализировать ЭЭС в допустимом формате. Однако более комплексные результаты могут быть получены в рамках оценки плановой надежности с учетом случайных процессов работы ЭЭС, где также необходимо проводить расчеты с использованием моделей минимизации дефицита мощности. В виду необходимости повышения адекватности расчетов минимизации дефицита мощности, а также повышению уровня детализации и соответствия реальной физике процесса существует потребность в разработке моделей с учетом активной и реактивной мощности. В рамках данных исследований был проведен анализ существующих моделей расчета установившегося режима и моделей минимизации затрат с учетом реактивной мощности [20-23]. В результате исследований была разработана и представлена в данной работе модель минимизации дефицита мощности с учетом активной и реактивной мощности для оценки плановой надежности. Также был проведен сравнительный эксперимент и первичный анализ ее работы относительно существующей модели с нелинейной постановкой для оценки балансовой надежности.

2. Постановка задачи

Определение минимума дефицита мощности входит в основу методики оценки плановой надежности ЭЭС методом Монте-Карло, где каждый сформированный случайный процесс работы ЭЭС имитирует её работу в соответствии с используемой моделью минимизации дефицита мощности (МДМ). Развитие подобных моделей МДМ при оценке балансовой надёжности происходило поэтапно в процессе их использования и по результатам обнаруженных проблем, таких как – физически неверное распределение перетоков мощности, наличие кольцевых перетоков, необходимость учета контролируемых сечений, активной и реактивной мощности. На сегодняшний день были разработаны и усовершенствованы несколько моделей МДМ [24-25], которые отличаются учетом тех или иных особенностей функционирования ЭЭС и имеют разную степень адекватности и соответствия физическим процессам в процессе определения дефицита мощности. Пользуясь данным опытом появляется возможность эффективной разработки новой модели МДМ с учетом специфики оценки плановой надежности.

В основе методики оценки плановой надежности лежит формирование задачи, исходя из модели минимизации дефицита мощности с использованием данных (ограничений) полученных из блока генерации конкретных случайных состояний системы с последующей их оптимизаций, т.е. поиском минимума дефицита мощности. В математических моделях в качестве параметров должны быть обозначены генераторные мощности и нагрузки всех зон надежности/узлов, пропускные способности связей, информация о потерях на перетоках мощности. В общем виде решается задача распределения мощности по всей системе для конкретного состояния таким образом, чтобы достичь минимума дефицита мощности, обеспечить максимум покрытой

нагрузки с учетом балансов мощности в зонах надежности/узлах и адекватным, физически корректным потокораспределением.

В настоящий момент в качестве модели минимизации дефицита мощности для оценки балансовой надежности используется модель с усовершенствованной постановкой в нелинейном виде, без учета реактивной мощности и контролируемых сечений, которая формулируется следующим образом: «*Для известных значений работоспособных генераторных мощностей, требуемых уровней нагрузок потребителей, пропускных способностей связей ЭЭС и коэффициентов потерь мощности в связях ЭЭС необходимо определить оптимальное потокораспределение в ЭЭС*».

Математически, проблема представляется так:

$$\min_{y,x,z} \sum_{i=1}^{n} (\bar{y}_i - y_i), \tag{1}$$

при соблюдении нелинейных балансовых ограничений:

$$x_i - y_i + \sum_{j=1}^{n} (1 - a_{ji} z_{ji}) z_{ji} - \sum_{j=1}^{n} z_{ij} = 0 \,, i = 1, \dots, n \,. \tag{2}$$

а также ограничений на оптимизируемые переменные:

$$0 \le y_i \le \bar{y}_i, i = 1, \dots, n, \,, \tag{3}$$

$$0 \le x_i \le \bar{x}_i, i = 1, \dots, n, \,, \tag{4}$$

$$0 \le z_{ij} \le \bar{z}_{ij}, i = 1, \dots, n, j = 1, \dots, n, i \ne j \,, \tag{5}$$

$$z_{ji} * z_{ij} = 0, i = 1, \dots, n, j = 1, \dots, n, \tag{6}$$

где: x_i - используемая мощность (МВт) в узле i, $\bar{x}_i$ – располагаемая генерирующая мощность (МВт) в узле i, y_i – покрываемая в узле i нагрузка (МВт), $\bar{y}_i$ – величина нагрузки в узле i (МВт), z_{ij} – поток мощности из узла i в узел j (МВт), $\bar{z}_{ij}$ – пропускная способность ЛЭП между узлами i и j (МВт), z_{ji} – поток мощности из узла j в узел i (МВт), $\bar{z}_{ji}$ – пропускная способность ЛЭП между узлами j и i (МВт) a_{ji} – заданные положительные коэффициенты удельных потерь мощности при ее передаче из узла j в узел i, $j \ne i$, $i = 1, \dots, n$, $j = 1, \dots, n$.

Коэффициенты удельных потерь мощности в ЛЭП при её передаче были определены в соответствии с [3] и рассчитываются следующим образом:

$$a_{ji} = \frac{r_{ji}}{U_{\text{ном}}^2 cos\varphi_{ji}^2}, j = 1, \dots J, i = 1, \dots I, i \ne j, \tag{7}$$

где: r_{ji} – активное сопротивление линии электропередачи между зонами надёжности j и i, Ом/км; $U_{\text{ном}}$ – номинальное напряжение линии электропередачи между зонами надёжности j и i, кВ; $cos\varphi_{ji}$ – усреднённый коэффициент мощности межзонных связей между зонами надёжности j и i (обычно принимается равным 0,9).

Представленная модель МДМ (1) - (6) основана на моделях потокораспределения в области оценки балансовой надежности и соответствует стандартной постановке транспортной задачи. Основная часть модели формируется за счет балансовых ограничений равенств, которые регулируют моделирование распределения активной мощности при перетоках. Также для корректного моделирования встречных перетоков мощности используется дополнительное ограничение (6), это позволяет внести однозначность в направление перетока мощности в каждом режиме работы системы.

Как показано в формуле (7) коэффициенты удельных потерь определяются только с учетом активного сопротивления и усредненным коэффициентом мощности межзонных связей, что несомненно упрощает расчеты. Однако модели такого плана не учитывают реактивное сопротивление, разные уровни напряжений и углы фаз что позволяет проводить расчеты для укрупненных схем с допущениями. Тем не менее, такой подход и другие упрощения могут привести к искажению результатов работы системы и их несоответствию реальным условиям, в том числе к накоплению погрешностей и ошибок при многократных расчетах различных состояний.

Постоянный процесс усложнения и укрупнения электроэнергетических систем, необходимость учета их особенностей и переход к более подробным схемам для корректного решения задачи оценки балансовой надежности ЭЭС, а также задачи планирования развития ЭЭС, требует обновления существующих моделей и подходов с учетом дополнительных параметров систем и низкоуровневых расчетов. В первую очередь повышение физической адекватности моделей должно быть проведено за счет добавления баланса активной и реактивной мощностей, сопротивлений и потерь на перетоках мощности в существующую модель (1) - (6). При этом целевая функция все также должна быть использована для расчета дефицитов активной мощности. Основные коррекции существующей модели должны коснуться балансовых ограничений, а именно расчетов мощности и потерь при перетоках с учетом возможной разницы напряжений, углов его фазы и сопротивлений.

Таким образом усовершенствованная модель должна быть сформирована в рамках моделей подобных Optimal Power Flow применяемых в области энергетики при различных расчетах. Разрабатываемая модель предполагает в качестве целевой функции обозначить минимум дефицита активной мощности:

$$\min \sum_{i=1}^{n} (P_{Di}^{max} - P_{Di}), \tag{8}$$

при выполнении балансовых ограничений с учетом активной и реактивной мощности на шине:

$$\sum_{g \epsilon G} P_i^g = \sum_{d \epsilon D} P_i^d + \sum_{j=1}^{n} P_{ij}^L, \tag{9}$$

$$\sum_{g \epsilon G} Q_i^g = \sum_{d \epsilon D} Q_i^d + \sum_{j=1}^{n} Q_{ij}^L, \tag{10}$$

вспомогательном расчете мощности при перетоках:

$$P_{ij}^L = \frac{1}{r_{ij}^2 + x_{ij}^2} \left[r_{ij} \left(v_j^2 - v_i v_j \cos(\theta_i - \theta_j) \right) + x_{ij} \left(v_i v_j \sin(\theta_i - \theta_j) \right) \right], \tag{11}$$

$$Q_{ij}^L = \frac{1}{r_{ij}^2 + x_{ij}^2} \left[x_{ij} \left(v_j^2 - v_i v_j \cos(\theta_i - \theta_j) \right) + r_{ij} \left(v_i v_j \sin(\theta_i - \theta_j) \right) \right], \tag{12}$$

и ограничений на переменные:

$$P_i^{D,min} \le P_i^D \le P_i^{D,max}, \tag{13}$$

$$P_i^{G,min} \le P_i^G \le P_i^{G,max}, \tag{14}$$

$$Q_i^{D,min} \le Q_i^D \le Q_i^{D,max}, \tag{15}$$

$$Q_i^{G,min} \le Q_i^G \le Q_i^{G,max}, \tag{16}$$

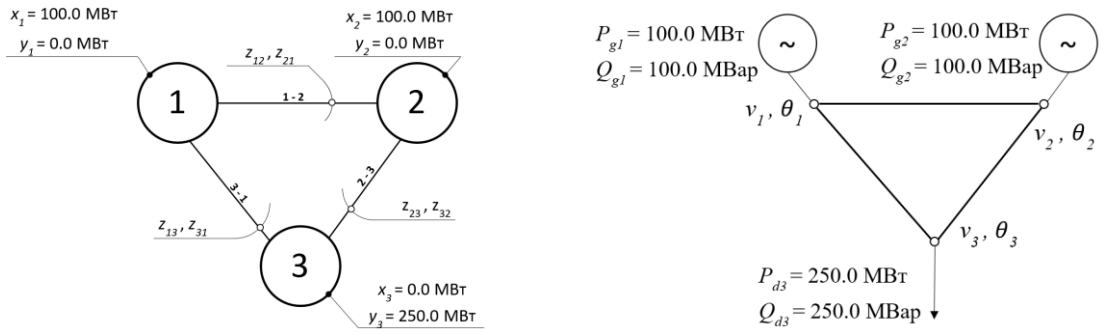
$$\theta_i^{min} \le \theta_i \le \theta_i^{max}, \tag{17}$$

$$v_i^{min} \le v_i \le v_i^{max}, \tag{18}$$

где D – набор потребителей на шине, нагрузка; G – набор генераторов на шине, генерация; L – набор связей; i, j – шины, узлы; P_i^G – суммарная активная мощность генераторов на шине i; P_i^D – суммарная активная мощность нагрузки на шине i; Q_i^G – суммарная реактивная мощность генераторов на шине i; Q_i^D – суммарная реактивная мощность нагрузки на шине i; θ – угол фазы напряжения на шине; v_i – модуль напряжения на шине i; $r_{i,j}$ - активное сопротивление линии между узлами/шинами i и j; $x_{i,j}$ – реактивное сопротивление линии между узлами/шинами i и j. Таким образом данная модель должна будет учитывать глубокую специфику различных ЭЭС, и в свою очередь потребует корректные наборы данных.

3. Экспериментальная часть

Экспериментальные расчеты проводились для систем различной конфигурации, с разными начальными параметрами. Проверка работоспособности модели, учитывающей баланс активной и реактивной мощности (8) - (18) и сравнение полученных результатов с моделью (1) - (6) проводилось на системе с 3 узлами. Однако в экспериментах использовалось несколько тестовых состояний системы, основные результаты были вынесены в данную работу на примерах двух состояний, где первое состояние показано на рисунке 1.



а) **б)**

Рисунок 1: Тестовая система №1 (а) – с 3 узлами для модели (1)-(6), (б) – с 3 узлами для модели (8)-(18)

На рисунке 1.а обозначена схема системы, используемая в паре с моделью (1) – (6), в данном случае считается, что каждый узел совмещает в себе располагаемую генерирующую мощность (x) и величину нагрузки в узле (y), а все потери внутри узла равны нулю, узлы также соединены связями каждая из них представлена перетоками (z) в прямом и обратном направлении. На рисунке 1.б представлена схема системы с двумя генераторами примыкающих к 1 и 2 узлам соответственно и одной нагрузкой на третьем узле, все узлы соединенными тремя шинами.

В результате вычислений были получены значения целевой функции и оптимизируемых переменных для задач, основанных на двух моделях. Как видно из таблицы 1 в результате оптимизации задачи по модели (1) – (6) дефицит составил 51,842 МВт, и были задействованы два перетока мощности z23 и z13. Такое потокораспределение мощности - корректно, так как предполагалось что за счет профицитных узлов 1 и 2 и связей с пропускной способностью равной 100 МВт будет обеспечена нагрузка в 3 – дефицитном узле. С другой стороны, задача, основанная на модели (8) – (18) показала минимум дефицита мощности на уровне 49,519 МВт что составляет 2,323 МВт разницы, при этом загружаются все три связи.

Таблица 1

Исходные параметры и результаты вычислений с использованием моделей (1) - (6) и (8) - (18) для тестовых систем №1 и №2

Параметры		Ограничения №1		Решение №1		Ограничения №2		Решение №2	
obj		0	+ INF	49.51	51.84	0	+ INF	19.80	20.42
Pg1	x1	0.00	100.00	100.00	100.00	0.00	100.00	100.00	100.00
Pg2	x2	0.00	100.00	100.00	100.00	0.00	100.00	100.00	100.00
Pg3	x3	0.00	0.00	0.00	0.00	0.00	100.00	100.00	100.00
Pd1	y1	0.00	0.00	0.00	0.00	0.00	50.00	30.19	50.00
Pd2	y2	0.00	0.00	0.00	0.00	0.00	20.00	20.00	20.00
Pd3	y3	0.00	250.00	200.48	198.15	0.00	250.00	250.00	229.57
PL12	z12	0.00	100.00	-23.00	0.00	0.00	100.00	-13.65	9.84
PL21	z21	0.00	100.00	23.02	0.00	0.00	100.00	13.66	0.00
PL13	z13	0.00	100.00	-76.99	100.00	0.00	100.00	-56.14	40.15
PL31	z31	0.00	100.00	77.18	0.00	0.00	100.00	56.22	0.00
PL23	z23	0.00	100.00	-123.02	100.00	0.00	100.00	-93.66	89.78
PL32	z32	0.00	100.00	123.29	0.00	0.00	100.00	93.77	0.00

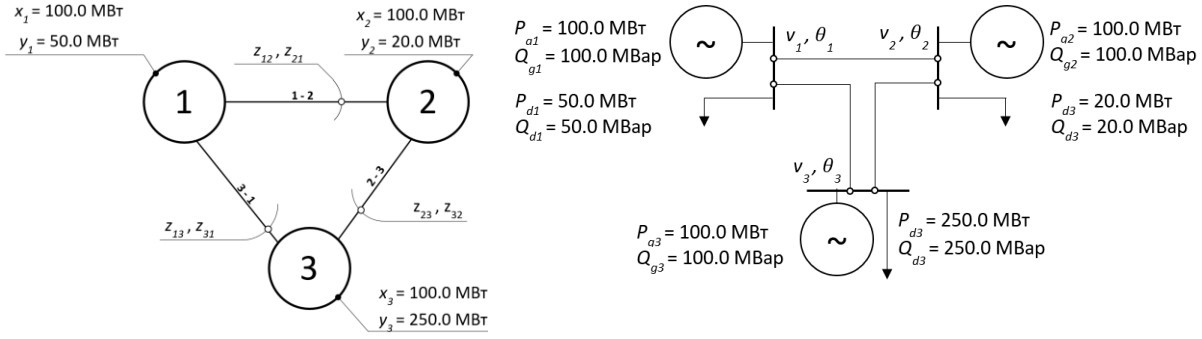
Данные результаты показывают работоспособность разрабатываемой модели (8) – (18) и разницу уровня минимума дефицита мощности в 4.48% для представленных на рисунке 1 подобных схем. Однако, существенным дополнением (8) – (18) является возможность учета реактивной мощности при расчетах, что влияет на потокораспределение, а анализ полученных данных для разрабатываемой модели показал загрузку дополнительных линий. Данная особенность результата могла считаться некорректным потокораспределением для задач модели (1) – (6), т.к. при аналитическом решении и корректном потокраспределении, передача мощности сквозь дополнительный узел не допускалась. В случае с моделью (8) - (18) наличие параметров уровня напряжения, углов фаз, активного и реактивного сопротивления, может влиять на потокораспределение подобным образом. Результаты оптимизации дополнительных параметров и реактивной мощности обозначены в таблице 2.

