ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

Information systems.
Computer sciences.
Issues of information security

Multiple robots (robotic centers) and systems.
Remote sensing and nondestructive testing

Modern radio engineering and telecommunication systems

Micro- and nanoelectronics.
Condensed matter physics

Analytical instrument engineering and technology

Mathematical modeling

Economics of knowledge-intensive and high-tech enterprises and industries.
Management in organizational systems

Product quality management. Standardization

Philosophical foundations of technology and society

12+
www.rtj-mirea.ru

12(5) 2024

ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

- Information systems. Computer sciences. Issues of information security
- Multiple robots (robotic centers) and systems. Remote sensing and nondestructive testing
- Modern radio engineering and telecommunication systems
- Micro- and nanoelectronics. Condensed matter physics
- Analytical instrument engineering and technology
- Mathematical modeling
- Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems
- Product quality management. Standardization
- Philosophical foundations of technology and society

- Информационные системы. Информатика. Проблемы информационной безопасности
- Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля
- Современные радиотехнические и телекоммуникационные системы
- Микро- и наноэлектроника. Физика конденсированного состояния
- Аналитическое приборостроение и технологии
- Математическое моделирование
- Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах
- Управление качеством продукции. Стандартизация
- Мировоззренческие основы технологии и общества

Russian Technological Journal
2024, Vol. 12, No. 5

Russian Technological Journal
2024, том 12, № 5

https://www.rtj-mirea.ru

https://doi.org/10.32362/2500-316X

ISSN 2500-316X (Online)

Russian Technological Journal
2024, Vol. 12, No. 5

Publication date September 27, 2024.

The peer-reviewed scientific and technical journal highlights the issues of complex development of radio engineering, telecommunication and information systems, electronics and informatics, as well as the results of fundamental and applied interdisciplinary researches, technological and economical developments aimed at the development and improvement of the modern technological base.

Periodicity: bimonthly.
The journal was founded in December 2013. The titles were «Herald of MSTU MIREA» until 2016 (ISSN 2313-5026) and «Rossiiskii tekhnologicheskii zhurnal» from January 2016 until July 2021 (ISSN 2500-316X).

Founder and Publisher:
Federal State Budget
Educational Institution of Higher Education
«MIREA – Russian Technological University»
78, Vernadskogo pr., Moscow, 119454 Russia.

The journal is included into the List of peer-reviewed science press of the State Commission for Academic Degrees and Titles of Russian Federation. The Journal is included in Russian State Library (RSL), Russian Science Citation Index, eLibrary, Socionet, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Editor-in-Chief:
Alexander S. Sigov, Academician at the Russian Academy of Sciences, Dr. Sci. (Phys.–Math.), Professor, President of MIREA – Russian Technological University (RTU MIREA), Moscow, Russia.
Scopus Author ID 35557510600, ResearcherID L-4103-2017, sigov@mirea.ru.

Editorial staff:
Managing Editor — Cand. Sci. (Eng.) Galina D. Seredina
Scientific Editor — Dr. Sci. (Eng.), Prof. Gennady V. Kulikov
Executive Editor — Anna S. Alekseenko
Technical Editor — Darya V. Trofimova
86, Vernadskogo pr., Moscow, 119571 Russia.
Phone: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

The registration number ПИ № ФС 77 - 81733 was issued in August 19, 2021 by the Federal Service for Supervision of Communications, Information Technology, and Mass Media of Russia.

The subscription index of *Pressa Rossii*: **79641**.

Russian Technological Journal
2024, том 12, № 5

Дата опубликования 27 сентября 2024 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радиотехнических, телекоммуникационных и информационных систем, электроники и информатики, а также результаты фундаментальных и прикладных междисциплинарных исследований, технологических и организационно-экономических разработок, направленных на развитие и совершенствование современной технологической базы.

Периодичность: один раз в два месяца.
Журнал основан в декабре 2013 года. До 2016 г. издавался под названием «Вестник МГТУ МИРЭА» (ISSN 2313-5026), а с января 2016 г. по июль 2021 г. под названием «Российский технологический журнал» (ISSN 2500-316X).

Учредитель и издатель:
федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет»
119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Журнал входит в Перечень ведущих рецензируемых научных журналов ВАК РФ, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук и доктора наук, индексируется в РГБ, РИНЦ, eLibrary, Соционет, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Главный редактор:
Сигов Александр Сергеевич, академик РАН, доктор физ.-мат. наук, профессор, президент ФГБОУ ВО МИРЭА – Российский технологический университет (РТУ МИРЭА), Москва, Россия.
Scopus Author ID 35557510600, ResearcherID L-4103-2017, sigov@mirea.ru.

Редакция:
Зав. редакцией — к.т.н. Г.Д. Середина
Научный редактор — д.т.н., проф. Г.В. Куликов
Выпускающий редактор — А.С. Алексеенко
Технический редактор — Д.В. Трофимова
119571, г. Москва, пр-т Вернадского, 86, оф. Л-119.
Тел.: +7 (499) 600-80-80 (#31288).
E-mail: seredina@mirea.ru.

Регистрационный номер и дата принятия решения о регистрации СМИ ПИ № ФС 77 - 81733 от 19.08.2021 г. СМИ зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Индекс по объединенному каталогу «Пресса России» **79641**.

https://www.rtj-mirea.ru

Editorial Board

Редакционная коллегия

Contents

Содержание

Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 025.4.03
https://doi.org/10.32362/2500-316X-2024-12-5-7-16
EDN CBEERK

RESEARCH ARTICLE

Automating the search for legal information in Arabic: A novel approach to document retrieval

Kamel S. Jafar [1, @], Ali A. Mohammad [2], Ali H. Issa [3], Alexander V. Panov [1, @]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *Higher School of Economics, Moscow, 109028 Russia*
[3] *Russian Biotechnological University, Moscow, 125080 Russia*
[@] *Corresponding authors, e-mail: zhafar.k@edu.mirea.ru, panov_a@mirea.ru*

Abstract

Objectives. The retrieval of legal information, including information related to issues such as punishment for crimes and felonies, represents a challenging task. The approach proposed in the article represents an efficient way to automate the retrieval of legal information without requiring a large amount of labeled data or consuming significant computational resources. The work set out to analyze the feasibility of a document retrieval approach in the context of Arabic legal texts using natural language processing and unsupervised clustering techniques.

Methods. The Topic-to-Vector (Top2Vec) topic modeling algorithm for generating document embeddings based on semantic context is used to cluster Arabic legal texts into relevant topics. We also used the HDBSCAN density-based clustering algorithm to identify subtopics within each cluster. Challenges of working with Arabic legal text, such as morphological complexity, ambiguity, and a lack of standardized terminology, are addressed by means of a proposed preprocessing pipeline that includes tokenization, normalization, stemming, and stop-word removal.

Results. The results of the evaluation of the approach using a dataset of legal texts in Arabic based on keywords demonstrated its superior effectiveness in terms of accuracy and memorability. The proposed approach provides 87% accuracy and 80% completeness. This circumstance can significantly improve the search for legal documents, making the process faster and more accurate.

Conclusions. Our findings suggest that this approach can be a valuable tool for legal professionals and researchers to navigate the complex landscape of Arabic legal information to improve efficiency and accuracy in legal information retrieval.

Keywords: search for documents, NLP, Top2Vec, HDBSCAN, Arabic legal documents, word embeddings, cosine similarity

• **Submitted:** 05.05.2023 • **Revised:** 04.04.2024 • **Accepted:** 11.07.2024

For citation: Jafar K.S., Mohammad A.A., Issa A.H., Panov A.V. Automating the search for legal information in Arabic: A novel approach to document retrieval. *Russ. Technol. J.* 2024;12(5):7–16. https://doi.org/10.32362/2500-316X-2024-12-5-7-16

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Автоматизация поиска юридической информации на арабском языке: подход к поиску документов

Камел С. Жафар [1, @],
Али А. Мохаммад [2],
Али Х. Исса [3],
А.В. Панов [1, @]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *Национальный исследовательский университет «Высшая школа экономики», Москва, 109028 Россия*
[3] *РОСБИОТЕХ – Российский биотехнологический университет, Москва, 125080 Россия*
[@] *Авторы для переписки, e-mail: zhafar.k@edu.mirea.ru, panov_a@mirea.ru*

Резюме

Цели. Поиск юридической информации, например, информации, связанной с различными юридическими вопросами, такими как наказание за преступления, является сложной задачей. Предлагаемый авторами подход может быть эффективным и действенным способом автоматизации поиска юридической информации без необходимости использования большого количества размеченных данных или значительных вычислительных ресурсов. Целью статьи является анализ возможности использования подхода к поиску документов в контексте юридических текстов на арабском языке, с применением методов обработки естественного языка и неконтролируемой кластеризации.

Методы. Использован подход Top2Vec – алгоритм моделирования темы, который создает вложения документов на основе семантического контекста, чтобы группировать юридические тексты на арабском языке в соответствующие темы. Использован алгоритм кластеризации на основе плотности для определения подтем внутри каждого кластера. Решаются проблемы работы с арабским юридическим текстом, такие как морфологическая сложность, двусмысленность и отсутствие стандартизированной терминологии. Предложен конвейер предварительной обработки, включающий в себя токенизацию, нормализацию, выделение корней и удаление стоп-слов.

Результаты. Результаты оценки подхода с использованием набора данных юридических текстов на арабском языке, основанного на ключевых словах, показали его эффективность и превосходство с точки зрения точности и запоминаемости. Предлагаемый подход обеспечивает точность поиска – 87% и полноту поиска – 80%. Применение этого подхода может значительно улучшить поиск юридических документов, сделав его более быстрым и точным.

Выводы. Предложенный подход может быть ценным инструментом для юристов и исследователей, которым необходимо ориентироваться в обширном и сложном ландшафте арабской юридической информации, повышая эффективность и точность ее поиска.

Ключевые слова: поиск документов, обработка естественного языка, Top2Vec, алгоритм кластеризации на основе плотности, арабские юридические документы, вложения слов, косинусное сходство

• **Поступила:** 05.05.2023 • **Доработана:** 04.04.2024 • **Принята к опубликованию:** 11.07.2024

Для цитирования: Жафар К.С., Мохаммад А.А., Исса А.Х., Панов А.В. Автоматизация поиска юридической информации на арабском языке: подход к поиску документов. *Russ. Technol. J*. 2024;12(5):7–16. https://doi.org/10.32362/2500-316X-2024-12-5-7-16

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The search and analysis of legal information is associated with certain difficulties due to specificities pertaining to the field of law. In recent years, there has been a growing interest in using natural language processing techniques to automate the process of accessing such information [1]. In particular, significant efforts have been expended to develop question and answer systems that can extract specific answers from legal documents [2, 3]. However, building a good QA system requires a large amount of marked-up data and is often computationally intensive. The present work proposes an alternative approach to automate the retrieval of legal information related to crimes, including criminal offenses, in regulations and legal documents in Arabic. In practice, it is often sufficient to retrieve the most relevant legal documents related to the user's query without extracting specific answers from the documents. The proposed approach consists of several steps including data collection, data preprocessing, document indexing, query processing, and document retrieval. The source dataset comprises standard Arabic grammatical and legal documents related to crimes and felonies.

1. LITERATURE REVIEW

Legal information retrieval is a field with a rich history and an extensive body of research. In the present section, key works and developments in this field are reviewed with a special focus on approaches to legal information retrieval and document retrieval in Arabic. Sansone et al. investigate state-of-the-art Artificial Intelligence (AI) techniques used for legal information retrieval systems [4]. With the advent of information and communication technologies, legal practitioners have faced a dramatic increase in digital information, which makes efficient retrieval techniques essential. This article discusses various approaches to AI, including natural language processing, machine learning, and knowledge extraction techniques, explaining how they can aid legal information retrieval. As well as describing the challenges faced by legal practitioners, especially when searching for similar cases, statutes, or paragraphs, the authors discuss open questions regarding legal information retrieval systems. Overall, the study emphasizes the importance of AI in the legal field and the need for continued research and development in legal information retrieval systems. Sartor et al. review nine significant sources published in the last decade [5]. Four of the articles focus on analyzing legal cases, introducing contextual considerations, predicting outcomes from natural language descriptions of cases, comparing different ways of presenting cases, and formalizing precedential reasoning. One article introduces the argumentation scheme method for analyzing arguments, which has subsequently become very widely used in AI and law. Two of the reviewed articles refer to ontologies for representing legal concepts, while the remaining two take advantage of the increasing availability of legal datasets in this decade to automate document summarization and arguments retrieval.

Zhong et al. present an overview of the development, current state and future directions of legal AI [6]. Legal AI applies natural language processing to assist lawyers in their work with the potential to increase efficiency by automating repetitive tasks. Illustrating perspectives of lawyers and natural language processing researchers through experimentation and analysis of existing work, the authors identify knowledge modeling, legal reasoning, and interpretability as three major problems in legal tasks that require further research. As a means of addressing these challenges, the paper proposes a combination of embedding-based and character-based methods to create large-scale and high-quality datasets, as well as to address ethical concerns such as gender bias and racial discrimination. Ultimately, legal AI should play a supporting role in the legal system, with professionals making decisions first and foremost. Shamma et al. describe the development of an Arabic system for extracting information from legal documents, which utilizes a hybrid approach of machine learning and rule-based methods [7]. The system is designed to extract important information from documents and present it in a structured form for complex queries. The described approach, which has been tested on a limited class of Arabic legal documents, has demonstrated good results. The authors

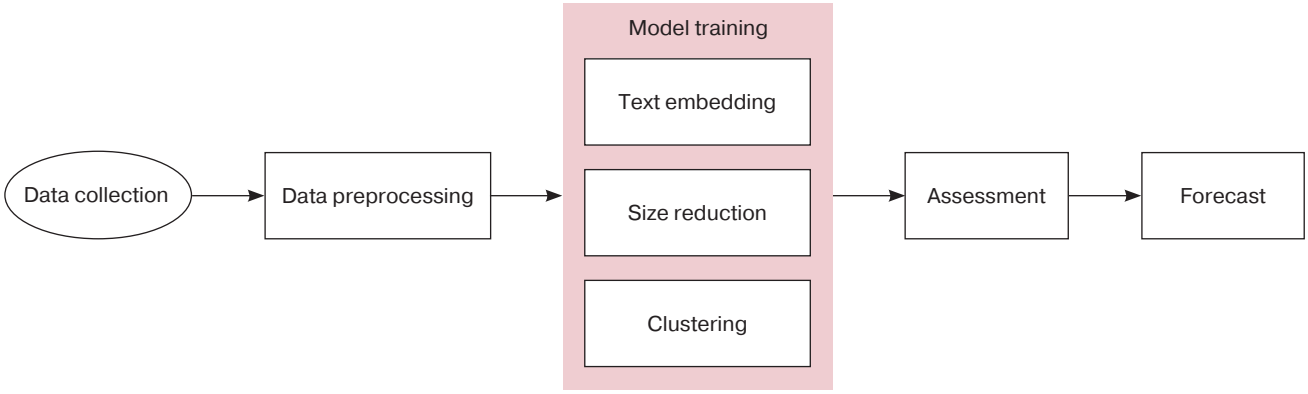


Fig. 1. Top2Vec approach to legal information retrieval

suggest several possible extensions to the system, such as considering different types of cases in a given legal system, using more advanced Arabic natural language processing techniques tools, exploring the use of deep learning, extracting more relationships, improving the presentation of results, and extending the system to other domains such as healthcare and finance.

2. PROPOSED APPROACH

The proposed Topic-to-Vector (Top2Vec) methodology for legal information retrieval comprises several steps (Fig. 1). The first step is data preparation, which involves collecting legal documents related to crimes and criminal offenses described in Arabic legal documents. In order to prepare it for the model, the data is then preprocessed using a pipeline that includes tokenization, normalization, root extraction, and stop word removal.

The next step is the model training, which involves creating document embeddings using the Top2Vec model. The embeddings are then reduced to a lower dimensional space using dimensionality reduction techniques such as a uniform manifold approximation and projection (UMAP)[1] in order to simplify document clustering. Clustering is performed using a density-based clustering algorithm such as the hierarchical density-based spatial clustering of applications with noise (HDBSCAN), which can efficiently identify document clusters based on their similarity [8].

Once the clustering is complete, the model can be evaluated with a set of real user queries. For each query, the model retrieves the most relevant document clusters, allowing the user to browse through the documents to find the relevant information. The performance of the model can be evaluated using metrics such as accuracy and completeness.

Finally, the model can be used for prediction, where the user enters a query and the model extracts the most relevant clusters of documents. The user then browses through the documents to find the relevant information. The prediction can be repeated for different queries, and the model can be continuously updated with new data to improve its performance. Overall, this methodology represents an efficient and effective means to automate the retrieval of legal information related to felonies, including criminal offenses, in Arabic legal documents.

2.1. Top2Vec approach

Top2Vec [9, 10] is a novel unsupervised document clustering and topic modeling technique that can discover topics in large-scale datasets without any prior knowledge of their number. The basic idea of Top2Vec is to first embed documents and topics in the same space and then cluster the embedded documents using a density-based clustering algorithm [11]. Top2Vec can also automatically determine the number of topics, allowing the topics to be represented as a set of words and a connected vector. The algorithm outperforms traditional topic modeling methods such as latent Dirichlet distribution and nonnegative matrix factorization, both in terms of clustering quality and scalability for large datasets. The potential of Top2Vec has been demonstrated in various domains including document similarity search, visualization, and anomaly detection.

Figure 2 shows an example of a semantic space. The purple dots are documents, while the green dots represent words. Words are closest to the documents they best represent; similar documents are presented close to each other.

[1] Uniform manifold approximation and projection is a machine learning algorithm that performs nonlinear dimensionality reduction.

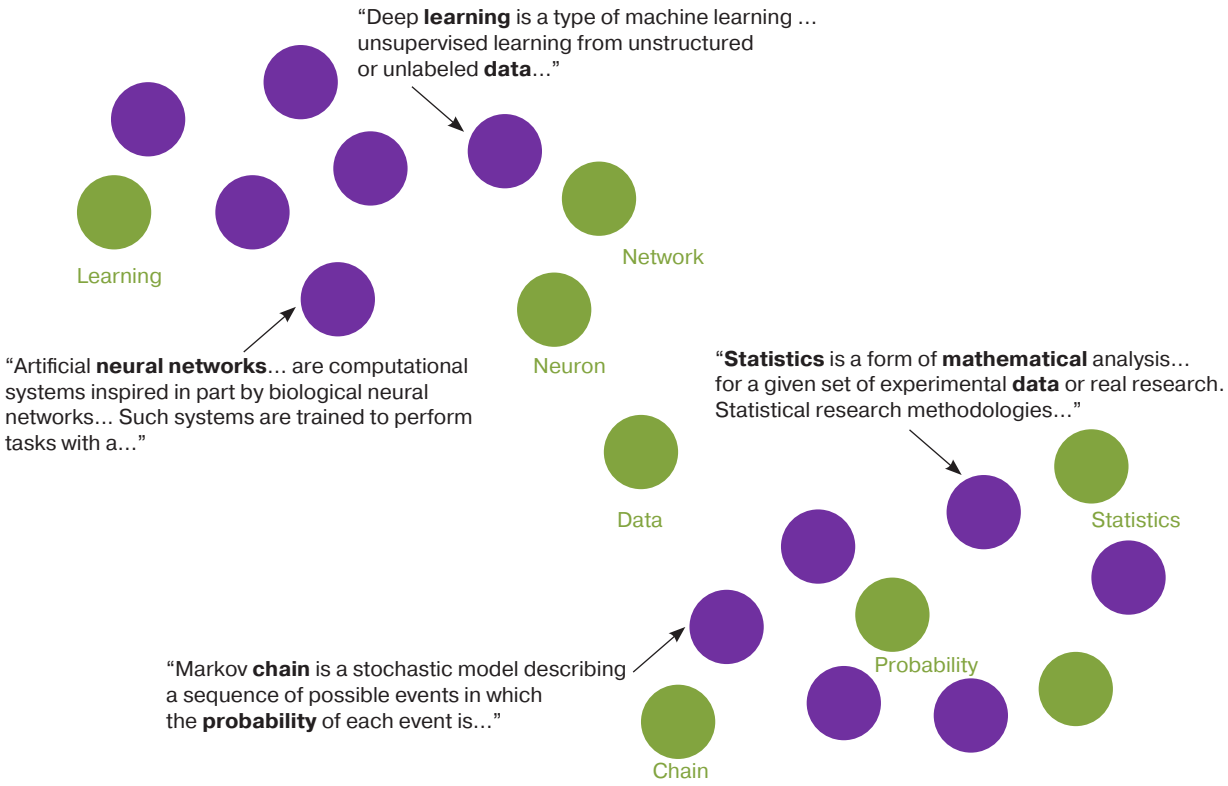


Fig. 2. Example of a semantic space

2.2. HDBSCAN clustering algorithm

The HDBSCAN application [12, 13] is a clustering algorithm that can identify clusters of different densities and shapes in multidimensional spaces. It uses a hierarchical approach to build a hierarchy of clusters and automatically determine the number of clusters. One of the advantages of HDBSCAN is that it can handle clusters of different sizes and shapes and identify noise points. The algorithm has several hyperparameters that can be tuned to optimize the clustering results. Some of the key hyperparameters are as follows:

- Minimum cluster size: this parameter sets the minimum number of points required to form a cluster. Increasing this parameter increases the number of clusters and decreases the number of clusters.
- Metric: this parameter defines the distance metric used to calculate the similarity between data points. Depending on the characteristics of the data, different metrics can be used.
- Cluster selection method: this parameter defines how the final set of clusters is selected from the hierarchy.

In the process of training the Top2Vec model, the HDBSCAN algorithm was used with the following hyperparameters: minimum cluster size—3, metric—Euclidean, cluster selection method—leaf.

These hyperparameters are selected based on the characteristics of the available dataset and tuned to optimize the clustering results. The Euclidean metric is a commonly-used metric for measuring the distance between points in multidimensional spaces; the leaf cluster selection method is suitable for large datasets due to providing a good balance between speed and accuracy. A minimum cluster size of 3 was chosen to prevent the formation of small clusters and the inclusion of noisy points.

The resulting visualization (Fig. 3) displays the clusters in two-dimensional space to permit a visual inspection of the relation of documents in each cluster. This approach is useful for exploring and navigating the complex landscape of legal information, helping legal professionals and researchers quickly identify relevant documents based on their content.

2.3. Data collection and preprocessing

In this phase, regulatory and legal documents in Arabic relating to various crimes are collected and pre-processed. The standard set of legal documents includes different types of legal texts, i.e., laws and regulations. These documents are collected in a single

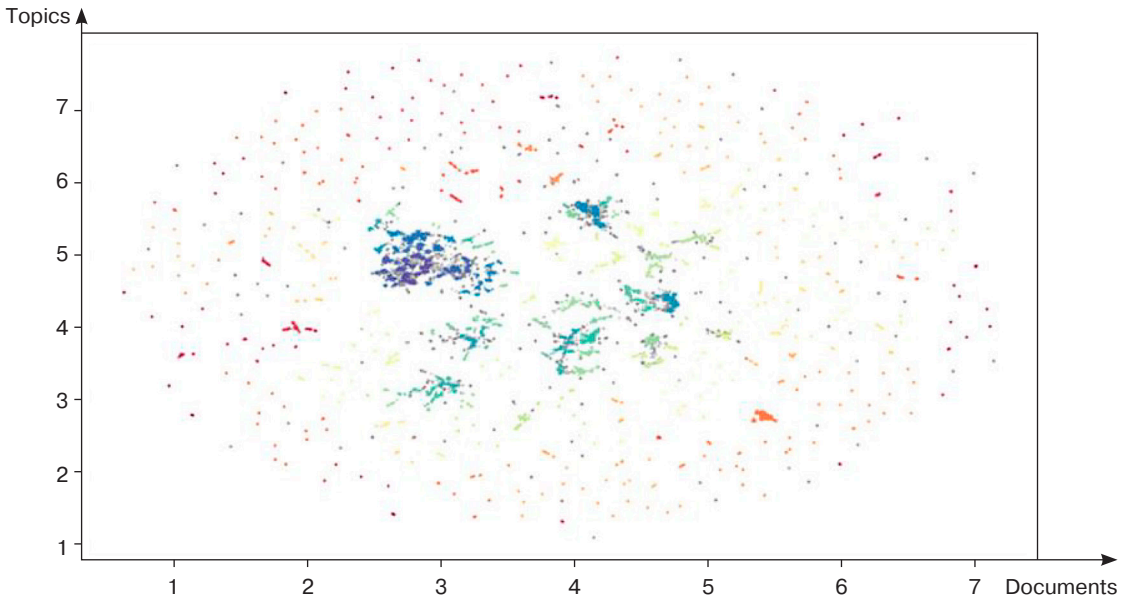


Fig. 3. Retrieving dense document regions using spatial clustering of applications based on hierarchical density with noise

csv file referred to as a corpus, in which each document represents one record (line).

Each row has several basic keys required for data organization:

1. Index: a unique sequential number for each document.
2. Identifier: a unique string for each document that represents a regulatory or legal document, serving as the actual identifier as it is referenced in the original source document.
3. Title: string name of the regulatory or legal document.
4. Summary: string brief description of the content of the regulation or document.
5. Details: string full description and content of the document as it appears in the original source.

For the purpose of supervised learning, the model is only trained on the data that are in the Details column from the corpus. These are then assembled into a single list, which is preprocessed to remove stop words, punctuation marks and diacritics, as well as to perform root detection and normalization.

Preprocessing stage includes the following sub-stages:

1. Tokenization: this step involves breaking the text into individual words or tokens. This is an important first step in any text processing pipeline since most algorithms work with individual words or tokens rather than whole sentences or documents. Tokenization can be performed using various methods: space-based tokenization, regular expressions, and rule-based tokenization.
2. Stop word removal. Stop words are common words that are often removed from a text because they do not make much sense and their presence may reduce the accuracy of text analysis. Examples of stop words are prepositions, conjunctions, and articles (and, the, of, in, etc.). By removing stop words, it is possible to reduce the dimensionality of the data and increase the efficiency of subsequent text processing steps.
3. Grounding (stemming). This step involves bringing words back to their root form, also known as the base form. The goal of stemming is to group similar words together, even if they are nonidentical. Stemming can be performed using Porter's algorithm, Snowball algorithm and Lancaster's algorithm.
4. Normalization. This step involves converting words to a standardized form to ensure consistency and reduce redundancy in the text. Examples of normalization include converting all words to lower case, converting all numbers to digits, and removing punctuation marks. By helping reduce noise in the data, normalization improves the accuracy of subsequent text processing steps.

Following this step, the necessary dataset of preprocessed documents is obtained for training the model.

2.4. Model training

Training the Top2Vec model involves several parameters that need to be set based on the characteristics of the dataset and the desired result. In this work, the Top2Vec model was trained using the following parameters:

- Documents: corpus of preprocessed documents for clustering.
- min count: the minimum number of times a word must appear in order to be included in the model dictionary.
- Embedding model: a type of embedding model used to create document attachments. In this case, the Document-to-Vector (Doc2Vec) model was used. This represents an unsupervised deep learning model that can learn vector representations of documents to enable efficient similarity computation and topic modeling.
- To split the documents: a Boolean value indicating whether the documents should be split for processing. In this work, the value "split documents" is true.
- Document blocker: a method used to separate documents. In this case, a sequential fragmenter was used. This is a method of dividing documents into smaller sequential fragments of fixed length to facilitate training and processing of machine learning models.
- Fragment length: the maximum length of each document fragment. In this case, the block length is set as 5, i.e., each document is split into blocks of 5 sentences. This parameter controls the level of detail of the splitting process: smaller values result in smaller fragments and potentially more detailed topics, while larger values result in larger fragments and potentially more general topics.
- Maximum number of fragments for each document. In this case, the maximum number of fragments is set as 2, i.e., each document is split into no more than 2 fragments of 5 sentences in length. Thus, if the document is longer than 10 words, it will be divided into 2 fragments of 5 words each, while, if the document is shorter than 6 words, it will be considered as one fragment.
- HDBSCAN arguments: hyperparameters of HDBSCAN clustering algorithm, minimum cluster size, distance measure and cluster selection method.
- Speed: the speed for model training. The paper establishes deep learning, which is the most sophisticated and efficient training method.

Once these parameters are set, the Top2Vec model is ready to be trained.

Optimum values are obtained after several experiments with different parameters and comparison of results. The model is trained on a large number of legal documents. Because of this and the deep learning capabilities, the training process takes several hours. The result of the training process is a model that has access to a set of clusters, each containing a representative document together with a list of similar documents. These clusters can be used to explore the corpus and identify patterns or themes in legal documents.

2.5. Queries processing

At this step, the user query is processed to match the preprocessed documents that were used to train the model. The same preprocessing methods are applied to the user query as for the documents. The preprocessed query is then used to create a set of candidate documents for the model.

2.6. Documents retrieval

In order to obtain the documents associated with a query using Top2Vec, the first step is to embed the query in the same vector space as the documents. This can be done by passing the query through the same neural network that was used to embed the documents. The result is a vector representing the query in the same multidimensional vector space.

Top2Vec then computes the cosine similarity between the query vector and all document vectors in the vector space. The cosine similarity score ranges from −1 to 1, where 1 indicates complete similarity, 0 indicates no similarity, while −1 indicates complete dissimilarity.

The cosine similarity between vectors **q** and **d** can be calculated as follows:

$$\cos(\mathbf{q},\mathbf{d}) = \frac{\mathbf{q}\cdot\mathbf{d}}{(\|\mathbf{q}\|\cdot\|\mathbf{d}\|)},$$

where $\|\mathbf{q}\|$ and $\|\mathbf{d}\|$ are the Euclidean norms of vectors **q** and **d**, and $\mathbf{q}\cdot\mathbf{d}$ is their scalar product [14].

3. ASSESSMENT AND DISCUSSION

The proposed document retrieval approach was evaluated using a set of real user queries related to felonies, including criminal offenses, in Arabic legal documents. First, a set of 100 user queries from lawyers working with such documents was generated. To evaluate the effectiveness of the approach, the 10 most popular documents for each query were compared with a set of relevant documents identified by legal experts. A relevant document is one that contains information relevant to the query topic even if it does not explicitly answer the query. For example, a document about a murder investigation will be considered relevant to a murder query. Two standard evaluation metrics are used to measure the performance of the proposed approach: precision and completeness. Accuracy measures the share of retrieved relevant documents, while completeness measures the share of retrieved relevant documents. Specifically, accuracy and completeness are defined as follows [15]:

$$\text{Precision} = \frac{\text{Relevant Documents Retrieved (RDR)}}{\text{Total Documents Retrieved (TDR)}},$$

$$\text{Recall} = \frac{\text{Relevant Documents Retrieved (RDR)}}{\text{Total Relevant Documents (TRD)}},$$

where RDR is the number of extracted relevant documents, TDR is the total number of extracted documents, and TRD is the total number of relevant documents.