Таблица 2

Дополнительные исходные параметры и результаты вычислений с использованием модели (8) - (18) для тестовых систем №1 и №2

Параметры	Ограничения №1		Решение №1	Ограничения №2		Решение №2
Qg1	0.00	100.00	100.00	0.00	100.00	0.00
Qg2	0.00	100.00	100.00	0.00	100.00	0.00
Qg3	0.00	0.00	0.00	0.00	100.00	77.77
Qd1	0.00	0.00	0.00	0.00	50.00	50.00
Qd2	0.00	0.00	0.00	0.00	20.00	20.00
Qd3	0.00	250.00	180.89	0.00	250.00	0.00
QL12	-INF	+INF	-22.68	-INF	+INF	18.57
QL21	-INF	+INF	22.16	-INF	+INF	-18.84
QL13	-INF	+INF	-77.31	-INF	+INF	31.42
QL31	-INF	+INF	69.10	-INF	+INF	-34.77
QL23	-INF	+INF	-122.16	-INF	+INF	38.84
QL32	-INF	+INF	111.79	-INF	+INF	-43.00
O1	-1.00	1.00	0.94	-1.00	1.00	0.95
O2	-1.00	1.00	0.95	-1.00	1.00	0.96
O3	-1.00	1.00	1.00	-1.00	1.00	1.00
v1	198.00	242.00	198.00	198.00	242.00	198.00
v2	193.50	236.50	200.19	193.50	236.50	196.06
v3	184.50	225.50	208.53	184.50	225.50	192.71

Еще один эксперимент был проведен с использованием других данных, представленных на рисунке 2, где каждому из узлов/шин были определены как нагрузка, так и генерация, что позволило протестировать модель на правильность потокораспределения в условиях наличия собственной генерации на узле/шине и возможности снижения дефицита в/на соседних узлах/шинах.



а) **б)**

Рисунок 2: Тестовая система №2 (а) – с 3 узлами для модели (1)-(6), (б) – с 3 шинами для модели (8)-(18)

По результатам оптимизации тестовой системы №2 (табл.1) были получены значения 20,428 МВт и 19,804 МВт для задач (1) – (6) и (8) – (18) моделей соответственно, где разница составляет 0,624 МВт или 3.05%. В процессе оптимизации задач по обеим моделям для минимизации дефицита в 3 узле были задействованы перетоки 3-1 и 2-3, однако также был использован переток 1-2, в чем нет прямой необходимости, так как генерации хватало для обеспечения внутриузловой нагрузки. Данная проблема модели (1) – (6) ранее решалась с помощью фиксации оптимизированного уровня покрытой нагрузки и дополнительной оптимизации по сумме нормы перетоков мощности. Также было обнаружено, что при решении задачи по модели (8) – (18) покрывается только 30,196 МВт из 50 МВт нагрузки. В дальнейшем планируется провести дополнительный анализ работы модели (8) – (18), а также внедрить необходимые уточнения и улучшения для повышения её точности и адекватности.

4. Заключение

В работе рассмотрена задача разработки модели минимизации дефицита активной мощности с учетом баланса активной и реактивной мощности. Разработанная модель сравнивалась с нелинейной (квадратичная) моделью минимизации дефицита активной мощности в электроэнергетических системах, но с учетом только баланса активной мощности. Разработанная модель учитывает несколько дополнительных параметров по сравнению с имеющейся, а именно - модули напряжений на шинах, углы фаз напряжений, активное и реактивное сопротивления. При разработке модели использован подход к формированию балансовых уравнений минимизации дефицита активной мощности, соответствующих моделям Optimal Power Flow. Для проведения экспериментальных исследований использовалась схема с 3-мя узлами. Для каждой системы проводилось несколько испытаний. По результатам проведенных экспериментов выявлено, что результаты решения задач на основании разработанной модели отличаются от результатов решения задач, на основе модели применяемой при оценке балансовой надёжности на 3,05-4,46%. Также выявлено различное потокораспределение в виде дополнительных перетоков мощности. Такое поведение объясняется наличием учета реактивной мощности и зависимости от дополнительных параметров уровня напряжения, углов фаз и сопротивлений. В дальнейшем будут проведены испытания на схемах большей размерности, где, как ожидается, разница будет более существенной.

5. Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 23-29-00435.

6. References

[1] Ковалев Г.Ф., Лебедева Л.М. Модель оценки надежности электроэнергетических систем при долгосрочном планировании их работы // Электричество. — 2000. — ¹ 11. — С. 17–24.
[2] Г.Ф. Ковалев. Надежность систем электроэнергетики / Г.Ф. Ковалев, Л.М. Лебедева; отв. ред. Н.И. Воропай. - Новосибирск: Наука, 2015. - 224 с.
[3] Ковалев Г.Ф., Лебедева Л.М. Комплекс моделей оптимизации режимов расчетных состояний при оценке надежности электроэнергетических систем. Иркутск: ИСЭМ СО РАН, 2000, 74 с.
[4] Review of the current status of tools and techniques for risk-based and probabilistic planning in power systems, Working Group 601 of Study Committee C4 //International Conference on Large High Voltage Electric Systems. March. 2010
[5] Fernandez Blanco Carramolino, R., Careri, F., Kavvadias, K., Hidalgo Gonzalez, I., Zucker, A. and Peteves, E., Systematic mapping of power system models: Expert survey, EUR 28875 EN,

Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-76462-2, doi:10.2760/422399, JRC109123.
[6] Antonopoulos G; Chondrogiannis S; Kanellopoulos K; Papaioannou I; Spisto A; Efthimiadis T; Fulli G., Assessment of underlying capacity mechanism studies for Greece, EUR 28611 EN, Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-68878-2, doi: 10.2760/51331, JRC106307
[7] RTE Antares, Antares Optimization problems formulation, [Электронный ресурс] URL: https://antares.rte-france.com, доступ: 09.05.2023
[8] A. Gaikwad, S. Agarwal, K. Carden, N. Wintermantel, S. Meliopoulos, M. Kumbale, A Study on Probabilistic Risk Assessment for Transmission and Other Resource Planning, Electric Power Research Institute For EISPC and NARUC (NARUC-2013-RFP027-DE0316), 2015.
[9] URREGO AGUDELO Lilliam, A novel method for the Approximation of risk of Blackout in operational conditions, Laboratoire Image- Signaux et Systèmes Intelligents / LISSI - EA 3956 (laboratoire) , 2016
[10] Milorad Papic, Survey of Tools for Risk Assessment of Cascading Outages, IEEE GM, 2011
[11] Ying-Yi Hong, Lun-Hui Lee Reliability assessment of generation and transmission systems using fault-tree analysis, Energy Conversion and Management 50 (2009) 2810–2817.
[12] Siemens AG and Siemens Industry, Inc., Model Management Module for PSS®E, 2020
[13] PSR – Energy Consulting and Analytics, OPTGEN User Manual, Version 7.4, 2019
[14] ENTSO-E, Mid-term Adequacy Forecast 2018, Appendix 1: Methodology and Detailed Results, 2018
[15] PLEXOS Market Simulation Software [Электронный ресурс], URL: https://energyexemplar.com/solutions/plexos/ доступ: 10.06.2023
[16] Kelvin Chu, MARS Multi-Area Reliability Simulation, EOP – On Demand Feature, General Electric Company, 2014
[17] Panida Jirutitijaroen, Chanan Singh, Reliability and Cost trade-off in Multi-Area Power System Generation Expansion Using Dynamic Programming and Global Decomposition, IEEE Transactions on power systems, Vol 21,No 3, August 2006
[18] J.McCalley, Module PE.PAS.U21.5 Multiarea reliability analysis // Electrical & Computer Engineering, Iowa State University, USA, – 81p.
[19] Simulate security-constrained unit commitment and economic dispatch in large-scale transmission networks // ABB GridView, 2016.
[20] Pandapower [Электронный ресурс], URL: https://pandapower.readthedocs.io/en/v1.4.3/powerflow/opf.html доступ: 04.05.2023
[21] Kim, Seong-Cheol & Salkuti, Surender Reddy. (2019). Optimal power flow based congestion management using enhanced genetic algorithms. International Journal of Electrical and Computer Engineering (IJECE). 9. 875. 10.11591/ijece.v9i2.pp875-883.
[22] Nusair, K.; Alasali, F. Optimal Power Flow Management System for a Power Network with Stochastic Renewable Energy Resources Using Golden Ratio Optimization Method. Energies 2020, 13, 3671. https://doi.org/10.3390/en13143671
[23] Telyatnik, Andrey & Vaskovskaya, Tatiana. (2019). Ускорение метода последовательного квадратичного программирования в задаче оптимизации установившихся режимов ЭЭС. Известия Российской академии наук. Энергетика. 3-15. 10.1134/S0002331019040149.
[24] Якубовский Д.В. «Анализ моделей минимизации дефицита мощности при оценке балансовой надежности электроэнергетических систем» // Системные исследования в энергетике/ Труды молодых ученых ИСЭМ СО РАН, Вып. 48. –Иркутск: ИСЭМ СО РАН, 2017. –131с.
[25] Якубовский Д.В., Крупенёв Д.С., Бояркин Д.А. Модель минимизации дефицита мощности электроэнергетических систем с учетом ограничений по контролируемым сечениям // Системы анализа и обработки данных. – 2021. – № 2 (82). – С. 95–120. – DOI: 10.17212/2782-2001-2021-2-95-120.

Анализ российского рынка IT-услуг

Михаил Баландин [1], Ольга Башарина [1]

[1] *Уральский государственный экономический университет», 8 марта 62, Екатеринбург, 620144, Россия*

Аннотация

В статье представлено исследование IT-рынка и, в частности, рынка IT-услуг за период 2011-2022 гг. Для проведения исследования использовались различные аналитические материалы. Подготовка, анализ и визуализация данных были выполнены при помощи инструмента бизнес-аналитики – MS Power BI. В результате исследования было определено, что российский рынок IT-услуг имеет положительную динамику на всём интервале 2011-2022 гг.

Ключевые слова

Рынок IT, рынок IT-услуг, анализ данных, визуализация данных, Power BI.

1. Введение

Социальные, экономические и геополитические события последних лет оказали существенное влияние на все сферы человеческой деятельности, в том числе внесли значительные изменения на рынок информационных технологий. Уход зарубежных IT-компаний с российского рынка, отток квалифицированных кадров, отказ иностранных компаний от привлечения российских разработчиков и многое другое – всё это повлекло за собой структурные изменения в одной из самых динамично развивающихся отраслей экономики России – IT-отрасли.

2. Методология и результаты исследования

Для проведения анализа необходим соответствующий набор данных – датасет. В открытом доступе нет готового датасета по объёмам рынка IT и IT-услуг, поэтому необходимо было агрегировать данные из доступных источников. В качестве таких источников использовалась статистические и аналитические материалы крупнейших агентств, специализирующихся в том числе на анализе российского рынка IT: TAdviser и Cnews **Ошибка! Источник ссылки не найден.**, 2]. Также были задействованы следующие ресурсы: Forbes, ИКС медиа и Рейтинг Рунета для получения информации о рынках облачных услуг, дата-центрах и веб-разработке. В результате была сформирована полноценная база данных для проведения аналитического исследования.

Для подготовки, анализа и визуализации данных использовался инструмент бизнес-аналитики – Microsoft Power BI. В таблице 1 представлены объемы IT-рынка в целом и выделен сектор IT-услуг за период 2011–2022 гг.

Анализ объёмов IT-рынка, показывает, что он пережил кризис в 2013-2014 годах и смог оправиться лишь к 2021 году. Уже в 2022 году мы вновь наблюдаем спад объёмов IT-рынка до 19,2 млрд. $ по сравнению с 31,2 млрд. $ в 2021 году (рисунок 1а). Как и в прошлый раз, текущий спад рынка вызван общеэкономическими и политическими факторами.

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: mikhailbalandin10a@gmail.com (A. 1); basharinaolga@mail.com (A. 2)
ORCID: 0000-0002-7151-782X (A. 2)

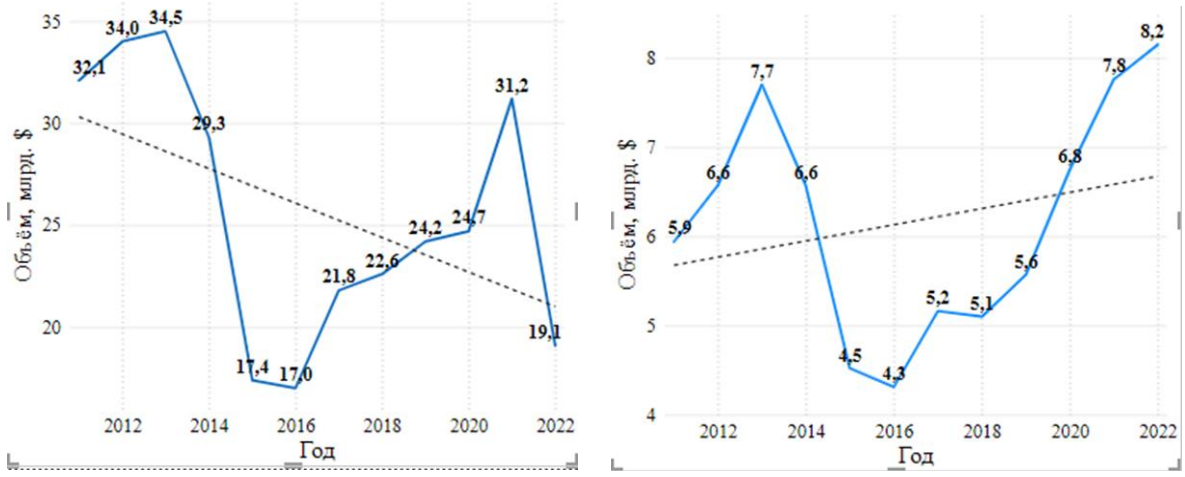
ICCS-DE 2023 Workshop Proceedings

Объём рынка IT-услуг так же пережил кризис в 2013–2014 гг. и смог выйти на докризисный уровень лишь к 2021 году (рисунок 1b). Однако 2022 год не стал для рынка IT-услуг кризисным: объём рынка в 2022 году вырос по сравнению с предыдущим периодом и составил 8,2 млрд. долларов.

Таблица 1

Динамика объемов рынков IT и IT-услуг

Год	IT-рынок		Рынок IT-услуг	
	Объём, млрд. $	Динамика, %	Объём, млрд. $	Динамика, %
2022	19,1	-39%	8,15	5%
2021	31,2	27%	7,76	15,0%
2020	24,7	2%	6,75	21,2%
2019	24,2	7%	5,57	9,2%
2018	22,6	4%	5,10	-1,2%
2017	21,8	28%	5,16	19,7%
2016	17,0	-2%	4,31	-4,6%
2015	17,4	-41%	4,52	-31,2%
2014	29,3	-15%	6,57	-14,7%
2013	34,5	1%	7,70	17,2%
2012	34,0	6%	6,57	10,6%
2011	32,1	15%	5,94	26,4%

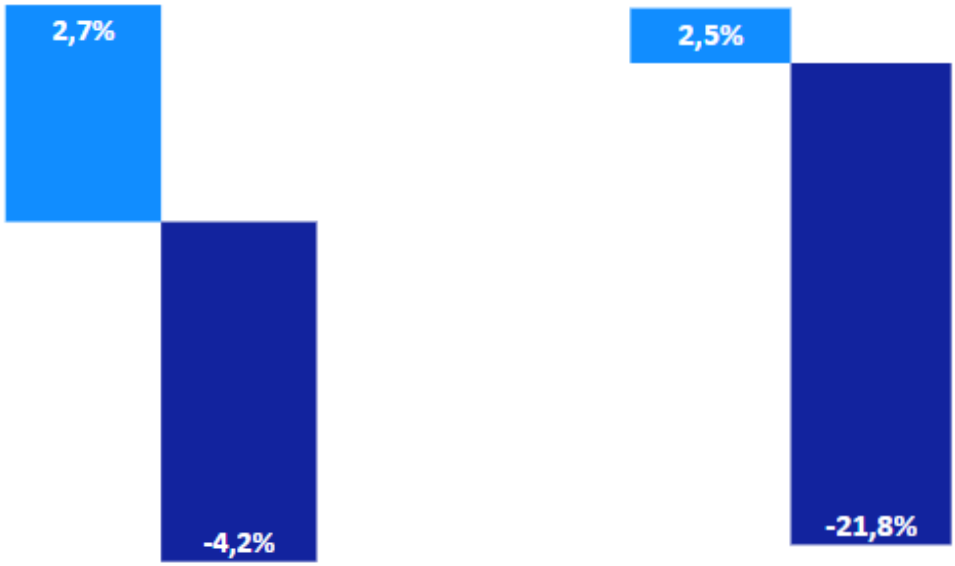


(a) (b)

Рисунок 1: Динамика объёмов рынка IT (a) и рынка IT-услуг (b)

Для оценки роста исследуемых рынков были рассчитаны значения совокупного среднегодового темпа роста CAGR (Compound Annual Growth Rate). Главное достоинство CAGR – он дает простую оценку в виде усредненного процента роста, и, соответственно может быть использован для быстрого анализа прошедшего периода и получения первого прогнозного приближения. Чаще всего этот расчетный показатель используется для работы с объектами, поведение которых выражается сложными зависимостями [3, 4].