In order to ensure a single summary measure of overall performance, an F1 score representing a harmonic mean of accuracy and completeness was evaluated. To compare the performance of the proposed approach with the baseline, a simple keyword-based approach retrieves documents containing at least one keyword from a user query. The same set of queries and evaluation metrics are used in the comparison. Experiments show that the proposed document retrieval approach outperforms the keyword-based approach in terms of both accuracy and completeness, achieving retrieval accuracy of 87% and retrieval completeness of 80% as compared to an accuracy of 66% and completeness of 62% for the keyword-based approach. The F1 score for the proposed approach is 0.83, while the F1 score for the keyword-based approach is 0.63.

These results show that utilizing natural language processing and machine learning techniques for document retrieval can significantly improve the performance of legal information retrieval systems. By processing user queries and documents using advanced algorithms, the proposed approach retrieves relevant legal documents with higher accuracy, reducing the burden on legal practitioners to manually search large document repositories. Nevertheless, it should be noted that this approach can still be improved. For example, it may not be effective in retrieving relevant documents for highly specialized legal domains that require a more nuanced understanding of legal language. Nevertheless, it is crucial to recognize that the current approach has room for improvement. In addition, potential distortions inherent in training data may lead to misrepresentations in the results obtained.

All in all, the experimental results demonstrate the promising potential of the proposed document retrieval approach in developing more efficient legal information retrieval systems. Further research and development in this area could lead to even more efficient algorithms that can better support legal practitioners in their work.

CONCLUSIONS

The proposed approach to legal document retrieval uses natural language processing techniques, including the Top2Vec model and the HDBSCAN clustering algorithm. A comparison of this approach with a keyword-based approach demonstrates its superiority in terms of accuracy and completeness. Specifically, the proposed approach achieves a retrieval accuracy of 87% and a retrieval completeness of 80%. This can significantly improve legal document retrieval by making it faster and more accurate. However, it is important to note its limitations and room for future research and development. In conclusion, the work underscores the relevance of using advanced natural language processing techniques in the legal domain, as well as successfully demonstrating their potential to improve the efficiency and accuracy of legal document retrieval.

Authors' contributions

K.S. Jafar—conducted the majority of the research from conception to execution, including the design of the methodology, implementation of the Topic-to-Vector algorithm, and evaluation of results.

A.A. Mohammad—contributed to the literature review, identifying relevant studies and background information related to legal information retrieval, and assisted in drafting the introduction.

A.H. Issa—made significant contributions to writing the discussion section, interpreting the findings within the context of existing literature, and contributed to the conclusion of the article.

A.V. Panov—supervised the research process, provided guidance on research direction, and assisted in refining the methodology and results.

REFERENCES

1. Sleimi A., Sannier N., Sabetzadeh M., Briand L., Dann J. Automated extraction of semantic legal metadata using natural language processing. In: *2018 IEEE 26th International Requirements Engineering Conference* (*RE*). IEEE; 2018. P. 124–135. https://doi.org/10.1109/RE.2018.00022
2. Rogers A., Gardner M., Augenstein I. QA dataset explosion: A taxonomy of NLP resources for question answering and reading comprehension. *ACM Comput. Surveys*. 2023;55(10):1–45. https://doi.org/10.1145/3560260
3. Alanazi S.S., Elfadil N., Jarajreh M., Algarni S. Question Answering Systems: A Systematic Literature Review. *International Journal of Advanced Computer Science and Applications* (*IJACSA*). 2021;12(3):359. https://doi.org/10.14569/IJACSA.2021.0120359
4. Sansone C., Sperl´ı G. Legal Information Retrieval systems: State-of-the-art and open issues. *Inform. Syst.* 2022;106:101967. https://doi.org/10.1016/j.is.2021.101967
5. Sartor G., Araszkiewicz M., Atkinson K., et al. Thirty years of Artificial Intelligence and Law: the second decade. *Artif. Intell. Law*. 2022;30(4):521–557. https://doi.org/10.1007/s10506-022-09326-7
6. Zhong H., Xiao C., Tu C., et al. How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence. 2020. *arXiv:2004.12158* [cs.CL]. https://arxiv.org/abs/2004.12158v5
7. Abu Shamma S., Ayasa A., Yahya A., et al. Information extraction from Arabic law documents. In: *2020 IEEE 14th International Conference on Application of Information and Communication Technologies* (*AICT*). IEEE; 2020;1–6. https://doi.org/10.1109/AICT50176.2020.9368577
8. Hammami E., Faiz R. Topic Modelling of Legal Texts Using Bidirectional Encoder Representations from Sentence Transformers. In: *Advances in Information Systems, Artificial Intelligence and Knowledge Management. Conference paper. International Conference on Information and Knowledge Systems.* Cham: Springer Nature Switzerland; 2023. V. 486. P. 333–343. https://doi.org/10.1007/978-3-031-51664-1_24
9. Angelov D. Top2Vec: Distributed Representations of Topics. 2020. *arXiv:2008.09470* [cs.CL]. https://arxiv.org/abs/2008.09470v1
10. Karas B., Qu S., Xu Y., Zhu Q. Experiments with LDA and Top2Vec for embedded topic discovery on social media data—A case study of cystic fibrosis. *Front. Artif. Intell*. 2022;5:948313. https://doi.org/10.3389/frai.2022.948313
11. Vianna D., de Moura E.S., da Silva A.S. A topic discovery approach for unsupervised organization of legal document collections. *Artif. Intell. Law*. 2023;Online First. https://doi.org/10.1007/s10506-023-09371-w
12. McInnes L., Healy J., Astels S. hdbscan: Hierarchical density-based clustering. *J. Open Source Softw*. 2017;2(11):205. https://doi.org/10.21105/joss.00205
13. Devlin J., Chang M.W., Lee K., Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. 2018. *arXiv*, *preprint arXiv:1810.04805*. https://arxiv.org/abs/1810.04805v2
14. Salton G., McGill M.J. *Introduction to Modern Information Retrieval*. N.Y.: McGraw-Hill; 1983. 472 p.
15. Manning C.D., Raghavan P., Schütze H. *Introduction to Information Retrieval*. Cambridge, England: Cambridge University Press; 2008. 492 p.

About the authors

Kamel S. Jafar, Postgraduate Student, Department of Corporate Information Systems, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: zhafar.k@edu.mirea.ru. Scopus Author ID 57552322300, https://orcid.org/0009-0004-1791-1406

Ali A. Mohammad, Master Student, Faculty of Computer Science, HSE University (11, Pokrovsky bulv., Moscow, 109028 Russia). E-mail: amokhammad_1@edu.hse.ru. https://orcid.org/0009-0002-3533-568X

Ali H. Issa, Postgraduate Student, Department of Automated Control Systems for Biotechnological Processes, BIOTECH University (11, Volokolamskoye sh., Moscow, 125080 Russia). E-mail: ali.issa.rus@gmail.com. https://orcid.org/0009-0001-3699-222X

Alexander V. Panov, Cand. Sci. (Eng.), Associate Professor, Department of Corporate Information Systems, Institute of Information Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: panov_a@mirea.ru. RSCI SPIN-code 8571-9729, https://orcid.org/0009-0003-0310-3638

Об авторах

Жафар Камел С., аспирант, кафедра корпоративных информационных систем, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: zhafar.k@edu.mirea.ru. Scopus Author ID 57552322300, https://orcid.org/0009-0004-1791-1406

Мохаммад Али А., магистрант, Факультет компьютерных наук, ФГАОУ ВО «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ) (109028, Россия, Москва, Покровский бульвар, д. 11). E-mail: amokhammad_1@edu.hse.ru. https://orcid.org/0009-0002-3533-568X

Исса Али Х., аспирант, кафедра автоматизированных систем управления биотехнологическими процессами, ФГБОУ ВО «Российский биотехнологический университет» (РОСБИОТЕХ) (125080, Россия, Москва, Волоколамское шоссе, д. 11). E-mail: ali.issa.rus@gmail.com. https://orcid.org/0009-0001-3699-222X

Панов Александр Владимирович, к.т.н. доцент, профессор кафедры корпоративных информационных систем, Институт информационных технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: panov_a@mirea.ru. SPIN-код РИНЦ 8571-9729, https://orcid.org/0009-0003-0310-3638

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.391.072
https://doi.org/10.32362/2500-316X-2024-12-5-17-32
EDN EBOWFT

RESEARCH ARTICLE

Noise immunity of QAM-OFDM signal reception using soft-decision demodulation in the presence of narrowband interference

Alexey A. Paramonov @**,**
Chu Van Vuong

MIREA – Russian Technological University, Moscow, 194454 Russia
@ *Corresponding author, e-mail: paramonov@mirea.ru*

Abstract

Objectives. The aim of this paper is to study the noise immunity of digital information transmission in systems with orthogonal frequency division multiplexing (OFDM) and quadrature amplitude modulation (QAM) of subcarriers in the presence of narrowband interference. As a way of managing this interference, the paper studies the use of a demodulator with soft outputs and subsequent decoding of the convolutional code and low-density parity-check (LDPC) code used in the system.

Methods. The results presented in the article were obtained using statistical radio engineering, mathematical statistics, encoding theory, and computer modeling.

Results. The paper presents a simple method for calculating soft bit estimates in the *M*-point signal QAM demodulator, where *M* is an even power of two. A considerable amount of numerical results were obtained which show the dependence of the transmitted information bit error rate on *M*, as well as on the signal-to-noise ratio, signal-to-narrowband interference, and code rates.

Conclusions. It can be concluded from the above results that the use of encoding with soft demodulator decisions significantly improves the noise immunity of OFDM signal reception, and enables narrowband interference to be managed efficiently. LDPC encoding is superior to convolutional encoding in increasing the noise immunity of OFDM signal reception both in the absence and in the presence of narrowband interference. Along with the use in QAM-OFDM systems, the proposed simple method for demodulating QAM signals with soft decisions can be used in any wireless communication system using *M*-position QAM signals, where *M* is 2 to an even power.

Keywords: OFDM, soft decision, encoding, narrowband interference, noise immunity, bit error rate

• **Submitted:** 23.04.2024 • **Revised:** 19.06.2024 • **Accepted:** 05.08.2024

For citation: Paramonov A.A., Chu V.V. Noise immunity of QAM-OFDM signal reception using soft-decision demodulation in the presence of narrowband interference. *Russ. Technol. J.* 2024;12(5):17–32. https://doi.org/10.32362/2500-316X-2024-12-5-17-32

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Помехоустойчивость приема сигнала OFDM с использованием квадратурной амплитудной модуляции с мягкими решениями при наличии узкополосных помех

А.А. Парамонов [@], **В.В. Чу**

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] *Автор для переписки, e-mail: paramonov@mirea.ru*

Резюме

Цели. Целью работы является исследование помехоустойчивости передачи цифровой информации в системах на основе мультиплексирования с ортогональным частотным разделением (orthogonal frequency division multiplexing, OFDM) и квадратурной амплитудной модуляцией (quadrature amplitude modulation, QAM) поднесущих в присутствии узкополосной помехи. В качестве способа борьбы с этой помехой исследовано применение демодулятора с мягкими выходами и последующее декодирование используемых в системе сверточного кода и кода LDPC (low-density parity-check code).

Методы. Представленные в статье результаты получены с использованием методов статистической радиотехники, математической статистики, теории кодирования и компьютерного моделирования.

Результаты. Представлен простой метод вычисления мягких оценок битов в демодуляторе *M*-ичных сигналов QAM, где *M* является четной степенью двойки. Получен большой объем численных результатов, показывающих зависимость вероятности ошибки на бит передаваемой информации от кратности *M*, от отношений сигнал/шум, сигнал/узкополосная помеха, от скорости кодов.

Выводы. Из полученных результатов можно сделать вывод, что использование кодирования с мягкими решениями демодулятора значительно улучшает помехоустойчивость приема OFDM-сигнала, позволяя эффективно бороться с узкополосными помехами. Кодирование LDPC показывает превосходство над сверточным кодированием в повышении помехоустойчивости приема сигнала OFDM как в отсутствие узкополосных помех, так и при их наличии. Наряду с использованием в системах QAM-OFDM, предложенный простой метод демодуляции сигналов QAM с мягкими решениями может применяться в любых системах беспроводной связи, использующих *M*-позиционные сигналы QAM, у которых *M* представляет собой число 2 в четной степени.

Ключевые слова: OFDM, квадратурная амплитудная модуляция, мягкое решение, кодирование, узкополосная помеха, помехоустойчивость, коэффициент битовых ошибок

• **Поступила:** 23.04.2024 • **Доработана:** 19.06.2024 • **Принята к опубликованию:** 05.08.2024

Для цитирования: Парамонов А.А., Чу В.В. Помехоустойчивость приема сигнала OFDM с использованием квадратурной амплитудной модуляции с мягкими решениями при наличии узкополосных помех. *Russ. Technol. J.* 2024;12(5):17–32. https://doi.org/10.32362/2500-316X-2024-12-5-17-32

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

In modern wireless communication networks where the radio spectrum efficiency is crucial, data transmission systems based on orthogonal frequency division multiplexing (OFDM) [1–3] and quadrature amplitude modulation of M multiplicity (M-QAM) [4–6] are widely used. However, under conditions of the active use of spectrum and the presence of numerous sources of interference including narrowband ones [7, 8], reliable data transmission is becoming an urgent issue.

The aim of the paper is to investigate the noise immunity of the QAM-OFDM system with encoding in the presence of narrowband interference [9–11]. Decoding methods using demodulator hard decisions cannot often compensate for the impact of narrowband interference effectively, thus significantly degrading the reception quality. In this context, the method of soft demodulator decisions is proposed. This demodulator can be used effectively in conjunction with an appropriate decoder for improving reception noise immunity. The proposed soft decision algorithm for demodulation has a lower computational complexity when compared to conventional methods. The proposed method is based on analyzing the quality of the received signal, which allows the reliability degree to be determined dynamically for each bit of the transmitted signal. Thus, greater efficiency of the decoding process may be achieved, which results in increased noise immunity of the system in the presence of narrowband interference.

This paper describes the proposed method of demodulation of M-QAM signals with soft decisions, compared with existing demodulation methods, and simulation results showing its efficiency. It shows that the use of decoding with soft decisions allows for a noticeable improvement in noise immunity reception in the presence of narrowband interference.

OFDM SYSTEM AND SOFT DECISION ALGORITHM

The OFDM system considered in the paper is shown in Fig. 1.

Figure 1a shows the structure of the OFDM transmitter. The bit stream is encoded before modulation. The system is assumed to use M-QAM. After the inverse fast Fourier transform (IFFT), a guard interval is added to the OFDM signal, in order to protect the received signal from intersymbol interference. The signal is then transmitted over the air.

Figure 1b shows the structure of the OFDM receiver. In the receiver, signal processing is carried out in the reverse order to the transmitter processing order. Firstly, time synchronization and division of the OFDM signal into OFDM symbols are performed. After removing the guard interval, the fast Fourier transform (FFT) is performed. Then the signal is demodulated and finally decoded to obtain the original data stream.

The M-QAM signal $x(t)$ with symbol duration equal to the OFDM signal symbol duration is transmitted on each orthogonal frequency of the OFDM signal. Signal $\tilde{x}(t)$ received by the receiver on this frequency can be described as follows:

$$\tilde{x}(t) = x(t) + w(t), \quad (1)$$

wherein $w(t)$ stands for additive white Gaussian noise.

In such a problem statement, it would be reasonable to develop the algorithm for receiving the M-QAM signal. Including any additive interference into the received signal $\tilde{x}(t)$, except for noise, implies the information transmission system should necessarily operate on frequencies occupied by interference. A more realistic scenario is when the transmission system is designed to operate on frequencies free of interference. However, interference may actually occur, so the transmission system immunity should be analyzed separately for this case. It is thus from the positions referred to below, that the study of noise immunity to reception of QAM-OFDM signal in the presence of narrowband interference is carried out.

Here, it is worth noting the particular impact of harmonic interference on OFDM signal. The following formulas relating to harmonic interference assume that OFDM signal is shifted to zero frequency. Figure 1 shows that the received oscillation $\tilde{x}(t)$ is subjected to FFT procedure performed digitally. If this oscillation contains harmonic interference in the region of some subcarrier, it can be written as a sequence of time samples in the following way:

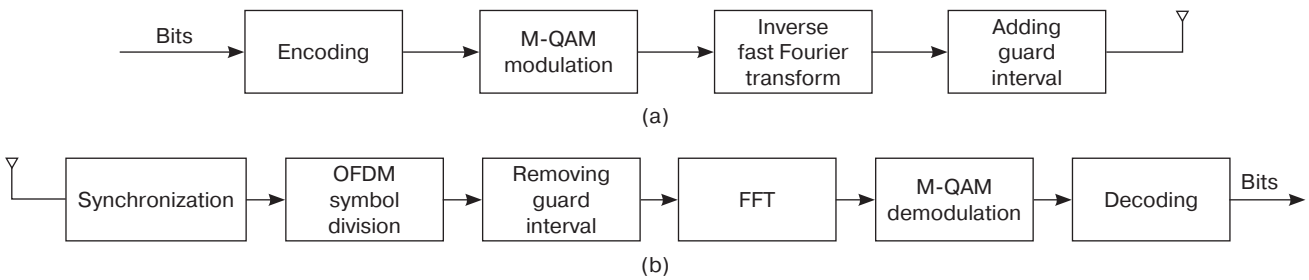


Fig. 1. The OFDM system structure: (a) transmitter, (b) receiver

$$g(n) = A_g e^{j(2\pi f_g n T_s + \theta)},$$

wherein A_g is the narrowband interference amplitude; f_g is the narrowband interference frequency; T_s is sampling interval equal to the symbol duration; and θ is the randomly distributed phase, $\theta \in (-\pi; \pi]$.

The narrowband interference frequency does not necessarily coincide with the subcarrier and is defined as follows:

$$f_g = \frac{m+\alpha}{N} f_s,\ 0 \le m \le N-1,\ -0.5 \le \alpha \le 0.5,$$

wherein f_s is frequency spacing between subcarriers; m is the number of the subcarrier closest to the interference; and N is the number of subcarriers in OFDM signal (and FFT dimensionality at reception).

Substituting f_g into the formula for harmonic interference and using the fact that $f_s = \frac{1}{T_s}$, the following can be obtained:

$$g(n) = A_g e^{j\left(\frac{2\pi}{N}(m+\alpha)n + \theta\right)}.$$

At $\alpha = 0$, the interference frequency coincides with the subcarrier frequency and the interference is orthogonal to other subcarriers. At $\alpha \neq 0$, the interference frequency does not coincide with the subcarrier frequency and this interference is not orthogonal to other subcarriers.

The expression for the narrowband interference spectrum after FFT is written as follows:

$$G(k) = \frac{1}{N}\sum_{n=0}^{N-1} g(n) e^{-\frac{j2\pi nk}{N}} = \frac{A_g}{N} e^{j\theta} \frac{1 - e^{j2\pi\alpha}}{1 - e^{j\frac{2\pi}{N}(m-k+\alpha)}}.$$

When the narrowband interference is orthogonal ($\alpha = 0$), it has the following form after FFT:

$$G(k) = \begin{cases} A_g e^{j\theta}, k = m, \\ 0, \quad k \neq m. \end{cases}$$

Thus, the narrowband interference coinciding in frequency with some mth subcarrier does not fall on frequencies of other subcarriers. If $\alpha \neq 0$, i.e., the narrowband interference is not orthogonal to other subcarriers, then the power of this interference is distributed over all subcarriers, i.e., a leakage of the interference spectrum occurs. The power of the narrowband interference leaked on some kth subcarrier due to non-orthogonality is determined by the following expression [7]:

$$\sigma^2_{G,k} = E[|G_k|^2] = \frac{A_g^2}{N^2} \cdot \frac{1 - \cos(2\pi\alpha)}{1 - \cos\frac{2\pi}{N}(m-k+\alpha)}.$$

Looking back to relation (1), its components can be represented as signal points in some signal space:

$$\tilde{X} = X + W. \tag{2}$$

The conditional probability density of the received signal $\tilde{X}$, given signal X, is transmitted can be written as follows:

$$f(\tilde{X} \mid X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{|\tilde{X}-X|^2}{2\sigma^2}\right)}. \tag{3}$$

Here, σ^2 is the distribution variance (3).

Each M-QAM signal symbols carries $\log_2 M$ bits of information. For example, for 16-QAM, the bits are b_0, b_1, b_2, and b_3. The soft decision for some ith bit is considered to be the logarithm of the likelihood ratio defined for a priori equal probability bits as follows [12–18]:

$$\begin{aligned} l(b_i) &= \ln \frac{\sum_{X \in S_i^-} f(X \mid \tilde{X})}{\sum_{X \in S_i^+} f(X \mid \tilde{X})} \approx \\ &\approx \ln \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{|\tilde{X}-X_{i,\text{opt}}^-|^2}{2\sigma^2}\right)}}{\frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{|\tilde{X}-X_{i,\text{opt}}^+|^2}{2\sigma^2}\right)}} = \\ &= \frac{1}{2\sigma^2}(|\tilde{X} - X_{i,\text{opt}}^-|^2 - |\tilde{X} - X_{i,\text{opt}}^+|^2), \end{aligned} \tag{4}$$

where S_i^+ and S_i^- stand for sets of symbols whose ith bit is 1 and 0, respectively.

In expression (4), $X_{i,\text{opt}}^+$ and $X_{i,\text{opt}}^-$ are signals X closest to the received oscillation $\tilde{X}$ whose ith bit is 1 and 0, respectively:

$$\begin{aligned} X_{i,\text{opt}}^+ &= \arg\min_{X \in S_i^+} |\tilde{X} - X|^2, \\ X_{i,\text{opt}}^- &= \arg\min_{X \in S_i^-} |\tilde{X} - X|^2. \end{aligned} \tag{5}$$

Determining soft decisions by using expressions (4) and (5) is a rather cumbersome task. The complexity of calculations increases significantly with increasing

modulation level M. The next section examines considerably simpler algorithms for determining soft decisions for signals QAM with M being 2 in even power. The algorithms for 16-QAM, 64-QAM, and 256-QAM are set out in more detail.

SIMPLE ALGORITHMS FOR SOFT DECISION-MAKING IN M-QAM DEMODULATION

Figure 2 shows constellations for, 64-QAM, and 256-QAM. A decimal number is located near each signal point, and when converted into binary form it shows the set of transmitted binary symbols (hereinafter referred to as bits) corresponding to this signal point. For example, for 16-QAM signal, the point marked by number 6 corresponds to the set of transmitted bits: 0, 1, 1, 0.

We first examine the demodulation of 16-QAM signal. Each point of the signal constellation corresponds to 4 transmitted bits: b_0, b_1, b_2, and b_3. The projections of signal points on the in-phase and quadrature axes for different combinations of transmitted bits are shown in Table 1.

Figure 3 explains the proposed simplified algorithm for computing the logarithm of the likelihood ratio. Let a given symbol of 16-QAM signal be taken, as shown by a blue circle on the IQ-diagram: $\tilde{X}_i = \Re(\tilde{X}_i) + j\Im(\tilde{X}_i)$, where $\Re(\tilde{X}_i)$, $\Im(\tilde{X}_i)$ are real and imaginary parts of oscillation $\tilde{X}$.

The projections of signal points of the transmitted 16-QAM signal on axes I and Q are marked with red circles.

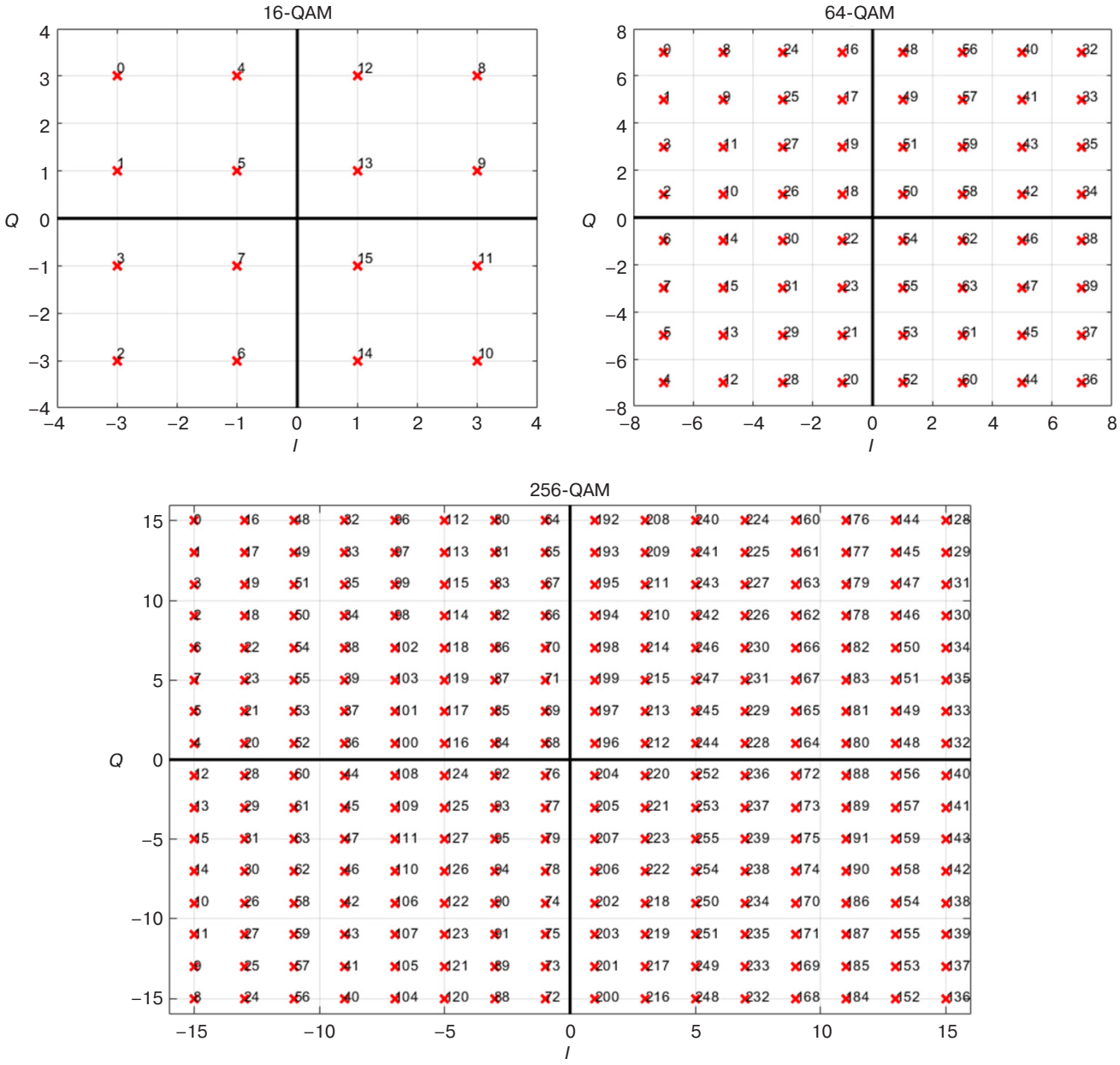


Fig. 2. M-QAM constellations

Table 1. Projections of signal points of 16-QAM signal on the *I* and *Q* axes

b_0b_1	I	b_2b_3	Q
00	−3	00	3
01	−1	01	1
11	+1	11	−1
10	+3	10	−3

It should be noted that the value of bit b_0 affects the projection of the signal point on the *I*-axis only, but does not affect the projection on the *Q*-axis. This can be clearly seen in Fig. 3a. When $b_0 = 0$, the real part of 16-QAM signal takes value −1 or −3. When $b_0 = 1$, the real part of the signal is +1 or +3.

The soft decision-making process for the first bit b_0 is shown in Fig. 4, representing the lower part of Fig. 3a.

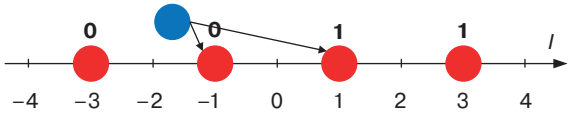
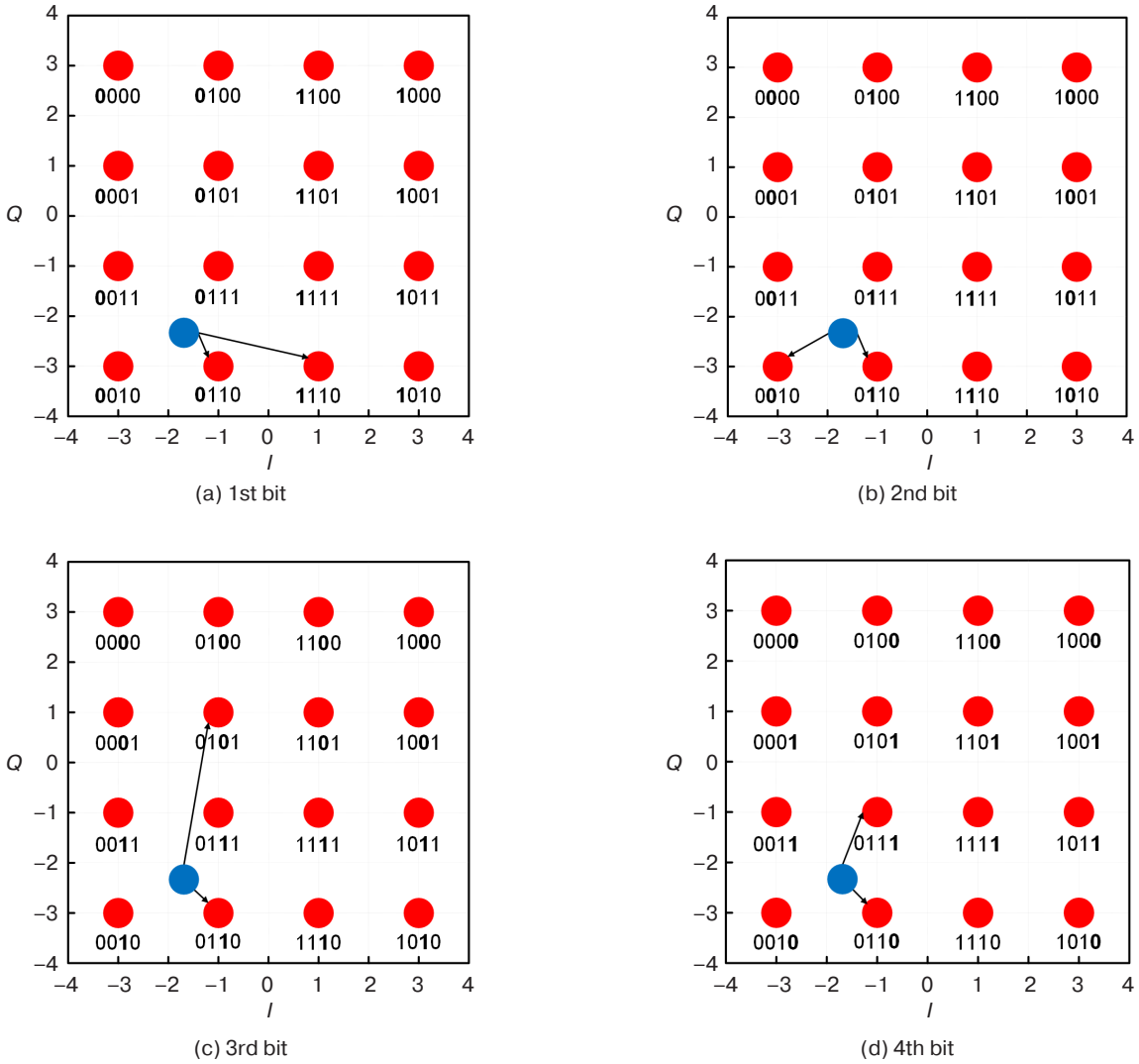


Fig. 4. Calculating the logarithm of the likelihood ratio for the 1st bit

The conditional probability of the received oscillation $\tilde{X}_i$ at $b_0 = 0$ is as follows:

$$P(\tilde{X}_i \mid b_0 = 0) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Re(\tilde{X}_i)+3)^2}{2\sigma^2}} + \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Re(\tilde{X}_i)+1)^2}{2\sigma^2}}. \quad (6)$$

The conditional probability of the received oscillation $\tilde{X}_i$ at $b_0 = 1$ is as follows:



(a) 1st bit

(b) 2nd bit

(c) 3rd bit

(d) 4th bit

Fig. 3. Example of calculating the logarithms of likelihood ratios for each bit of 16-QAM signal

$$P(\tilde{X}_i \mid b_0 = 1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Re(\tilde{X}_i)-3)^2}{2\sigma^2}} + \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Re(\tilde{X}_i)-1)^2}{2\sigma^2}}. \quad (7)$$

The likelihood ratio is described by the following expression:

$$\frac{P(\tilde{X}_i \mid b_0 = 0)}{P(\tilde{X}_i \mid b_0 = 1)} = \frac{e^{-\frac{(\Re(\tilde{X}_i)+1)^2}{2\sigma^2}} + e^{-\frac{(\Re(\tilde{X}_i)+3)^2}{2\sigma^2}}}{e^{-\frac{(\Re(\tilde{X}_i)-1)^2}{2\sigma^2}} + e^{-\frac{(\Re(\tilde{X}_i)-3)^2}{2\sigma^2}}}. \quad (8)$$

Using expression (8) is inconvenient since it requires that exponents be calculated. However, it can be simplified considering that the projection of the received oscillation in practice usually appears closer to one of the two possible signal points corresponding to a certain decision. For example, the projection of the received oscillation shown in Fig. 4 is closer to the value −1, rather than to −3, on which basis the decision on symbol $b_0 = 0$ is made. Given that expression (8) includes exponents of the squares of projection differences, it may be assumed that the first summand in denominator (8) is significantly larger than the second. Thus, the second exponent can be neglected. Similar reasoning is valid for evaluating numerator (8). Consequently, the logarithm of the likelihood ratio at value $\tilde{X}$ represented by the blue circle can be calculated quite accurately using the following expression:

$$l(b_0, \tilde{X}_i) \approx \frac{1}{2\sigma^2}((\Re(\tilde{X}_i)-1)^2 - (\Re(\tilde{X}_i)+1)^2) = -\frac{2}{\sigma^2}\Re(\tilde{X}_i). \quad (9)$$

When making the decision on bit b_1, reference should be made to Fig. 3b. The lower part of this figure is shown in Fig. 5. As in the previous case, the value of bit b_0 affects the projection of the signal point on the *I*-axis only and does not affect the projection on the *Q*-axis. At $b_1 = 0$, the real part of 16-QAM signal takes –3. When $b_1 = 1$, the real part of the signal takes –1, as shown in Fig. 5.

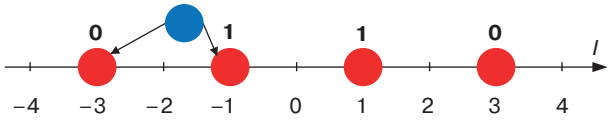


Fig. 5. Calculating the logarithm of the likelihood ratio for the 2nd bit

The likelihood ratio for the second bit b_1 of the received oscillation $\tilde{X}$ is the following:

$$l(b_1, \tilde{X}_i) \approx \frac{1}{2\sigma^2}((\Re(\tilde{X}_i)+1)^2 - (\Re(\tilde{X}_i)+3)^2) = -\frac{2}{\sigma^2}(\Re(\tilde{X}_i)+2). \quad (10)$$

For 16-QAM signals, only the projection of the signal point on the *Q*-axis varies depending on the value of bit b_2 (Fig. 3c). When $b_2 = 0$, the imaginary part of 16-QAM signal takes 1. When $b_2 = 1$, the imaginary part of the signal takes −3, as shown in Fig. 6.

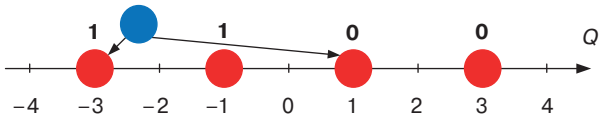


Fig. 6. Calculating the logarithm of the likelihood ratio for the 3rd bit

The logarithm of the likelihood ratio for b_2 of the received oscillation $\tilde{X}$ is the following:

$$l(b_2, \tilde{X}_i) \approx \frac{1}{2\sigma^2}((\Im(\tilde{X}_i)+3)^2 - (\Im(\tilde{X}_i)-1)^2) = \frac{2}{\sigma^2}(\Im(\tilde{X}_i)+1). \quad (11)$$

It follows from Fig. 3d that the value of b_3 affects the projection of the signal point on the *Q*-axis only. When $b_3 = 0$, this projection takes −3. When $b_3 = 1$, the projection takes −1, as shown in Fig. 7.

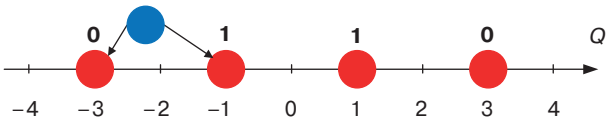


Fig. 7. Calculating the logarithm of the likelihood ratio for the 4th bit

The likelihood ratio for the fourth bit b_3 of the received oscillation $\tilde{X}$ is as follows:

$$l(b_3, \tilde{X}_i) \approx \frac{1}{2\sigma^2}((\Im(\tilde{X}_i)+1)^2 - (\Im(\tilde{X}_i)+3)^2) = -\frac{2}{\sigma^2}(\Im(\tilde{X}_i)+2). \quad (12)$$

Previously, the procedure of soft decision-making has been detailed on the basis of the specific example of the input oscillation depicted as a blue circle in Fig. 3. Similarly, after considering all possible positions of the input oscillation, the following is obtained:

$$l(b_0,\tilde{X}_i)=\begin{cases}-\frac{2}{\sigma^2}[\Re(\tilde{X}_i)+1], & \Re(\tilde{X}_i)<-2,\\ -\frac{2}{\sigma^2}\Re(\tilde{X}_i), & -2\le\Re(\tilde{X}_i)<0,\\ -\frac{2}{\sigma^2}\Re(\tilde{X}_i), & 0\le\Re(\tilde{X}_i)<2,\\ -\frac{2}{\sigma^2}[\Re(\tilde{X}_i)-1], & 2\le\Re(\tilde{X}_i),\end{cases} \quad (13)$$

$$l(b_1,\tilde{X}_i)=\frac{2}{\sigma^2}[|\Re(\tilde{X}_i)|-2] \quad \forall\Re(\tilde{X}_i), \quad (14)$$

$$l(b_2,\tilde{X}_i)=\begin{cases}\frac{2}{\sigma^2}[\Im(\tilde{X}_i)+1], & \Im(\tilde{X}_i)<-2,\\ \frac{2}{\sigma^2}\Im(\tilde{X}_i), & -2\le\Im(\tilde{X}_i)<0,\\ \frac{2}{\sigma^2}\Im(\tilde{X}_i), & 0\le\Im(\tilde{X}_i)<2,\\ \frac{2}{\sigma^2}[\Im(\tilde{X}_i)-1], & 2\le\Im(\tilde{X}_i),\end{cases} \quad (15)$$

$$l(b_3,\tilde{X}_i)=\frac{2}{\sigma^2}[|\Im(\tilde{X}_i)|-2] \quad \forall\Im(\tilde{X}_i). \quad (16)$$

The demodulation process for 64-QAM signals can be considered in a similar way to the soft decision-making for 16-QAM signal. The signal constellation, the received oscillation and signal points closest to it, corresponding to 0 and 1 for each of the six bits $b_0b_1b_2b_3b_4b_5$, contained in one symbol of 64-QAM signal are shown in Fig. 8, similar to Fig. 3.

Figure 8 shows that the first three bits affect the real part of the signal only, i.e., they determine the signal projection on the *I*-axis. The remaining three bits affect its imaginary part only, thus determining the projection on the *Q*-axis. According to these projections, the logarithms of likelihood ratios for all six bits can be unambiguously determined. The corresponding formulas are given in Tables 2 and 3.

For 256-QAM modulation, each point of this constellation corresponds to eight bits of the transmitted information: $b_0b_1b_2b_3b_4b_5b_6b_7$. A thorough examination of this constellation indicates that the first four bits affect the real part of the signal only, i.e., signal projections on the *I*-axis. The remaining four bits determine the imaginary part of the signal only, i.e., signal projections on the *Q*-axis. The calculation results for the logarithms of the likelihood ratios for all eight bits are summarized in Tables 4 and 5.

Table 2. Values of $l(b_k,\tilde{X}_i)$ for k = 0, 1, 2 for 64-QAM signal

$\Re(\tilde{X}_i)$	$l(b_0,\tilde{X}_i)$	$l(b_1,\tilde{X}_i)$	$l(b_2,\tilde{X}_i)$
$\Re(\tilde{X}_i)<-6$	$-8[\Re(\tilde{X}_i)+3]$	$-4[\Re(\tilde{X}_i)+5]$	$-2[\Re(\tilde{X}_i)+6]$
$-6\le\Re(\tilde{X}_i)<-4$	$-6[\Re(\tilde{X}_i)+2]$	$-2[\Re(\tilde{X}_i)+4]$	$-2[\Re(\tilde{X}_i)+6]$
$-4\le\Re(\tilde{X}_i)<-2$	$-4[\Re(\tilde{X}_i)+1]$	$-2[\Re(\tilde{X}_i)+4]$	$2[\Re(\tilde{X}_i)+2]$
$-2\le\Re(\tilde{X}_i)<0$	$-2\Re(\tilde{X}_i)$	$-4[\Re(\tilde{X}_i)+3]$	$2[\Re(\tilde{X}_i)+2]$
$0\le\Re(\tilde{X}_i)<2$	$-2\Re(\tilde{X}_i)$	$4[\Re(\tilde{X}_i)-3]$	$-2[\Re(\tilde{X}_i)-2]$
$2\le\Re(\tilde{X}_i)<4$	$-4[\Re(\tilde{X}_i)-1]$	$2[\Re(\tilde{X}_i)-4]$	$-2[\Re(\tilde{X}_i)-2]$
$4\le\Re(\tilde{X}_i)<6$	$-6[\Re(\tilde{X}_i)-2]$	$2[\Re(\tilde{X}_i)-4]$	$2[\Re(\tilde{X}_i)-6]$
$\Re(\tilde{X}_i)\ge6$	$-8[\Re(\tilde{X}_i)-3]$	$4[\Re(\tilde{X}_i)-5]$	$2[\Re(\tilde{X}_i)-6]$

Table 3. Values of $l(b_k,\tilde{X}_i)$ for k = 3, 4, 5 for 64-QAM signal

$\Im(\tilde{X}_i)$	$l(b_3,\tilde{X}_i)$	$l(b_4,\tilde{X}_i)$	$l(b_5,\tilde{X}_i)$
$\Im(\tilde{X}_i)<-6$	$8[\Im(\tilde{X}_i)+3]$	$-4[\Im(\tilde{X}_i)+5]$	$-2[\Im(\tilde{X}_i)+6]$
$-6\le\Im(\tilde{X}_i)<-4$	$6[\Im(\tilde{X}_i)+2]$	$-2[\Im(\tilde{X}_i)+4]$	$-2[\Im(\tilde{X}_i)+6]$
$-4\le\Im(\tilde{X}_i)<-2$	$4[\Im(\tilde{X}_i)+1]$	$-2[\Im(\tilde{X}_i)+4]$	$2[\Im(\tilde{X}_i)+2]$
$-2\le\Im(\tilde{X}_i)<0$	$2\Im(\tilde{X}_i)$	$-4[\Im(\tilde{X}_i)+3]$	$2[\Im(\tilde{X}_i)+2]$
$0\le\Im(\tilde{X}_i)<2$	$2\Im(\tilde{X}_i)$	$4[\Im(\tilde{X}_i)-3]$	$-2[\Im(\tilde{X}_i)-2]$
$2\le\Im(\tilde{X}_i)<4$	$4[\Im(\tilde{X}_i)-1]$	$2[\Im(\tilde{X}_i)-4]$	$-2[\Im(\tilde{X}_i)-2]$
$4\le\Im(\tilde{X}_i)<6$	$6[\Im(\tilde{X}_i)-2]$	$2[\Im(\tilde{X}_i)-4]$	$2[\Im(\tilde{X}_i)-6]$
$\Im(\tilde{X}_i)\ge6$	$8[\Im(\tilde{X}_i)-3]$	$4[\Im(\tilde{X}_i)-5]$	$2[\Im(\tilde{X}_i)-6]$

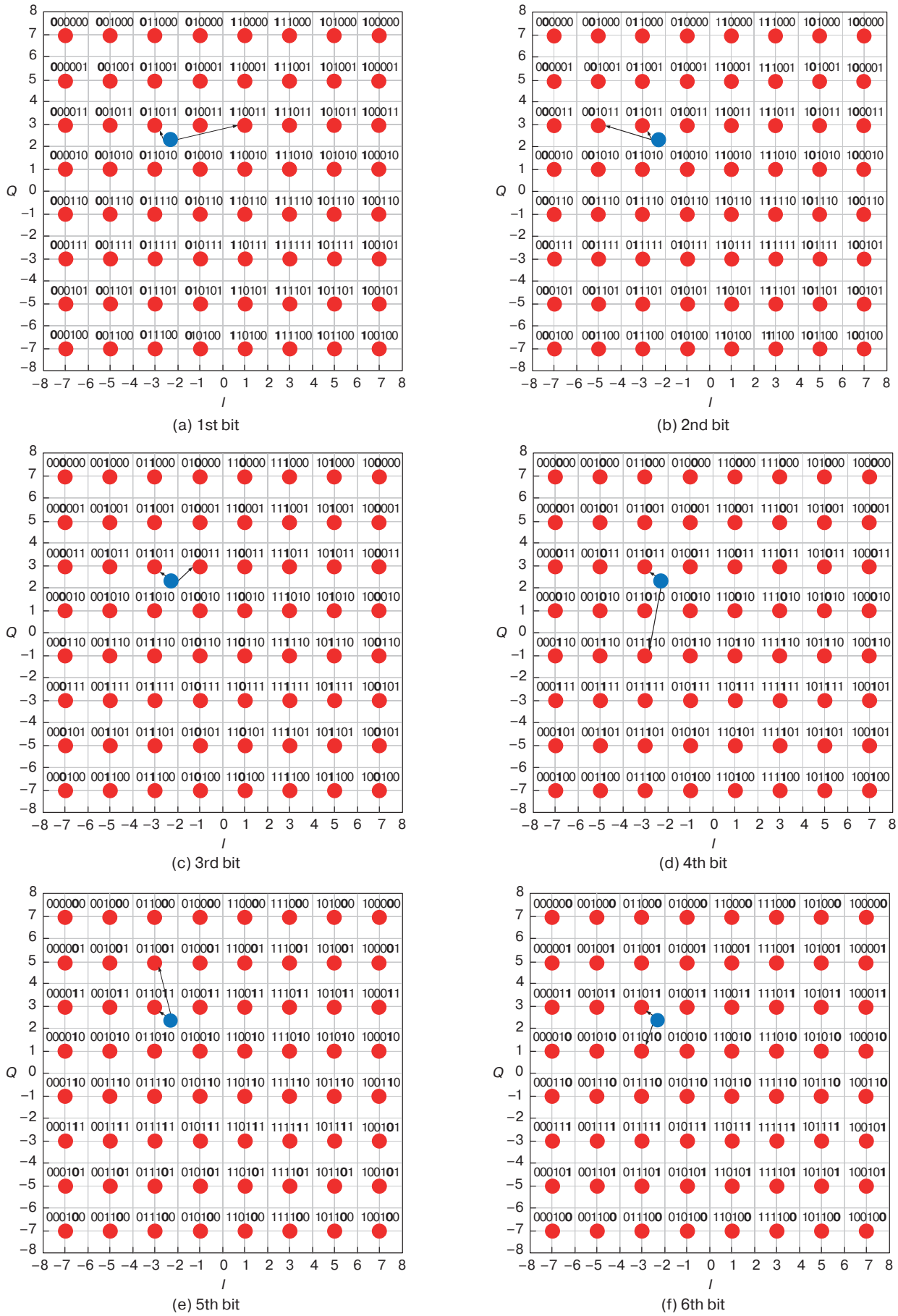


Fig. 8. Example of calculating the logarithms of likelihood ratios for each bit of the 64-QAM signal

Table 4. Values of $l(b_k,\tilde{X}_i)$ for k = 0, 1, 2, 3 for 256-QAM signal

$\Re(\tilde{X}_i)$	$l(b_0,\tilde{X}_i)$	$l(b_1,\tilde{X}_i)$	$l(b_2,\tilde{X}_i)$	$l(b_3,\tilde{X}_i)$
$\Re(\tilde{X}_i) \le -14$	$-16[\Re(\tilde{X}_i)+7]$	$-8[\Re(\tilde{X}_i)+11]$	$-4[\Re(\tilde{X}_i)+13]$	$-2[\Re(\tilde{X}_i)+14]$
$-14 \le \Re(\tilde{X}_i) < -12$	$-14[\Re(\tilde{X}_i)+6]$	$-6[\Re(\tilde{X}_i)+10]$	$-2[\Re(\tilde{X}_i)+12]$	$-2[\Re(\tilde{X}_i)+14]$
$-12 \le \Re(\tilde{X}_i) < -10$	$-12[\Re(\tilde{X}_i)+5]$	$-4[\Re(\tilde{X}_i)+9]$	$-2[\Re(\tilde{X}_i)+12]$	$2[\Re(\tilde{X}_i)+10]$
$-10 \le \Re(\tilde{X}_i) < -8$	$-10[\Re(\tilde{X}_i)+4]$	$-2[\Re(\tilde{X}_i)+8]$	$-4[\Re(\tilde{X}_i)+11]$	$2[\Re(\tilde{X}_i)+10]$
$-8 \le \Re(\tilde{X}_i) < -6$	$-8[\Re(\tilde{X}_i)+3]$	$-2[\Re(\tilde{X}_i)+8]$	$4[\Re(\tilde{X}_i)+5]$	$-2[\Re(\tilde{X}_i)+6]$
$-6 \le \Re(\tilde{X}_i) < -4$	$-6[\Re(\tilde{X}_i)+2]$	$-4[\Re(\tilde{X}_i)+7]$	$2[\Re(\tilde{X}_i)+4]$	$-2[\Re(\tilde{X}_i)+6]$
$-4 \le \Re(\tilde{X}_i) < -2$	$-4[\Re(\tilde{X}_i)+1]$	$-6[\Re(\tilde{X}_i)+6]$	$2[\Re(\tilde{X}_i)+4]$	$2[\Re(\tilde{X}_i)+2]$
$-2 \le \Re(\tilde{X}_i) < 0$	$-2\Re(\tilde{X}_i)$	$-8[\Re(\tilde{X}_i)+5]$	$4[\Re(\tilde{X}_i)+3]$	$2[\Re(\tilde{X}_i)+2]$
$0 \le \Re(\tilde{X}_i) < 2$	$-2\Re(\tilde{X}_i)$	$8[\Re(\tilde{X}_i)-5]$	$-4[\Re(\tilde{X}_i)-3]$	$-2[\Re(\tilde{X}_i)-2]$
$2 \le \Re(\tilde{X}_i) < 4$	$-4[\Re(\tilde{X}_i)-1]$	$6[\Re(\tilde{X}_i)-6]$	$-2[\Re(\tilde{X}_i)-4]$	$-2[\Re(\tilde{X}_i)-2]$
$4 \le \Re(\tilde{X}_i) < 6$	$-6[\Re(\tilde{X}_i)-2]$	$4[\Re(\tilde{X}_i)-7]$	$-2[\Re(\tilde{X}_i)-4]$	$2[\Re(\tilde{X}_i)-6]$
$6 \le \Re(\tilde{X}_i) < 8$	$-8[\Re(\tilde{X}_i)-3]$	$2[\Re(\tilde{X}_i)-8]$	$-4[\Re(\tilde{X}_i)-5]$	$2[\Re(\tilde{X}_i)-6]$
$8 \le \Re(\tilde{X}_i) < 10$	$-10[\Re(\tilde{X}_i)-4]$	$2[\Re(\tilde{X}_i)-8]$	$4[\Re(\tilde{X}_i)-11]$	$-2[\Re(\tilde{X}_i)-10]$
$10 \le \Re(\tilde{X}_i) < 12$	$-12[\Re(\tilde{X}_i)-5]$	$4[\Re(\tilde{X}_i)-9]$	$2[\Re(\tilde{X}_i)-12]$	$-2[\Re(\tilde{X}_i)-10]$
$12 \le \Re(\tilde{X}_i) < 14$	$-14[\Re(\tilde{X}_i)-6]$	$6[\Re(\tilde{X}_i)-10]$	$2[\Re(\tilde{X}_i)-12]$	$2[\Re(\tilde{X}_i)-14]$
$\Re(\tilde{X}_i) \ge 14$	$-16[\Re(\tilde{X}_i)-7]$	$8[\Re(\tilde{X}_i)-11]$	$4[\Re(\tilde{X}_i)-13]$	$2[\Re(\tilde{X}_i)-14]$

Table 5. Values of $l(b_k,\tilde{X}_i)$ for k = 4, 5, 6, 7 for 256-QAM signal

$\Im(\tilde{X}_i)$	$l(b_4,\tilde{X}_i)$	$l(b_5,\tilde{X}_i)$	$l(b_6,\tilde{X}_i)$	$l(b_7,\tilde{X}_i)$
$\Im(\tilde{X}_i) \le -14$	$16[\Im(\tilde{X}_i)+7]$	$-8[\Im(\tilde{X}_i)+11]$	$-4[\Im(\tilde{X}_i)+13]$	$-2[\Im(\tilde{X}_i)+14]$
$-14 \le \Im(\tilde{X}_i) < -12$	$14[\Im(\tilde{X}_i)+6]$	$-6[\Im(\tilde{X}_i)+10]$	$-2[\Im(\tilde{X}_i)+12]$	$-2[\Im(\tilde{X}_i)+14]$
$-12 \le \Im(\tilde{X}_i) < -10$	$12[\Im(\tilde{X}_i)+5]$	$-4[\Im(\tilde{X}_i)+9]$	$-2[\Im(\tilde{X}_i)+12]$	$2[\Im(\tilde{X}_i)+10]$
$-10 \le \Im(\tilde{X}_i) < -8$	$10[\Im(\tilde{X}_i)+4]$	$-2[\Im(\tilde{X}_i)+8]$	$-4[\Im(\tilde{X}_i)+11]$	$2[\Im(\tilde{X}_i)+10]$
$-8 \le \Im(\tilde{X}_i) < -6$	$8[\Im(\tilde{X}_i)+3]$	$-2[\Im(\tilde{X}_i)+8]$	$4[\Im(\tilde{X}_i)+5]$	$-2[\Im(\tilde{X}_i)+6]$
$-6 \le \Im(\tilde{X}_i) < -4$	$6[\Im(\tilde{X}_i)+2]$	$-4[\Im(\tilde{X}_i)+7]$	$2[\Im(\tilde{X}_i)+4]$	$-2[\Im(\tilde{X}_i)+6]$
$-4 \le \Im(\tilde{X}_i) < -2$	$4[\Im(\tilde{X}_i)+1]$	$-6[\Im(\tilde{X}_i)+6]$	$2[\Im(\tilde{X}_i)+4]$	$2[\Im(\tilde{X}_i)+2]$
$-2 \le \Im(\tilde{X}_i) < 0$	$2\Im(\tilde{X}_i)$	$-8[\Im(\tilde{X}_i)+5]$	$4[\Im(\tilde{X}_i)+3]$	$2[\Im(\tilde{X}_i)+2]$
$0 \le \Im(\tilde{X}_i) < 2$	$2\Im(\tilde{X}_i)$	$8[\Im(\tilde{X}_i)-5]$	$-4[\Im(\tilde{X}_i)-3]$	$-2[\Im(\tilde{X}_i)-2]$
$2 \le \Im(\tilde{X}_i) < 4$	$4[\Im(\tilde{X}_i)-1]$	$6[\Im(\tilde{X}_i)-6]$	$-2[\Im(\tilde{X}_i)-4]$	$-2[\Im(\tilde{X}_i)-2]$
$4 \le \Im(\tilde{X}_i) < 6$	$6[\Im(\tilde{X}_i)-2]$	$4[\Im(\tilde{X}_i)-7]$	$-2[\Im(\tilde{X}_i)-4]$	$2[\Im(\tilde{X}_i)-6]$
$6 \le \Im(\tilde{X}_i) < 8$	$8[\Im(\tilde{X}_i)-3]$	$2[\Im(\tilde{X}_i)-8]$	$-4[\Im(\tilde{X}_i)-5]$	$2[\Im(\tilde{X}_i)-6]$
$8 \le \Im(\tilde{X}_i) < 10$	$10[\Im(\tilde{X}_i)-4]$	$2[\Im(\tilde{X}_i)-8]$	$4[\Im(\tilde{X}_i)-11]$	$-2[\Im(\tilde{X}_i)-10]$
$10 \le \Im(\tilde{X}_i) < 12$	$12[\Im(\tilde{X}_i)-5]$	$4[\Im(\tilde{X}_i)-9]$	$2[\Im(\tilde{X}_i)-12]$	$-2[\Im(\tilde{X}_i)-10]$
$12 \le \Im(\tilde{X}_i) < 14$	$14[\Im(\tilde{X}_i)-6]$	$6[\Im(\tilde{X}_i)-10]$	$2[\Im(\tilde{X}_i)-12]$	$2[\Im(\tilde{X}_i)-14]$
$\Im(\tilde{X}_i) \ge 14$	$16[\Im(\tilde{X}_i)-7]$	$8[\Im(\tilde{X}_i)-11]$	$4[\Im(\tilde{X}_i)-13]$	$2[\Im(\tilde{X}_i)-14]$

Tables 2–5 show the equations required to calculate a small number of logarithms of likelihood ratios sufficient for a simplified decision-making algorithm for each of the bits defining any signal point of the QAM constellation. The equations are finalized in a form convenient for their practical use. The proposed algorithm requires significantly fewer calculations than the maximum likelihood algorithm which involves calculating logarithms of likelihood ratios for all possible combinations of bits.

The soft decision-making algorithm was developed for channels with white Gaussian noise. The efficiency of this algorithm for receiving OFDM signals in the presence of narrowband interference using soft decision encoding requires further analysis. These issues are discussed in the next section.

NOISE IMMUNITY OF QAM-OFDM SIGNAL RECEPTION

Below are the results of studying noise immunity when receiving QAM-OFDM signals in the presence of noise interference or a mixture of noise and narrowband interference. The study investigated the efficiency of using encoding to handle narrowband interference with demodulator soft decisions obtained using the algorithms described above.

The OFDM system was modeled using the *MATLAB*[1] tool. The number of FFT points in the formation of the OFDM signal amounts to 128, the guard interval length is –32, and the methods for modulating subcarriers are 16-QAM, 64-QAM, and 256-QAM.

The focus of the study was on modeling the signal and interference. According to Fig. 2, complex envelopes of signals are represented by numbers $(a + jb)$, where $a,b \in \{\pm1,\pm3,...,\pm(\sqrt{M}-1)\}$. In this case, the average signal strength is dependent on modulation multiplicity M. For comparable simulation results at different modulation multiplicities, coefficient l depending on M need to be introduced into the signal representation. Then complex envelopes of signals may be represented by numbers, as follows:

$$l(a+jb), \text{ where } a,b \in \{\pm1,\pm3,...,\pm(\sqrt{M}-1)\}.$$

Coefficient l should be selected so that the signal energy per transmitted symbol E_s does not depend on modulation multiplicity M. The average energy of a single symbol of M-QAM signal with duration T_s is as follows:

$$E_s = \frac{T_s}{M}\sum_{i=1}^{M}\frac{l^2(a_i^2+b_i^2)}{2}.$$

For certainty, $E_s = 1$ and $T_s = 1$ are assumed in modeling, while the desired signal-to-noise ratio (SNR) is provided by selecting the noise variance. Then

$$1 = \frac{l^2}{M}\sum_{i=1}^{M}\frac{(a_i^2+b_i^2)}{2}.$$

According to this relationship, it follows that $l = \frac{1}{\sqrt{5}}$ for 16-QAM signal, $l = \frac{1}{\sqrt{21}}$ for 64-QAM signal, and $l = \frac{1}{\sqrt{85}}$ for 256-QAM signal.

The energy per bit of the transmitted information is as follows:

$$E_b = E_s / \log_2 M = \frac{1}{\log_2 M}.$$

In the OFDM system, subcarriers are located on the frequency axis at distance $\frac{1}{T_s}$ apart. The noise variance in this band is $\sigma_n^2 = \frac{N_0}{T_s} = N_0$ (N_0 is single-sided noise spectral density).

The SNR, understood as the ratio of the average signal energy per bit of transmitted information to the noise spectral density:

$$\frac{E_b}{N_0} = \frac{1}{\sigma_n^2 \log_2 M}.$$

Hence, it may be written in the following way:

$$\sigma_n^2 = \frac{1}{\frac{E_b}{N_0}\log_2 M}.$$

This means that since the signal energy is assumed equal to one for all types of M-QAM modulation, the Gaussian noise with variance $\sigma_n^2 = \frac{1}{\frac{E_b}{N_0}\log_2 M}$ should be modeled to obtain the desired SNR $\frac{E_b}{N_0}$.

Figures 9–11 show the dependencies of bit error rate (BER) P_{eb} on SNR per bit of transmitted information for 16-QAM, 64-QAM, and 256-QAM. The simulation is performed for a channel with white Gaussian noise in the absence and in the presence of narrowband interference for the signal to interference ratio (SIR) equal to 0 dB in power. The transmission methods without encoding,

[1] https://www.mathworks.com/products/matlab.html. Accessed March 31, 2024.

with low-density parity-check (LDPC) encoding [19], and with convolutional encoding under soft and hard demodulator decisions were investigated. The code rate (R) is 1/2 for all types of encoding.