В рамках проведенного исследования были рассчитаны значения CAGR для IT-рынка и рынка IT-услуг за два периода: 2011–2022 гг. и 2021–2022 гг. Полученные значения подтверждают положительную динамику рынка IT-услуг (рисунок 2).



(a) (b)

Рисунок 2: Совокупный среднегодовой темп роста за период 2011–2022 гг. (a) и период 2021–2022 гг. (b). ● Рынок IT-услуг, ● Рынок IT.

3. Заключение

Результаты проведенного исследования позволяют подтвердить мнение, что российский IT-рынок действительно подвергся значительному влиянию общего экономического кризиса и находится в процессе структурных изменений. Многие компании России оказались в ситуации потери старых налаженных экономических связей и поставщиков, оттока квалифицированных кадров и пр. Всё это в полной мере отразилось и на IT-отрасли России.

Рынок IT-услуг подвергся экономическому кризису в значительно меньшей мере, нежели рынок IT. Не зря всё большее число компаний обращают своё внимание именно на рынок IT-услуг, занимая освободившиеся или просто перспективные ниши в следующих направлениях: заказной разработке; технической поддержке; внедрении CRM и ERP-систем; облачных услугах; тестировании; веб-разработке и прочее.

Стабильность данного сегмента IT-рынка вызвана рыночной конъюнктурой (повышенным спросом на IT-услуги); трендом на импортозамещение иностранного ПО (как следствие, рост сегмента заказной аутсорс разработки); государственной поддержкой (гранты на системную интеграцию, разработку ПО, льготы для IT-компаний).

4. Список использованных источников

[1] TAdviser – Портал выбора технологий и поставщиков. URL: https://www.tadviser.ru
[2] CNews – Интернет-издание о высоких технологиях. URL: https://www.cnews.ru/
[3] Д. С. Козлова, В. А. Быков, И.Н. Якшилов, Разработка модели оценки эффективности финансово-хозяйственной деятельности организации на основе сценарного подхода, 2022. URL: https://cyberleninka.ru/article/n/razrabotka-modeli-otsenki-effektivnosti-finansovo-hozyaystvennoy-deyatelnosti-organizatsii-na-osnove-stsenarnogo-podhoda
[4] С. В. Черемушкин, Методология расчёта совокупной акционерной доходности Экономический анализ, 2008. – URL: https://cyberleninka.ru/article/n/metodologiya-rascheta-sovokupnoy-aktsionernoy-dohodnosti

UAV Mathematical Modeling Using Differential Equations

Alexandr Vasichenko [1]

[1] *IDSTU 134, Lermontova, Irkutsk, 664033, Russia*

Abstract

In connection with the growth of scientific and applied interest in the subject of the use of unmanned aerial vehicles (UAVs) of a helicopter type, there is an increasing need to obtain and use universal mathematical models that describe the movement of UAVs with varying degrees of completeness of these descriptions. The purpose of this article is to briefly review recent results in this area. First, the description of the UAV motion with the help of differential equations is considered. We will focus on describing the motion of the UAV with the help of systems of linearized equations, since the analytical synthesis of the control system is difficult for nonlinear equations. After a brief review of the mathematical model of the UAV movement using the Newton–Euler equations, generalized coordinates and the Lagrange method, the problem of the effectiveness of the UAV flight properties is studied, for which various mathematical models found in the literature are described. Then the problem of mathematical modeling is investigated, that is, under what conditions it is possible to obtain a system close to the real one by correctly developing control laws. Please note that the problem of correctly describing the mathematical model was one of the most common problems when trying to model the system in a virtual environment.

Keywords

Review of works, mathematical modeling, drones, differential equations.

1. Introduction

Due to the growing application of scientific and applied interest in the subject of multi-rotor unmanned aerial vehicles (UAVs), the requirement for the acceptance and assembly of universal mathematical models that describe the movement of UAVs with most of the completeness of these descriptions. When analyzing the movement of a UAV and its control, it is important to understand its rules for the position and orientation of various coordinate systems. The choice of a coordinate system depends on the problem being solved for studying dynamics. For example, with the help of sensors (particularly GPS) it is possible to control the speed and speed of an aircraft relative to the Earth, so the coordinate system associated with the Earth's surface preferably uses the registration of translational motion. On the other hand, inertial sensors (for example, a gyroscope) are evaluated relative to the fuselage of the UAV on which they are installed. accordingly, the equation of rotational motion is most easily described by an associative coordinate system.

Over the past 20 years, works have been published dedicated to the dedication dedicated to the modeling of quadropters. The paper discusses the issues of dynamics simulation, sliding mode control, efficient control and various nonlinear control laws.

2. Simulation of Quadcopter Dynamics

Quadcopter modeling begins with a description of the mathematical model. The quadrocopter has 4 valve motors on board, located at an equal distance from the center and from each other. When considering the issues of mathematical modeling of the movement of a quadrocopter, its calculation

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: svasichenko@mail.ru (A. 1)
ORCID: 0000-0002-2411-6783 (A. 1)

scheme and a description of differential equations based on general theorems of dynamics are given [1]. The mathematical model of the aircraft has 6 degrees of freedom and is described respectively by a system of six differential equations, they are also 2 coordinate systems: relative to the ground, relative to the quadcopter. The equations of dynamics of a quadrocopter as a mechanical system are obtained using Newton's laws and the Euler–Lagrange equations. Quadcopter motion is described by a system of six non-linear differential equations [6]:

$$\begin{cases} \ddot{x} = (sin\psi \cdot sin\varphi + cos\psi \cdot cos\varphi)\dfrac{U_1}{m} \\ \ddot{y} = (-cos\psi \cdot sin\varphi + sin\psi \cdot cos\varphi)\dfrac{U_1}{m} \\ \ddot{z} = -g + (cos\theta \cdot cos\varphi)\dfrac{U_1}{m} \\ \ddot{\varphi} = \dfrac{I_y - I_z}{I_x} \cdot \dot{\theta} \cdot \dot{\psi} - \dfrac{U_2}{I_x} \\ \ddot{\theta} = \dfrac{I_z - I_x}{I_y} \cdot \dot{\varphi} \cdot \dot{\psi} - \dfrac{U_3}{I_y} \\ \ddot{\psi} = \dfrac{I_x - I_y}{I_z} \cdot \dot{\varphi} \cdot \dot{\theta} - \dfrac{U_4}{I_z} \end{cases}, \quad (1)$$

where x, y, z – the Cartesian coordinates of the quadrocopter; φ, θ, ψ – the Euler angles (φ – the yaw angle; θ – the pitch angle; ψ – the roll angle); I_x, I_y, I_z – diagonal elements of the quadcopter inertia tensor; m – the mass of the quadcopter; g – free fall acceleration; $U = (U_1, U_2, U_3, U_4)$ – virtual control forces associated with the motor control forces by equations.

Description of the mathematical model of UAV movement using Newton-Euler, considering the cross-links presented in [3,7]. With the use of generalized coordinates and Lagrange methods, he forms a mathematical model of the UAV movement in [6]. Linearized and simplified models of UAV motion underlie the synthesis of controllers and filters: LQR controllers [8], Kalman filter [9], L1 optimization [10]; sliding mode control [11, 12]. In most cases, after obtaining control systems synthesized using simplified models, an increased assessment of the selection score is not maintained. When synthesizing the UAV automatic motion control system using simplified models, it is necessary to use their adequacy.

3. Acknowledgements

The article is based on Hai Lin's work "Survey of systems with switching". This article was created while studying the topic of UAV modeling using differential equations. Unfortunately, it was not possible to fit all the work.

4. References

[1] Popov, N.I. Modeling the flight dynamics of a quadrocopter [Electronic resource] / N.I. Popov, O.V. Emelyanova // Fire engineering. - 2014. - No. 4. - P. 1–7.

[2] Zagordan, A.M. Elementary helicopter theory / A.M. Zagordan // Voennoe izdatelstvo Ministerstva oboronyi Soyuza USSR, 1955, – 216 p.

[3] Gen, K. Stabilization algorithms for automatic control of quadrocopter trajectory [Electronic resource] / K. Gen, N.A. Chulin // Science and education MSTU. - 2015. - No. 5. - P. 218-235. – URL: http://engineering-science.ru/doc/771076.html.

[4] Karpunin, A.A. Simplification and linearization of the mathematical model of the movement of unmanned aerial vehicles in space and in the vertical plane / A.A. Karpunin, I.P. Titkov // Modern science-intensive technologies. - 2019. - No. 2. - P. 69–76.

[5] Kalyagin, M.Yu. Modeling of a quadrocopter flight control system in Simulink and Simscape Multibody / M.Yu. Kalyagin, D.A. Voloshin // Proceedings of the MAI. – 2020. – No. 112. – P. 1– 27. – URL: https://trudymai.ru/published.php?ID=116625.
[6] Margun, A.A. Control system for unmanned aircraftequipped with robotics arm / A.A. Margun, K.A. Zimenko, D.N. Bazylev, A.A. Bobtsov, A.S. Kremlev, D.D. Ibraev, M. Cech // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. – 2014. – No. 6. – P. 54 – 62.
[7] Belkheiri, M. Different linearization control techniques for a quadrotor system / M. Belkheiri, A. Rabhi, A. EL Hajjaji, C. Pegard // CCCA12, 2012. – P. 1–6. – DOI: 10.1109/CCCA.2012.6417914.
[8] Fessi R. Modelling and Optimal LQG Controller Design for a Quadrotor UAV / R. Fessi, S. Bouall`egue // Proceedings of the 3th Internation conference on automation, control engineering and computer science, 2016. – P. 264–270.
[9] Satici A.C., Poonawala H., Spong M.W. Robust Optimal Control of Quadrotor UA Vs. IEEE Access. – 2013. – Vol. 1. – P. 79–93. DOI: 10.1109/ACCESS.2013.2260794.
[10] Gherouat O., Matouk D., Hassam A., Abdessemed F. Sliding Mode Control for a Quadrotor Unmanned Aerial Vehicle. J. Automation & Systems Engineering. – 2017. – Vol. 10, No. 3. – P. 150–157.
[11] Xu Rong, Umit Ozguner. Sliding Mode Control of a Quadrotor Helicopter. Proceedings of the 45th IEEE Conference on Decision and Control, 2006. – P. 4957–4962. DOI: 10.1109/ CDC.2006.377588.
[12] Lee D., Kim H.J., Sastry S. Feedback linearization vs. adaptive sliding mode control for a quadrotor helicopter. International Journal of Control, Automation and Systems. 2009. No. 73, P. 419–428. DOI: 10.1007/s12555-009-0311-8.

Исследование инвариантности алгоритмов поиска ключевых точек

Елизавета Викулова [1]

[1] *ИДСТУ имени В.М. Матросова СО РАН, ул. Лермонтова, 134, а/я 292, Иркутск, 664033, Россия*

Аннотация

В данной работе рассмотрены алгоритмы поиска и описания ключевых точек, также их сравнение между собой, проверялась инвариантность алгоритмов к слабо освещенным в литературе преобразованиям, таким как Гауссов шум, высветление, размытие по Гауссу. В результате проведенной серии экспериментов был выявлен наиболее успешный по точности и скорости работы алгоритм для каждого из целевых объектов. Результаты будут использованы в рамках комплекса TEMAR.

Ключевые слова

Компьютерное зрение, ключевая точка, инвариантность

1. Введение

Тема компьютерного зрения является актуальной, так как она имеет множество применений в различных областях, включая медицину, автомобильную промышленность, робототехнику, безопасность и видеонаблюдение, игровую индустрию и многое другое. Компьютерное зрение позволяет компьютерам анализировать и интерпретировать изображения и видео, что может помочь улучшить качество жизни людей и повысить эффективность различных процессов. Благодаря постоянному развитию технологий компьютерного зрения, эта тема остается актуальной и важной для научных исследований и промышленного применения.

Методы, которые дают лучшую стабильность обнаружения и распознания объектов основываются на нахождении и описании особых точек. Особой точкой изображения называется точка на изображении, окрестность которой отличается от окрестности любой другой точки изображения. У каждой такой точки есть дескриптор, это числовое представление(описание) особой точки на изображении, которое позволяет ее уникально идентифицировать и использовать для поиска соответствий в других изображениях. Для поиска таких особенностей разработано большое количество методов, у которых есть свои достоинства и недостатки. Чтобы результат был хорошим, важна инвариантность этих методов к различным видам преобразований.

2. Исследуемые алгоритмы и условия экспериментов

Для сравнения использовались следующие методы:

1. SIFT (Scale-Invariant Feature Transform) – на данный момент является наиболее популярным алгоритмом благодаря тому, что имеет ряд преимуществ, в частности, гарантированную инвариантность к масштабированию и повороту; однако, алгоритм запатентован [1];
2. ORB (Oriented FAST and Rotated BRIEF) – имеет открытый исходный код, также, как и SIFT, обладает инвариантностью к масштабу и повороту; работает быстрее, чем многие другие алгоритмы, не ограничен в использовании для любых проектов [2];
3. AKAZE (Accelerated-KAZE) – является относительно новым методом, его эффективность и применимость в различных задачах компьютерного зрения до сих пор не полностью исследованы [3];

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia

EMAIL: vikulizavet85@gmail.com;

ICCS-DE 2023 Workshop Proceedings

4. BRISK (Binary Robust Invariant Scalable Keypoints) – считается наиболее простым из перечисленных методов, обладает высокой скорости работы и не высокой ресурсоемкостью [4].

Применение этих методов было успешно продемонстрировано в различных областях, включая компьютерное зрение, робототехнику, биометрию и медицинскую диагностику.

Анализ проводится на нескольких изображениях, исследуются объекты, относящихся к следующим типам: архитектура, как объект, который имеет четкие границы и форму, что делает его идеальным для анализа и сравнения алгоритмов; объекты биологической природы, такие как животные, обладают множеством деталей и текстур, что позволяет проверить способность алгоритмов обрабатывать и анализировать сложные изображения; объекты техногенной природы с меньшем количеством деталей и простыми формами.

В рамках данного исследования проверялась инвариантность алгоритмов к мало освещенным в литературе преобразованиям [5], которые наиболее распространены в комплексе TEMAR [6], помощью в разработке которого я в настоящий момент занимаюсь.

1. Гауссов шум – является наиболее распространенным типом шума в изображениях и может возникать из-за различных факторов;
2. Высветление – может происходить из-за неправильной экспозиции камеры или яркого источника света в кадре;
3. Размытие по Гауссу – крайне распространен из-за дефектов оптики, сжатия и движения объектов в кадре.

Для чистоты эксперимента предварительная обработка изображений не производилась. В качестве критерия оценки сравнения использовалось расстояние Хемминга между дескрипторами совпавших точек и общее количество найденных пар связей.

3. Результаты

Для проведения тестов была написана универсальная программа, которая запускалась несколько раз, результаты усреднены. Методы дали примерно одинаковые результаты на всех типах объектов. Частный пример изменения расстояния Хемминга от силы преобразования, отражающий наиболее частое поведение алгоритмов представлен на рисунке 1.