The convolutional code considered in this case is based on generating polynomials $G_1(X) = 1 + X + X^2 + X^3 + X^4$ and $G_2(X) = 1 + X + X^3 + X^4$. This code is decoded using the Viterbi algorithm. Using the convolutional code at rate 3/4 is also discussed below. The code is obtained from the code at rate 1/2 by poking every third output bit according to the pattern $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

At LDPC encoding at rate 1/2, the codeword length totals 648 bits, 324 of them being information ones. Soft decoding was performed according to the Belief propagation algorithm (or sum of products).

It can be seen in each figure that when encoding is used, BER P_{eb} is significantly reduced when compared to cases without encoding.

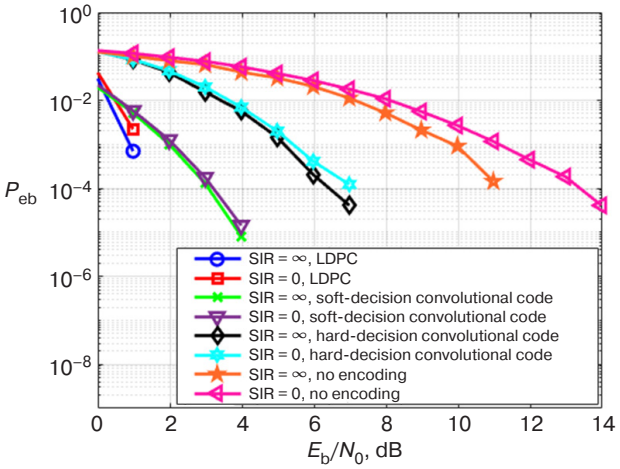


Fig. 9. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 16-QAM modulation and different types of encoding

Figure 9 shows that in the presence of narrowband interference and for $P_{eb} = 10^{-3}$, LDPC encoding gives 10 dB better results than no encoding, 4 dB better results than in the case of convolutional encoding with hard decision, and 1 dB better results than convolutional encoding with soft decision.

It follows from Fig. 10 that in the presence of narrowband interference and for $P_{eb} = 10^{-3}$, LDPC encoding gives 16 dB better results than no encoding, 6 dB better results than in the case of convolutional encoding with hard decision, and 1.5 dB better results than convolutional encoding with soft decision.

Figure 11 shows that in the presence of narrowband interference and at $P_{eb} = 10^{-3}$, LDPC encoding gives 10 dB better result than convolutional encoding with hard decision, and 2 dB better result than in the case of convolutional encoding with soft decision.

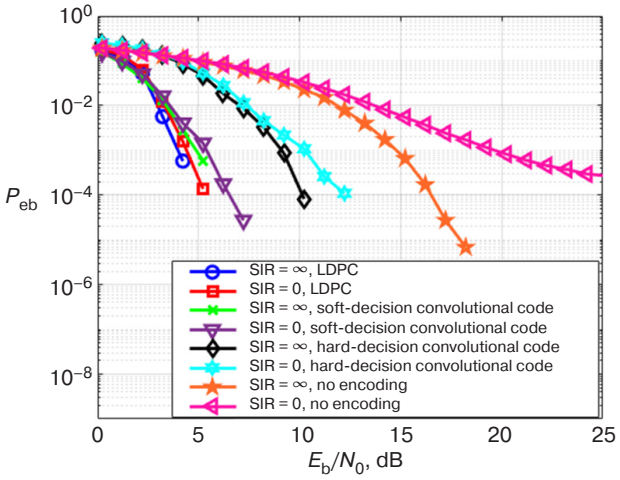


Fig. 10. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 64-QAM modulation and different types of encoding

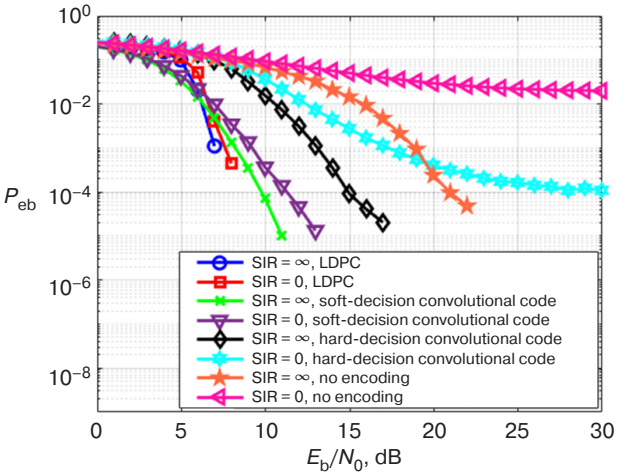


Fig. 11. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 256-QAM modulation and different types of encoding

Figures 9–11 suggests that in the presence of narrowband interference, the encoding reduces BER significantly compared to the case of no encoding. Encoding with soft demodulator decisions gives better results than with hard decisions. LDPC encoding with soft decisions gives better results than in the case of convolutional encoding with soft decisions.

Figures 12–14 show simulation results comparing the efficiencies of convolutional encoding and LDPC encoding with soft demodulator decisions and with $R = 1/2$ and 3/4 in the presence of narrowband interference: SIR is 0 dB in power.

It can be seen from Fig. 12 that at $P_{eb} = 10^{-3}$, LDPC encoding with soft demodulator decisions at $R = 1/2$ is 3.5 dB better than at $R = 3/4$. Convolutional encoding with soft demodulator decisions at $R = 1/2$ is 5 dB better than at $R = 3/4$. When $R = 1/2$, LDPC encoding with soft demodulator decisions is 0.5 dB better than in the case of convolutional encoding with soft demodulator decisions.

When R = 3/4, LDPC encoding with soft demodulator decisions is 5 dB better than in the case of convolutional encoding with soft demodulator decisions.

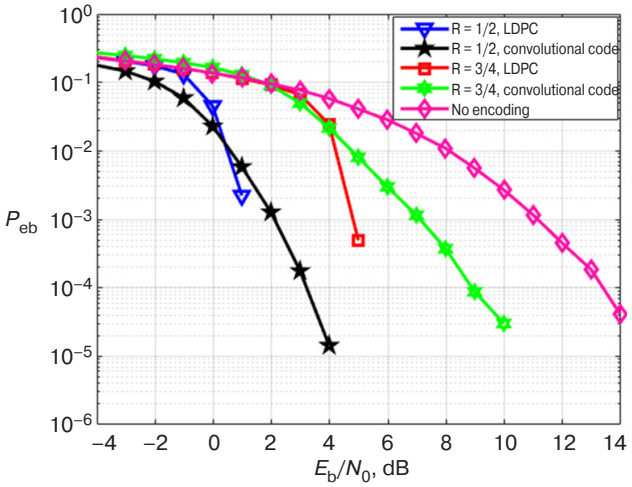


Fig. 12. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 16-QAM modulation and different code rates

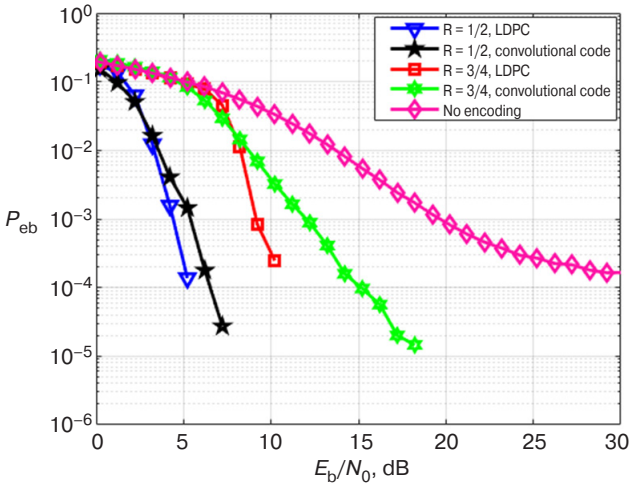


Fig. 13. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 64-QAM modulation and different code rates

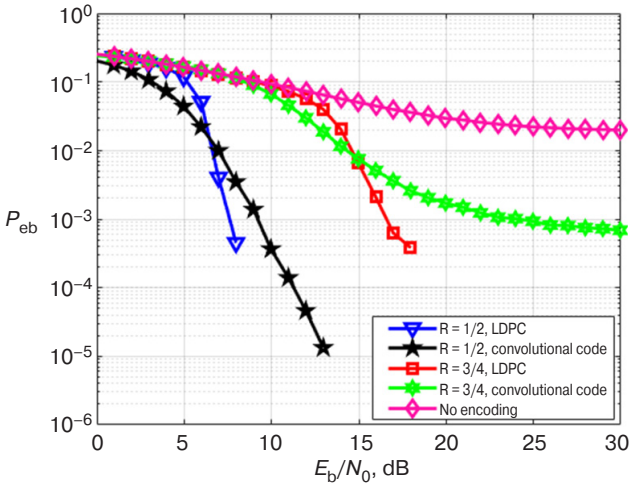


Fig. 14. Dependence of BER P_{eb} on SNR E_b/N_0 for OFDM system with 256-QAM modulation and different code rates

Figure 13 shows that at $P_{eb} = 10^{-3}$, LDPC encoding with soft demodulator decisions at R = 1/2 is energetically 4.5 dB better than at R = 3/4. The convolutional encoding with soft demodulator decisions at R = 1/2 is 6.5 dB better than at R = 3/4. At R = 1/2, LDPC encoding with soft demodulator decisions is 1 dB better than in the case of convolutional encoding with soft demodulator decisions. At R = 3/4, LDPC encoding with soft demodulator decisions is 3 dB better than convolutional encoding with soft demodulator decisions.

Figure 14 shows that at $P_{eb} = 10^{-3}$, LDPC encoding with soft demodulator decisions at R = 1/2 is 9 dB better than at R = 3/4. The convolutional encoding with soft demodulator decisions at R = 1/2 is 14 dB better than at R = 3/4. At R = 1/2, LDPC encoding with soft demodulator decisions is 1.5 dB better than convolutional encoding with soft demodulator decisions. At R = 3/4, LDPC encoding with soft demodulator decisions is 6.5 dB better than in the case of convolutional encoding with soft demodulator decisions.

Figures 12–14 suggest that LDPC encoding gives better results than convolutional encoding with soft demodulator decisions. In this case, R = 1/2 produces better reception immunity than rate 3/4.

Figures 15–17 show the results of evaluating the reception immunity of OFDM signal with QAM subcarrier modulation using soft demodulator decisions at SIR values: R = 1/2 for all types of encoding.

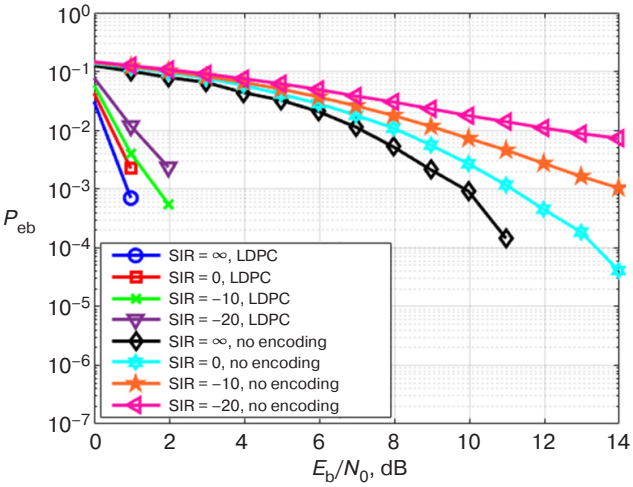


Fig. 15. Dependence of BER on SNR E_b/N_0 for OFDM system with 16-QAM in the presence of narrowband interference with different levels

Figure 15 shows that at $P_{eb} = 10^{-3}$, when there is no narrowband interference (SIR = ∞), LDPC encoding gives 9 dB better results than when no encoding is used. At SIR = 0 dB, LDPC encoding gives 10 dB better results than no encoding. At SIR = −10 dB,

LDPC encoding gives 11 dB better results than no encoding. At $P_{eb} = 10^{-3}$, when SIR = −20 dB, LDPC encoding gives at least 20 dB better results than transmission without encoding.

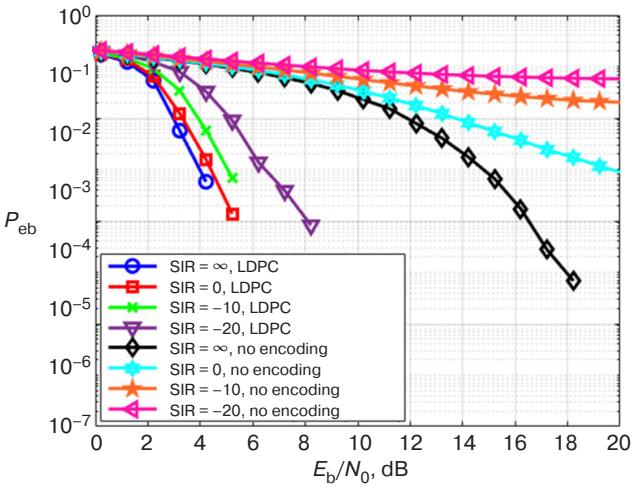


Fig. 16. Dependence of BER on SNR E_b/N_0 for OFDM system with 64-QAM in the presence of narrowband interference with different levels

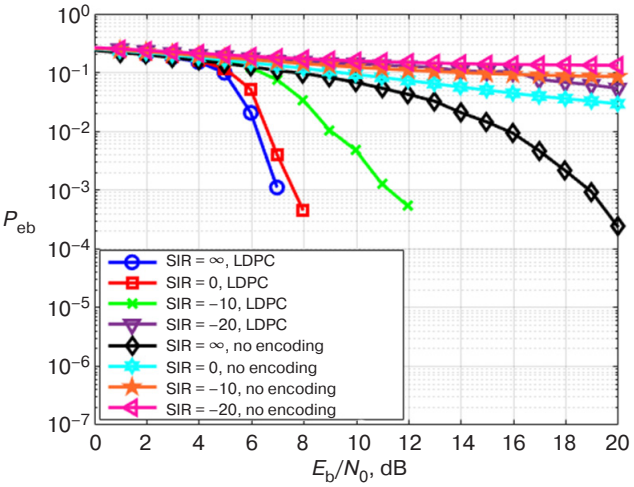


Fig. 17. Dependence of BER on SNR E_b/N_0 for OFDM system with 256-QAM in the presence of narrowband interference with different levels

Figure 16 shows that at $P_{eb} = 10^{-3}$, when there is no narrowband interference (SIR = ∞), LDPC encoding gives 11 dB better results than no encoding. At SIR = 0 dB, LDPC encoding gives 15.5 dB better results than no encoding. At SIR = –10 dB, BER drops below 10^{-3} with E_b/N_0 higher than 5 dB. At SIR = –20 dB. This occurs at E_b/N_0 higher than 8 dB.

Figure 17 shows that at $P_{eb} = 10^{-3}$, when there is no narrowband interference (SIR = ∞), LDPC encoding gives 12 dB better results than no encoding. At SIR = 0 dB, BER drops lower than 10^{-3} when E_b/N_0 is greater than 8 dB. At SIR = –10 dB, BER drops lower than 10^{-3} when E_b/N_0 is higher than 12 dB.

The results shown in Figs. 15–17 describe the noise immunity of OFDM signal reception in the presence of narrowband interference at different SIR with and without LDPC encoding. It can be observed that for all three types of modulation, the narrowband interference significantly degrades the noise immunity of the transmission system in the absence of encoding, and the noise immunity drops more significantly with increasing QAM multiplicity. The LDPC code reduces the impact of narrowband interference on the system significantly, being especially noticeable at small QAM multiplicities.

CONCLUSIONS

Based on these findings, it can be concluded that the use of encoding with soft demodulator decisions significantly improves the noise immunity of OFDM signal reception, while mitigating the impact of narrowband interference on the transmission system. LDPC encoding is superior to convolutional encoding in improving the noise immunity of OFDM signal reception, including in the presence of narrowband interference. In addition to use in QAM-OFDM systems, the proposed simple demodulation method for soft decision QAM signals can be applied to any wireless communication system using M-QAM signals, where M is number 2 to an even power.

Authors' contributions. All authors equally contributed to the research work.

REFERENCES

1. Cimini Jr. L.J. Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Trans. Commun*. 1985;33(7):665–675. Available from URL: https://doi.org/10.1109/TCOM.1985.1096357
2. Mosse P.H. A technique for orthogonal frequency division multiplexing frequency offset correction. *IEEE Trans. Commun*. 1994;42(10):2908–2914. https://doi.org/10.1109/26.328961
3. Van Nee R., Prasad R. *OFDM for Wireless Multimedia Communications*. Boston: Artech House; 2000. 260 p.
4. Pandey S., Bharti M., Agrawal A.K. Analysis of M-ary QAM-Based OFDM Systems in AWGN Channel. In: Bansal R.C., Agrawal A., Jadoun V.K. (Eds.). *Advances in Energy Technology: Select Proceedings of EMSME.* 2022;(766):223–235. https://doi.org/10.1007/978-981-16-1476-7_22
5. Van Wyk J., Linde L. Bit error probability for a M-ary QAM OFDM-based system. In: *AFRICON 2007*. IEEE; 2007. https://doi.org/10.1109/AFRCON.2007.4401578
6. Fuqin Xiong. M-ary amplitude shift keying OFDM system. *IEEE Trans. Commun*. 2003;51(10):1638–1642. https://doi.org/10.1109/TCOMM.2003.818103
7. Batra A., Zeidler J.R. Narrowband interference mitigation in OFDM systems. In: *MILCOM 2008. 2008 IEEE Military Communications Conference.* 2008. https://doi.org/10.1109/MILCOM.2008.4753296
8. Coleri S., Ergen M., Puri A., Bahai A. Channel estimation techniques based on pilot arrangement in OFDM systems. *IEEE Trans. Broadcast*. 2002;48(3):223–229. http://doi.org/10.1109/TBC.2002.804034
9. Lu B., Yue G., Wang X.D. Performance analysis and design optimization of LDPC-coded MIMO OFDM systems. *IEEE Trans. Signal Process.* 2004;52(2):348–361. https://doi.org/10.1109/TSP.2003.820991
10. Lu B., Wang X. Space-time code design in OFDM systems. In: *Globecom '00 – IEEE. Global Telecommunications Conference. Conference Record* (*Cat. No.00CH37137*). IEEE; 2000. V. 2. P. 1000–1004. https://doi.org/10.1109/GLOCOM.2000.891288
11. Wang Q., Onotera L.Y. Coded QAM using a binary convolutional code. *IEEE Trans. Commun.* 1995;43(6):2001–2004. https://doi.org/10.1109/26.387437
12. Mosleh M.F. Log-Likelihood Ratio to Improve Hard Decision Viterbi Algorithm. *Eng. & Tech. J.* 2013;31(9):1779–1790. https://doi.org/10.30684/etj.2013.82189
13. Hagenauer J., Hoeher P. A Viterbi algorithm with soft-decision outputs and its applications. In: *1989 IEEE Global Telecommunications Conference and Exhibition "Communications Technology for the 1990s and Beyond*." 1989;3:1680–1686. https://doi.org/10.1109/GLOCOM.1989.64230
14. Cao S., Kam P.Y., Yu C. Pilot-Aided Log-Likelihood Ratio for LDPC Coded MPSK-OFDM Transmission. *IEEE Photon. Technol. Lett.* 2013;25(6):594–597. https://doi.org/10.1109/LPT.2013.2246563
15. Cao S., Kam P.Y., Yu C. Pilot-aided log-likelihood ratio for LDPC coded M-QAM CO-OFDM system. *OFC 2014.* 2014, W3. https://doi.org/10.1364/OFC.2014.W3J.1
16. Chen J., Dholakia A., Ftheriou E., et al. Reduced-complexity decoding of LDPC codes. *IEEE Trans. Commun*. 2005;53(8):1288–1299. https://doi.org/10.1109/TCOMM.2005.852852
17. Jiabin T., Yue X., Lilin D., Wei X., et al. Efficient LLR Approximation for Coded Constant Envelope OFDM. *IEEE Trans. Vehicular Technol.* 2023;72(5):6194–6208. https://doi.org/10.1109/TVT.2022.3231912
18. Zhenyu Z., Caihong G., Hua L., et al. Soft-Input Soft-Output Detection via Expectation Propagation for Massive Spatial Modulation MIMO Systems. *IEEE Commun. Lett.* 2021;25(4):1173–1177. https://doi.org/10.1109/LCOMM.2020.3047081
19. Chu V.V. Noise immunity of OFDM transmission system when using LDPC code. In: *Fundamental, Exploratory, Applied Research and Innovation Projects: Proceedings of the National Scientific and Practical Conference*. Moscow: RTU MIREA; 2022. P. 389–392 (in Russ.).

About the authors

Alexey A. Paramonov, Dr. Sci. (Eng.), Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: paramonov@mirea.ru. Scopus Author ID 57208923552, RSCI SPIN-code 5605-9459, http://orcid.org/0000-0002-4537-4626

Chu Van Vuong, Postgraduate Student, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: muadem1110@gmail.com. http://orcid.org/0009-0003-0143-0168

Об авторах

Парамонов Алексей Анатольевич, д.т.н., профессор, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: paramonov@mirea.ru. Scopus Author ID 57208923552, SPIN-код РИНЦ 5605-9459, http://orcid.org/0000-0002-4537-4626

Чу Ван Вуонг, аспирант, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: muadem1110@gmail.com. http://orcid.org/0009-0003-0143-0168

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Dr. David Mossop

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.391.072
https://doi.org/10.32362/2500-316X-2024-12-5-33-41
EDN ELQHEK

RESEARCH ARTICLE

Noise immunity of signal reception with multiple frequency-shift keying against retransmitted interference

Alexandra E. Troitskaya,
Yuriy A. Polevoda [@],
Gennady V. Kulikov

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: polevoda@mirea.ru

Abstract

Objectives. Radio engineering information transmission systems are widely used in robotic systems employed in various military and civilian services. If such systems are used in a harsh environment, where a large amount of retransmitted interference occurs, for example, if a complex is buried under rubble or is located in reinforced concrete pipes or other utility facilities, communication with the command post may be lost. Thus, the task of maintaining reliable communications under difficult conditions of radio wave propagation is very urgent. In the field of telecommunications, multiposition types of modulation are widely used, which, despite their good spectral characteristics, provide low noise immunity under conditions of nonfluctuating interference, especially in cases of retransmitted interference. Therefore, it is relevant to explore the possibility of using multiple frequency-shift keying (M-FSK) signals in radio systems with complex interference environments. The paper sets out to analyze the noise immunity of coherent reception of M-FSK signals against the background of retransmitted interference.

Methods. Statistical radio engineering and mathematical modeling methods are used according to the theory of optimal signal reception.

Results. A model of the M-FSK signal and retransmitted interference is provided. The statistical parameters of the distributions of random processes occurring in a multichannel coherent receiver of M-FSK signals against retransmitted interference are obtained; based on this, the bit error rate is calculated when receiving M-FSK signals of different positionality M against retransmitted interference with different intensities.

Conclusions. The impact of retransmitted interference is shown to result in a decrease in the noise immunity of M-FSK signal reception, which is greater the higher its intensity. With increasing positionality of M-FSK signals at low intensity of retransmitted interference, the noise immunity of reception is significantly improved; however, high-intensity interference significantly increases the bit error rate. The presence of interference with a relative intensity of 0.5 causes energy losses from 4 to 6 dB depending on the positionality. When $M > 4$, M-FSK signals gain significantly in terms of noise immunity over signals with multiple phase-shift keying, quadrature amplitude modulation, and amplitude and phase-shift keying.

Keywords: multiple frequency-shift keying, retransmitted interference, noise immunity, bit error rate

• **Submitted:** 12.03.2024 • **Revised:** 03.04.2024 • **Accepted:** 12.07.2024

For citation: Troitskaya A.E., Polevoda Yu.A., Kulikov G.V. Noise immunity of signal reception with multiple frequency-shift keying against retransmitted interference. *Russ. Technol. J.* 2024;12(5):33–41. https://doi.org/10.32362/2500-316X-2024-12-5-33-41

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Помехоустойчивость приема сигналов с многопозиционной частотной манипуляцией на фоне ретранслированной помехи

А.Е. Троицкая,
Ю.А. Полевода [@],
Г.В. Куликов

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: polevoda@mirea.ru

Резюме

Цели. Радиотехнические системы передачи информации широко применяются в роботизированных комплексах для использования военными и гражданскими службами. При попадании такого комплекса в сложную окружающую среду, в которой возникает большое количество ретранслированных помех, например, при попадании комплекса под завал, в железобетонные трубы или различные коммунальные объекты, связь с командным пунктом может быть потеряна. Задача поддержания надежной связи в сложных условиях распространения радиоволн является весьма актуальной. В области телекоммуникаций широко используются многопозиционные виды модуляции, которые, несмотря на их хорошие спектральные характеристики, обеспечивают невысокую помехоустойчивость в условиях нефлуктуационных помех, особенно в случае ретранслированных помех. Представляется целесообразным исследовать возможность применения сигналов с многопозиционной частотной манипуляцией (М-ЧМ) в радиосистемах со сложной помеховой обстановкой. Целью работы является анализ помехоустойчивости когерентного приема сигналов М-ЧМ на фоне ретранслированной помехи.

Методы. Использованы методы статистической радиотехники, теории оптимального приема сигналов и математического моделирования.

Результаты. Приведена модель сигнала М-ЧМ и ретранслированной помехи. Получены статистические параметры распределений случайных процессов в многоканальном когерентном приемнике сигналов М-ЧМ на фоне ретранслированной помехи. На этой основе рассчитана вероятность битовой ошибки при приеме сигналов М-ЧМ разной позиционности M на фоне ретранслированной помехи с различной интенсивностью.

Выводы. Показано, что воздействие ретранслированной помехи приводит к снижению помехоустойчивости приема сигналов М-ЧМ, которое тем больше, чем выше интенсивность помехи. С возрастанием позиционности сигналов М-ЧМ при небольшой интенсивности ретранслированной помехи помехоустойчивость приема значительно улучшается, но помеха большой интенсивности сильно увеличивает вероятность битовой ошибки. Наличие помехи с относительной интенсивностью 0.5 вызывает энергетические потери от 4 до 6 дБ в зависимости от позиционности. При $M > 4$ сигналы М-ЧМ значительно выигрывают в помехоустойчивости у сигналов с многопозиционной фазовой, квадратурной амплитудной и амплитудно-фазовой манипуляцией.

Ключевые слова: многопозиционная частотная манипуляция, ретранслированная помеха, помехоустойчивость, вероятность битовой ошибки

• **Поступила:** 12.03.2024 • **Доработана:** 03.04.2024 • **Принята к опубликованию:** 12.07.2024

Для цитирования: Троицкая А.Е., Полевода Ю.А., Куликов Г.В. Помехоустойчивость приема сигналов с многопозиционной частотной манипуляцией на фоне ретранслированной помехи. *Russ. Technol. J.* 2024;12(5):33–41. https://doi.org/10.32362/2500-316X-2024-12-5-33-41

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Radio engineering information transmission systems are widely used in robotic complexes for military and civilian applications such as civil defense, emergencies and disaster relief[1] [1–4]. In a harsh environment involving a large amount of retransmitted interference[2] [5–7], e.g., when the complex gets under rubble, reinforced concrete pipes, or various utilities, communication with the command center may be lost. Therefore, maintaining reliable communication with such robotic complexes under complex conditions of radio wave propagation is highly relevant.

Currently, multiposition modulation types are widely used in telecommunications. It is known that signals with multiple phase-shift keying (M-PSK), quadrature amplitude manipulation (QAM), and amplitude and phase-shift keying (APSK), despite their good spectral characteristics, have low noise immunity under conditions of nonfluctuating interference, especially in cases of retransmitted interference [8–12]. Therefore, it makes sense to investigate the possibility of using M-FSK signals [13–15] in radio systems operating in a complex interference environment. The paper aims to analyze the noise immunity of the coherent reception of M-FSK signals against retransmitted interference.

MODELING M-FSK SIGNAL AND RETRANSMITTED INTERFERENCE

Multiple frequency-shift keying is a method for transmitting digital information using multiple frequencies. In communication technology, M-FSK signals of different M positionality (ranging from 4 to 64) are used, which have different bandwidths at a given rate of information transmission.

For an M-FSK signal, the ith version of sending a channel symbol with duration T_s, carrying information about $k = \log_2 M$ information bits, can be written as follows:

$$s_i(t) = A_0 \cos\left[\left(\omega_0 + \left(i - \frac{M+1}{2}\right)\right)\Delta\omega t\right], \quad i = \overline{1, M},\ t \in (0, T_s], \tag{1}$$

where $A_0 = \sqrt{2E_s / T_s}$ is the signal nominal amplitude, $E_s = kE_b$ is channel symbol energy, E_b is the energy per one bit of information; ω_0 is the center frequency of the signal spectrum (carrier frequency); $\Delta\omega = 2\Delta\omega_d$, $\Delta\omega_d$ is frequency deviation (the minimum deviation that fulfills the orthogonality condition is equal to $\pi/2T_s$); and t is time.

By introducing the notations of shift keying index $h = \Delta\omega_d T_s/\pi$ and multiposition channel symbol $C_i = \pm 1, \pm 3, ..., \pm(M-1)$, expression (1) can be represented in the following form [15]:

$$s_i(t) = A_0 \cos\left(\omega_0 t + C_i h \frac{\pi}{T_s} t\right).$$

An example of frequency distribution for 8-FSK signal is shown in Fig. 1. The signal points and their corresponding channel symbols are arranged according to the Gray code.

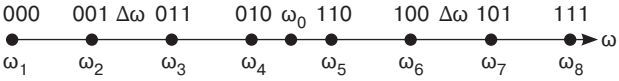


Fig. 1. Example of frequency arrangement for 8-FSK signal

[1] Order No. 633 dated December 26, 2018. On Approval and Enactment of the Radio Communications Manual of the Ministry of the Russian Federation for Civil Defense, Emergencies and Elimination of Consequences of Natural Disasters. https://base.garant.ru/72152196/ (in Russ.). Accessed November 25, 2023.

[2] ITU-R Recommendation P.1238-6 (10/2009). *Radio wave propagation data and prediction methods for planning indoor radio communication systems and local area radio networks in the frequency range 900 MHz to 100 GHz*. Series R. Radio wave propagation (in Russ.).

We use signal (1) with time delay $\tau \le T_s$ and random phase φ_{int} as the retransmitted interference arising due to M-FSK signal reflecting from some obstacle:

$$s_{int}(t) = \begin{cases} \mu A_0 \cos\left[\left(\omega_0 + \left(j - \frac{M+1}{2}\right)\Delta\omega\right)(t-\tau) + \varphi_{int}\right], \\ 0 < t \le \tau, \\ \mu A_0 \cos\left[\left(\omega_0 + \left(i - \frac{M+1}{2}\right)\Delta\omega\right)(t-\tau) + \varphi_{int}\right], \\ \tau < t \le T_s, \end{cases} \quad (2)$$

where μ is the relative intensity of the retransmitted interference, i and j are running signal numbers in the transmitted sequence.