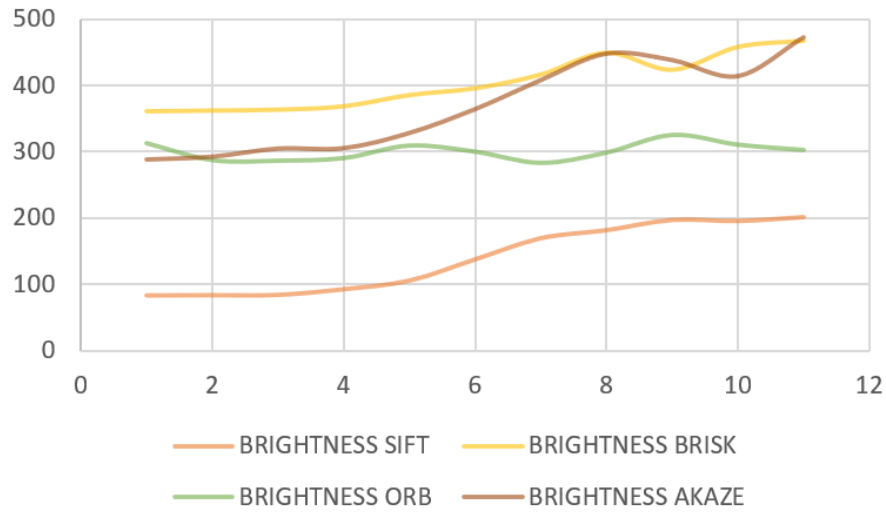


Рисунок 1: Частный пример результатов тестирования (архитектурный объект, преобразование – высветление)

У метода SIFT самые хорошие показатели средней дистанции и большое количество найденных пар, ошибка растет линейно с изменением уровня преобразования изображения. Исходя из графиков, ORB имеет наименьшую скорость роста метрики относительно уровня накладываемого преобразования, но при этом он находит очень малое количество пар точек, по сравнению с остальными алгоритмами. AKAZE не имеет такого большого количества пар

найденных особых точек, но, несмотря на высокие значения в критериях, найденные связи имеют одну из самых низких вероятностей ошибки второго рода. BRISK находит большое количество связей, но значение метрики выше, чем у предыдущих методов.

На основе полученных результатов, на данном наборе тестовых изображений, можно сделать вывод, что наибольший потенциал для идентификации объектов в рамках комплекса TEMAR имеет метод SIFT. Однако, стоит отметить, что в некоторых случаях у каждого метода значение метрики остальных алгоритмов сильно растет за счет ложно-положительно выбранных пар ключевых точек, несколько искажая результаты. Природа происхождения данного аспекта требует более подробного изучения. Помимо этого, ведется добавление новых изображений и анализ взаимосвязи между типами объектов и результатами работы алгоритмов.

Литература

[1] D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, Int. J. Comput. Vision 60 (2004) 91–110.

[2] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, ORB: An Efficient Alternative to SIFT or SURF, in: Proceedings of the 2011 International Conference on Computer Vision, IEEE Computer Society, Washington, DC, USA, 2011, pp. 2564–2571.

[3] P. Fernández Alcantarilla, Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces, in: Proc. of British Machine Vision Conference, 2013, pp. 1–11.

[4] S. Leutenegger, M. Chli, R. Y. Siegwart, BRISK: Binary Robust invariant scalable keypoints, in: Proc. of 2011 International Conference on Computer Vision, Barcelona, Spain, 2011, pp. 2548–2555.

[5] E. Karami, S. Prasad, M. Shehata, Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images, in: Proc. of the 2015 Newfoundland Electrical and Computer Engineering Conference, St. johns, Canada, 2015, pp. 1–5.

[6] D. Kostylev, A. Tolstikhin, S. Ul'yanov, Development of the Complex Modelling System for Intelligent Control Algorithms Testing, in: Proc. of 42nd Intern. Convention on Information and Communication Technology, Croatian Society for Information and Communication Technology, Electronics and Microelectronics, Opatija, 2019, pp. 1091–1096.

СРАВНЕНИЕ ПОПУЛЯЦИОННЫХ АЛГОРИТМОВ ОПТИМИЗАЦИИ, ВДОХНОВЛЁННЫХ ЖИВОЙ ПРИРОДОЙ

Надежда Душкина [1]

[1] *Институт динамики систем и теории управления Им. В.М. Матросова, СО РАН, ул. Лермонтова, 134, Академгородок м-н, Свердловский район, Иркутск, 664033, Россия*

Аннотация

В данной статье производится сравнение популяционных алгоритмов по точности и скорости нахождения решения. Алгоритмы были выбраны по принципу усложнения биологических организмов, послуживших вдохновением для их создания. Реализованные на языке Java, алгоритмы были испытаны на нескольких классах тестовых функций.

Ключевые слова

Роевой интеллект, роевой алгоритм, оптимизация, популяционный алгоритм

1. Введение

Во многих областях для получения близкого к оптимальному решения за время, допустимое для функционирования системы, успешно применяются стохастические эвристические методы оптимизации, такие как алгоритм имитации отжига, эволюционные и роевые алгоритмы. Они относятся к методам локального поиска и основаны на идеях, взятых из природы. В настоящее время большинство нейронных сетей и значительный блок оптимизации работает именно на роевых алгоритмах. Термин «Роевой интеллект» был введен Ван Цзином и Херардо Бени в 1989 году, а уже в 1995 году Джеймс Кеннеди и Рассел Эберхарт предложили метод для оптимизации непрерывных нелинейных функций. Наука не стоит на месте и с тех пор было придумано и реализовано несколько сотен роевых алгоритмов.

В данной работе было произведено сравнение некоторых часто встречающихся базовых популяционных алгоритмов. Целями сравнения являются: оценка и сравнение скорости нахождения и точности решения, а также обнаружение зависимости результатов от обследуемого типа функций.

2. Оцениваемые алгоритмы и условия сравнения

Для рассмотрения были выбраны следующие алгоритмы, взятые по принципу усложнения биологических организмов:

- Алгоритм роя частиц (Particle Swarm Optimization, PSO). Наблюдение за птицами вдохновило Крейга Рейнольдса на создание в 1986 году компьютерной модели, которую он назвал Boids. Он использовал три простых принципа, где каждая птица стремилась избежать столкновений с другими птицами, двигалась в том же направлении, что и находящиеся неподалеку птицы, и все птицы стремились двигаться на одинаковом расстоянии друг от друга. В 1995 году Джеймс Кеннеди и Рассел Эберхарт предложили метод, названный алгоритмом роя частиц. Вдохновением послужила модель Рейнольдса. Алгоритм моделирует многоагентную систему, где агенты-частицы двигаются к оптимальным решениям,

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: dushckina.nad@yandex.ru

ICCS-DE 2023 Workshop Proceedings

обмениваясь при этом информацией с соседями. Данный алгоритм был выбран потому, что является одним из первых роевых алгоритмов, служит своеобразной основой других популяционных алгоритмов.

- Алгоритм искусственной иммунной системы (Artificial Immune Systems, AIS). Первые попытки разработки AIS относятся к 70-м гг. XX века. Однако представляющие практический интерес результаты были получены только в 1990-х гг. Биологическим прототипом AIS является иммунная система человека и обработка информации в ней молекулами белков (пептидов). Эта система представляет собой сложную адаптивную структуру, использующую комбинацию различных механизмов защиты от внешних патогенов - любых микроорганизмов (включая вирусы и бактерии), способных вызывать патологическое состояние (болезнь) человека. Многие внешние черты сближают иммунную сеть с нейронными сетями. Подобно нейронным сетям, иммунные сети обладают способностью к обучению, прогнозированию и принятию решений в незнакомой ситуации. Как и нейронные сети, иммунные сети не нуждаются в заранее известной модели задачи, а строят эту модель на основе полученной информации. Интерес к алгоритму искусственной иммунной системы вызван по причине его широкого использования для построения многоагентных и самоорганизующихся систем.
- Алгоритм сорняковой оптимизации (Invasive Weed Optimization, IWO). Вдохновлен таким общераспространенным явлением, как колонизация сельскохозяйственных угодий сорняками. Алгоритм предложен в 2006 г. иранскими учеными Мехрабианом и Лукасом и основан на моделировании таких свойств, как посев, рост и конкуренция в колонии сорняка. Основным механизмом, определяющим динамику сообщества любых растений, является естественный отбор, из которого выделяют два крайних типа в основе которых лежат реальные стратегии отбора: r-отбор (живи быстро, размножайся быстро, умирай молодым) и k-отбор (живи медленно, размножайся медленно, умирай в старости). Этот алгоритм является одним из известных представителей популяционных алгоритмов.
- Алгоритм пчелиной колонии (Artificial Bee Colony, ABC). В 2005 г. Дервис Карабога опубликовал научную статью и описал в ней модель роевого интеллекта, на создание которой его вдохновили пчелиные танцы. Модель получила название «алгоритм пчелиной колонии». Алгоритм основан на имитации поведения колонии медоносных пчел при сборе нектара в природе. Основной целью работы пчелиной колонии в природе является разведка пространства вокруг улья с целью поиска нектара с последующим его сбором. Для этого в составе колонии существуют различные типы пчел: пчелы-разведчики и рабочие пчелы-фуражиры. Разведчики ведут исследование окружающего улей пространства и сообщают информацию о перспективных местах, в которых было обнаружено наибольшее количество нектара (для обмена информацией в улье существует специальный механизм, именуемый танцем пчелы). Представленный алгоритм выбран потому, что является одним из широко известных роевых алгоритмов.
- Алгоритм кукушки (Cuckoo Search, CS) предложен и разработан Янгом и Дебом в 2009 г. Вдохновением для его создания послужил гнездовой паразитизм некоторых видов кукушек, что подкладывают свои яйца в гнезда других птиц (других видов птиц). Некоторые из владельцев гнезд могут вступить в прямой конфликт с кукушками, что врываются к ним. Например, если владелец гнезда обнаружит, что яйца не его, то он или выбросит эти чужие яйца, или просто покинет гнездо и создаст новое где-то в другом месте. Алгоритм кукушки стремительно набирает популярность в настоящее время, был рассмотрен для лучшего понимания его работы.

В данной работе эти алгоритмы были реализованы на языке Java и испытаны на широком спектре классически используемых тестовых функций, включающих следующие классы: невыпуклые, многоэкстремальные, а также функции нескольких переменных и с различной областью допустимых значений. Для чистоты эксперимента все тестирования приведенных выше алгоритмов проводились с учетом зафиксированного количества популяций, выделенного на проведение вычислений, а также количества агентов в каждой популяции.

Далее представлена часть исследуемых функций. Во всех запусках были заранее известны координаты экстремума и его абсолютное значение, а также ограничения обследуемой области.

Исследуемые функции:

$$F_1(x) = \sum_{i=1}^{n} x_i^2 \quad (1)$$

$$F_2(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right] \quad (2)$$

$$F_3(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 \quad (3)$$

$$F_4(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad (4)$$

$$F_5(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10 \quad (5)$$

Усредненные результаты проведенных экспериментов представлены в таблицах 1-2.

Таблица 1

Среднее значение

	F1	F2	F3	F4	F5
AIS	0,0024	14,0166	-1,0316	0,2164	0,3979
Bees	0,0000	0,0000	-0,3093	0,9667	0,3979
Cuckoo	0,3160	0,6327	0,2396	0,5886	0,4954
InvasiveWeed	0,0000	0,0000	-0,3693	0,0000	1,2961
PSO	0,0827	29,3280	-0,9497	0,0493	0,4195
Экстремум	0,0000	0,0000	-1,0316	0,0000	0,398

Таблица 2

Среднее время

	F1	F2	F3	F4	F5
AIS	464,1	522	440,2	3040	422,2
Bees	75,2	90,1	92,2	88,6	70,6
Cuckoo	26,5	20,6	20,3	20,9	20,1
InvasiveWeed	195,8	204,9	331,3	212	399
PSO	13,6	69,6	35,4	59,8	17

Первый класс тестируемых функций – унимодальные. Все алгоритмы, за исключением алгоритма кукушки, справились хорошо, превосходно себя показали сорняковый алгоритм и алгоритм пчелиной колонии.

Следующий класс - многомерные мультимодальные. Отлично справился сорняковый алгоритм. Чуть хуже алгоритм поиска кукушки. Алгоритмы роя частиц и искусственной иммунной системы не прошли испытание на первой функции, но хорошо справились с испытанием на второй функции, с другой стороны, у алгоритма пчелиной колонии наблюдается полностью обратный результат.

И последний рассмотренный класс функций - мультимодальные с фиксированной размерностью. С этим испытанием справился алгоритм искусственной иммунной системы Сорняковый алгоритм, хорошо показавший себя в прошлых испытаниях, с этим типом функций справился несколько хуже. Неоднозначно себя показал алгоритм роя частиц, хорошо справившись с первой функцией, но показав средний результат для второй функции.

3. Заключение

В рамках проведенного исследования были выявлены лидеры по каждой заявленной категории: лучшим по точности стал сорняковый алгоритм, а вот лучших по времени оказалось двое, а именно алгоритмы поиска кукушки и роя частиц. Исходя из этих данных были сделаны выводы, что дополнительных исследований в контексте применимости для решения задачи сшивки изображений требуют алгоритм поиска кукушки и сорняковый алгоритм.

Полученные данные будут использоваться для последующего решения задачи сшивки изображений, а также проведения их дальнейшего исследования и модификации при возникновении необходимости.

Литература

[1] B. Basturk, D. Karaboga, An artificial bee colony (ABC) algorithm for numeric function optimization, in: IEEE swarm intelligence symposium, 2006, pp. 12–14.

[2] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of IEEE International Conference on Neural Networks, 1995, pp. 1942–1948

[3] X.-S. Yang, S. Deb, Cuckoo search via L´evy flights, in: Proc. of World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), December 2009, India. IEEE Publications, USA, 2009, pp. 210-214.

[4] N. Rathod, S. Wankhade, Investigation of optimized ELM using Invasive Weed-optimization and Cuckoo-Search optimization, Nonlinear Engineering 1 (2022) 568–581.

[5] А.П. Карпенко, Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учебное пособие 2-е изд., Издательство МГТУ им. Н. Э. Баумана, Москва, 2017.

[6] А.П. Карпенко, Е.Ю. Селиверстов, Обзор методов роя частиц для задачи глобальной оптимизации, Машиностроение и компьютерные технологии (2009) 1–26.

Методика продвижения сайта

Михаил Жданов [1], Ольга Башарина [1]

[1] *Уральский государственный экономический университет», 8 марта 62, Екатеринбург, 620144, Россия*

Аннотация

В статье представлена разработанная методика продвижения сайта, в основе которой заложена модель маркетинговых стратегий SOSTAC. Методика включает в себя шесть этапов: анализ текущего состояния сайта, определение цели, стратегический и тактические планы для каждого инструмента продвижения; определение действий по их реализации; контроль и измерение эффективности исполнения действий. Главным преимуществом представленной методики является комплексное применение различных инструментов для продвижения сайта. Апробация методики проведена для сайта интернет-магазина техники. В результате чего составлен план мероприятий по продвижению данного сайта, проведены необходимые оценки и расчеты.

Ключевые слова

Продвижение сайта, анализ сайта, модель SOSTAC, SMM, SEO.

1. Введение

В наше время интернет стал неотъемлемой частью бизнеса и экономики в целом. Уровень конкуренции и количество веб-ресурсов растет с каждым днем, и у компаний все чаще появляется необходимость заниматься продвижением собственного сайта, что позволило бы выделиться среди конкурентов, привлечь к себе большое число новых потребителей, повысить уровень лояльности клиентов и снизить затраты на интернет–рекламу.

Компании используют различные методы продвижения сайта, основываясь на своих целях, задачах и потребностях, что делает задачу продвижения сайта в сети интернет достаточно сложной и актуальной для многих компаний. Продвижение сайта – это комплекс мероприятий, направленный на улучшение позиций сайта в поисковых системах и привлечение большего количества целевых посетителей [1]. В этот комплекс входят множество методов, например, поисковая оптимизация, продвижение с помощью социальных сетей, реклама в поисковых системах, контент–маркетинг и другие.

2. Методика продвижения сайта по модели SOSTAC

Для качественного и эффективного продвижения сайта в сети интернет можно использовать множество различных инструментов и методов, каждый из которых эффективен для достижения различных целей и задач. В данной работе предложена адаптированная методика продвижения сайта, основанная на модели маркетинговых стратегий SOSTAC.