The temporal relationship between the useful radio signal and retransmitted interference is shown in Fig. 2.

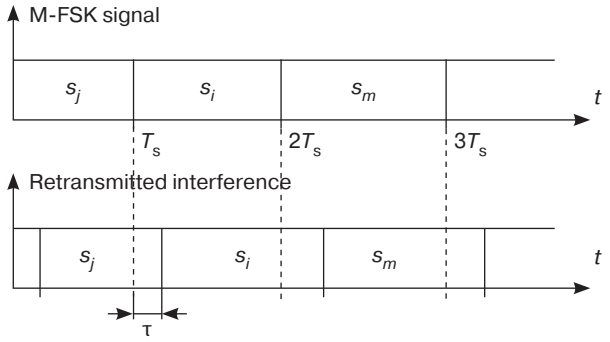


Fig. 2. Temporal relations between useful signal and retransmitted interference

We consider the presence of white Gaussian noise $n(t)$ with single-sided spectral density N_0 and δ correlation function at the receiver input of the information transmission system which are additional to the useful signal and retransmitted interference:

$$R(\tau) = \frac{N_0}{2}\delta(\tau).$$

CALCULATING THE BIT ERROR RATE

We consider a multichannel coherent demodulator [14, 15] as a receiver of M-FSK signals. In the presence of white Gaussian noise, the demodulator computes M integrals of convolution I_i of the received oscillation and M reference signals, as well as making a maximum likelihood decision on the channel symbol.

The bit error rate (BER) is calculated using the method described in [10–12]. For this, random parameters of retransmitted interference (2), such as intensity μ, phase φ_{int}, delay τ, and information symbol in the delayed copy, are assumed as initially fixed. This allows the probability of erroneous reception of a channel symbol conditioned on them to be determined:

$$P_{es}^* = 1 - \prod_{i=1}^{M} p_i(I_m > I_i)\Big|_m, \; m \ne i. \quad (3)$$

Given this condition, and assuming the presence of white Gaussian noise at the receiver input, the distributions of random processes being convolution integrals in the receiver channels and their linear combinations "*mi*" can be considered as normal. To obtain probabilities $p_i(I_m > I_i)$ included in (3) and defined through the standard probability integral (error function), it would be sufficient to calculate their conditional statistical characteristics:

- mathematical expectation

$$m_{mi}^* = \frac{2E_s}{N_0}\left\{1 + \mu\frac{\sin x}{x} \times \right.$$
$$\times \frac{\tau}{T_s}\cos\left[\left(\frac{M+1}{2} - \frac{j+m}{2}\right)\Delta\omega T_s\frac{\tau}{T_s} + \eta\right] +$$
$$+ \mu\left(1 - \frac{\tau}{T_s}\right)\cos\left[\left(\frac{M+1}{2} - m\right)\Delta\omega T_s\frac{\tau}{T_s} + \eta\right] - \frac{\sin y}{y} -$$
$$- \mu\frac{\tau}{T_s}\cdot\frac{\sin z}{z}\cos\left[\left(\frac{M+1}{2} - \frac{(j+i)}{2}\right)\Delta\omega T_s\frac{\tau}{T_s} + \eta\right] -$$
$$- \left(1 - \frac{\tau}{T_s}\right)\mu\frac{\sin g}{g}\cos\left[\frac{(m-i)\Delta\omega T_s}{2} - \right.$$
$$\left.\left. - \left(\frac{m+i}{2} + \frac{M+1}{2}\right)\Delta\omega T_s\frac{\tau}{T_s} + \eta\right]\right\},$$

- variance

$$D_{mi}^* = \frac{4E_s}{N_0}\left(1 - \frac{\sin y}{y}\right),$$

where

$$\eta = \varphi_{int} - \omega_0\tau,\; x = (j-m)\frac{\Delta\omega T_s}{2}\cdot\frac{\tau}{T_s},$$
$$y = (m-i)\Delta\omega T_s,\; z = (j-i)\frac{\Delta\omega T_s}{2}\cdot\frac{\tau}{T_s},$$
$$g = (m-i)\frac{\Delta\omega T_s\left(1 - \frac{\tau}{T_s}\right)}{2}.$$

For obtaining the unconditional error probability P_{es}, the averaging of conditional probability P_{es}^* for

random interference parameters, such as the random phase φ_{int}, is carried out using a numerical method. Parameters μ and τ are assumed to be fixed. At the same time, the calculations are performed with an enumeration of all possible combinations of channel symbols. When calculating the BER P_{eb}, recommendation [14] describing the relationship between its value and the error probability when receiving channel symbol P_{es} for multiposition orthogonal signals is taken into account:

$$P_{eb} = P_{es}/((M-1)/(M/2)).$$

CALCULATION RESULTS

Based on the derived formulas, the noise immunity of the M-FSK signal reception with positionalities M = 2, 4, 8, 16, 32 is calculated. During the study, the following radio channel parameters vary: signal-to-noise ratio (SNR) E_b/N_0 lies within the range from 1 to 13 dB; the relative intensity of retransmitted interference μ lies in the range [0; 1]; the relative interference delay τ/T_s = 0.5. The calculation results are presented in Figs. 3 and 4.

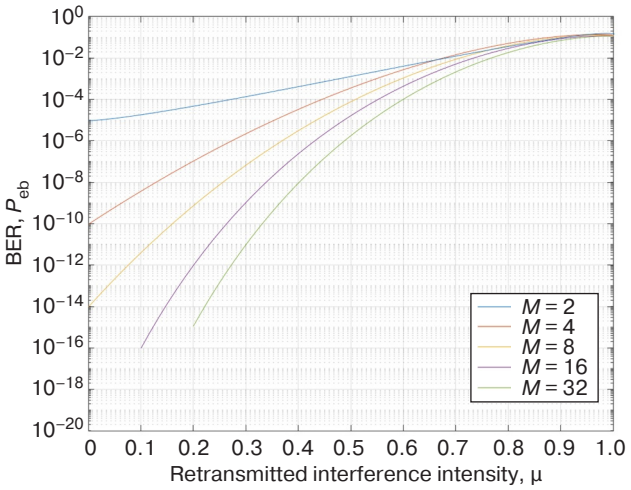


Fig. 3. Dependencies of BER P_{eb} on the relative intensity of the retransmitted interference μ at E_b/N_0 = 13 dB

The exposure to retransmitted interference results in reduced noise immunity: the higher the interference intensity, the greater the reduction. From this perspective, signals with higher positionality ($M \geq 8$) are preferable; at low relative interference intensity (μ < 0.3), these provide a BER several orders of magnitude lower than low-position 2-FSK and 4-FSK signals.

Even at high relative intensity of retransmitted interference (μ ∈ [0.4; 0.5]), the use of signals with multiple frequency-shift keying allows a fairly low bit error rate (10^{-5}–10^{-4}) to be maintained, which is not available when using other types of QAM [10], M-PSK [11], and APSK [12] signals at the same positionality.

The energy losses due to the impact of retransmitted interference during reception of M-FSK signals can be estimated by plotting the dependence of BER P_{eb} on SNR E_b/N_0 at different relative interference intensities μ (Fig. 4).

From the graphs, it can be seen that increasing the M-FSK signal positionality from 2 to 32 significantly improves the reception noise immunity. For example, at a low intensity of retransmitted interference (μ ≤ 0.1) and SNR (E_b/N_0 = 10 dB), the BER is seven times lower.

The presence of interference with relative intensity μ = 0.5 at $P_{eb} = 10^{-4}$ causes an energy loss of 5 to 6 dB depending on positionality. Nevertheless, the reception noise immunity of M-FSK signals of high positionality remains generally high.

COMPARING RECEPTION NOISE IMMUNITY OF M-FSK SIGNALS WITH OTHER TYPES OF MULTIPOSITION MODULATION

For comparison, the graphs of BER dependence on the relative intensity of retransmitted interference while receiving M-FSK and QAM signals [10] of the same positionality 4, 16, and 32 are shown in Fig. 5.

It can be seen that, at positionality M = 4, the reception noise immunity of these signals is approximately the same; however, with an increasing value of M, signals with M-FSK improve significantly (by several times for BER) over QAM signals across a wide range of retransmitted interference intensities (μ ∈ [0; 0.6]). A similar conclusion can be made with respect to APSK signals [12], whose reception immunity approximately corresponds to QAM and, even more so, with respect to M-PSK signals [11] being lost to QAM. However, when mentioning the advantages of M-FSK signals, their spectral characteristics, according to which significant losses to other signals may occur, should be taken into account [13, 15].

CONCLUSIONS

In the present work, the reception noise immunity of signals with multiple frequency-shift keying of M positionality from 2 to 32 is analyzed in the presence of retransmitted interference in the communication channel for different SNR and relative interference intensity values. The research allows the following conclusions to be drawn:

1. Increasing positionality improves noise immunity of M-FSK signal reception while decreasing BER by several orders of magnitude at a low intensity of retransmitted interference.

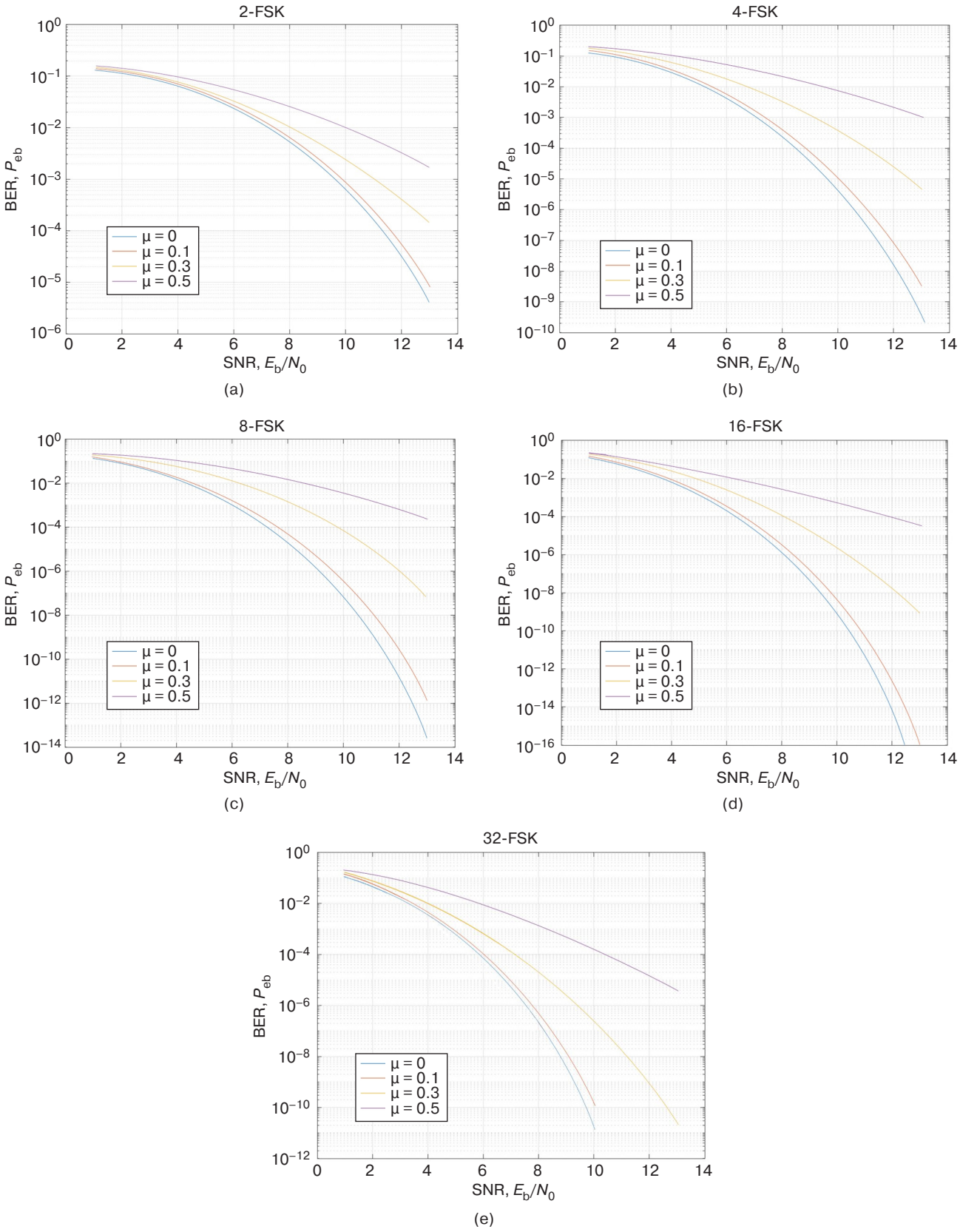


Fig. 4. Dependencies of BER P_{eb} on SNR E_b/N_0 under the influence of retransmitted interference for signals:
(a) 2-FSK,
(b) 4-FSK,
(c) 8-FSK,
(d) 16-FSK,
(e) 32-FSK

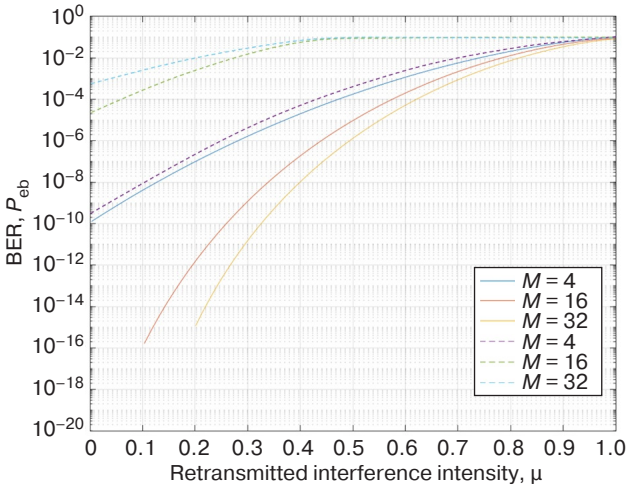


Fig. 5. Dependencies of BER P_{eb} on the relative intensity μ of retransmitted interference while receiving M-FSK (solid lines) and QAM (dashed lines) signals of the same positionality

2. The impact of retransmitted interference reduces M-FSK signal reception noise immunity, which is greater the higher its intensity. For example, the presence of interference with a relative intensity μ = 0.5 at $P_{eb} = 10^{-4}$ causes energy losses from 4 to 6 dB depending on the M-FSK signal positionality.
3. At $M > 4$, M-FSK signals significantly gain in noise immunity over M-PSK, QAM, and APSK signals across a wide range of retransmitted interference intensities.
4. The use of M-FSK signals with such high energy characteristics is justified in radio channels without significant frequency limitations.

Authors' contributions
A.E. Troitskaya—conducting computer calculations.
Yu.A. Polevoda—processing the results.
G.V. Kulikov—the research idea, consultations on the issues of conducting all stages of the study.

REFERENCES

1. Askhadeev A.I., Pavlov E.V., Barannik A.Y., Lagutina A.V., Kozlov V.I., Penkov I.A., Chirko O.V. Emercom of Russia Robotics System. State and Prospects of Development. *Tekhnologii grazhdanskoi bezopasnosti = Civil Security Technologies*. 2022;19(2–72):41–47 (in Russ.). Available from URL: https://www.vniigochs.ru/upload/medialibrary/6b3/ufktj81owcrv39y864piuh8dtzmw0njo/p_8_Emercom_Russia_Robotics_System_tgb_2_2022.pdf
2. Antokhin E.A., Panasenko N.N., Chernova A.D. Main requirements to wireless communication channels of military ground robotic complexes. *Robototekhnika i tekhnicheskaya kibernetika = Robotics and Technical Cybernetics*. 2017;4(17):10–14 (in Russ.).
3. Andreev V.P. Ensuring the safety of the work of film and TV journalists by means of extreme robotics. In: *Proceedings of the International Scientific and Technological Conference EXTREME ROBOTICS*. 2022;1(1):25–34 (in Russ.). Available from URL: https://er.rtc.ru/images/docs/2022/Proceedings_ER-2022.pdf
4. Manko S.V., Lokhin V.M., Diane S.K. Prototype multi-agent robotic debris removal system: principles of development and experimental studies. *Russ. Technol. J*. 2022;10(6):28–41 (in Russ.). https://doi.org/10.32362/2500-316X-2022-10-6-28-41
5. Gureev A.V., Gureev I.A., Pavlyuk Y.M., Thurain T. The effect of multipath propagation on characteristics of indoor broadband wireless systems. *Elektronnye informatsionnye sistemy = Electronic Information Systems*. 2017;4(15):17–23 (in Russ.).
6. Gureev A.V. Energetic characteristics of indoor wave propagation. *Izvestiya vysshikh uchebnykh zavedenii. Elektronika = Proceedings of Universities. Electronics*. 2015;20(4):421–430 (in Russ.).
7. Gureev A.V., Gureev I.A., Pavlyuk Y.M., Zhmylev V.A., Kachurenko D.S. Investigation of characteristics of electromagnetic waves propagation in buildings. *Elektronnye informatsionnye sistemy = Electronic Information Systems*. 2019;4(23):95–101 (in Russ.).
8. Bashkirov A.V., Krisilov A.V., Mashin V.V., Muratov A.V., Khoroshailova M.V. Analysis of symbol error rate for 16-QAM signals in the presence of QPSK signals and 16-QAM signals. *Radiotekhnika = Radioengineering*. 2020;84(6–12):5–9 (in Russ.). https://doi.org/10.18127/j00338486-202006(12)-01
9. Sidelnikov G.M. Comparative analysis of efficiency of spaced reception of signals with quadrature amplitude and with phase modulation in a channel with discrete multipath. *Vestnik Povolzhskogo gosudarstvennogo tekhnologicheskogo universiteta. Seriya: Radiotekhnicheskie i infokommunikatsionnye sistemy = Vestnik of Volga State University of Technology. Series: Radio Engineering and Infocommunication Systems*. 2020;2(46):18–30 (in Russ.). https://doi.org/10.25686/2306-2819.2020.2.18
10. Kulikov G.V., Lelyukh A.A., Grachenko E.N. Noise immunity of coherent signal receiver with quadrature amplitude modulation in the presence of relayed interference. *J. Commun. Technol. Electron*. 2020;65(8):934–938. https://doi.org/10.1134/S1064226920070074
    [Original Russian Text: Kulikov G.V., Lelyukh A.A., Grachenko E.N. Noise immunity of coherent signal receiver with quadrature amplitude modulation in the presence of relayed interference. *Radiotekhnika i Elektronika*. 2020;65(8):804–808 (in Russ.). https://doi.org/10.31857/S0033849420070074]
11. Nguyen V.D. Noise immunity of correlation receiver of signal with multi-position phase shift keying in the presence of retransmitted interference. *Zhurnal radioelektroniki = J. Radio Electronics*. 2019;3 (in Russ.). https://doi.org/10.30898/1684-1719.2019.3.4

12. Kulikov G.V., Dang X.Kh. Noise immunity of reception of signal with amplitude-phase shift keying in a two-path communication channel. *Voprosy radioelektroniki. Seriya: Tekhnika televideniya = Questions of Radio Electronics. Series: TV Technique*. 2022;2:43–49 (in Russ.).
13. Haykin S. *Communication Systems*. John Wiley & Sons. Inc.; 2001. 816 p.
14. Proakis J. *Tsifrovaya svyaz'* (*Digital Communication*): transl. from Engl. Moscow: Radio i svyaz'; 2000. 800 p. (in Russ.). [Proakis J. *Digital Communications.* N.Y.: McGraw-Hill; 1995. 928 p.]
15. Xiong F. *Digital Modulation Techniques.* 2nd ed. Boston, London: Artech House, Inc.; 2006. 1039 p.

СПИСОК ЛИТЕРАТУРЫ

1. Асхадеев А.И., Павлов Е.В., Баранник А.Ю., Лагутина А.В., Козлов В.И., Пеньков И.А., Чирко О.В. Система робототехники МЧС России. Состояние и перспективы развития. *Технологии гражданской безопасности.* 2022;19(2–72): 41–47. URL: https://www.vniigochs.ru/upload/medialibrary/6b3/ufktj81owcrv39y864piuh8dtzmw0njo/p_8_Emercom_Russia_Robotics_System_tgb_2_2022.pdf
2. Антохин Е.А., Панасенко Н.Н., Чернова А.Д. Основные требования к беспроводным каналам связи наземных робототехнических комплексов военного назначения. *Робототехника и техническая кибернетика*. 2017;4(17):10–14.
3. Андреев В.П. Обеспечение безопасности работы кино- и тележурналистов средствами экстремальной робототехники. *Экстремальная робототехника*. *Труды 33-й Международной научно-технической конференции*. 2022;1(1):25–34. URL: https://er.rtc.ru/images/docs/2022/Proceedings_ER-2022.pdf
4. Манько С.В., Лохин В.М., Диане С.К. Принципы построения и экспериментальные исследования прототипного образца многоагентной робототехнической системы для разбора завалов. *Russian Technological Journal*. 2022;10(6): 28–41. https://doi.org/10.32362/2500-316X-2022-10-6-28-41
5. Гуреев А.В., Гуреев И.А., Павлюк Ю.М., Тхурайн Т. Влияние многолучевости распространения радиоволн на характеристики широкополосных беспроводных систем внутри зданий. *Электронные информационные системы*. 2017;4(15):17–23.
6. Гуреев А.В. Энергетические характеристики распространения электромагнитных волн внутри зданий. *Известия высших учебных заведений. Электроника*. 2015;20(4):421–430.
7. Гуреев А.В., Гуреев И.А., Павлюк Ю.М., Жмылев В.А., Качуренко Д.С. Исследование характеристик распространения электромагниных волн внутри зданий. *Электронные информационные системы*. 2019;4(23):95–101.
8. Башкиров А.В., Крисилов А.В., Машин В.В., Муратов А.В., Хорошайлова М.В. Анализ вероятностей символьных ошибок для 16-QAM-сигналов в присутствии QPSK-сигналов и 16-QAM-сигналов. *Радиотехника*. 2020;84(6–12): 5–9. https://doi.org/10.18127/j00338486-202006(12)-01
9. Сидельников Г.М. Сравнительный анализ эффективности разнесенного приема сигналов с квадратурной амплитудной и с фазовой модуляцией в канале с дискретной многолучевостью. *Вестник Поволжского государственного технологического университета. Серия: Радиотехнические и инфокоммуникационные системы*. 2020;2(46):18–30. https://doi.org/10.25686/2306-2819.2020.2.18
10. Куликов Г.В., Лелюх А.А., Граченко Е.Н. Помехоустойчивость когерентного приемника сигналов с квадратурной амплитудной манипуляцией при наличии ретранслированной помехи. *Радиотехника и электроника*. 2020;65(8): 804–808. https://doi.org/10.31857/S0033849420070074
11. Нгуен В.З. Помехоустойчивость корреляционного приемника сигналов с многопозиционной фазовой манипуляцией при наличии ретранслированной помехи. *Журнал радиоэлектроники.* 2019;3. https://doi.org/10.30898/1684-1719.2019.3.4
12. Куликов Г.В., Данг С.Х. Помехоустойчивость приема сигналов с амплитудно-фазовой манипуляцией в двухлучевом канале связи. *Вопросы радиоэлектроники. Серия: Техника телевидения*. 2022;2:43–49.
13. Haykin S. *Communication Systems*. John Wiley & Sons. Inc.; 2001. 816 p.
14. Прокис Дж. *Цифровая связь*: пер. с англ.; под ред. Д.Д. Кловского. М.: Радио и связь; 2000. 800 с.
15. Xiong F. *Digital Modulation Techniques.* 2nd ed. Boston, London: Artech House, Inc.; 2006. 1039 p.

About the authors

Alexandra E. Troitskaya, Student, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: troitskaya.a.e@mirea.ru. https://orcid.org/0009-0002-9315-8829

Yuriy A. Polevoda, Postgraduate Student, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: polevoda@mirea.ru. https://orcid.org/0009-0007-6327-9685

Gennady V. Kulikov, Dr. Sci. (Eng.), Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kulikov@mirea.ru. Scopus Author ID 36930533000, RSCI SPIN-code 2844-8073, http://orcid.org/0000-0001-7964-6653

Об авторах

Троицкая Александра Евгеньевна, студент, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: troitskaya.a.e@mirea.ru. https://orcid.org/0009-0002-9315-8829

Полевода Юрий Александрович, аспирант, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: polevoda@mirea.ru. https://orcid.org/0009-0007-6327-9685

Куликов Геннадий Валентинович, д.т.н., профессор, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kulikov@mirea.ru. Scopus Author ID 36930533000, SPIN-код РИНЦ 2844-8073, http://orcid.org/0000-0001-7964-6653

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Dr. David Mossop

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.391, 621.396
https://doi.org/10.32362/2500-316X-2024-12-5-42-49
EDN EMMZPA

RESEARCH ARTICLE

Method for limiting the peak factor using an additional compensation signal in a system with orthogonal frequency division multiplexing for a Gaussian channel

Pham Thanh Tuan [@], Olga V. Tikhonova

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: anhsequayve.ru@gmail.com

Abstract

Objectives. Orthogonal frequency division multiplexing (OFDM) has become the standard for various high-speed wireless communication systems due to its several advantages, one of which is the efficient use of bandwidth. The main disadvantage of OFDM is the high peak-power-to-average-power ratio (PAPR), which is indicated by an increase in the bit error rate due to the nonlinearity of the power amplifier. The paper sets out to evaluate the possibility of reducing the PAPR value using a limitation method developed by the authors involving an additional compensation signal in a channel with white Gaussian noise, as well as to analyze its main parameters.

Methods. Statistical radio engineering and computer modeling methods are used according to optimal signal reception theory.

Results. The effect of the OFDM signal limitation level and the number of additional signals when using the limitation method with an additional compensation signal on the quality (reduction of transmission losses) of the OFDM signal is analyzed. The results show a decrease in the OFDM signal PAPR value, along with the dependencies of the bit error rate on the signal-to-noise ratio at fixed limitation values and a determined number of additional signals in the channel with white Gaussian noise.

Conclusions. The proposed limitation method with an additional signal when transmitting an OFDM signal in a channel with white Gaussian noise provides compensation for information losses due to signal level limitations in the transmission channel. Increasing the limitation level is shown to increase the PAPR value, while varying the number of additional signals changes PAPR insignificantly. In order to ensure the effective implementation of the limitation method with an additional compensation signal, the parameters of the threshold limitation level and number of additional signals should be selected depending on the predicted signal-to-noise ratio in the system.

Keywords: limitation method with additional signal, peak power to average power ratio, bit error rate, orthogonal frequency division multiplexing, white Gaussian noise

• **Submitted:** 23.01.2024 • **Revised:** 29.03.2024 • **Accepted:** 12.07.2024

For citation: Pham Th.T., Tikhonova O.V. Method for limiting the peak factor using an additional compensation signal in a system with orthogonal frequency division multiplexing for a Gaussian channel. *Russ. Technol. J*. 2024;12(5):42–49. https://doi.org/10.32362/2500-316X-2024-12-5-42-49

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Метод ограничения пик-фактора с дополнительным сигналом компенсации в системе с ортогональным частотным разделением каналов для гауссовского канала

Фам Тхань Туан [@], О.В. Тихонова

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: anhsequayve.ru@gmail.com

Резюме

Цели. Мультиплексирование с ортогональным частотным разделением каналов (orthogonal frequency division multiplexing, OFDM) стало стандартом для различных систем беспроводной связи с высокой скоростью передачи данных благодаря нескольким преимуществам, одним из которых является эффективное использование полосы частот. Главным недостатком OFDM является высокое значение отношения пиковой мощности к средней мощности – пик-фактору, на что указывает повышение частоты битовых ошибок из-за нелинейности усилителя мощности. Цель статьи – оценка возможности снижения значения пик-фактора с помощью разработанного авторами метода ограничения с дополнительным сигналом компенсации в канале с белым гауссовским шумом и анализ его основных параметров.

Методы. Использованы методы статистической радиотехники, теории оптимального приема сигналов и компьютерного моделирования.

Результаты. Проведен анализ влияния уровня ограничения сигнала OFDM и числа дополнительных сигналов при использовании метода ограничения с дополнительным сигналом компенсации на качество (уменьшение потерь при передаче) сигнала OFDM. Получены результаты, показывающие снижение величины пик-фактора OFDM-сигнала, и зависимости вероятности битовой ошибки от отношения сигнал/шум при фиксированных значениях ограничения и числа дополнительных сигналов в канале с белым гауссовским шумом.

Выводы. Предложен метод ограничения с дополнительным сигналом при передаче OFDM-сигнала в канале с белым гауссовским шумом, обеспечивающий компенсацию потерь информации, обусловленных ограничениями уровня сигнала в канале передачи. Установлено, что увеличение уровня ограничения увеличивает величину пик-фактора, а при изменении числа дополнительных сигналов пик-фактор изменяется незначительно. Для эффективной реализации метода ограничения с дополнительным сигналом компенсации необходимо выбирать параметры уровня ограничения порогового значения и число дополнительных сигналов в зависимости от прогнозируемого отношения сигнал/шум в системе.

Ключевые слова: метод ограничения с дополнительным сигналом, пик-фактор, вероятность битовой ошибки, ортогональное частотное разделение каналов, белый гауссовский шум

• **Поступила:** 23.01.2024 • **Доработана:** 29.03.2024 • **Принята к опубликованию:** 12.07.2024

Для цитирования: Фам Т.Т., Тихонова О.В. Метод ограничения пик-фактора с дополнительным сигналом компенсации в системе с ортогональным частотным разделением каналов для гауссовского канала. *Russ. Technol. J.* 2024;12(5):42–49. https://doi.org/10.32362/2500-316X-2024-12-5-42-49

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is currently applied to various high-speed digital information transmission systems [1–4], including high-speed mobile communications, wireless local networks, video broadcasting, and high-speed cellular data transmission. For the wider application of OFDM systems, one of the problems requiring a solution is how to reduce high peak-to-average power ratios (PAPR) that increase the bit error rate (BER) of the system.

In the present work, the possibility of reducing the PAPR value using a limitation method with an additional signal for an OFDM system in a channel with white Gaussian noise is evaluated along with the main method parameters.

OFDM SIGNAL

A discrete-time OFDM signal can be mathematically represented as a set of time samples of a signal following the inverse discrete Fourier transform:

$$S_{\text{OFDM}}(n)=\frac{1}{N}\sum_{k=0}^{N-1}X(k)\mathrm{e}^{\frac{j2\pi kn}{N}},\quad n=\overline{0,N-1}, \quad (1)$$

where N is the number of subcarriers, $X(k)$ is the transmitted information on the kth subcarrier, n is the OFDM signal sample number, and j is an imaginary unit.