Модель SOSTAC в 1990-х годах разработал британский эксперт Королевского института маркетинга Пол Смит. По версии Chartered Institute of Marketing модель входит в Топ-3 мировых бизнес-моделей [2]. Она основана на 6 обязательных компонентах: Situation analysis (анализ текущей ситуации); Objectives (цель, к которой нужно прийти); Strategy (стратегия, как достичь цели); Tactics (тактика, что использовать для достижения цели); Action

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: lipeckprotiv19@gmail.com (A. 1); basharinaolga@mail.com (A. 2)
ORCID: 0000-0002-7151-782X (A. 2)

ICCS-DE 2023 Workshop Proceedings

(конкретные действия, задачи и сроки); Control (контроль, как понять, что цель достигнута, по каким показателям).

Целью данной работы было создание комплексной и универсальной (насколько возможно) методики продвижения сайта за счет использования нескольких инструментов продвижения, которые способны наиболее эффективно повлиять на увеличение трафика сайта, повысить узнаваемость бренда и лояльность клиентов, вывести сайт в топ у поисковых систем, и т.п. Такими инструментами были определены: SMM (Social Media Marketing), SEO (Search Engine Optimization) и контекстная реклама.

Разработанная методика является универсальной и подходит для решения множества целей в сфере продвижения сайтов в сети интернет (рисунок 1).

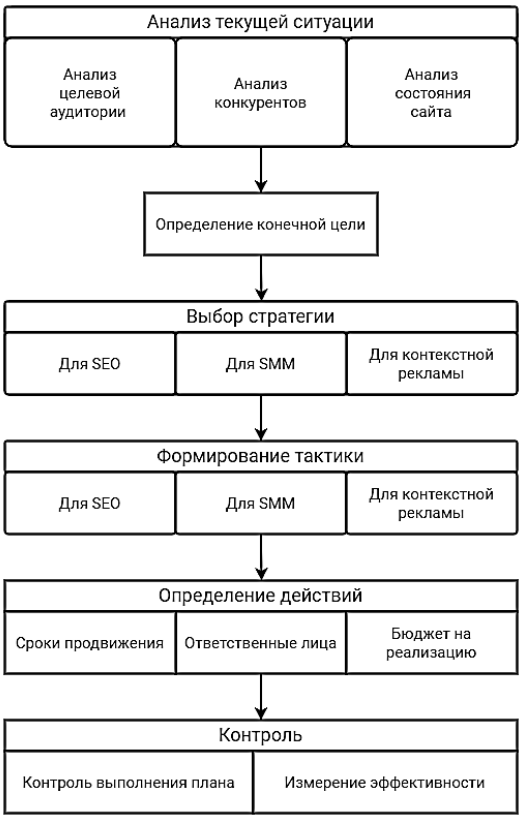


Рисунок 1: Методика продвижения сайта в сети интернет

Апробация методики была проведена для реальной компании, занимающейся продажей техники и оказанием сопутствующих услуг как в традиционном, так и электронном формате. Далее кратко представим основные результаты. Анализ целевой аудитории показал, что основной аудиторией являются мужчины (63,6%) и женщины (36,4%) в возрасте от 18 до 65 лет.

Анализ состояния сайта, произведенный с помощью инструмента «Parsesite» [3] выявил несколько проблемных мест: большое количество символов в заголовке (103 символа); низкая скорость загрузки главной страницы сайта (2.08 с); отсутствие alt-атрибутов у некоторых изображений; ошибки html-кода, наличие ключевых слов с низким трафиком и др.

В качестве рекомендации стратегией для SEO было выбрано локальное продвижение, так как основной поток клиентов находится на территории Свердловской и Челябинской областей. Способом продвижения с помощью социальной сети «ВКонтакте» выбрана таргетированная реклама, способная максимально точно определить целевую аудиторию, и как следствие, сократить затраты на рекламу и увеличить конверсию. Для контекстной рекламы выбрана стратегия конкурентные преимущества в тексте объявления.

Расчет бюджета на продвижение исследуемого сайта достаточно затруднителен. По результатам анализа текущего состояния сайта и рынка труда SEO- и SMM-специалистов, а так

же руководствуясь экспертной оценкой планируемых работ, была определена сумма 870 тыс. руб. на период 6 месяцев.

На протяжении всего периода необходимо проводить контроль выполнения мероприятий, анализ полученных результатов и, в случае необходимости, корректировать стратегию, перераспределять ресурсы, изменять приоритеты.

На основе диаграммы прогнозируемой динамики входящего трафика (рисунок 2) можно сказать, что SEO-продвижение способно значительно увеличить трафик посетителей на сайт уже через 3-4 месяца. Данный график был построен с помощью онлайн калькулятора SEO [4], основываясь на дынных о возрасте сайта, органическому трафику за месяц и его тематике.

Рисунок 2: Прогноз динамики поискового трафика сайта

3. Заключение

В рамках представленной исследовательской работы были изучены основные методы продвижения сайта и их принципы работы. Предложенная методика продвижения сайта, позволяет оценить текущее состояние сайта, учесть стратегические цели компании, спланировать тактические шаги для их реализации, требуемые ресурсы и инструменты и провести анализ и оценку эффективности проведенных мероприятий.

Проведена практическая работа по анализу сайта интернет-магазина техники с использованием различных сервисов. На основе результатов анализа и согласно предложенной в работе методике разработан план продвижения этого сайта, произведены расчеты требуемых трудовых и финансовых ресурсов.

4. Список использованных источников

[1] Курочкин М.Е. Каналы и инструменты продвижения в интернете в контексте концепции маркетинговых коммуникаций // Инновации и инвестиции. – 2020. – №9. – URL: https://cyberleninka.ru/article/n/kanaly-i-instrumenty-prodvizheniya-v-internete-v-kontekste-kontseptsii-marketingovyh-kommunikatsiy (дата обращения: 26.05.2023).

[2] 100% эффективное продвижение сайта по модели SOSTAC // 1ps: сайт. – URL: https://1ps.ru/blog/dirs/2017/100-effektivnoe-prodvizhenie-sajta-po-modeli-sostac/ (дата обращения: 16.05.2023).

[3] Parsesite – анализ сайта онлайн. – URL: https://parsesite.ru/ (дата обращения: 26.05.2023).

[4] Яндекс.Wordstat. – URL: https://wordstat.yandex.ru/ (дата обращения: 26.05.2023).

Нейросетевой подход в проблеме распознавания техногенных шумов

Оксана Копылова [1]

[1] *Институт вычислительной математики и математической геофизики СО РАН, пр. акакд. Лаврентьева, 6, 630090, Новосибирск, Россия*

Аннотация

Проблема геоэкологического мониторинга окружающих техногенных шумов в связи с их возрастающим воздействием на социальную среду приобретает все более высокую актуальность. Одной из важных составляющих этой проблемы является распознавание источников транспортных шумов в условиях реальной фоновой обстановки. Целью работы является разработка методов и алгоритмов распознавания движущихся транспортных средств, предназначенных для использования в интеллектуальных системах обнаружения и распознавания. В качестве источников рассматриваются различные виды железнодорожного, тяжелого колесного, гусеничного видов транспорта. Решение задачи основано на использовании сверточной нейронной сети с предварительной обработкой данных в спектральной области. Рассматриваются вопросы распознавания источников транспортных колебаний на фоне изменяющегося уровня внешних шумов и пространственного положения транспорта по отношению к пунктам регистрации колебаний. Приводятся результаты работы алгоритма распознавания на данных полевых экспериментов.

Ключевые слова

Транспортные вибрации, сейсмические колебания, техногенные шумы, распознавание транспортных средств, сверточная нейронная сеть, полевой эксперимент.

1. Введение

Проблема геоэкологического мониторинга окружающих техногенных шумов в связи с возрастающим шумовым загрязнением городов в условиях непрерывно растущей автомобилизации в мире связана с изучением воздействия техногенных шумов на окружающую социальную среду и, прежде всего, на человека. Особо остро стоит мониторинг на низких – инфранизких частотах, которые являются особо угрожающими для живых организмов [1], а также наиболее разрушительными для крупных сооружений (мостов, зданий, производственных помещений и т.д.). Последнее определяется тем, что в области инфранизких частот находятся собственные частоты колебаний сооружений [2, 3]. Одной из важных задач геоэкологического мониторинга является распознавание источников транспортных колебаний. В работе анализируются колебания от тяжелых видов транспорта – железнодорожного, тяжелого колесного и гусеничного. Такие колебания распространяются в виде сейсмических в земле и акустических в атмосфере.

Решение задачи распознавания движущихся транспортных средств по записям сейсмических сигналов может быть построено на сравнении значений информативных признаков регистрируемых колебаний от различных источников с их эталонами, полученными статистическими методами на этапе предварительного обучения. [4, 5]. В работе [6] распознавание объектов транспортных средств проводилось с помощью вероятностной

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: oksana@opg.sscc.ru (A. 1)

ICCS-DE 2023 Workshop Proceedings

нейронной сети. В работе [7] предлагалось использовать трехслойную нейронную сеть прямого распространения и метод k-ближайших соседей. В работе [8] использовался метод опорных векторов. В работе [9] распознавание основывалось на сверточной нейронной сети.

Отличительная особенность настоящей работы увязывается с вопросами обнаружения и распознавания на предельных расстояниях, превосходящих известные в условиях воздействия внешних шумов. При этом алгоритмы и программы ориентированы на обработку данных в режиме реального времени.

2. Постановка задачи распознавания

В качестве критерия распознавания используется процент правильно распознанных объектов из заданного набора классов. В общем случае критерий качества определяется следующим образом:

$$P=(F, N, A) \quad (1)$$

где P вероятность правильного распознавания, F – вектор признаков, N – составляющая внешнего шума, A – алгоритм распознавания.

Вектор признаков F формируется из функции регистрируемого волнового поля от транспортных источников. Она описывается суммой квазигармонических функций с переменными параметрами и широкополосной шумовой составляющей, отражающей взаимодействие источника со средой в процессе его движения, а также внешнего мешающего шума. Модель такого колебания описывается функцией

$$x_k(t_i) = A_x(k) h_k L\left[\cos\left(\omega_0(t_i - \triangle t_k) + \varphi_x(k)\right) + n_k(t_i)\right] \quad (2)$$

где ω_0 – преобладающая в спектре колебаний частота вращения двигателя; $\varphi_x(k)$ – фаза ее в k-ой точке регистрации, t_i меняется в пределах от t_{0i} до $t_{0i.} + T$, где T – длительность сигнала, а t_{oi} соответствует моменту появления полезного колебания, $n_k(t)$ – широкополосная шумовая составляющая, отражающая взаимодействие источника со средой в процессе его движения и внешний шум; A_x – амплитуда колебания на k-ом датчике; h_k – чувствительность датчика; L – оператор фильтрации сигнала.

С учетом принятой модели колебания ставятся задачи анализа, обнаружения и распознавания источника.

2.1. Предобработка данных и выделение информативных признаков

Анализ смеси полезного колебания с шумом (2) распадается на два этапа: вычисление амплитудного спектра с целью выделения транспортных колебательных составляющих на фоне внешнего шума; вычисление спектрально-временной функции с целью изучения динамики спектров транспортных колебаний на интервале времени T.

Решение задачи выделения информативных признаков осуществляется на основе использования оконного спектрального анализа Фурье для прослеживания динамики изменения спектра во времени:

$$F(k,l) = \sum_{n=0}^{N-1} S_l(t_n) \exp(-i\frac{2\pi nk}{N}), l = 1,...,L \quad (3)$$

Здесь L - количество секций длительностью $\Delta T = N \cdot \Delta t$ каждая, на которые разделяется искомый сигнал $S(t)$, Δt - интервал выборки дискретных значений сигнала.

Нормировка является необходимым этапом предобработки тренировочных и тестовых данных, подаваемых на вход алгоритма распознавания для формирования спектральных функций, инвариантных по расстоянию и уровню внешних шумов. Преобразование исходных данных результата спектрального анализа выполняется так, чтобы все амплитудные значения бинов (отсчетов) лежали в диапазоне [0;1] по формуле: $x_i = x_i / \max(x)$, где x_i – i-ое значение бина спектральной функции, x – вектор значений бинов спектральной функции.

2.2. Методика проведения экспериментов

С целью исследования вопроса о выборе информативных параметров сейсмических и акустических колебаний от различного вида железнодорожного и автомобильного движущегося транспорта на шумных магистралях были проведены эксперименты с целью их записи. Также была выполнена запись таких колебаний от тяжелого колесного и гусеничного транспорта, Эксперименты проводились в утреннее, дневное, вечернее и ночное время суток в разные сезоны года. Полученные данные являются исходным материалом для предварительного анализа и выделения информативных признаков, а также обучения и тестирования нейронной сети. Ниже приведен пример нормированной спектрально-временной функции, полученной для грузового поезда (см. Рисунок 1).

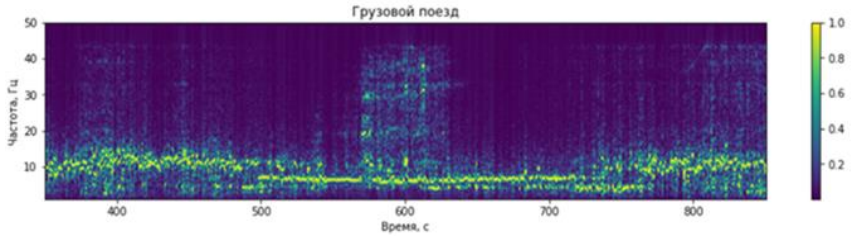


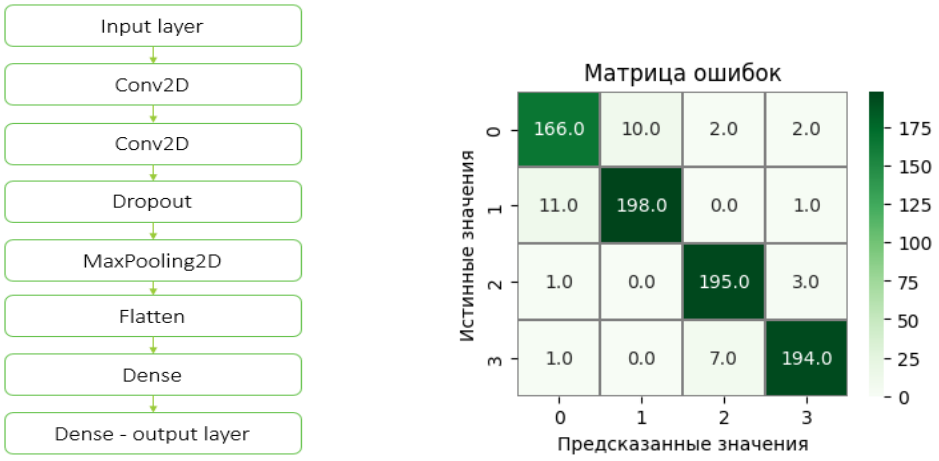
Рисунок 1*:* Нормированная СВФ записей сейсмических колебаний от движущегося грузового поезда на удалении 30 м от железнодорожного полотна

2.3. Нейросетевой алгоритм распознавания транспортных источников.

Набор исходных данных представляется в виде набора матриц с размером M ×N, где M – количество секунд в окне, в пределах которого осуществляется анализ, N – количество бинов спектральной функции.

Классы кодировались следующим образом: 0 – электропоезд, 1 – грузовой поезд, 2 – гусеничный транспорт, 3 –тяжелый колесный транспорт. Полученный набор данных был перемешан случайным образом и разделен на тренировочную и тестовую выборки в соотношении 3:1. Такое разделение обеспечивает достаточное количество экземпляров в обучающей и тестовой выборках для обучения модели и ее тестирования.

Структурная схема предлагаемой нейронной сети представлена на рисунке ниже (Рисунок 2 (а)). В качестве функции потерь использовалась категориальная кроссэнтропия, которая описывает, насколько близко предсказанное распределение к истинному. В качестве алгоритма оптимизации подбора весов и повышения скорости обучения используется оптимизатор «Adam» [10].



а) б)

Рисунок 2: а) Структурная схема нейронной сети б) Матрица ошибок классификации

Проверка точности работы классификатора, обученного на тестовой выборке, производилась на данных, не участвовавших в процессе обучения нейронной сети, что обеспечивает оценку обобщающей способности модели к обучению и исключает влияние эффекта переобучения на значение показателя точности. Процент правильного распознавания рассчитывается как отношение числа правильно классифицированных образцов к числу всех образцов. На тестовой выборке такой показатель составил 95%. Ниже на рисунке (Рисунок 2 (б)) представлена матрица ошибок, показывающая распределение случаев правильного и ошибочного распознавания по классам.

3. Заключение

В работе предложен и реализован нейросетевой алгоритм распознавания транспортных объектов по сейсмическим колебаниям на основе сверточной нейронной сети. На группе из 4-х разнотипных транспортных объектов (электропоезд, грузовой поезд, тяжелый колесный и гусеничный виды транспорта) достигнут результат их правильного распознавания около 95%. Алгоритм работает на данных, полученных при различных дальностях движения транспорта, начиная от 700 метров и выше. Полученные при этом результаты одновременно по дальности обнаружения при той же точности распознавания превосходят ранее полученные другими авторами.

4. Благодарности

Работа выполнена в рамках госзадания FWNM–2022–0004.

5. Литература

[1] Васильев А.В. Анализ инфразвукового излучения в условиях территории жилой застройки городского округа Самара. Известия Самарского научного центра Российской академии наук, т. 22, № 5, 2020, с. 60-68.
[2] Экспериментальная динамика сооружений. Мониторинг транспортной вибрации: Монография / Е.К. Борисов, С.Г. Алимов, А.Г. Усов и др. – Петропавловск-Камчатский: КамчатГТУ, 2007. – 128 с. ISBN 978–5–328–00160–1
[3] Prediction of Ground Vibrations Induced by Urban Railway Traffic: An Analysis of the Coupling Assumptions Between Vehicle, Track, Soil, and Buildings. International Journal of Acoustics and Vibration, Vol. 18, No. 4, 2013, pp. 163-172
[4] Левковская Т.В., Козлов Э.В., Мурашко Н.И. Обработка сейсмических сигналов в интеллектуальных системах пассивной локации. Информатика. 2010;(3(27)) с. 89-96.
[5] Хайретдинов М.С., Авроров С.А. Обнаружение и распознавание взрывных источников. // Вестник НЯЦ РК. — 2012. — №12. — с.17–24.
[6] Алямкин С.А. Классификация объектов в сейсмической системе обнаружения с учетом параметров их движения. Автореф… дис. кан. тех. наук. – Новосибирск.: 2014. – 21с.
[7] P. Khunarsal, C. Lursinsap, and T. Raicharoen, ‘Very short time environmental sound classification based on spectrogram pattern matching’, Information Sciences, vol. 243, no. Complete, 2013, pp. 57–74, doi:10.1016/j.ins.2013.04.014.
[8] Q. Zhou, X. Yao, C. Wang, J. Hu, P. Liu and J. Lin, "Adaptive Moving Ground-Target Detection Method Based on Seismic Signal," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 2503705, doi: 10.1109/LGRS.2022.3153368.
[9] Y. Wang, X. Cheng, X. Li, B. Li and X. Yuan, "Powerset Fusion Network for Target Classification in Unattended Ground Sensors," in IEEE Sensors Journal, vol. 21, no. 12, pp. 13466-13473, 15 June15, 2021, doi: 10.1109/JSEN.2021.3067648.
[10] D. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. In Proc. of the Int. Conf. on Learning Representations (ICLR), 2015.

Optimization of Network Interaction on a High-Performance Cluster Using Graph Scheduling

Ilya S.Timokhin[1], Denis I.Shaikhislamov[2] and Aleksey M.Teplov[2]

[1]HSE University, Moscow, Russia
[2]Advanced Software Technologies Laboratory of Huawei Moscow Research Center, Moscow, Russia

Abstract
Data access time become large problem for data processing in distributed environments with a growth of system scale and data size and source of software optimization. In this paper, we research the problem of optimizing I/O and processing operations for a distributed Hadoop cluster that process data in HPC paradigm framework. To increase the efficiency of the framework that processes queries in HDFS, a graph algorithm were developed with respect to optimal use of resources using HDFS data locality on a processing graph. The graph uses data file blocks and replicas, hosts and workers as a nodes, and links between them as edges. The results of reading stage optimization phase was implemented and performance improvements measured comparing baseline and Spark as Industry standard framework.

Keywords
HPC, graph, scheduling, multiprocessing, network, HDFS, reading utilization

1. Introduction

Distributed computations performance and efficiency are highly controlled metrics as far as distributed computing resources are expensive but obligatory to solve large-scale problems. The computations overhead reduction become the high priority task for developers in distributed environments of any type. Big Data clusters and HPC [1] clusters have many common features in the architecture but historically inherit very different approaches. High-performance clusters are designed to perform complex heavy calculations and simulations, handle processing of large-scale data, and require high-speed data transfer rates to keep latency and bandwidth between computing nodes and from storage area network (SNA).

Big Data clusters store and process data in the same nodes and more focused on fault tolerance and availability. Last years these two independent branches converging heavily due to increased data size for scientific processing and requirements to increase computation speed and density for Big Data clusters. Many projects like [2, 3, 4] were developed to converge HPC computations stile with benefits of Big Data infrastructure. One of the most difficult problems to solve is data access and I/O in general and especially for Big Data clusters as far the hardware use more

5^{th} International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
is.timokhin@hse.ru (I. S.Timokhin); denis.shaikhislamov@huawei.com (D. I.Shaikhislamov); aleksey.teplov@huawei.com (A. M.Teplov)
0000-0003-1893-906X (I. S.Timokhin); 0000-0002-9279-6397 (D. I.Shaikhislamov); 0009-0001-7286-6466 (A. M.Teplov)

ICCS-DE 2023 Workshop Proceedings (iccs-de.icc.ru)

typical storage and networking devices than HPC and computation time takes smaller portion of total execution time providing larger portion to I/O. High Performance clusters are based on more advanced hardware and use more advanced approaches and algorithms to reach top level of computations efficiency, so they are used as a source of technologies and approaches to enhance software in Big Data to ensure maximum performance and scalability.

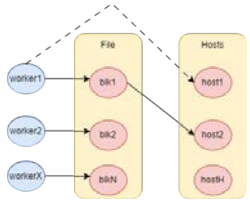
Apache Hadoop is an open-source software framework used for distributed storage and processing of large datasets. It enables users to store, manage, and process big data across clusters of commodity hardware. Hadoop Distributed File System (HDFS) is a distributed file system that is designed to store large amounts of data reliably and provide high bandwidth access to that data. However, I/O (input/output) performance can sometimes become a bottleneck [5] in HDFS due to factors such as network latency, disk speed, and data locality. This paper [6] compares the I/O performance of different file systems used in Hadoop, including HDFS, Amazon S3,and GlusterFS. The authors analyze the impact of various parameters, such as block size and replication factor, on HDFS I/O performance. In [7] authors investigate the performance of HDFS using different benchmarks and workload patterns. Problems were also identified in network bandwidth.One of the possible solutions for optimizing network interactions (including reading, writing and forwarding blocks) can be the use of graph data locality algorithms [8]. When trying to read data relative to the current node without using unnecessary access to the network, the performance of the computing system is significantly increased [9]. The authors of the article developed an algorithm that significantly reduces query processing time in a high-performance cluster by using a graph algorithm that minimizes the number of transfers in the network and maximizes the number of HDFS blocks located locally on a given cluster node.

2. Background

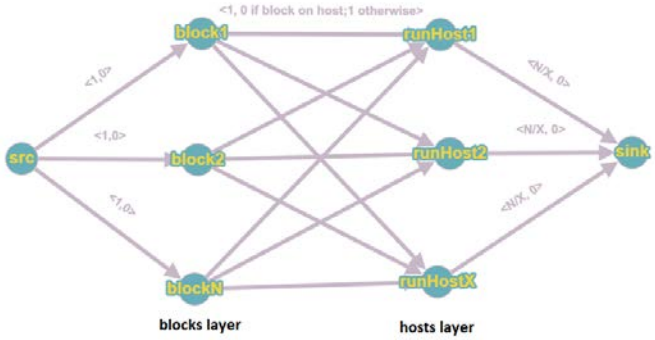
Previously, the graph algorithm was implemented by the authors in [10] for a similar data locality problem. The previous research problem was as follows: let there be some SQL-like query that operates on data in a particular file. This file contains N blocks, is distributed across H hosts in HDFS, and has a replication factor of R (that is, the number of copies of each block on different nodes will be $R-1$). The task is to redistribute blocks on hosts in such way that the blocks are most evenly distributed. The solution for such a task would be to build a "transport network" type graph with vertices representing the blocks and hosts of the system, as well as finding a path from the entry point ("source") to the exit point ("sink"), which implies finding a connection between each block and the optimal host for this block. For this was proposed the max-load value L ("each host can process no more than L blocks") with Dinic's algorithm.

Finding the value of L can be obtained by minimizing the value of the maximum flow (binary search on values from N to N/H). Thus, an optimal graph with well-distributed blocks will reflect the following situation: each host processes a uniform number of N/H blocks, all of these blocks are in the most local positions.

Graph $G = (V, E, C, F)$ is a transport network (V – set of vertices, E – set of edges, C – set of capacities between edges, F – set of flow between edges), which means the HDFS-file representation to optimize paths to specific host and choose the optimal number of blocks for processing on each host.



(a) Workers I/O problem



(b) Graph $\widetilde{G}$ with network optimization

Figure 1: Architecture of optimization

However, this task by itself does not solve the problem of I/O in HDFS or any other file system, because it does not take into account the actual computing resources of the system. Next, let's consider another problem based on a graph approach, but taking into account the problems of reading and forwarding, as well as the locality of data in the actual execution of queries.

3. Contribution

When using an experimental framework for processing big data, system resources are used, which the authors in this work called "workers". These can be cores (physical or virtual), coroutines, or other entities that perform part of the whole task of processing an SQL-like query. The problem with a real computing system is that there is no method of reading blocks in a distributed order in query operations. Thus, the entire file is read sequentially, and each worker is allocated its own data block (or set of blocks) for reading.

Based on this, worker located physically on one host can work with a block located on another host. In this case, such a block will need to be implicitly forwarded to the first host, which entails unjustified network operations, which can be avoided if data blocks are localized in advance by the corresponding workers. Such situation is represented in Figure 1(a).

Let's describe a graph $\widetilde{G} = (V, C, W, E, F)$. Each edge has two properties $< W, C >$ (weight W and capacity C). All the weights for edges are equals to zero (except for the edges between block and the host that needs network transferring). Maximum flow in this current graph is the number of blocks N, but the blocks' distribution is optimized with F value (as it was for G graph's algorithm). So the usage of the edges with weight 1 is minimum. Vertices $\{V_1, ..., V_N\}$ are related to blocks from 1 to N and vertices $\{V_{N+1}, ..., V_{N+H+1}\}$ are related to hosts from 1 to H. Let's also denote **GetBlockLocation** function which returns a set Q of hosts that

contains current block physically (obviously, $|Q| = R$). Next, Algorithm 1 for constructing a new graph will be described.

Algorithm 1 Graph $\widetilde{G}$ construction for network optimization

Let G has $N + X + 2$ vertices and $N(R-1) + X$ edges;
$V_0 =$ source, $V_{N+X+2} =$ sink;
for $i \in \{1, ..., N\}$ **do**
  $E(\text{source}, V_i) =< 1, 0 >$;
  $Q =$ **GetBlockLocation**(V_i)
  for each $q \in Q$ **do**
    $E(V_i, q) =< 1, 1 >$
  end for
  for each $q \notin Q$ **do**
    $E(V_i, q) =< 0, 1 >$
  end for
end for
for $j \in \{N + 1, ..., N + H + 1\}$ **do**
  $(V_j, \text{sink}) =< N/X, 0 >$
end for

Once again, it must be emphasized that the process of building this graph is a modification of the graph G using control over the transfer of blocks to hosts that are not related to this worker. The structure of the $\widetilde{G}$ graph is shown in Figure 1(b).

In other words, one can build two tables - "worker-block(s)" (with a 1:N relationship) and "worker-host" (with a 1:1 relationship), and then maximize the number of blocks that are located on the host of the current worker to build a third table "blocks-host" (with an N:M relationship). This will minimize the number of network transfers associated with additional overhead when trying to process a query.

Thus, thanks to the use of graph strategies, the authors obtained a reduction in data processing time on specific hosts, as well as an increase in fault tolerance due to the use of only the data on which the current blocks are stored.

4. Results

To demonstrate the effectiveness of the obtained data locality algorithm on the $\widetilde{G}$ graph, testing was carried out using an experimental framework for distributed computing of big data, on which 43 SQL-queries from the TPC-DS benchmark were processed. The authors carried out tests on a limited set of TPC-DS tests (43 out of 99), because it is for these queries that the correct execution plan is generated in experimental framework. Same tests were provided for Spark (versions 3.1.3 and 3.4.0). 5 iterations of tests were carried out for each query, and the mean among all iterations was chosen as the final value. The **"Total"** line contains the total execution time, as well as the mean value among all acceleration ratio values.

The acceleration ratio metric for each case was calculated as $a_{xi} = \frac{t_{xi}}{o_i}$, where o_i is i-th query

execution time in the case of a graph optimized framework, t_{Xi} is execution time (i-th query, respectively) in x-th framework (Spark v.3.1.3, Spark v.3.4.0, baseline).

To analyze the uniformity of tasks, a data skew metric has been introduced. For each query, the distribution of file blocks in HDFS is known. Let each host have $\{b_1, b_2, ..., b_H\}$ of file blocks. Introduce the skew metric as follows:

$$\sigma = HM / \sum_{i=1}^{H} b_i, \; M = \max_{i=\overline{1,H}} b_i.$$

The data is evenly distributed if $\sigma = 1$. Obviously, if the distribution is uniform, hence $M = \sum_{i=1}^{H} b_i/H$, and $\sigma = H \left(\sum_{i=1}^{H} b_i/H \right) \Big/ \sum_{i=1}^{H} b_i = \frac{H}{H} = 1$.

The tests were carried out on a cluster using 3 physical nodes, 10 workers per node, CPU with Intel Xeon Gold 6230N, RAM: 8*32G DDR4 ECC. In distributed computing, the MPI standard (an implementation of MVAPICH3.2) was used, as well as hash table optimization using the **robin_hood** library. Total amount of processed files contains 96.3 GBs of data.

First, consider a subset of tasks with $\sigma \leq 1.6$, as well as tasks whose data contains more than 9000 bytes of information. Further, the authors will call such a set the "common". Figure 2 shows a histogram of acceleration for 30 tasks that form a "common" dataset. It can be seen that for all tasks the acceleration of optimization on the graph exceeds one, which means that there is an acceleration for all "common" tasks. Table 1 below shows detailed execution time and speedup values for each query in the "common" group.