For a discrete-time OFDM signal $S_{\text{OFDM}}(n)$, the value of the peak-to-average power ratio (PAPR) is defined as the ratio of the signal maximum power to its average power [5, 6], which can be represented in the following form:

$$\text{PAPR}(S_{\text{OFDM}}(n))=\frac{\max_n\left|S^2_{\text{OFDM}}(n)\right|}{P_{\text{avg}}\left\{\left|S^2_{\text{OFDM}}(n)\right|\right\}}, \quad (2)$$

where $\max_n\left|S^2_{\text{OFDM}}(n)\right|$ is the OFDM signal maximum power, $P_{\text{avg}}\left\{\left|S^2_{\text{OFDM}}(n)\right|\right\}$ is the OFDM signal average power, and $\left|S_{\text{OFDM}}(n)\right|$ is the nth sample amplitude of the OFDM signal.

The PAPR parameter represents the value of the peak-to-average power ratio, i.e., its higher value; thus, a high PAPR means higher maximum power.

LIMITATION METHOD WITH ADDITIONAL SIGNAL

One of the methods for reducing PAPR is the limitation method with additional signal for OFDM signal [7, 8]. According to this method, the amplitudes of OFDM signal samples are limited when a certain threshold value is exceeded while preserving the original phase of the nth sample. This can be represented in the following form:

$$S_{\text{lim}}(n)=\begin{cases}S_{\text{init}}(n)\ \text{at}\ \left|S_{\text{init}}(n)\right|\le C,\\ C\mathrm{e}^{j\phi_n}\ \text{at}\ \left|S_{\text{init}}(n)\right|>C,\end{cases} \quad (3)$$

where S_{lim} is the signal following OFDM signal limitation, C is the threshold value or signal limitation level, $\left|S_{\text{init}}(n)\right|$ is the amplitude of the initial nth signal sample, and ϕ_n is the phase angle of the nth signal sample.

According to the limitation method with an additional signal for signal $S_{\text{OFDM}}(n)$, a search is carried out for the positions (position numbers are denoted by p) where the signal amplitude is greater than the threshold value C. In this case, the positions are limited in accordance with expression (3). When transmitting a signal limited in this way, a loss of information may occur in some positions. In order to compensate for these losses, the transmitted signal S_{lim} is augmented with samples containing information about position numbers p and additional data $S_{\text{add}}(p)$ as the difference between the initial and limited samples, which can be described by the following expression:

$$S_{\text{add}}(p)=S_{\text{init}}(p)-S_{\text{lim}}(p). \quad (4)$$

Thus, the transmitted signal S_{lim} does not contain samples whose amplitude exceeds the threshold value C; however, no loss of information occurs since it is possible to restore the original samples using their numbers p and additional data $S_{\text{add}}(p)$. The number of additional samples reserved for transmitting position

numbers p and values of $S_{add}(p)$ is denoted as the number of additional signals m.

In order to achieve the optimal system performance, it is necessary to select the level of limitation C and the number of additional signals m. When the level of signal limitation C is significant (a large number of signals are subjected to transformation (3)) while keeping the possible number of additional signals m low, full compensation cannot be achieved resulting in a high number of bit errors occurring in the system. A high value of m results in an increased length of the transmitted information and consequent loss of energy and power. At an insignificant level of limitation C and high value of m, many additional signals do not carry any information since the number of signal samples S_{init} exceeding the limitation is insignificant.

The impact of limiting the level of C parameters and number of additional signals m on the OFDM signal compensation efficiency was studied using mathematical modeling based on the above formulas. The block diagram of the algorithm using *MATLAB*[1] mathematical package is shown in Fig. 1. Following modification using the limitation method with an additional signal at given values of C and m, OFDM signals are transmitted through a channel with white Gaussian noise. On the receiving side, the received data array is compared with the original array following the OFDM signal recovery process (demodulation with allowance for additional compensating information) to obtain the BER value.

[1] https://www.mathworks.com/products/matlab.html. Accessed March 29, 2024.

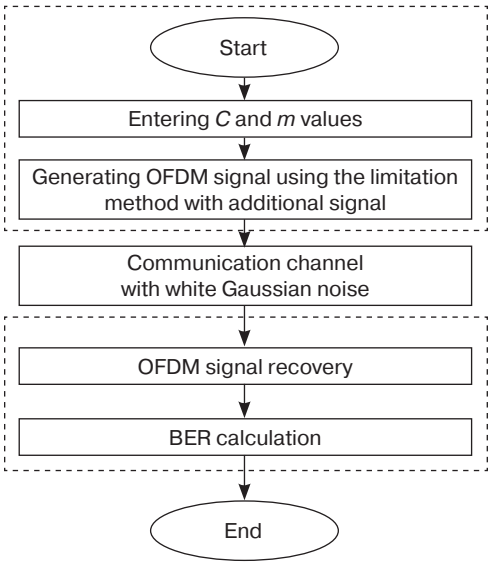


Fig. 1. Block diagram of the algorithm for modeling the impact of parameters C and m on the quality of OFDM signal compensation

The block diagram of the algorithm for calculating PAPR is presented in Fig. 2. The results are obtained for two cases. In the first case, when $m = 16$ for each value of C is fixed, OFDM signals are limited according to (3); here, the PAPR value after limiting the OFDM signal is calculated using formula (2). In the second case, similar calculations are performed at the fixed value $C = 0.18$ for each value of m. The resulting PAPR dependencies are shown in Figs. 3a and 3b.

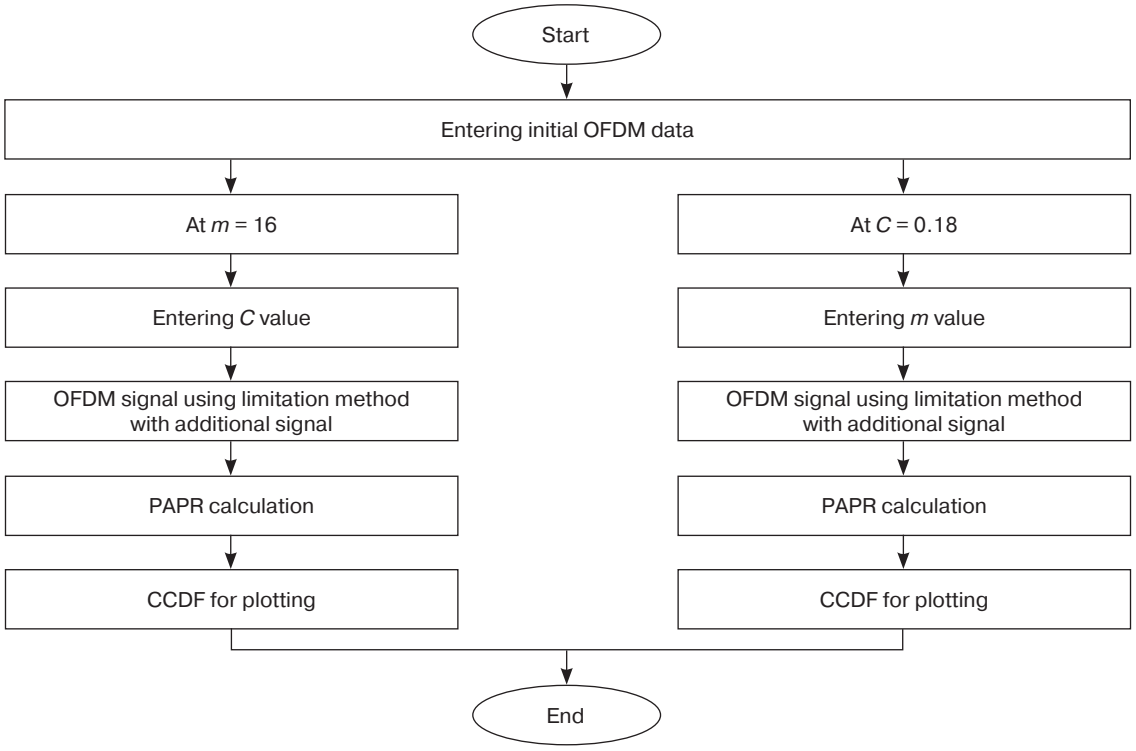


Fig. 2. Block diagram of the algorithm for calculating PAPR

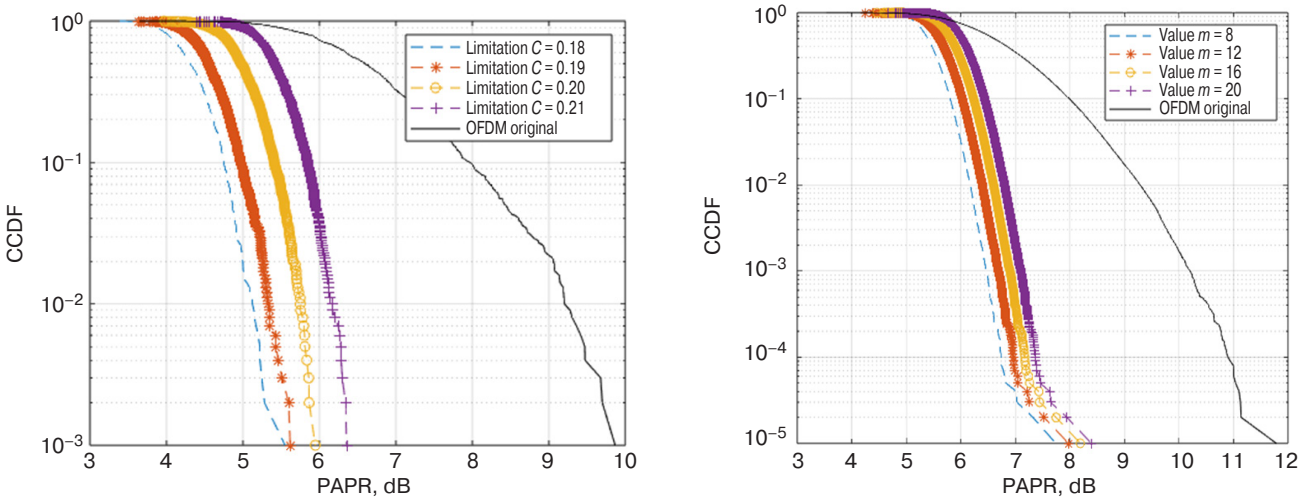


Fig. 3. Results of the OFDM signal PAPR reduction using the limitation method with an additional signal: (a) at $m = 16$, (b) at $C = 0.18$. CCDF is the complementary cumulative distribution function

As a result of modeling, the change in the BER value when changing the limitation level C and fixed m is estimated. BER dependencies are similarly obtained when changing the value of m at a fixed level of C. Since the maximum amplitude of OFDM signal amounts to 0.4229 (dimensionless value normalized to 1 V) in the model with $N = 64$ subcarriers, values of C are chosen less than this value.

The reduction of the OFDM signal peak power when limited with an additional signal at different values of parameters C and m is estimated using the PAPR parameter from relation (2). In order to estimate the PAPR reduction efficiency, the complementary cumulative distribution function (CCDF) is used to indicate the probability of the OFDM signal PAPR exceeding the specified threshold level PAPR0.

The results of reducing PAPR with the number of additional signals $m = 16$ at different values of C are shown in Fig. 3a; the results of PAPR reduction at threshold value $C = 0.18$ when varying m are shown in Fig. 3b.

It is evident from Fig. 3 that the stronger the limitation (low value of C), the smaller the PAPR value when preserving the number of additional signals m. When maintaining the threshold value and changing the number of additional signals, the PAPR value changes insignificantly.

The main parameter determining the quality of a digital communication system is BER during signal recovery. For evaluating the OFDM system performance, the BER value should be estimated depending on three main parameters: the limitation value C, the number of additional signals m, and the signal-to-noise ratio (SNR) representing the ratio of signal power to noise power in the bandwidth in the presence of white Gaussian noise in the channel [9–11].

The BER calculation results determined when modulating 16-QAM OFDM signal in a Gaussian channel when fixing threshold value $C = 0.18$ with different values of m depending on the noise level are shown in Fig. 4. For comparison, BER for the standard OFDM limitation method is evaluated [12–18].

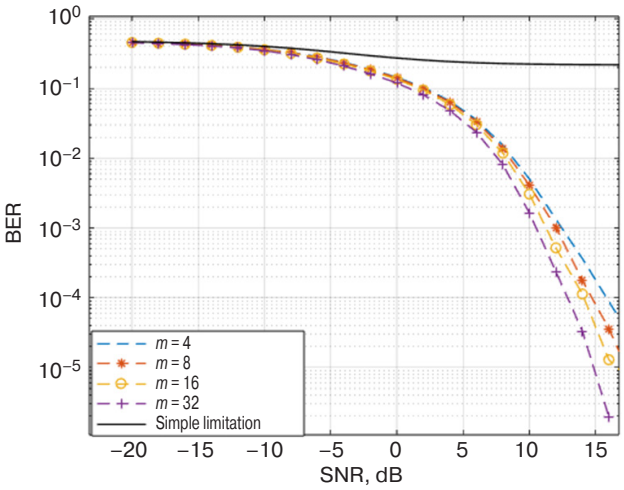


Fig. 4. BER dependence on SNR at fixed value of $C = 0.18$

The advantage of the limitation method with additional signal (at limitation level $C = 0.18$ when the maximum amplitude of the unlimited signal amounts to 0.4229) over the conventional PAPR limitation method is evident from Fig. 4. When the value of m increases, BER decreases.

The BER calculation results at a fixed number of additional signals $m = 16$ with different values of limitation threshold C depending on the noise level is shown in Fig. 5.

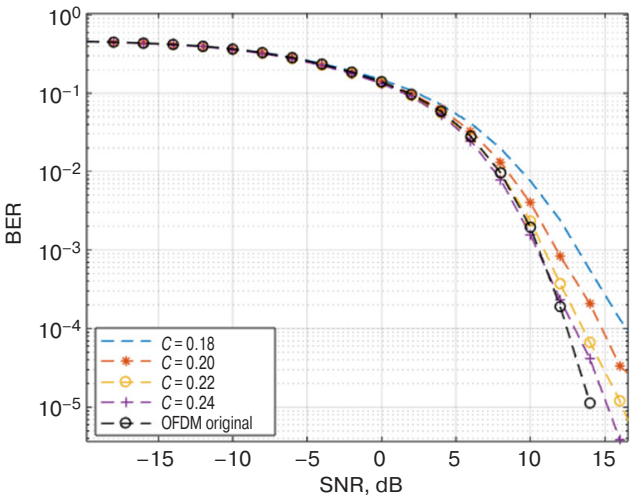


Fig. 5. Dependence of BER on SNR at fixed value of $m = 16$

It can be seen from Fig. 5 that BER decreases with the increased limitation level of C. For comparison, the results for the OFDM signal without limitations (original OFDM) are given. Although these are obviously better, their hardware implementation is physically impossible.

The BER dependencies on the threshold value in the Gaussian channel at SNR = 10 dB and SNR = 15 dB are shown in Fig. 6.

According to the analysis of the results, BER decreases with increased limitation level C and number of additional signals m.

CONCLUSIONS

From the study of the limitation method with additional compensation signal for an OFDM signal in the channel with white Gaussian noise and analysis of its main parameters, the following conclusions can be drawn:

1. At a higher limitation level C, the signal PAPR decreases, whereas it changes insignificantly when the number of additional signals m is changed.
2. In the channel with white Gaussian noise, the effective implementation of the limitation method with an additional compensation signal requires selection of the limitation level C and the number of additional signals m depending on the predicted SNR in the system.

Authors' contributions

Th.T. Pham—making calculations, processing the results.

O.V. Tikhonova—the research idea, consultations on the issues of conducting all stages of the study.

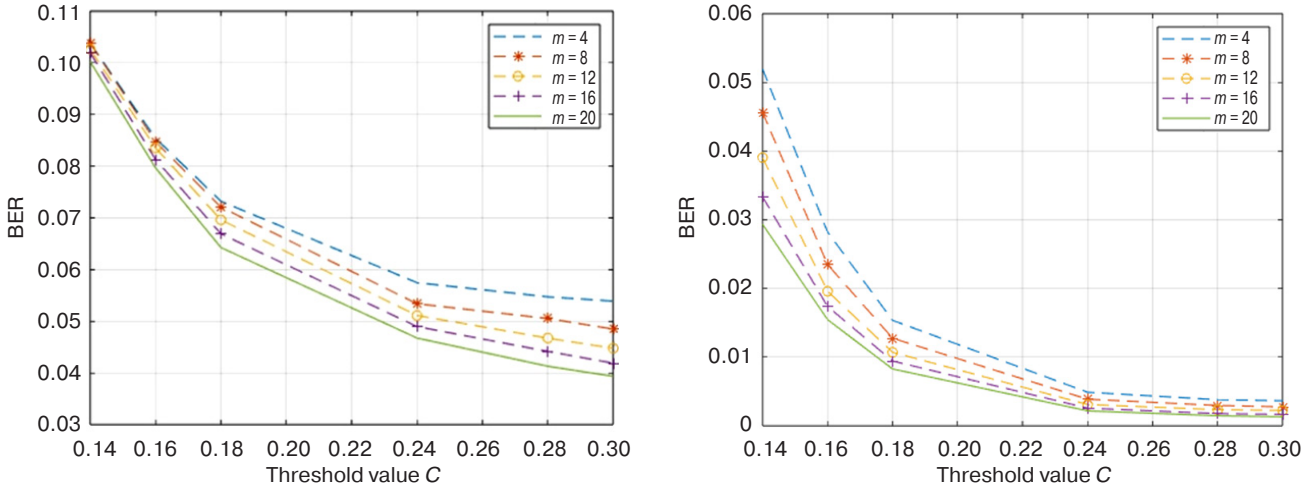


Fig. 6. BER dependence on the limitation level C: (a) SNR = 10 dB, (b) SNR = 15 dB

REFERENCES

1. Kovalev V.V., Seletskaya O.Yu., Pokamestov D.A. Formation and processing of OFDM signals. *Molodoi uchenyi = Young scientist*. 2016;14(118):151–154 (in Russ.). Available from URL: https://moluch.ru/archive/118/32800/
2. Galustov G.G., Meleshkin S.N. *Mul'tipleksirovanie s ortogonal'nym chastotnym razdeleniem signalov* (*Multiplexing with Orthogonal Frequency Division of Signals*): textbook. Taganrog: TTI SFU; 2012. 80 p. (in Russ.).
3. Van Nee R., Prasad R. *OFDM for Wireless Multimedia Communications*. Boston; London: Artech House; 2000. 260 p.
4. Wu Y., Zou W.Y. Orthogonal frequency division multiplexing: A multi-carrier modulation scheme. *IEEE Trans. Consumer Electronics*. 1995;41(3):392–399.
5. Luferchik P.V., Konev A.N., Bogatyrev E.V., Galeev R.G. Methods for improving the energy characteristics of OFDM modems in frequency selective fading communication channels. *Sibirskii aerokosmicheskii zhurnal = Siberian Aerospace Journal*. 2022;23(2):189–196 (in Russ.). https://doi.org/10.31772/2712-8970-2022-23-2-189-196
6. Hassan G.M., Mukred M., Gumaei A.H. Modified Method of PAPR Reduction using Clipping and Filtering for Image Transmission with OFDM. *Al-Mustansiriyah Journal of Science*. 2023;34(4):75–86. https://doi.org/10.23851/mjs.v34i4.1400
7. Pham Th.T., Tikhonova O.V. Limitation method with an additional signal to PAPR reduction in the system with orthogonal frequency division of channels. *Voprosy elektromekhaniki. Trudy VNIIEM = Electromechanical Matters. VNIIEM Studies.* 2023;193(2):34–38 (in Russ.).
8. Pham Th.T., Tikhonova O.V. Method for limiting the peak factor of an OFDM signal with an additional signal. In: *Actual Problems and Prospects for the Development of Radio Engineering and Infocommunication Systems Radioinfocom-2023: Collection of Scientific Articles of the 7th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2023. P. 140–142 (in Russ.).
9. Sklar B. Tsifrovaya svyaz'. *Teoreticheskie osnovy i prakticheskoe primenenie* (*Digital Communication. Theoretical Foundations and Practical Application*): transl. from Engl. Moscow: Vil'yams; 2017. 1100 p. (in Russ.).
[Sklar B. *Digital Communication: Foundamentals and Applications*. Prentice-Hall PTR; 2001. 1079 p.]
10. Proakis J. *Tsifrovaya svyaz'* (*Digital Communication*): transl. from Engl. Moscow: Radio i svyaz'; 2000. 800 p. (in Russ.).
[Proakis J. *Digital Communications.* N.Y.: McGraw-Hill Publ.; 1995. 928 p.]
11. Bairwa M.K., Porwal M.K. BER Performance of OFDM System over AWGN channel with Different Modulation Schemes. *Int. J. Eng. Tech. Res.* (*IJETR*). 2014;2(8):117–119.
12. Shatrughna P.Y., Subhash C.B. PAPR Reduction using Clipping and Filtering Technique for Nonlinear Communication Systems. In: *International Conference on Computing, Communication and Automation*. IEEE; 2015. P. 1220–1225. https://doi.org/10.1109/CCAA.2015.7148590
13. Kannadhasan S., Karthikeyan G., Indumathi G. PAPR Reduction in OFDM Systems using Adaptive Active Constellation Extension Algorithm. *Int. J. Innov. Res. Computer and Commun. Eng*. (*IJIRCCE*). 2013;1(4):950–957.
14. Puksa A.O. Reduction of peak-factor of OFDM signal by methods based on signal limitation. *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal = Int. Res. J.* 2017;12(66):124–127 (in Russ.). https://doi.org/10.23670/IRJ.2017.66.145
15. Wang L., Tellambura C. A simplified clipping and filtering technique for PAPR reduction in OFDM systems. *IEEE Signal Processing Letters*. 2005;12(6):453–456. https://doi.org/10.1109/LSP.2005.847886
16. Tellambura C., Jayalath A.D.S. PAR reduction of an OFDM signal using partial transmit sequences. In: *IEEE 54th Vehicular Technology Conference. NTC Fall 2001. Proceedings*. IEEE; 2001. P. 465–469. https://doi.org/10.1109/VTC.2001.956642
17. Kumar P., Ahuja A.K., Chakka R. BCH/hamming/cyclic coding techniques: comparison of PAPR-reduction performance in OFDM systems. In: *International Conference on Intelligent Computing and Applications*. Springer; 2018. P. 557–566. https://doi.org/10.1007/978-981-10-5520-1_50
18. Han S.H., Lee J.H. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *IEEE Wireless Commun.* 2005;12(2):56–65. https://doi.org/10.1109/MWC.2005.1421929

СПИСОК ЛИТЕРАТУРЫ

1. Ковалев В.В., Селецкая О.Ю., Покаместов Д.А. Формирование и обработка OFDM сигналов. *Молодой ученый*. 2016;14(118):151–154. URL: https://moluch.ru/archive/118/32800/
2. Галустов Г.Г., Мелешкин С.Н. *Мультиплексирование с ортогональным частотным разделением сигналов:* учебное пособие. Таганрог: Изд-во ТТИ ЮФУ; 2012. 80 с.
3. Van Nee R., Prasad R. *OFDM for Wireless Multimedia Communications*. Boston; London: Artech House; 2000. 260 p.
4. Wu Y., Zou W.Y. Orthogonal frequency division multiplexing: A multi-carrier modulation scheme. *IEEE Trans. Consumer Electronics*. 1995;41(3):392–399.
5. Луферчик П.В., Конев А.Н., Богатырев Е.В., Галеев Р.Г. Методы повышения энергетической эффективности OFDM модемов в каналах связи с частотно-селективными замираниями. *Сибирский аэрокосмический журнал*. 2022;23(2):189–196. https://doi.org/10.31772/2712-8970-2022-23-2-189-196
6. Hassan G.M., Mukred M., Gumaei A.H. Modified Method of PAPR Reduction using Clipping and Filtering for Image Transmission with OFDM. *Al-Mustansiriyah Journal of Science*. 2023;34(4):75–86. https://doi.org/10.23851/mjs.v34i4.1400

7. Фам Т.Т., Тихонова О.В. Метод ограничения с дополнительным сигналом для уменьшения значения пик-фактора в системе с ортогональным частотным разделением каналов. *Вопросы электромеханики. Труды ВНИИЭМ*. 2023;193(2):34–38.
8. Фам Т.Т., Тихонова О.В. Метод ограничения пик-фактора сигнала OFDM с дополнительным сигналом. В сб.: Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем «*Радиоинфоком-2023*»: *Сборник научных статей VII Международной научно-практической конференции*. М.: РТУ МИРЭА; 2023. С. 140–142.
9. Скляр Б. *Цифровая связь. Теоретические основы и практическое применение*: пер. с англ. М.: Вильямс; 2017. 1104 с.
10. Прокис Дж. *Цифровая связь*: пер. с англ.; под ред. Д.Д. Кловского. М.: Радио и связь; 2000. 800 с.
11. Bairwa M.K., Porwal M.K. BER Performance of OFDM System over AWGN channel with Different Modulation Schemes. *Int. J. Eng. Tech. Res.* (*IJETR*). 2014;2(8):117–119.
12. Shatrughna P.Y., Subhash C.B. PAPR Reduction using Clipping and Filtering Technique for Nonlinear Communication Systems. In: *International Conference on Computing, Communication and Automation*. IEEE; 2015. P. 1220–1225. https://doi.org/10.1109/CCAA.2015.7148590
13. Kannadhasan S., Karthikeyan G., Indumathi G. PAPR Reduction in OFDM Systems using Adaptive Active Constellation Extension Algorithm. *Int. J. Innov. Res. Computer and Commun. Eng*. (*IJIRCCE*). 2013;1(4):950–957.
14. Пукса А.О. Уменьшение пик-фактора OFDM-сигнала с помощью методов, основанных на ограничении сигналов. *Международный научно-исследовательский журнал*. 2017;12(66):124–127. https://doi.org/10.23670/IRJ.2017.66.145
15. Wang L., Tellambura C. A simplified clipping and filtering technique for PAPR reduction in OFDM systems. *IEEE Signal Processing Letters*. 2005;12(6):453–456. https://doi.org/10.1109/LSP.2005.847886
16. Tellambura C., Jayalath A.D.S. PAR reduction of an OFDM signal using partial transmit sequences. In: *IEEE 54th Vehicular Technology Conference. NTC Fall 2001. Proceedings*. IEEE; 2001. P. 465–469. https://doi.org/10.1109/VTC.2001.956642
17. Kumar P., Ahuja A.K., Chakka R. BCH/hamming/cyclic coding techniques: comparison of PAPR-reduction performance in OFDM systems. In: *International Conference on Intelligent Computing and Applications*. Springer; 2018. P. 557–566. https://doi.org/10.1007/978-981-10-5520-1_50
18. Han S.H., Lee J.H. An overview of peak-to-average power ratio reduction techniques for multicarrier transmission. *IEEE Wireless Commun.* 2005;12(2):56–65. https://doi.org/10.1109/MWC.2005.1421929

About the authors

Pham Thanh Tuan, Postgraduate Student, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: anhsequayve.ru@gmail.com. https://orcid.org/0009-0000-7430-1779

Olga V. Tikhonova, Dr. Sci. (Eng.), Senior Researcher, Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: o_tikhonova@inbox.ru. Scopus Author ID 57208923772, RSCI SPIN-code 3362-9924, https://orcid.org/0009-0009-4013-9182

Об авторах

Фам Тхань Туан, аспирант, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: anhsequayve.ru@gmail.com. https://orcid.org/0009-0000-7430-1779

Тихонова Ольга Вадимовна, д.т.н., старший научный сотрудник, профессор, кафедра радиоэлектронных систем и комплексов, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: o_tikhonova@inbox.ru. Scopus Author ID 57208923772, SPIN-код РИНЦ 3362-9924, https://orcid.org/0009-0009-4013-9182

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Micro- and nanoelectronics. Condensed matter physics

Микро- и наноэлектроника. Физика конденсированного состояния

UDC 532.6, 53.06, 535.016
https://doi.org/10.32362/2500-316X-2024-12-5-50-62
EDN LYYDJX

RESEARCH ARTICLE

Effect of surface electromagnetic wave treatment on the refractive properties of thin films based on indium tin oxides with laser-deposited single-walled carbon nanotubes

Andrei S. Toikka [1, 2, @], Natalia V. Kamanina [1, 2, 3, 4]

[1] *St. Petersburg Electrotechnical University, St. Petersburg, 197022 Russia*
[2] *Petersburg Nuclear Physics Institute, National Research Center "Kurchatov institute", Gatchina, 188300 Russia*
[3] *Scientific and Production Corporation "S.I. Vavilov State Optical Institute", St. Petersburg, 192171 Russia*
[4] *Vavilov State Optical Institute, St. Petersburg, 199053 Russia*
[@] *Corresponding author, e-mail: astoikka.nano@gmail.com*

Abstract

Objectives. The article investigates the effect of surface electromagnetic wave (SEW) treatment on the refractive properties of thin conducting films based on indium tin oxide (ITO) with laser-deposited single-walled carbon nanotubes (CNTs). The effective thickness of the layer of laser-deposited CNTs before and after SEW treatment is evaluated.

Methods. A laser-oriented deposition method employing a CO_2 laser (λ = 10.6 μm) was used to form the structures. Diagnostics of modifications of ITO thin films were carried out using an ellipsometer operating in the spectral range of 300–1000 nm. The Cauchy model was used to describe the optical properties of K8 crown substrates and ITO thin films. To interpret the ellipsometry results of ITO modifications with CNTs, an effective-thickness virtual layer model was introduced. During post-processing of the surface, a CO_2 marker (λ = 10.6 μm) was used to generate SEW. The influence of SEW treatment on the thickness of the virtual layer was assessed using ellipsometry and atomic force microscopy in contact mode.

Results. Based on the ellipsometry data, the effective thickness of the CNT layer was in the range of 24–26 nm. Following SEW treatment, the thickness of the effective CNT layer decreased to 4–8 nm, indicating the possibility of precision processing of the ITO surface with CNTs using SEW. When CNTs are deposited on an ITO surface with subsequent SEW treatment of the surface, reflection losses for p-polarized radiation are reduced. In a spectral range of 400–750 nm at an angle of incidence relative to the normal to the plane of structures α = 65°, a decrease in reflection is observed from 18.5% to 13.5% relative to ITO without CNTs and SEV treatment; at α = 71°, a decrease from 6.4% to 4.7% is observed; at α = 77°, a decrease from 1.8% to 1.2%.

Conclusions. For ITO-based thin films with laser-deposited CNTs, the described SEW treatment method provides a precise reduction in the thickness of the composite structure while preserving the antireflective properties of the CNTs. These capabilities make it possible to use the studied ITO modifications in solving problems in optical electronics, microfluidics, and biomedicine.