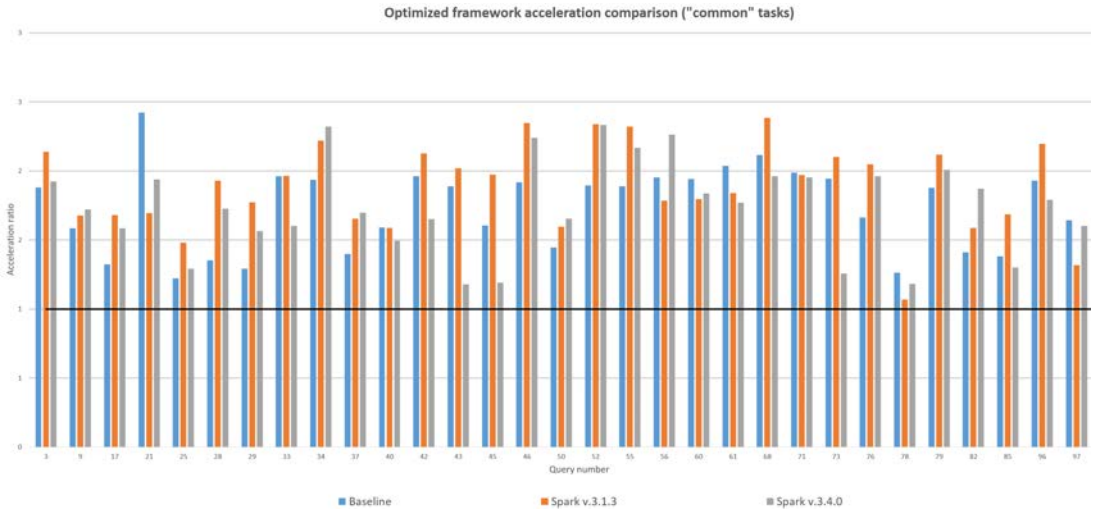


Figure 2: Acceleration comparison ("common" tasks)

Further, we will consider only tasks in which the initial data size does not exceed 9000 bytes ("small" tasks). In total, there are 5 such requests in this set. In Figure 3(a), the one can see that Spark v.3.4.0 is more efficient on such tasks – the speedup in comparison with this framework is always less than one. This is due to the presence of an overhead for processing small tasks in

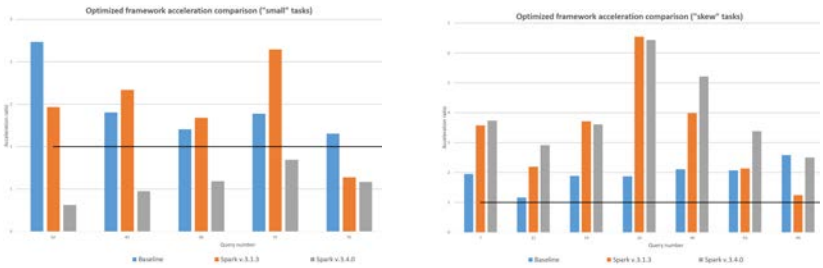
Table 1
Performance comparison (baseline TPC-DS and graph optimization) on "common" tasks

Query #	Execution time (sec.)				Acceleration		
	Spark		Experimental framework		Graph optimization comparison		
	v.3.1.3	v.3.4.0	Baseline	Graph optimized	Baseline	Spark v.3.1.3	Spark v.3.4.0
3	11.549	10.391	10.142	**5.397**	1.879	2.140	1.925
9	110.468	113.433	104.448	**65.856**	1.586	1.677	1.722
17	29.178	27.555	22.998	**17.371**	1.324	1.680	1.586
21	7.366	8.414	10.516	**4.342**	2.422	1.697	1.938
25	27.961	24.433	23.130	**18.905**	1.223	1.479	1.292
28	60.890	54.480	42.698	**31.553**	1.353	1.930	1.727
29	30.995	27.310	22.569	**17.464**	1.292	1.775	1.564
33	21.875	17.840	21.843	**11.132**	1.962	1.965	1.603
34	12.889	13.483	11.241	**5.807**	1.936	2.220	2.322
37	12.881	13.220	10.892	**7.785**	1.399	1.655	1.698
40	12.691	11.933	12.695	**7.987**	1.590	1.589	1.494
42	10.881	8.443	10.039	**5.117**	1.962	2.127	1.650
43	11.307	6.606	10.581	**5.599**	1.890	2.019	1.180
45	8.463	5.100	6.878	**4.287**	1.604	1.974	1.190
46	16.046	15.322	13.122	**6.838**	1.919	2.346	2.241
50	14.948	15.500	13.537	**9.363**	1.446	1.596	1.655
52	12.246	12.208	9.924	**5.234**	1.896	2.340	2.332
55	12.102	11.294	9.848	**5.212**	1.890	2.322	2.167
56	19.955	25.310	21.827	**11.181**	1.952	1.785	2.264
60	20.312	20.766	21.946	**11.302**	1.942	1.797	1.837
61	24.222	23.305	26.807	**13.163**	2.037	1.840	1.770
68	14.999	12.329	13.290	**6.285**	2.115	2.387	1.962
71	21.132	20.955	21.329	**10.722**	1.989	1.971	1.954
73	11.337	6.780	10.490	**5.394**	1.945	2.102	1.257
76	20.363	19.504	16.535	**9.940**	1.664	2.049	1.962
78	45.485	50.313	53.735	**42.582**	1.262	1.068	1.182
79	14.149	13.400	12.529	**6.675**	1.877	2.120	2.008
82	16.493	19.430	14.644	**10.383**	1.410	1.588	1.871
85	7.387	5.690	6.043	**4.377**	1.381	1.688	1.300
96	10.887	8.879	9.567	**4.956**	1.930	2.197	1.792
97	18.353	22.309	22.847	**13.913**	1.642	1.319	1.603
Total	669.810	645.935	618.692	**386.121**	1.879	1.930	1.727

the experimental framework. However, the execution time and acceleration for "small" tasks, generally speaking, is not critical in operation, since their execution time does not exceed 5 seconds on average. However, for such tasks, an acceleration of 40% relative to the base version of the framework was obtained, and also an acceleration of 56% relative to Spark v.3.1.3. The Table 2 shows the detailed results of the study for each request from the group of "small".

Table 2
Performance comparison (baseline TPC-DS and graph optimization) on "small" tasks

Query #	Execution time (sec.)				Acceleration		
	Spark		Experimental framework		Graph optimization comparison		
	v.3.1.3	v.3.4.0	Baseline	Graph optimized	Baseline	Spark v.3.1.3	Spark v.3.4.0
62	4.966	1.055	7.562	**3.383**	2.235	1.468	0.312
83	4.884	1.388	4.107	**2.925**	1.404	1.670	0.475
84	2.412	1.066	2.169	**1.799**	1.206	1.341	0.593
91	2.764	1.088	1.788	**1.289**	1.388	2.145	0.844
95	16.772	15.396	30.326	**26.274**	1.154	0.638	0.586
Total	26.832	18.938	38.390	**32.286**	1.396	1.569	0.534



(a) "Small" tasks (b) "Skewed" tasks

Figure 3: Acceleration comparison for special cases

In this experiment, the focus is on data with skew value $\sigma > 1.6$. This group of data is called "skewed" data, and it is important to note that their distribution in the system is uneven. Results of experiments on such data are shown in Figure 3(b) and Table 3.

Table 3
Performance comparison (baseline TPC-DS and graph optimization) on "skewed" tasks

Query #	Execution time (sec.)				Acceleration			Skewness σ
	Spark		Experimental framework		Graph optimization comparison			
	v.3.1.3	v.3.4.0	Baseline	Graph optimized	Baseline	Spark v.3.1.3	Spark v.3.4.0	
7	21.535	22.499	11.744	**6.016**	1.952	3.580	3.740	1.678
15	19.515	25.943	10.383	**8.905**	1.166	2.192	2.913	1.820
19	21.485	20.841	10.881	**5.773**	1.885	3.722	3.610	1.715
26	30.076	29.551	8.618	**4.591**	1.877	6.550	6.436	1.772
90	8.026	10.507	4.239	**2.014**	2.105	3.985	5.216	1.601
93	21.021	33.303	20.396	**9.844**	2.072	2.135	3.383	1.622
99	7.415	14.950	15.411	**5.959**	2.586	1.244	2.509	1.750
Total	129.073	157.594	81.672	**43.102**	1.952	3.580	3.610	1.899

Thus, the total execution time on Spark v.3.1.2 was 825.715 seconds, and for Spark v.3.4.0, a subset of TPC-DS benchmark queries runs in 809.756 seconds. An optimization for the new version of the Spark framework is the presence of the BloomFilter Join [11] operation, which is the main innovation at the moment. For the experimental framework, the total execution time was 707.250 seconds, which is 15% faster (on average) than running a similar batch on Spark. However, with the presence of graph optimization, a time of 444.793 seconds was obtained. Relative to the experimental framework, a 60% speedup was obtained by building a graph.

5. Conclusion

In this paper, an algorithm for optimizing calculations on a high-performance cluster with a decrease in the number of inter-network interactions was demonstrated, since this is exactly the bottleneck both in HDFS and in many other distributed file systems. This algorithm uses the solution of a two-parameter problem related to minimizing the penalty for data transfer and minimizing the flow in the transport network $\widetilde{G}$.

This optimization significantly affected the execution of all test queries of the TPC-DS benchmark, and therefore we can unambiguously conclude that locality is a key point when working with distributed data. In real systems, one should take into account not only the

distribution of the file, but also the context units (cores, executors, threads) and build a detailed relationship between how the blocks of the file are distributed and how these blocks are executed on physical resources. In future studies, the authors will reflect a more detailed method for solving this problem with the introduction of weights not only on the edges of the graph, but also in the switch device. It is also necessary to develop the idea of a graph approach in a larger number of dimensions, taking into account the features of the system and other factors that affect performance degradation. Experiments have shown that optimizing only I/O operations leads to an increase in the efficiency of the entire framework by 20%, so building more complex models to take into account a different type of operations will qualitatively affect the operation of the system.

References

[1] A. M. Middleton, Hpcc systems: Introduction to hpcc (high-performance computing cluster), 2011.

[2] J. Rivadeneira, F. Garcia-Carballeira, J. Carretero, J. Garcia-Blas, Exposing data locality in hpc-based systems by using the hdfs backend, in: 2020 IEEE 27th International Conference on High Performance Computing, Data, and Analytics (HiPC), 2020, pp. 243–250. doi:10.1109/HiPC50609.2020.00038.

[3] H.-b. Chen, Z. Qiao, S. Fu, Applying sdn based data network on hpc big data computing – design, implementation, and evaluation, in: 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 6007–6009. doi:10.1109/BigData47090.2019.9006039.

[4] H.-b. Chen, S. tang, S. Fu, A middle-ware approach to leverage the distributed data de-duplication capability on hpc and cloud storage systems, in: 2020 IEEE International Conference on Big Data (Big Data), 2020, pp. 2649–2653. doi:10.1109/BigData50022.2020.9377841.

[5] Y. Xie, A. Farhan, M. Zhou, Performance analysis of hadoop distributed file system writing file process, 2018, pp. 116–120. doi:10.1109/ICoIAS.2018.8494199.

[6] W. Inoubli, S. Aridhi, H. Mezni, A. Jung, Big data frameworks: A comparative study (2016).

[7] S. Devulapalli, R. Lakshmi, Performance evaluation of hadoop distributed file system in pseudo distributed mode and fully distributed mode (2015).

[8] K. Wang, X. Zhou, T. Li, D. Zhao, M. Lang, I. Raicu, Optimizing load balancing and data-locality with data-aware scheduling, in: 2014 IEEE International Conference on Big Data (Big Data), 2014, pp. 119–128. doi:10.1109/BigData.2014.7004220.

[9] L. Cheng, Y. Wang, Q. Liu, D. H. Epema, C. Liu, Y. Mao, J. Murphy, Network-aware locality scheduling for distributed data operators in data centers, IEEE Transactions on Parallel and Distributed Systems 32 (2021) 1494–1510. doi:10.1109/TPDS.2021.3053241.

[10] I. Timokhin, A. Teplov, Study of scheduling approaches for batch processing in big data cluster, in: Supercomputing, Springer International Publishing, Cham, 2022, pp. 533–547.

[11] S. Sengupta, A. Rana, Role of bloom filter in analysis of big data, in: 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 6–9. doi:10.1109/ICRITO48877.2020.9197859.

Разработка среды пиктографического программирования киберфизических систем в парадигме машин состояний

Михаил Андреевич Чекан [1]

[1] *Институт динамики систем и теории управления им. В.М. Матросова Сибирского отделения Российской академии наук, Россия, 664033, г. Иркутск, ул. Лермонтова, 134*

Аннотация

В докладе обсуждается текущий статус разработки среды программирования, создаваемой в рамках национальной киберфизической платформы. Актуальность создания обусловлена снижением входного порога для пользователей за счёт применения пиктографического языка и парадигмы иерархических машин состояний. При этом использование последних связано с переходом к новой парадигме проектирования киберфизических систем.

Ключевые слова

Автоматное программирование, киберфизические системы, программная инженерия

1. Введение

Согласно [1], киберфизическая система – это комплексная система физических элементов и их цифровых двойников в вычислительном слое управления, которая постоянно получает данные из окружающей их среды и использует их для оптимизации процессов достижения установленных целей. Киберфизика существует на стыке физики, информационных технологий и математики, и несмотря на отсутствие в настоящий момент массовой известности, охватывает широкий спектр задач от программирования отдельных электронных устройств до проектирования «умных» инфраструктур в энергетике, на транспорте, в сельском хозяйстве и других важных сферах человеческой деятельности.

В текущих экономических условиях актуальным является достижение технологического и кадрового суверенитета. Старт формирования национальной киберфизической платформы (НКФП) вызван прежде всего потребностью в инженерных кадрах, способных проектировать и создавать микроэлектронные и киберфизические системы [2]. Разработчики НКФП планируют решить этот вопрос через популяризацию технического творчества и технологического образования в сфере микроэлектроники и киберфизики, создавая траекторию от школьной программы и технологических кружков до профильных курсов и решения промышленных задач.

Первым массовым решением в рамках НКФП стала платформа «Берлога» [3], представляющая собой набор образовательных игр, существующих в едином сеттинге. В настоящее время доступна игра «Защита пасеки» в жанре Tower Defence, где логика всех игроков задаётся изменяемыми программами, и для прохождения более сложных этапов необходимо совершенствование этих программ. На этой платформе также планируется выход ролевой игры-квеста с набором мини-игр. Главная особенность «Берлоги» – это двусторонняя интеграция достижений в игре и реальной жизни. Принимая участие в инженерных соревнованиях и интенсивах, школьник или студент формирует портфолио на платформе «Талант» Кружкового движения Национальной технологической инициативы (НТИ), которое позволяет открывать новые функции и предметы в играх. Успешное прохождение игры также

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: chekoopa@mail.ru (М.А. Чекан)
ORCID: 0000-0001-7622-421X (М.А, Чекан)

ICCS-DE 2023 Workshop Proceedings

отмечается в «Таланте» и проявляется в профиле пользователя, что играет мотивирующую и профориентационную роль. При этом в технологической плоскости НКФП разрабатываются новые стандарты и инструменты, призванные поднять качество создаваемых систем и доступность их создания более широкому кругу пользователей. Одним из таких инструментов призвана стать, если приводить официальную формулировку, среда графического интерфейса программирования, основанного на диаграммах иерархических машин состояний. Разработка этой среды осуществляется в рамках проекта «Пиктограмма», главная цель которого – создание IDE для разработки киберфизических систем.

2. Среда Lapki IDE

В настоящий момент среда под названием Lapki IDE активно разрабатывается студенческой командой с мая 2023 года под руководством автора. В конце июля 2023 года планируется вывод среды в закрытый бета-тест, а в сентябре – публикация первого релиза на платформе для проектов с открытыми исходным кодом.

В первой итерации Lapki IDE рассчитана на разработку схем для игры «Защита пасеки» на платформе «Берлога», а также для платы Arduino Uno. В дальнейшем планируется расширение списка платформ, в том числе внешними командами и энтузиастами. Одной из таких платформ станет симулятор космических аппаратов «Орбита» [4], используемый в профиле «Спутниковые системы» Национальной технологической олимпиады. Очевидно, в планах также расширение аппаратных платформ, включая основанные на STM32, ESP8266 и др.

Главной, основополагающей чертой среды Lapki IDE является использование визуальной парадигмы программирования на основе расширенных иерархических машин состояний (РИМС). Это промышленный стандарт программирования встраиваемых систем, основанный на UML и широко применяемый в широком классе устройств от датчиков до марсоходов. За рубежом парадигма РИМС распространена благодаря М. Самеку [5] и его компании Quantum Leaps. В отечественной среде существует родственная парадигма автоматного программирования, обозначенная и развиваемая А.А. Шалыто [6]. Применение РИМС обеспечивает низкий порог входа в создание систем для начинающих разработчиков. Продвинутым же пользователям такая парадигма дает возможность создавать сложнейшие алгоритмы, сохраняя простую и прозрачную визуальную структуру. В аппаратных системах управления (например, контроллерах умного дома или роботах) чаще всего требуются не постоянные вычисления, а реакция на внешние события. Поэтому парадигма событийного программирования для таких систем является естественной. Схема РИМС наглядна, поэтому становится легче разрабатывать программу итеративно, последовательно усложняя ее логику и добавляя новые состояния. Диаграмма становится для себя самой элементом документации, что значительно облегчает совместную разработку и поддержку программ.

Кроме этого, в РИМС предлагается применение пиктографического языка, где каждая пиктограмма соотносится с элементом языка целевой платформы. При этом пользователь не ограничен пиктограммами и может использовать сегменты кода, написанные вручную. Так, общие принципы построения диаграмм и их графическое представление сочетаются с возможностью генерировать код на любом традиционном языке программирования.

Проект «Пиктограмма» призван реализовать в среде ряд принципов, способствующих более быстрому входу в разработку с одной стороны, а с другой – повышению осознанности и понимания процесса разработки. Например, принцип «открытого капота» декларирует прозрачность разработки на всех этапах. Это достигается вынесением операций над платформой (будь то этапы компиляции или прошивка устройства) в поле прямого доступа пользователя, а также снабжением интерфейсов настройки, выходных журналов и артефактов «переводом» на более доступный неспециалистам язык. Вдобавок к этому, принцип «автостопом по контроллеру» предполагает методики и документацию как неотъемлемую часть среды. В идеале пользователю не требуется самостоятельно подыскивать обучающие и информационные материалы по целевой платформе, т. к. они доступны непосредственно при разработке.

Принцип «разработка в состоянии потока» обеспечивает комфортное пребывание в среде за счёт тщательной проработки пользовательского опыта (англ., user experience, UX) с целью

сократить число лишних и контринтуитивных действий. Принцип «ориентация на совместную работу» призван способствовать переходу разработчиков, особенно ещё обучающихся, к работе в команде за счёт встроенной поддержки инструментов совместной одновременной разработки. Принцип «модульность и транс-облачность» определяет архитектуру Lapki IDE как клиентское приложение и набор сопровождающих его модулей, размещаемых вместе с клиентом или на удалённом сервере (в облаке). Например, размещение в облаке модуля компиляции позволит распределённой команде производить сборку выходного файла прошивки в одинаковых воспроизводимых условиях, а также снизить требования к рабочей станции. Вынесение модуля загрузчика также позволяет облегчить процесс разработки, убирая необходимость каждому разработчику владеть прошиваемым устройством.

Архитектура разрабатываемой IDE представлена на рисунке 1. Зелёным цветом выделены модули, планируемые в следующей итерации разработки.

Рисунок 1: Схема архитектуры Lapki IDE

3. Заключение

В работе рассмотрен текущий статус разработки среды программирования, создаваемой в рамках национальной киберфизической платформы под руководством автора. В настоящий момент готовы доказательные прототипы (proof-of-concept) всех ядерных модулей: клиент с редактором иерархических машин состояний, модуль компилятора, выполняющий сборку x86-приложений на C++, и загрузчик для Arduino. Ведётся интеграция модулей и подготовка закрытой бета-версии.

4. Список использованных источников

[1] Е.И. Громаков, А.А. Сидорова, Современные технологии. Киберфизические системы, Изд-во Томского политехнического университета, г. Томск, 2021.

[2] Кружковое движение, Владимир Путин поддержал запуск Национальной киберфизической платформы «Восток», 21 июля 2022. URL: https://team.kruzhok.org/news/post/vladimir-putin-podderzhal.

[3] Национальная киберфизическая платформа, 2023. URL: https://platform.kruzhok.org.

[4] Образовательная игра «Орбита», 2023. URL: http://orbitagame.ru

[5] M. Samek, Practical UML statecharts in C/C++: event-driven programming for embedded systems, 2nd. ed., Elsevier Inc., Oxford, UK, 2009.

[6] Н. И. Поликарпова, А. А. Шалыто, Автоматное программирование, Издательство «Питер», г. Санкт-Петербург, 2009.

Сбор и визуализация метеорологических данных, получаемых из научных веб-сервисов

Еделев Ярослав[1]

[1] *МБОУ г.Иркутска СОШ с углубленным изучением отдельных предметов №19, ул. Лермонтова 279, Иркутск, 664033, Россия*

Аннотация

Зачастую в результате научных исследований образуются большие объемы необработанных данных, содержащих временные ряды с набором различных характеристик (параметров). Научные работники-предметники нуждаются в программных средствах для исследования таких временных рядов. Параметры временных рядов варьируются в зависимости от предметной области и, как правило, доступ к ним предоставляется в сервис-ориентированной парадигме. При этом известные средства общего назначения не пригодны для удовлетворения данной потребности, т.к. наборы характеристик и источники данных имеют большое число вариаций, что затрудняет дальнейшие исследования. Возникает необходимость в разработке универсального программного средства, обеспечивающего сбор выбранных временных рядов с последующей их визуализацией анализа. В работе реализовано такое программное средство и протестировано на получении данных из научных веб-сервисов на примере метеостанции.

Ключевые слова

Микросервисы, визуализация, API, базы данных

1. Введение

На данный момент, в сфере научных исследований, возникают большие объёмы необработанных данных. Как следствие, научные работники нуждаются в средствах для исследования полученной информации, однако, получаемые данные имеют большое количество вариаций, что затрудняет их анализ существующими программными решениями.

Целью работы является разработка системы сбора, визуализации и анализа данных из научных веб-сервисов на примере метеостанции.

В рамках поставленной цели были определены следующие задачи:

1. Анализ литературы
2. Выбор средств реализации
3. Написание программного обеспечения
4. Тестирование программного обеспечения

2. Реализация проекта

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: yarvaleev07@bk.ru

ICCS-DE 2023 Workshop Proceedings

2.1. Выбор средств реализации

В качестве основного средства реализации проекта был выбран язык программирования Python, т.к. он обладает хорошей документацией, простым синтаксисом и поддерживает библиотеки, необходимые для работы с API.
Средой разработки являлся Jupyter notebook, т.к. данная среда позволяет запускать и тестировать несколько программ одновременно, а так же имеет встроенную библиотеку виджетов, необходимых для создания интерфейса приложения.
Данные запрашивались через API при помощи библиотеки Requests.
Для систематизации информации была использована база данных на основе SQLite и библиотека взаимодействия с ней sqlite3, т.к. SQLite является простым в управлении и изменении.

Для визуализации полученных данных была использована библиотека Matplotlib, т.к. она обладает широким функционалом и подробной документацией.

2.2. Написание программного обеспечения

1. Программа обновления базы данных состоит из 3 блоков. В 1 блоке, к API посылается запрос значений за последнюю неделю. Запрос состоит из 2 частей- URL источника и параметров запроса, в которые включен временной отрезка запрашиваемых данных.
   Далее, во втором блоке, полученные значения программа преобразует в массив словарей. В 3 блоке, через метод REPLACE алгоритм загружает новые значения в базу данных. Если таковые уже существуют, программа обновляет имеющиеся записи, но не добавляет новые.

2. Программа создания графиков из имеющихся значений состоит из 8 блоков. В первом блоке, происходит импорт библиотек, далее- объявление и вывод основных элементов интерфейса. После, с помощью метода SELECT из базы данных программа получает массив значений за временной период, выбранный пользователем. После, полученные значения соединяются со своими обозначениями в словарь для более удобной обработки. В 5 блоке, программа, согласно выбору пользователя, отсеивает значения и формирует из них массив переменных, по которым будет строиться график. В зависимости от выбора пользователя, график будет построен с 1 переменной, или 2. В случае с 1 переменной, программа создает координатную плоскость и далее строит на ней график. В случае с 2 переменными, программа создает координатную плоскость с 2 независимыми осями Y и создает график. Далее, график выводится на экран (рис. 1). Согласно выбору пользователя, полученный график программа может сохранить в формате PNG.

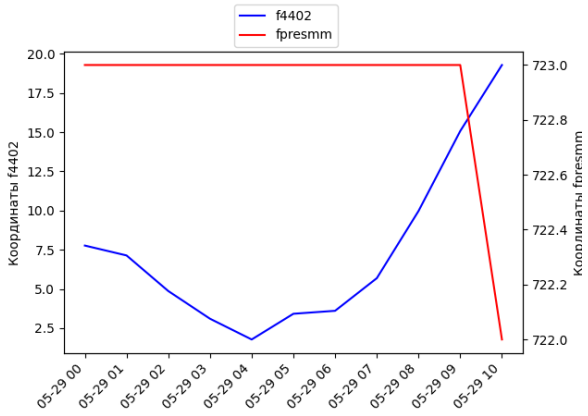


Рисунок 1: Результат работы программы создания графиков из имеющихся значений

3. Программа конструктора запросов к API и визуализации полученных данных состоит из 10 блоков. В первых двух блоках, программа импортирует библиотеки и выводит основные элементы интерфейса. Далее, через API, программа конструирует запрос с тем URL источника и диапазоном дат, что указал пользователь. Т.к. API метеостанции не может передать за раз более 10 000 значений, программа получает данные пакетами по одному дню. Далее, JSON-пакет преобразуется в массив, в котором, помимо необходимых значений, могут содержаться лишние данные, которые после отсеиваются. Из-за специфики получения данных- небольшими пакетами, в полученном массиве могут содержаться повторяющиеся элементы, которые удаляются пятым блоком. После, из массива значений программа получает список дат и сохраняет их в отдельную переменную. Далее, алгоритм выводит пользователю интерфейс выбора интересующих его значений из предложенных. После выбора, программа формирует массив значений координат и создает график.

3. Выводы

В результате проделанной работы, было создано средство систематизации и визуализации данных, полученных из научных веб-сервисов. Созданное программное решение позволяет собирать данные и визуализировать их в виде линейного графика или гистограммы. В отличие от существующих продуктов, созданные программы позволяют визуализировать данные из разных источников с отличающимся количеством переменных и параметров, что обеспечивает универсальность продукта.

Литература

[1] Ipywidgets.readthedocs.io, Jupyter Widgets 8.0.6 documentation. URL: https://ipywidgets.readthedocs.io/en/latest/.
[2] Matplotlib.org, Matplotlib 3.7.1 documentation. URL: https://matplotlib.org.
[3] Работа с датой и временем в Python. URL: https://python-scripts.com/datetime-time-python.
[4] Василий Волгин, Про API, Rest API для начинающего тестировщика. Какой запрос быстрее?, 2023 URL: https://habr.com/ru/articles/734900/.
[5] Руководство по SQLite в Python. URL: https://pythonru.com/osnovy/sqlite-v-python.
[6] Requests в Python – Примеры выполнения HTTP запросов. URL: https://python-scripts.com/requests.

Разработка экосистемы “умного” дома: программная и аппаратная составляющие

Данила Кашкарев

Лицей ИГУ, ул. Академика Курчатова 13а, Иркутск, 664074, Россия

Руководители
Костромин Роман Олегович, кандидат технических наук;
Рудых Александр Николаевич, ПДО МБУДО г. Иркутска ЦДТТ;
Лавлинский Максим Викторович, учитель информатики МАОУ Лицей ИГУ г. Иркутска.

1. Тема

Создание самодельной обособленной *экосистемы*[2] для удобного использования домашних *IoT*[4] устройств.

2. Цель

Изучить открытые стандарты передачи информации, организовать грамотный процесс *логинизации*[1] и реализовать на их основе прототип программно-аппаратной *экосистемы*[2] умного дома на примере взаимодействия клиентов как с микро-сервисами, так и с собственным медиа-устройством.

3. Задачи

3.1. Изучить существующие разработки и технологии по организации умного дома, выделить их достоинства и недостатки,
3.2. Провести анализ существующих аппаратных систем,
3.3. Определить способы устранения критических недостатков, сформулировать свои предложения,
3.4. Спроектировать свою установку (аппаратная часть),
3.5. Спроектировать программную и сетевую части,
3.6. Реализовать серверную часть (виртуализация, Linux-сервер, контейнеры и изолированная сеть),
3.7. Реализовать программную и аппаратную части на базе созданной сети,
3.8. Интегрировать готовое устройство в систему,
3.9. Провести тестирование созданных систем,
3.10. Проанализировать полученные результаты,

4. Проблемы

4.1. Безопасный доступ к микро-сервисам из внешней сети: многие люди хотят подключаться к своему серверу из вне, но они не задумываются о правильном процессе идентификации, аутентификации, авторизации.
4.2. Удобное устройство, которое легко интегрируется в *экосистему*[2] дома: на рынке “умных” устройств присутствует большое количество медиа-установок, но они не предоставляют открытого *API*[3] для взаимодействия с ними.

5th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2023), July 03–07, 2023, Irkutsk, Russia
EMAIL: danila.kashkarev@yandex.ru (A. 1);

ICCS-DE 2023 Workshop Proceedings

4.3. Любой клиент в домашней сети может получить доступ к микро-сервису, который не имеет защиты от brute force: Многие серверные программы и компоненты умного дома обладают посредственной системой логинизации[1], которой может не быть вовсе.
4.4. Данные хранятся на чужих серверах: китайские устройства транслируют через себя видео с камер и служебную информацию от устройств автоматизации
4.5. Низкая отказоустойчивость: все вычислительные операции зависят от тонкого кабеля, который уходит к провайдеру.

5. Решения

5.1. Для грамотного подключения к микро-сервисам нужно использовать связку из Authentik (сервер идентификации, аутентификации и авторизации) и Nginx Reverse Proxy (обратный proxy-сервер). Обеспечивают защиту серверных программ, которые не предусматривают систему логинизации. Authentik выступает в роли *SSO* [5] (системы единого входа), он помогает облегчить процесс входа для пользователя. Nginx же перенаправляет через себя весь трафик, он является пограничным элементом сети.
5.2. Многие создатели “умного” дома хотят добавить голосовое управление через станции от Яндекса или Мэйла, но их проприетарные ОС не предоставляют эту возможность. Поэтому мы использовали открытую ОС OSMC, которая основывается на Debian.
5.3. Для защиты IoT устройств, серверов и пользователей сеть разбивается на зоны. Зонирование сети позволяет настроить правила в *firewall`a* [6] для разграничения доступа между клиентами. Основой сети является программируемый *Nettop* [7] (мрашутизатор) под управлением Pfsense.
5.4. Пользователи не хотят делиться секретной информацией с “третьей стороной”, так как данная информация может быть конфиденциальной, поэтому мы решили использовать собственный сервер.
5.5. В эпоху *SaaS* [8] люди поняли, что множество зависимостей как финансовых, так и технических ведут к низкой отказоустойчивости домашней системы. Поэтому мы концентрируем вычислительные процессы внутри сети.

6. Сравнение возможностей устройства с аналогами

№	Характеристика	Собственная установка	Яндекс станция МАКС
1	Звук	20w + 20w средние частоты, 10 w + 10w высокие частоты	40w низкие частоты, 10w + 10w средние частоты, 15w + 15w высокие частоты
2	Матрица	64x128px, RGB LED	25x16 px, монохромный, белый
3	Одноплатный компьютер	4x1,5 Ггц ARM cortex-a72, RAM 8GB, GPU Video Core VI 500МГц, H.265	4x1,8 Ггц ARM cortex-a52, RAM 2GB, GPU Mali-G31 MP2, HEVC
4	ПО	OSMC Linux, полностью открытое	Yandex Linux, закрыта
5	Подписки	Отсутствуют	Обязательны

7. Вывод

Работая над проектом, были изучены концепции виртуализации и контейнеризации. Мы поработали с 3D принтером, который дал нам возможность напечатать собственный корпус для станции. Продумывая внутреннюю электронику, мы познакомились с миром одноплатных компьютеров. Для создания грамотной сети мы познакомились с сетевым оборудованием.

8. Термины

Логинизация [1] – процесс входа в некоторый ресурс сети интернет.
Экосистема [2] – единая сеть устройств, участвующих в работе домашнего медиа пространства.

API [3] (application programming interface) – описание способов общения ПО с данной программой.
IoT [4] (internet of things) – идея общения электронных устройств в сети
SSO [5] (single sign on) – система единого входа
Firewall [6] – межсетевой экран, компонент защищенной сети
Nettop [7] – небольшой персональный компьютер
SaaS [8] (software as a service) – идея предоставления программного обеспечения в виде услуги

9. Литература

1. Кирченко П. Г. Электроника для начинающих, 2019 г.
2. Айсберг Е. Транзистор… Это очень просто, 1976 г.
3. Клайн Л. Fusion 360. 3D-моделирование для мейкеров, 2021 г.
4. Керриск М. Linux API, 2019 г.
5. Флёнов М. Е. Linux глазами хакера, 2018 г.
6. Тарренбаум Э. С., Уэзеролл Д. Компьютерные сети, 2022 г.

Научно-организационный отдел
Федерального государственного бюджетного учреждения науки
Института динамики систем и теории управления имени В.М. Матросова
Сибирского отделения Российской академии наук
664033, Иркутск, ул. Лермонтова, д. 134
E-mail: rio@icc.ru

Подписано к печати 23.08.2023 г.
Формат бумаги 60×84 1/16, объем 7,79 п.л.
Заказ 1. Тираж 1 экз.

Отпечатано в ИДСТУ СО РАН