Keywords: indium tin oxides, carbon nanotubes, laser exposure, surface electromagnetic wave treatment

• **Submitted:** 05.02.2024 • **Revised:** 14.03.2024 • **Accepted:** 12.07.2024

For citation: Toikka A.S., Kamanina N.V. Effect of surface electromagnetic wave treatment on the refractive properties of thin films based on indium tin oxides with laser-deposited single-walled carbon nanotubes. *Russ. Technol. J.* 2024;12(5):50–62. https://doi.org/10.32362/2500-316X-2024-12-5-50-62

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Влияние обработки поверхностными электромагнитными волнами на рефрактивные свойства тонких пленок на основе оксидов индия и олова с лазерно-осажденными одностенными углеродными нанотрубками

А.С. Тойкка [1, 2, @], **Н.В. Каманина** [1, 2, 3, 4]

[1] *Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», Санкт-Петербург, 197022 Россия*
[2] *НИЦ Курчатовский институт – Петербургский институт ядерной физики «ПИЯФ», Гатчина, 188300 Россия*
[3] *НПО ГОИ им. С.И. Вавилова, Санкт-Петербург, 192171 Россия*
[4] *ГОИ им. С.И. Вавилова, Санкт-Петербург, 199053 Россия*
[@] *Автор для переписки, e-mail: astoikka.nano@gmail.com*

Резюме

Цели. Цель работы – исследование влияния обработки поверхностными электромагнитными волнами (ПЭВ) тонких проводящих пленок на основе оксидов индия и олова (indium tin oxides, ITO) с лазерно-осажденными одностенными углеродными нанотрубками (УНТ) на рефрактивные свойства, оценка эффективной толщины слоя лазерно-осажденных УНТ до и после ПЭВ-обработки.

Методы. Для формирования структур использовался метод лазерно-ориентированного осаждения с применением CO_2-лазера (λ = 10.6 мкм). Диагностика модификаций тонких пленок ITO осуществлялась при помощи эллипсометра в спектральном диапазоне 300–1000 нм. Для описания оптических свойств подложек крон К8 и тонких пленок ITO использовалась модель Коши. Для интерпретации результатов эллипсометрии модификаций ITO с УНТ была введена модель виртуального слоя с эффективной толщиной. При постобработке поверхности использовался CO_2-маркер (λ = 10.6 мкм) для генерации ПЭВ. Оценка влияния ПЭВ-обработки на толщину виртуального слоя проводилась при помощи эллипсометрии и атомно-силовой микроскопии в контактном режиме.

Результаты. На основе данных эллипсометрии установлено, что эффективная толщина слоя УНТ находилась в диапазоне 24–26 нм. После ПЭВ-обработки толщина эффективного слоя УНТ снизилась до 4–8 нм. При осаждении УНТ на поверхность ITO и последующей ПЭВ-обработке поверхности снижаются потери на отражение для p-поляризованного излучения. В спектральном диапазоне 400–750 нм при угле падения относительно нормали к плоскости структур 65° наблюдается снижение отражения с 18.5% до 13.5% относительно ITO без УНТ и ПЭВ-обработки, при 71° – снижение с 6.4% до 4.7%, при 77° – снижение с 1.8% до 1.2%.

Выводы. Для тонких пленок на основе ITO с лазерно-осажденными УНТ доступен метод ПЭВ-обработки, которая позволяет сохранить просветляющие свойства УНТ и обеспечивает прецизионное снижение толщины композитной структуры. Указанные возможности позволяют использовать исследуемые модификации ITO в задачах оптической электроники, микрофлюидики и биомедицины.

Ключевые слова: оксиды индия и олова, углеродные нанотрубки, лазерное воздействие, обработка поверхностными электромагнитными волнами

• **Поступила:** 05.02.2024 • **Доработана:** 14.03.2024 • **Принята к опубликованию:** 12.07.2024

Для цитирования: Тойкка А.С., Каманина Н.В. Влияние обработки поверхностными электромагнитными волнами на рефрактивные свойства тонких пленок на основе оксидов индия и олова с лазерно-осажденными одностенными углеродными нанотрубками. *Russ. Technol. J.* 2024;12(5):50–62. https://doi.org/10.32362/2500-316X-2024-12-5-50-62

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Indium tin oxides (ITO) are degenerate semiconductors with *n*-type conductivity [1–3]. Their properties depend largely on the stoichiometric composition $(In_2O_3)_x$–$(SnO_2)_{1-x}$. Here, SnO_2 is used to increase electron concentration (N). It has been shown [1] that, in a pure In_2O_3 matrix ($x = 0$), electron concentration $n_e = 1 \cdot 10^{20}$ cm^{-3}. In the range $0.02 < x < 0.15$, the N value increases up to $9 \cdot 10^{20}$ cm^{-3}. Further increase of SnO_2 content is not reasonable due to the decreased mobility of charge carriers μ [1].

The criterion for comparing the parameters of ITO films depends on the application. When considering ITO as electrical contacts, it is necessary to minimize the resistivity ρ. In the case of ITO, the value of this parameter varies in the range of 10^{-4}–10^{-3} Ohm·cm [1–3]. In addition to varying the stoichiometric composition of ITO, the electrical properties optimized by altering the heating rate of substrates during deposition [2], selecting the types of working gases and their partial pressures [3–5], and using different targets [6].

In some cases, for example, in photovoltaics [7] and liquid crystal optics [8], high optical transmittance is required in addition to low electrical resistivity. In the visible and near-infrared (near-IR) spectral regions, this requirement is fulfilled due to the relatively low extinction coefficient k of ITO in this range; thus, optical losses are mainly due to reflection [9].

Indium and tin oxides also represent sought-after structures in IR optics due to the presence of plasmonic electron gas resonance in this range. Here, the plasma frequency ω_{pl} and damping frequency γ depend on the concentration and mobility of charge carriers, respectively [10]. By varying the above parameters, it is possible to obtain spectral regions with a negative value of dielectric permittivity, which finds a number of applications, e.g., in modulators [11], solar cells [12], and sensors [13].

Thus, ITO applications are largely determined by the concentration and mobility of electrons. One of the strategies to improve the ITO performance is the use of composite materials with nanostructures, by which means it is possible to reduce the electrical resistance and tune the restricted bandgap width [14–16].

Significant results were achieved by laser-oriented deposition of carbon nanotubes on the ITO surface [17–19]. Carbon nanotubes (CNTs) have been shown to contribute to an increase in optical transmittance, improved mechanical and laser strength, and reduced electrical resistance of ITO thin films. Based on atomic force microscopy data [19], CNTs are understood to be deposited in the form of clusters.

In order to model the developed electro-optical devices, in which one of the key elements is a structured ITO layer, it is necessary to take into account the basic properties of the CNT layer, including its thickness. Since this material is not deposited as a continuous layer, there is a need for a separate study devoted to the determination of the CNT layer thickness. In the present work, approach nondestructive ellipsometry was chosen as the diagnostic method due to the possibility of examining relatively large apertures. The current study also investigated the precision variation of CNT layer thickness. Surface electromagnetic wave (SEW) treatment was selected as the mechanism for providing thickness modification.

1. MATERIALS AND METHODS

The formation of ITO thin films on K8 crown glass substrates (Scientific and Production Corporation S.I. Vavilov State Optical Institute, Russia) was carried out by laser-oriented deposition [17] using a CO_2 laser (Laser Center, Russia; wavelength $\lambda = 10.6$ μm;

power $P \approx 30$ W; beam diameter $d = 5$ mm). Cerac, Inc. (USA) provided ITO pellets with stoichiometric composition $(In_2O_3)_{0.9}$–$(SnO_2)_{0.1}$. The average thickness of the considered ITO films was 80–100 nm with mean-square roughness *Sq* less than 2 nm. Next, Aldrich (USA) brand single-walled carbon nanotubes (CNTs, No. 704121, chirality <7.6>) were deposited on a part of the samples with ITO films by laser-oriented deposition.) A control field with an average strength of 100 V/cm was used in the CNT deposition process. This is necessary to increase the kinetic energy of CNTs to provide their subsequent implantation on the material surface [17–19].

The thickness and roughness of the films were controlled using an atomic force microscope NT-MDT Solver Next (NT-MDT, Russia) in contact mode with a scanning frequency of 1 Hz. Refractive properties were investigated on a J.A. Woolam M-2000RCE ellipsometer (J.A. Woolam, USA) with built-in *CompleteEASE* software (version 4.91)[1]. For this purpose, the complex reflectance indices for s- and p-polarized radiation were r_p and r_s, respectively. Then, according to the ellipsometry equation (1), the ellipsometric parameters ψ and Δ were determined on the basis of the measured complex reflection indices $\dot{r}_p$ and $\dot{r}_s$ and phase difference in reflection from the interface δ_p and δ_s [20][2]:

$$\frac{\dot{r}_p}{\dot{r}_s} = \frac{|\dot{r}_p|}{|\dot{r}_s|} e^{j(\delta_p - \delta_s)} = \operatorname{tg}\psi e^{j\Delta}. \quad (1)$$

This expression is used in the Jones formalism; the parameters r_p and r_s correspond to the diagonal elements of the Jones matrix [20], which are valid for fully polarized radiation. In order to avoid unnecessary artifacts in the interpretation of ellipsometry results, the Stokes–Mueller formalism was used to take into account the degree *P* in the interaction of radiation with the studied samples [20]. In this approach, the Stokes matrices of the output and input radiation are related using the Mueller matrix $\mathbf{M}$[2]:

$$\mathbf{M} = R\begin{bmatrix} 1 & -PN & 0 & 0 \\ -PN & 1 & 0 & 0 \\ 0 & 0 & PC & PS \\ 0 & 0 & -PS & PC \end{bmatrix}. \quad (2)$$

Parameters *N*, *C*, *S* are related to the ellipsometric angle ψ and Δ[2]:

$$\begin{cases} N = \cos 2\psi, \\ C = \sin 2\psi \cos\Delta, \\ S = \sin 2\psi \sin\Delta. \end{cases} \quad (3)$$

The abovementioned parameters are normalized by reflection coefficient by intensity:

$$R = \frac{1}{2}(\dot{r}_s \dot{r}_s^* + \dot{r}_p \dot{r}_p^*) = \frac{R_s + R_p}{2}. \quad (4)$$

In expression (3), R_s and R_p correspond to the intensity reflection coefficients for s- and p-type polarization, respectively. When analyzing the spectral dependencies $\dot{r}_p$ and $\dot{r}_s$ along with their complex conjugate values $\dot{r}_p^*$ and $\dot{r}_s^*$, it is necessary to take into account the complex refractive indices $(\dot{n}_i)$ of the corresponding media [20]:

$$\begin{cases} \dot{n}_i = n_i + jk_i, \\ \dot{r}_p = \dfrac{\dot{n}_1 \cos\varphi_0 - \dot{n}_0 \cos\varphi_1}{\dot{n}_1 \cos\varphi_0 + \dot{n}_0 \cos\varphi_1} = |\dot{r}_p| e^{j\delta_p}, \\ \dot{r}_s = \dfrac{\dot{n}_1 \cos\varphi_0 - \dot{n}_1 \cos\varphi_1}{\dot{n}_1 \cos\varphi_0 + \dot{n}_1 \cos\varphi_1} = |\dot{r}_s| e^{j\delta_s}. \end{cases} \quad (5)$$

Parameter *n* is the refractive index of the medium, while *k* is the extinction coefficient. The radiation propagates in the medium with $i = 0$ and is reflected from the interface with $i = 1$, where *i* is the order number of the medium. Thus, the complex index of the studied structures is related to the amplitudes and phases of the complex reflection coefficients (4), which are measured using an ellipsometer. Using expression (1), the ellipsometric parameters ψ and Δ are determined to obtain the components of the Mueller matrix (2)–(3). In order to determine the refractive indices of K8 crown substrates, ITO thin films and their modifications, it is necessary to solve the inverse ellipsometry problem. This problem is reduced to the selection of those values of the parameters of refractive indices, extinction coefficients, and material thicknesses at which the values of *N, C, S* obtained with the help of the selected model will have the least divergence from the experimental data. For this purpose, we used the minimization of mean square error (*MSE*)[2]:

$$MSE = \sqrt{\frac{1}{3n-m}} \times \sqrt{\sum_{l=1}^{n}\left[(N_l^{\text{mod}} - N_l^{\text{exp}})^2 + (C_l^{\text{mod}} - C_l^{\text{exp}})^2 + (S_l^{\text{mod}} - S_l^{\text{exp}})^2\right]}. \quad (6)$$

[1] https://www.jawoollam.com/ellipsometry-software/completeease. Accessed April 21, 2024.

[2] *Complete EASE: Data Analysis Manual* (*version 4.63*). J.A. Wollam Co., Inc. 2011. 410 p.

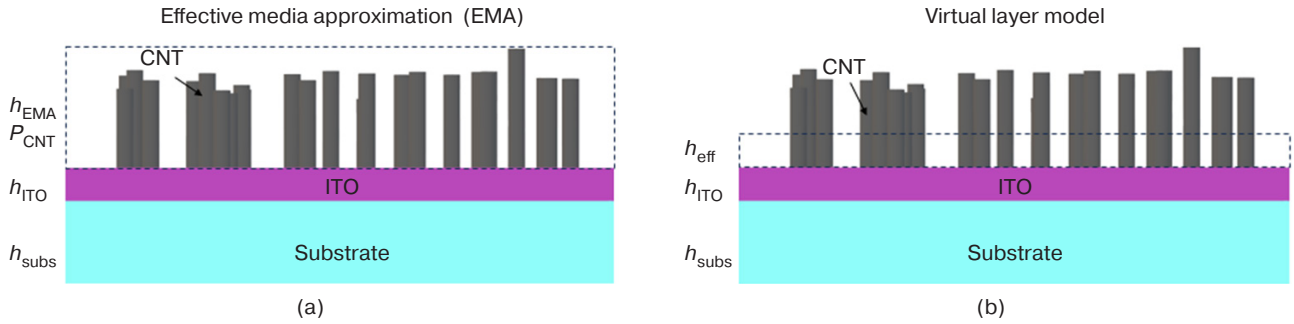


Fig. 1. Interpretation of layers based on CNT:
(a) solid medium model; (b) virtual layer model.
h_{ITO} is the thickness of the ITO, h_{subs} is the thickness of the substrate

Parameter n corresponds to the number of iterations (in the spectral range 300–1000 nm the sources on the ellipsometer have 710 wavelength iterations), m is the number of fitting parameters. For the characterization of K8 crown substrates it is necessary to know the refractive index and extinction dependencies, as well as the roughness ($m = 3$). For ITO modifications, in addition to the above parameters, the thickness is also unknown ($m = 4$). Index l is the iteration number, *mod* and *exp* denote data based on approximation (model) and experiment, respectively.

The Cauchy model [21] was chosen to approximate the refractive index of K8 and ITO crown substrates in the spectral range of 300–1000 nm:

$$n(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4}. \quad (7)$$

This model describes the refractive index of ITO in the selected range with a high degree of confidence [22].

When CNT are deposited on the ITO surface, it is necessary to take into account the fact that these nanoparticles are deposited in clusters rather than as continuous layers [19]. However, to interpret the results, a value of the CNT layer thickness is necessary. For this purpose, we introduce the value of the effective thickness h_{eff} of the CNT layer according to the following assumptions:

1. Effective thickness parameter h_{eff} increases with increasing deposition density and with decreasing distance between deposited CNT clusters;
2. The extinction $k = f(\lambda)$ and refractive index $n = f(\lambda)$ of this layer are fixed and do not depend on the treatment;
3. Optical properties at the 'virtual layer–ITO' interface change discontinuously.

The abovementioned approach can be compared with Bruggemann's effective medium approximation (EMA), in which the layer thickness h_{EMA} is fixed and the variable parameter is the percentage content of components; in particular, component 1 is air, while component 2 is the material [20]. In the effective thickness approximation, the component content is taken as 100% ($P_{CNT} = 1$), while the varying parameter is the effective thickness h_{eff}. A visual comparison of the approaches is presented in Fig. 1.

Results of [23] were used as reference data for the description of n and k CNT. Interpretation of the results of ellipsometry of CNT/ITO/K8 crown structures was performed after determining the fitting parameters of K8 crown and ITO/K8 crown structures. When working with the CNT layer, the virtual layer model was used. The refractive index and extinction coefficients in the range of 300–1000 nm were imported from the data [23], while the fitting parameter was the thickness of the effective layer.

After diagnosing the structures based on ITO with CNTs, a part of the samples was subjected to laser treatment at CO_2 marker (λ = 10.6 μm, modulation frequency 1 kHz, beam diameter 150 μm, processing frequency 50 mm/s, power 21 W).

2. RESULTS AND DISCUSSION

According to the previously described procedures, the spectral dependencies of the refractive index of K8 crown substrates and ITO thin films deposited on K8 crown substrates were sequentially determined (Fig. 2). In the spectral range of 300–1000 nm, the Cauchy fitting parameters for the K8 crown substrates were as follows: $A = 1.550$, $B = 0.00541$ and $C = -9.153 \cdot 10^{-7}$ (Fig. 2a). The parameters for the ITO thin films were as follows: $A = 1.506$, $B = 0.085$ and $C = -3.72 \cdot 10^{-8}$ (Fig. 2b). From Fig. 2a, it can be concluded that the refractive index dispersion is negative; in the visible range, the refractive index varies between 1.51 and 1.54, which is typical for crown glasses[3]. Here, the optical parameters of ITO depend on the stoichiometric composition, deposition methods and modes, as well as the annealing temperature and post-processing procedures. In spite of this, the data

[3] GLASS—glasses, Refractive index database. https://refractiveindex.info/. Accessed April 21, 2024.

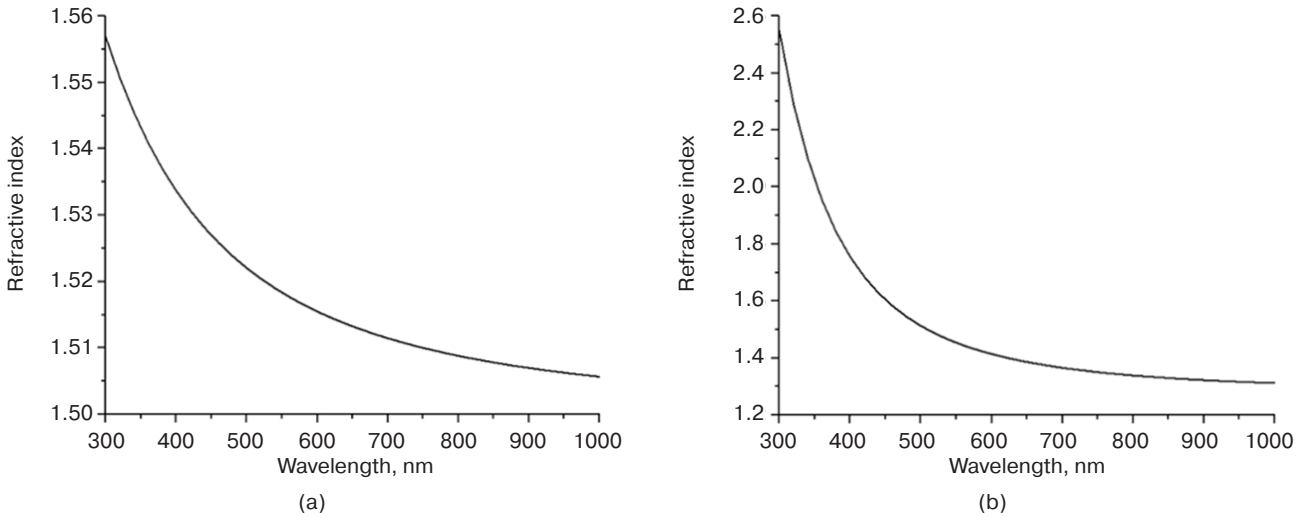


Fig. 2. Spectral dependencies of the refractive index: (a) K8 crown substrates, (b) ITO thin films

in Fig. 2b can be qualitatively compared with the results obtained in [22] to obtain a close correspondence.

The studied ITO structures are thin-film structures as manifested in the form of interference extrema (Fig. 3). The spectral dependencies were measured at three angles of incidence value (α = 65°, 71°, and 77°), which exceed the Brewster angle θ_{Br}. In the case of ITO, the value of θ_{Br} in the visible spectral range varies between 60.6° and 64.1°. Such a choice of incidence angles allows us to work with a steep section of the $R_p = f(\alpha)$ dependence, to reduce the noise/reflected signal ratio, and thus to interpret the ellipsometry results more reliably.

The data in Fig. 3 illustrate only the amplitude values of reflection coefficients, which are used to compare structures in the context of their application. To determine the thickness of ITO films by the optical method, we use the ellipsometric parameters ψ and Δ (Fig. 4), then follow the procedures described in step 1.

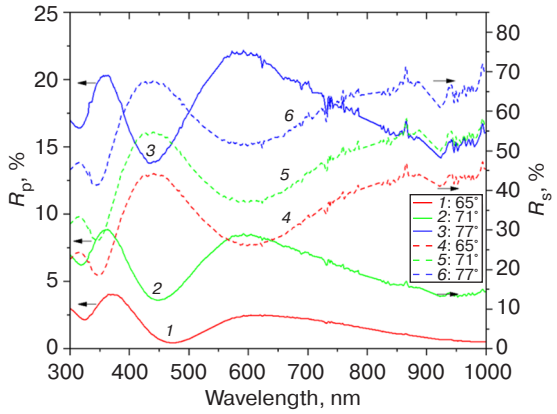


Fig. 3. Spectral dependencies of reflection coefficients of ITO thin films on glass in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

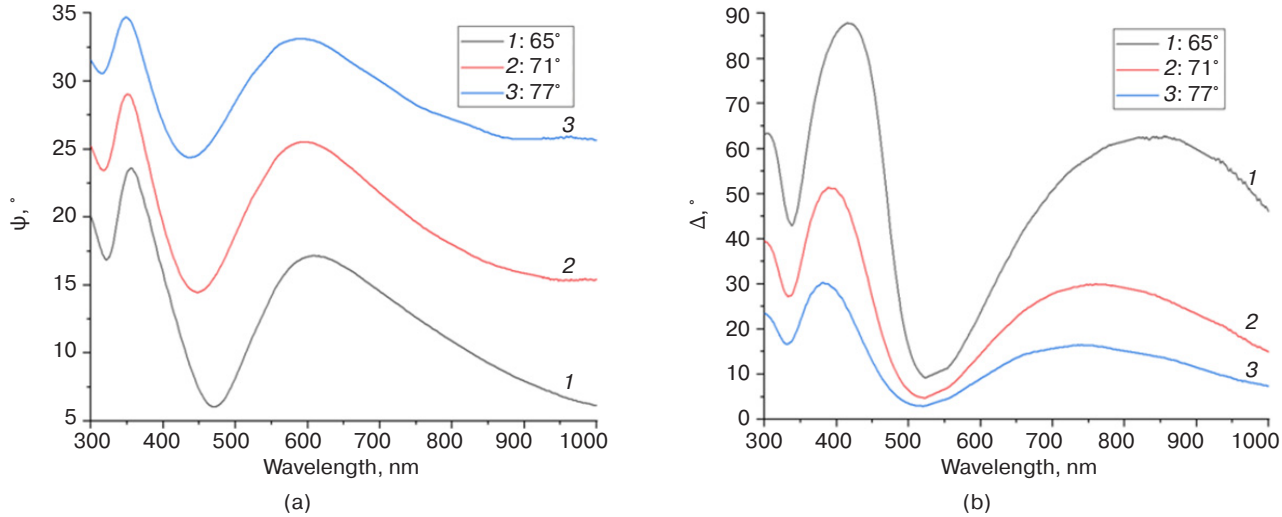


Fig. 4. Spectral dependencies of ellipsometric parameters ψ (a) and Δ (b) of ITO films on K8 crown substrate in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

Based on the measured dependencies of reflection coefficients (Fig. 3), calculated ellipsometric parameters (Fig. 4), and known refractive properties (Fig. 2), it was found that the minimum *MSE* corresponds to an ITO thickness equal to 87 nm.

During CNT deposition, the number of extrema in the spectral dependencies of reflectance coefficients increases (Fig. 5), which is associated with an increase in optical length.

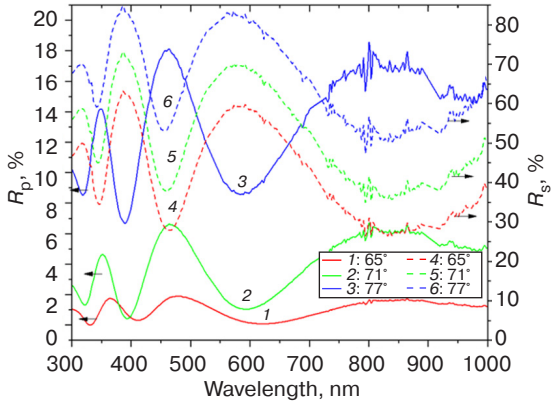


Fig. 5. Spectral dependencies of reflection coefficients of ITO thin films with laser-deposited CNTs on the K8 crown substrate in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

It is noteworthy that the value of the reflection coefficient decreases during CNT deposition. For example, the values of local maxima of R_p at α = 65° in the case of ITO were 20.3% (λ = 358.3 nm) and 22.0% (λ = 591.4 nm), and 14.2% (λ = 348.5 nm) and 18.1% (λ = 463.3 nm) in the case of ITO with CNT. A more detailed comparison is given in the table.

In order to define the effective thickness of the CNT layer, it is necessary to take the ellipsometric parameters into account (Fig. 6). Then, also considering the thickness of the ITO layer (87 nm) calculated at the previous step, the procedure is performed as described in step 1. In this case, the minimum *MSE* corresponds to the effective thickness of the CNT layer in a range of 22–24 nm.

The effective thickness of the CNT layer depends on the density of cluster deposition and their sizes, which are determined by the mode of their deposition on the ITO surface. The required effect of the CNT layer on the ITO surface can involve the rearrangement of spectral properties, the formation of an orienting effect or a change in the free surface energy [17, 24]. As a rule, several parameters are changed at once, for example, surface energy and transmittance (including due to interference and formation of charge transfer complexes). Sometimes the reverse procedure is required, i.e., reducing the thickness of thin films, for example, based on the requirements for optical coordination of functional layers of optoelectronic devices.

In some cases, it becomes possible to treat the ITO surface via SEW [25]. In this treatment method, the SEW is distributed along the interface between the ITO surface and air. If we denote the dielectric permittivity of the ITO modification as ε_2 and of the boundary layer (air) as ε_1, the depth of the SEW attenuation l_2 in ITO can be estimated as follows [25]:

$$l_2 = \frac{c}{\omega}\sqrt{-\frac{(\varepsilon_1+\varepsilon_2)}{\varepsilon_2^2}}. \quad (8)$$

Here c and ω are the velocity of propagation in vacuum and frequency of radiation of electromagnetic waves, respectively. This approach imposes requirements on the dielectric permittivity of the treated material: $|\varepsilon_2| > \varepsilon_1$ and $\varepsilon_2 < 0$. The possibility of satisfying this

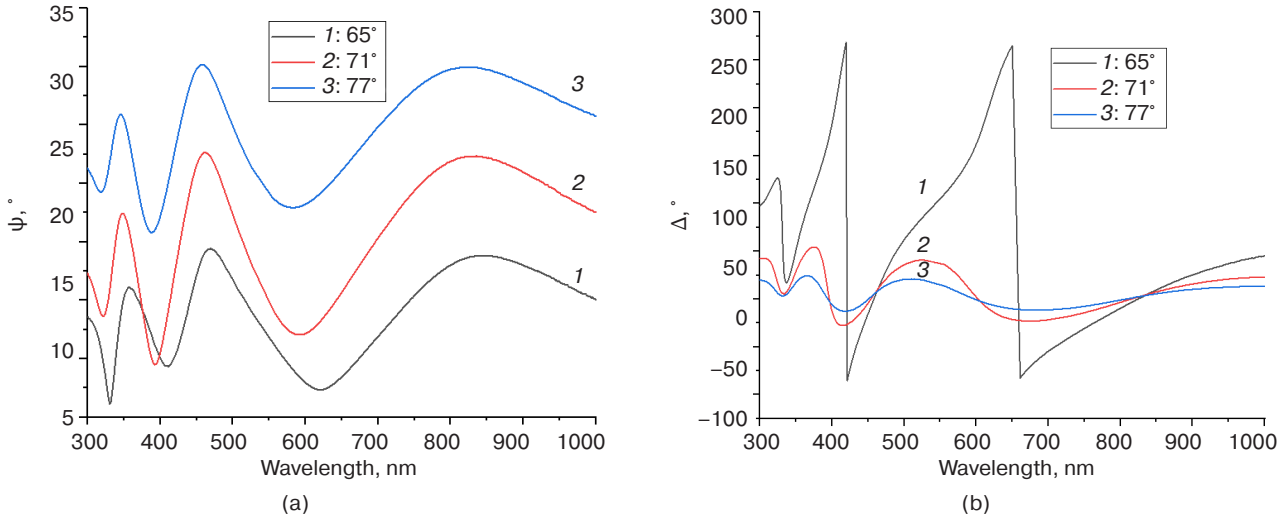


Fig. 6. Spectral dependencies of ellipsometric parameters ψ (a) and Δ (b) of ITO films with CNTs on K8 crown substrate in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

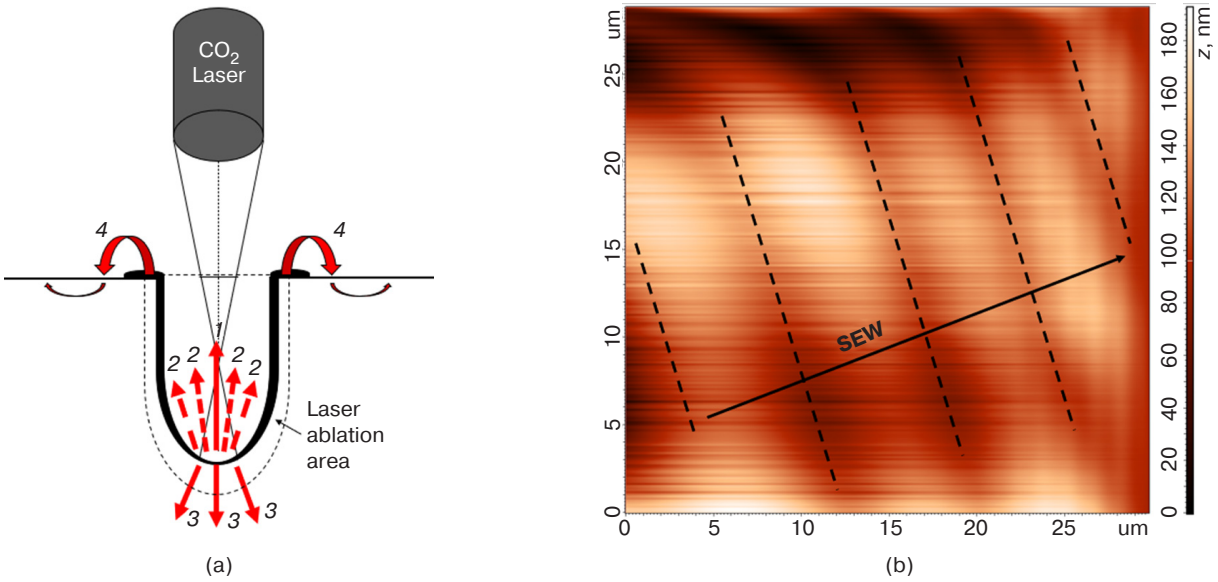


Fig. 7. SEW treatment:
(a) treatment scheme: *1* mirror-reflected flux, *2* diffuse-reflected flux, *3* refracted radiation flux, *4* SEW;
(b) atomic force microscopy profile of ITO surface in the area of SEW treatment

condition near the resonant frequencies ω_{pl} was already discussed in the introduction section.

SEW treatment should be distinguished from laser ablation by direct hit of radiation. The input radiation from the CO_2 laser was used to generate SEW, a significant part of which was absorbed by the ITO-based thin film with the formation of caverns, while a small part was diffusely reflected. Consequently, the SEW treatment exposed regions **around** the direct laser ablation zones, with energy densities for SEW exposure substantially lower than those required for laser ablation activation. The schematic differences are shown in Fig. 7 along with the surface profile of ITO modifications following SEW treatment.

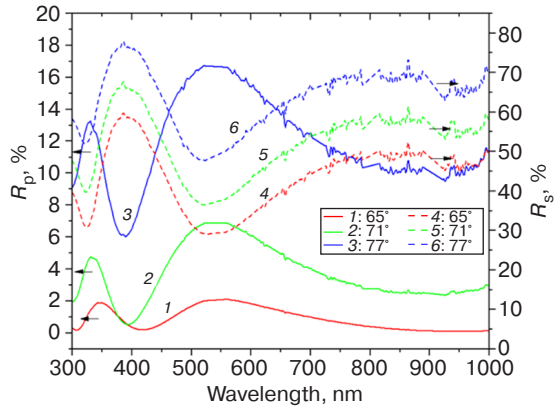


Fig. 8. Spectral dependencies of reflection coefficients of ITO thin films with laser-deposited CNT on the K8 crown substrate after SEW treatment in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

The spectral dependencies of the surface reflectance coefficients of ITO thin films with CNT in the SEW treatment regions are shown in Fig. 8.

According to the position of the interference peaks, the optical length it can be seen to have decreased relative to the ITO films with CNT. At the same time, the numerical values of reflection coefficients are lower than in the case of the original ITO. The comparative data are presented in the table.

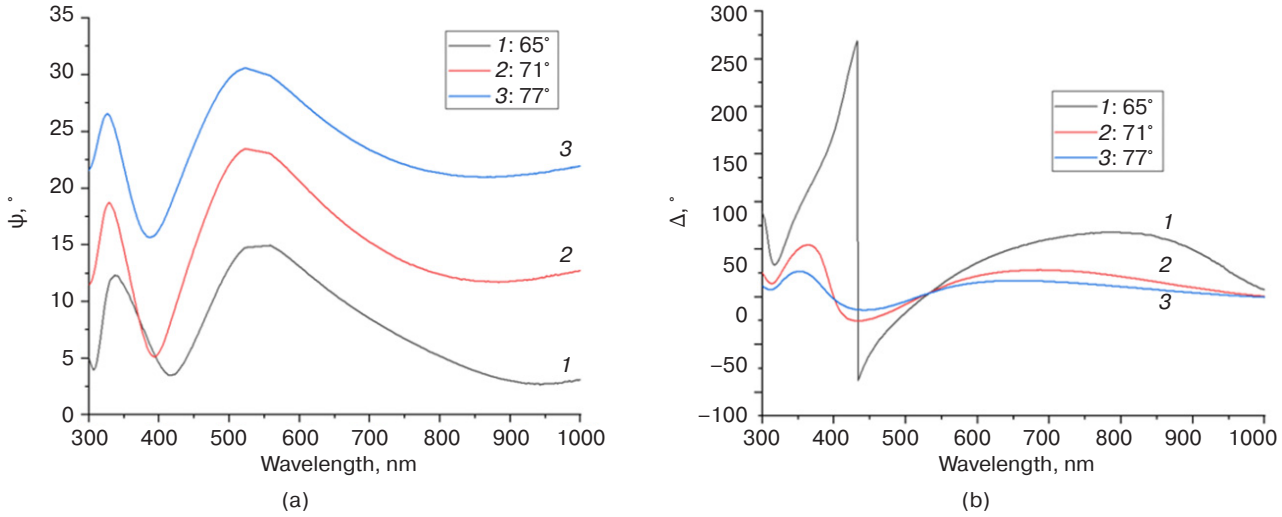
Spectral dependencies of ellipsometric parameters (Fig. 9) were also used to estimate the effective CNT layer thickness after SEW treatment, followed by the statistical processing described in step 1.

Based on the data shown in Figs. 8 and 9, it can be stated that the minimum *MSE* corresponds to $h_{eff} \approx 6$ nm. Similar procedures were carried out for 10 regions of ITO/CNT and ITO/CNT/SEW structures. It was found that h_{eff} decreases from 22–24 nm to 4–8 nm during the SEW treatment (Fig. 10).

At first glance (Fig. 10), there is some doubt about the presence of CNT after SEW treatment. On a macro-scale, the markers of CNT represent the reduced reflection loss (related to the luminescent effect of CNT with respect to ITO) and the increased contact wetting angle (related to the formation of Wenzel and Cassie–Baxter states on the CNT framework) [19, 24]. The comparison of reflection losses given earlier can be visualized in the table. A detailed analysis of the wetting mechanisms of the original structures was considered separately in [19]. ITO surfaces with CNT and SEW treatment preferentially possess contact angles θ in the range of 100° to 110°, while this range is 80°–90° for ITO surfaces with SEW, and 115°–125° for ITO with CNT.

Table. Comparison of refractive properties of ITO modifications for p-polarization

Parameter	ITO modification		
	ITO	ITO with CNT	ITO with CNT and SEW
Parameters of extrema at α = 77°: λ (nm) / R_p (%)			
Maximum 1	358.3 nm / 20.3%	348.5 nm / 14.2%	330.6 nm / 13.2%
Minimum 1	437.3 nm / 13.9%	390.9 nm / 6.8%	387.2 nm / 6.1%
Maximum 2	593.3 nm / 22.2%	463.3 nm / 18.1%	535.7 nm / 16.7%
Minimum 2	The data are not applicable for the specified conditions	590.6 nm / 8.6%	The data are not applicable for the specified conditions
Maximum 3	The data are not applicable for the specified conditions	832.9 nm / 16.9%	The data are not applicable for the specified conditions
Average R_p value (%) in the visible range (400–750 nm) at different α			
α = 65°	18.5%	12.8%	13.5%
α = 71°	6.4%	3.4%	4.7%
α = 77°	1.8%	0.8%	1.2%



(a) (b)

Fig. 9. Spectral dependencies of ellipsometric parameters ψ (a) and Δ (b) of ITO films with CNT on K8 crown substrate after SEW-treatment in the range of 300–1000 nm at different angles of incidence: 65°, 71°, and 77°

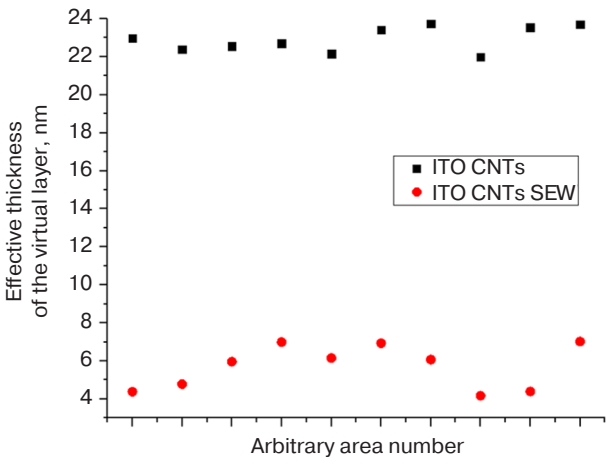


Fig. 10. Effective CNT layer thickness before and after SEW treatment determined on the basis of the ellipsometry method

Thus, the data on the change of refractive properties presented in the current study, as well as in a separate study devoted to wetting mechanisms during laser ablation and SEW treatment [26], indicate that a part of CNT clusters remains on the ITO surface at the selected mode of SEW treatment, resulting in thickness variation. Additional information on the effect of SEW treatment can be obtained from atomic force microscopy.

Figure 11a shows the surface profile of CNT on the ITO film prior to SEW treatment. The mean-square roughness *Sq* is 41.2 nm, while the maximum height h_{max} and h_{min} depth with respect to the zero-level line are 573.9 and −65.9 nm, respectively. It should be understood that CNTs are deposited in clusters of different height and density; thus, the parameters *Sq*, h_{max}, h_{min} are different in different surface areas. However, a tendency of decreasing roughness can be distinguished in SEW

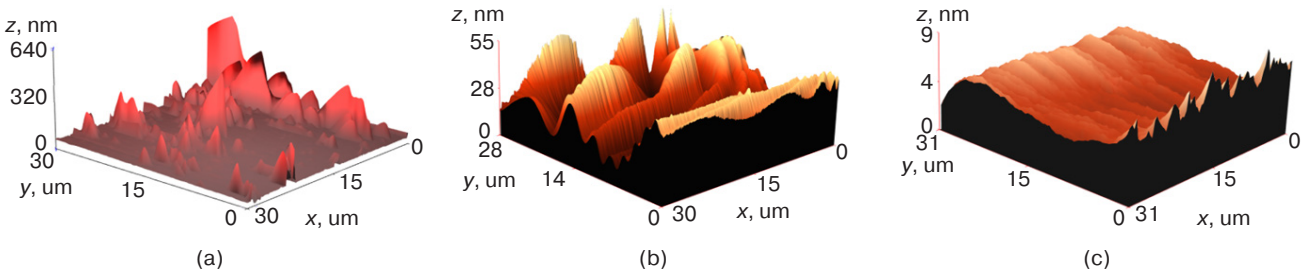


Fig. 11. Effect of SEW treatment on the surface of CNT deposited on ITO films: (a) before treatment; (b) SEW treatment of the region with high density of CNT deposition; (c) SEW treatment of the region with low density of CNT deposition

treatment. The case shown in Fig. 11b corresponds to SEW treatment of CNT with high deposition density: here, depressions are observed, which can be related both to a thinning of CNT clusters and a partial removal of ITO material. A decrease in the *Sq* value to 11.0 nm at h_{max} = 32.2 nm and h_{min} = −22.9 nm is observed. In cases where the density of deposited CNT is lower, such as when single tubes of small size are deposited on the ITO surface, the SEW treats the ITO surface to a greater extent than CNT (Fig. 11c). These regions are even smoother: *Sq* = 0.9 nm, h_{max} = 32.2 nm, h_{min} = −22.9 nm.

When evaluating the effect of SEW on the CNT layer thickness, the atomic force microscopy method has a limitation related to the ratio between the sizes of probes (up to 3–5 μm) and scanned objects (diameter of single CNT 1–2 nm). It is also necessary to take into account the adhesion of probes to CNT clusters, which causes additional measurement errors. Therefore, the noncontact ellipsometry method is preferable for solving this task.

CONCLUSIONS

Two important scientific and methodological results were obtained in the course of the performed research:

1. The effective thickness h_{eff} of the layer of laser-deposited single-walled carbon nanotubes on the ITO surface was determined. At the value of the control field strength of 100 V/cm in the process of CNT deposition on the ITO surface by the optical nondestructive method, the values of h_{eff} were determined to be in the range of 22–24 nm. Knowing the numerical value of this parameter allows us to perform calculations more efficiently in the development of electro-optical devices.
2. It is shown that the effective thickness of the CNT layer on the ITO surface can be precisely reduced by SEW treatment. For example, in the current study, it was possible to reduce h_{eff} from 22–24 nm to 4–8 nm. This is an important application of nanoplasmonics in the context of ITO surface treatment, since the accuracy of direct laser ablation directly depends on the **technical parameters** of the optical system (beam divergence, focal length, diameter, scanner pitch, etc.), while the accuracy of SEW treatment is determined by the **material properties** (n, μ, ω_{pl}).

The performed research and analysis contribute to the expansion of the material science database of photonics systems and functional materials of optoelectronics.

Thus, the effective thickness of the CNT layer on the ITO surface can be both increased (by means of additional oriented deposition) and precisely reduced by means of SEW treatment. This makes it possible to rearrange the interference peaks of thin films based on ITO with CNT as a means of optimizing the parameters of electro-optical devices with respect to the required spectral range. At the same time, the useful properties of CNT are preserved, i.e., an increase in coating strength, rearrangement of wetting mechanisms and optical luminescence, as well as an increase in electrical conductivity.

Aknowledgments

This research was partially supported by the Russian Science Foundation, grant No. 24-23-00021.

Authors' contributions

A.S. Toikka—setting up and conducting experiments, interpretation of the results, conceptualizing, drafting and editing the article.

N.V. Kamanina—interpretation of the results, conceptualization, editing the article.

REFERENCES

1. Kim H., Gilmore C.M., Piquie A., Horwitz J.S., Mattoussi H., Murata H., Kafafi Z.H., Chrisey D.B. Electrical, optical, and structural properties of indium–tin–oxide thin films for organic light-emitting devices. *J. Appl. Phys*. 1999;86(11): 6451–6461. https://doi.org/10.1063/1.371708
2. Adurodija F.O., Izumi H., Ishihara T., Yoshioka H., Motoyama M., Murai K. Influence of substrate temperature on the properties of indium oxide thin films. *J. Vac. Sci. Technol. A*. 2000;18:814–818. https://doi.org/10.1116/1.582260
3. Zhang K., Zhu F., Huan C.H.A., Wee T.S. Effect of hydrogen partial pressure on optoelectronic properties of indium tin oxide thin films deposited by radio frequency magnetron sputtering method. *J. Appl. Phys*. 1999:86(2):974–980. https://doi.org/10.1063/1.370834
4. Kerkache L., Layadi A., Mosser A. Effect of oxygen partial pressure on the structural and optical properties of dc sputtered ITO thin films. *J. Alloys Compd*. 2019;18(1):46–50. https://doi.org/10.1016/j.jallcom.2009.06.103
5. Kim J.-H., Lee J.-H., Heo Y.-W., Kim J.-J., Park J.-O. Effects of oxygen partial pressure on the preferential orientation and surface morphology of ITO films grown by RF magnetron sputtering. *J. Electroceram*. 2009;23:169–174. https://doi.org/10.1007/s10832-007-9351-8
6. Yang S., Sun B., Liu Y., Zhu J., Song J., Hao Z., Zeng X., Zhao X., Shu Y., Chen J., Yi J., He J. Effect of ITO target crystallinity on the properties of sputtering deposited ITO films. *Ceram. Int*. 2020;46(5):6342–6350. https://doi.org/10.1016/j.ceramint.2019.11.110
7. Chen Y., Du C., Sun L., Fu T., Zhang R., Rong W., Cao S., Li X., Shen H., Shi D. Improved optical properties of perovskite solar cells by introducing Ag nanoparticles and ITO AR layers. *Sci. Rep*. 2011;11:14550 https://doi.org/10.1038/s41598-021-93914-1
8. Chu F., Wang D., Liu C., Li L., Wang W.H. Multi-View 2D/3D Switchable Display with Cylindrical Liquid Crystal Lens Array. *Crystals*. 2021;11(6):715. https://doi.org/10.3390/cryst11060715
9. Rasheed M., Barille R. Optical constants of DC sputtering derived ITO, TiO_2 and TiO_2:Nb thin films characterized by spectrophotometry and spectroscopic ellipsometry for optoelectronic devices. *J. Non. Cryst. Solids*. 2017;476:1–14. https://doi.org/10.1016/j.jnoncrysol.2017.04.027
10. Losego M.D., Efremenko A.Y., Rhodes C.L., Cerruti M.G., Franzen S., Maria J.P. Conductive oxide thin films: Model systems for understanding and controlling surface plasmon resonance. *J. Appl. Phys*. 2009;106(2):024903. https://doi.org/10.1063/1.3174440
11. Amin R., Maiti R., Gui Y., Suer C., Miscuglio M., Heidari E., Khurgin J.B., Chen R.T., Dalir H., Sorger V.J. Heterogeneously integrated ITO plasmonic Mach–Zehnder interferometric modulator on SOI. *Sci. Rep.* 2021;11:1287. https://doi.org/10.1038/s41598-020-80381-3
12. Dong W.J., Yu H.K., Lee J.L. Abnormal dewetting of Ag layer on three-dimensional ITO branches to form spatial plasmonic nanoparticles for organic solar cells. *Sci. Rep.* 2020;10:12819. https://doi.org/10.1038/s41598-020-69320-4
13. Liu C., Wang J., Wang F., Su W., Yang L., Lv J., Fu G., Li X., Liu Q., Sun T., Chu P.K. Surface plasmon resonance (SPR) infrared sensor based on D-shape photonic crystal fibers with ITO coatings. *Opt. Commun.* 2020;464:125496. https://doi.org/10.1016/j.optcom.2020.125496
14. El Nahrawy A.M., Abou Hammad A.B., Youssef A.M., Mansour A.M., Othman A.M. Thermal, dielectric and antimicrobial properties of polystyrene-assisted/ITO:Cu nanocomposites. *Appl. Phys. A.* 2019;125:46. https://doi.org/10.1007/s00339-018-2351-5
15. Mei F., Huang J., Yuan T., Li R. Effect of cerium doping on the microstructure and photoelectric properties of Ce-doped ITO films. *Appl. Surf. Sci.* 2020;509:144810. https://doi.org/10.1016/j.apsusc.2019.144810
16. Taha H., Jiang Z.T., Yin C.Y., Henry D.J., Zhao X., Trotter G., Amri A. A Novel Approach for Fabricating Transparent and Conducting SWCNTs/ITO Thin Films for Optoelectronic Applications. *J. Phys. Chem. C.* 2018;122(5):3014–3027. https://doi.org/10.1021/acs.jpcc.7b10977
17. Kamanina N.V., Vasil'ev P.Ya., Studeonov V.I., Usanov Yu.E. Strengthening transparent conductive coatings and "soft" materials of the IR range when nanotechnologies are used. *J. Opt. Technol.* 2008;75(1):67–68. https://doi.org/10.1364/JOT.75.000067
[Original Russian Text: Kamanina N.V., Vasil'ev P.Ya., Studeonov V.I., Usanov Yu.E. Strengthening transparent conductive coatings and "soft" materials of the IR range when nanotechnologies are used. *Opticheskii zhurnal*. 2008;75(1):83–84 (in Russ.).]
18. Kamanina N., Toikka A., Gladysheva I. ITO conducting coatings properties improvement via nanotechnology approach. *Nano Express*. 2021;2(1):010006. https://doi.org/10.1088/2632-959X/abd90c
19. Kamanina N., Toikka A., Valeev B., Kvashnin D. Carbon Nanotubes Use for the Semiconductors ZnSe and ZnS Material Surface Modification via the Laser-Oriented Deposition Technique. *C – Journal of Carbon Research*. 2021;7(4):84. https://doi.org/10.3390/c7040084
20. Garcia-Caurel E., De Martino A., Gaston J.P., Yan L. Application of Spectroscopic Ellipsometry and Mueller Ellipsometry to Optical Characterization. *Appl. Spectroscopy*. 2013;67(1):1–21. https://doi.org/10.1366/12-06883
21. Fujiwara H. *Spectroscopic Ellipsometry.* The Atrium, Chichester, West Sussex, England: John Wiley & Songs; 2007. 369 p.
22. König T.A.F., Ledin P.A., Kerszulis J., Mahmoud M.A., El-Sayed M.A., Reynolds J.R., Tsukruk V.V. Electrically Tunable Plasmonic Behavior of Nanocube–Polymer Nanomaterials Induced by a Redox-Active Electrochromic Polymer. *ACS Nano*. 2014;8(6):6182–6192. https://doi.org/10.1021/nn501601e

23. Ermolaev G.A., Tsapenko A.P., Volkov V.S., Anisimov A.S., Gladush Y.G., Nasibulin A.G. Express determination of thickness and dielectric function of single-walled carbon nanotube films. *Appl. Phys. Lett*. 2020;116:231103. https://doi.org/10.1063/5.0012933
24. Toikka A.S., Fedorova L.O., Kamanina N.V. Influence of laser-deposited carbon-containing nanoparticles on the orienting properties of indium-tin-oxide-based conducting layers for liquid crystal devices. *J. Opt. Technol.* 2024;91(1):55–60. https://doi.org/10.1364/JOT.91.000055
[Original Russian Text: Toikka A.S., Fedorova L.O., Kamanina N.V. Influence of laser-deposited carbon-containing nanoparticles on the orienting properties of indium-tin-oxide-based conducting layers for liquid crystal devices. *Opticheskii zhurnal*. 2024;91(1):91–100 (in Russ.).]
25. Bonch-Bruevich A.M., Libenson M.N., Makin V.S., Trubaev V.V. Surface electromagnetic waves in optics. *Opt. Eng*. 1992;31(4):718–730. https://www.doi.org/10.1117/12.56133
26. Toikka A.S., Kamanina N.V. Formation of the anisotropic ITO-based orienting layers for the liquid crystal devices. *St. Petersburg State Polytechnical University Journal. Physics and Mathematics*. 2023;16(3.2):244–248. https://www.doi.org/10.18721/JPM.163.242

About the authors

Andrei S. Toikka, Postgraduate Student, Photonics Department, St. Petersburg Electrotechnical University "LETI" (5, ul. Professora Popova, St. Petersburg, 197022 Russia); Junior Researcher, Advanced Development Division, B.P. Konstantinov Petersburg Nuclear Physics Institute, National Research Center "Kurchatov Institute" (1, Orlova Roshcha, Gatchina, Leningradskaya oblast, 188300 Russia). E-mail: astoikka.nano@gmail.com. Scopus Author ID 57216272706, RSCI SPIN-code 1261-2571, https://orcid.org/0000-0002-8694-8497

Natalia V. Kamanina, Dr. Sci. (Phys.-Math.), Head of the Laboratory of Photophysics of Nanostructured Materials and Devices, Research and Production Corporation "S.I. Vavilov State Optical Institute" (36, Babushkina ul., St. Petersburg, 192171 Russia); Head of the Laboratory of Photophysics of Media with Nanoobjects, Vavilov State Optical Institute (5, Kadetskaya Liniya V.O., St. Petersburg, 199053 Russia), Professor, Photonics Department, St. Petersburg Electrotechnical University "LETI" (5, ul. Professora Popova, St. Petersburg, 197022 Russia); Lead Researcher of Advanced Development Division, B.P. Konstantinov Petersburg Nuclear Physics Institute, National Research Center "Kurchatov Institute" (1, Orlova Roshcha, Gatchina, Leningradskaya oblast, 188300 Russia). E-mail: nvkamanina@mail.ru. Scopus Author ID 55980751700, RSCI SPIN-code 1231-5045, https://orcid.org/0000-0002-2903-2685

Об авторах

Тойкка Андрей Сергеевич, аспирант, кафедра фотоники, ФГАОУ ВО «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина) (СПбГЭТУ «ЛЭТИ») (197022, Россия, Санкт-Петербург, ул. Профессора Попова, д. 5); младший научный сотрудник, отдел перспективных разработок, ФГБУ «Петербургский институт ядерной физики им. Б.П. Константинова «ПИЯФ» Национального исследовательского центра «Курчатовский институт» (188300, Россия, Ленинградская обл., Гатчина, Орлова роща, д. 1). E-mail: astoikka.nano@gmail.com. Scopus Author ID 57216272706, SPIN-код РИНЦ 1261-2571, https://orcid.org/0000-0002-8694-8497

Каманина Наталия Владимировна, д.ф.-м.н., заведующая отделом фотофизики наноструктурированных материалов и устройств, АО «Научно-производственное объединение «Государственный оптический институт им. С.И. Вавилова» (192171, Россия, Санкт-Петербург, ул. Бабушкина, д. 36); заведующая лаборатории фотофизики сред с нанообъектами, АО «Государственный оптический институт им. С.И. Вавилова» (199053, Россия, Санкт-Петербург, Кадетская линия В.О., д. 5); профессор, кафедра фотоники, ФГАОУ ВО «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина) (СПбГЭТУ «ЛЭТИ») (197022, Россия, Санкт-Петербург, ул. Профессора Попова, д. 5); ведущий научный сотрудник, отдел перспективных разработок, ФГБУ «Петербургский институт ядерной физики им. Б.П. Константинова «ПИЯФ» Национального исследовательского центра «Курчатовский институт» (188300, Россия, Ленинградская обл., Гатчина, Орлова роща, д. 1). E-mail: nvkamanina@mail.ru. Scopus Author ID 55980751700, SPIN-код РИНЦ 1231-5045, https://orcid.org/0000-0002-2903-2685

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 681.5.015
https://doi.org/10.32362/2500-316X-2024-12-5-63-76
EDN OIBKMA

RESEARCH ARTICLE

On identification of interconnected systems

Nikolay N. Karabutov @

MIREA – Russian Technological University, Moscow, 119454 Russia
@ *Corresponding author, e-mail: karabutov@mirea.ru*

Abstract

Objectives. Interconnected control systems are widely used in various technical contexts, generally involving multichannel systems. However, due to the complexity of their description, the problem of identifying interconnected systems has received insufficient attention. As a result, simplified models are commonly used, which do not always reflect the specifics of the object. Thus, the synthesis of mathematical models for the description of interconnected control systems becomes a relevant endeavor. The paper sets out to develop an approach to obtaining models under conditions of incomplete a priori information. A mathematical model is developed on the example of two-channel systems (TCSs) having cross-connections and identical channels. The case of asymmetric cross-connections is considered, along with estimates of their influence on the quality of the adaptive identification system. The problem of estimating the identifiability of the parameters of a TCS is formulated on the basis of available experimental information and subsequent synthesis of the adaptive system. The proposed approach is then generalized to the case of an interconnected system.

Methods. The adaptive system identification and Lyapunov vector function methods are used along with implicit identification representation for the model.

Results. The influence of excitation constancy on estimates of the TCS parameters is demonstrated on the basis of the proposed approach for estimating the identifiability of TCS with cross-connections. The synthesis of adaptive algorithms of parameter estimation for TCSs with cross-connections based on input-output data is generalized to the case of interconnected systems. The results are applied to building models of tracking system and two-channel corrector for automatic control systems.

Conclusions. The features of adaptive identification of TCSs with identical channels, cross-connections and feedbacks are considered. The conditions for the TCS identifiability are obtained. Adaptive algorithms for estimating TCS parameters are synthesized. The proposed approach is generalized to the case of nonidentical channels and multi-connected systems. The exponential dissipativity of the adaptive identification system is verified. The proposed methods can be used in the development of systems for identification and control of complex dynamic systems.

Keywords: adaptive identification, identifiability, stability, two-channel system, Lyapunov vector function, multiconnected system, excitation constant

• **Submitted:** 25.01.2024 • **Revised:** 29.03.2024 • **Accepted:** 15.07.2024

For citation: Karabutov N.N. On identification of interconnected systems. *Russ. Technol. J*. 2024;12(5):63–76. https://doi.org/10.32362/2500-316X-2024-12-5-63-76

Financial disclosure: The author has no financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Об идентификации взаимосвязанных систем

Н.Н. Карабутов [@]

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] *Автор для переписки, e-mail: karabutov@mirea.ru*

Резюме

Цели. Проблеме идентификации взаимосвязанных систем до настоящего времени уделялось недостаточно внимания. Взаимосвязанные системы управления широко применяются в различных технических системах. Как правило, применяются многоканальные системы. Из-за сложности их описания применяют упрощенные модели, которые не всегда отражают специфику объекта. Поэтому задача синтеза математических моделей является актуальной. Целью настоящей работы является разработка подхода к получению моделей в условиях неполной априорной информации. Для решения задачи применяется адаптивный подход. На примере двухканальных систем (ДС) с перекрестными связями и идентичными каналами разрабатывается метод получения математической модели. Рассматривается случай асимметричных перекрестных связей, и получены оценки их влияния на качество работы адаптивной системы идентификации. В рамках предлагаемой постановки ставится задача оценки идентифицируемости параметров двухканальной системы на основе имеющейся экспериментальной информации и последующем синтезе адаптивной системы. Дается обобщение предлагаемого подхода на случай многосвязной системы.

Методы. Применяются метод адаптивной идентификации системы, неявное идентификационное представление для модели, метод векторных функций Ляпунова.

Результаты. Предложен подход к оценке идентифицируемости двухканальных систем с перекрестными связями. Показано влияние постоянства возбуждения на оценки параметров двухканальной системы. Предложен метод синтеза адаптивных алгоритмов оценки параметров для двухканальных систем с перекрестными связями по данным «вход-выход». Дано обобщение подхода на случай взаимосвязанных систем. Результаты применены для построения моделей системы слежения и двухканального корректора для систем автоматического регулирования.

Выводы. Рассмотрены особенности адаптивной идентификации двухканальных систем с идентичными каналами, перекрестными и обратными связями. Получены условия идентифицируемости ДС. Синтезированы адаптивные алгоритмы оценивания параметров ДС. Дано обобщение предлагаемого подхода на случай неидентичных каналов и многосвязных систем. Доказана экспоненциальная диссипативность адаптивной системы идентификации. Предлагаемые методы могут использоваться при разработке систем идентификации и управления сложными динамическими системами.

Ключевые слова: адаптивная идентификация, идентифицируемость, устойчивость, двухканальная система, векторная функция Ляпунова, многосвязная система, постоянство возбуждения

• **Поступила:** 25.01.2024 • **Доработана:** 29.03.2024 • **Принята к опубликованию:** 15.07.2024

Для цитирования: Карабутов Н.Н. Об идентификации взаимосвязанных систем. *Russ. Technol. J.* 2024;12(5):63–76. https://doi.org/10.32362/2500-316X-2024-12-5-63-76

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

Interconnected systems (ICS) are widely used in control systems [1–5]. They are most commonly used to control robot and manipulator drives [6, 7] as well as forming the basis for various technical systems [8, 9].

In [10], the possibility of applying adaptive identification to ICS is considered. Here, the application of iterative-probabilistic method is proposed. In [11], a black-box model identification method for interconnected nonstationary dynamic objects with uncertainty is proposed. The development of adaptive algorithms for decentralized robust control with a reference model for ICS with time delay is discussed in [12]. Here, the asymptotic stability of the system is justified. The identifiability of a closed-loop interconnected stochastic system is considered in [13] along with the proposed decomposition of the system into subsystems. The identifiability both of separate elements of the system and separate closed loops without simultaneous identification of other elements and of loops of the system are considered. Sufficient conditions for almost certain convergence estimates of likelihood parameters are determined. The high-modular normalized adaptive lattice algorithm for multichannel filtering proposed in [14] is based on the least squares method.

In [15], an algorithm is presented for model identification using a neural network in the form of transfer functions for two-dimensional spatial ICS, which are causal for both open and closed loop. In [16], decentralized robust adaptive stabilization with output feedback is considered. The synthesis of control laws is based on adaptive nonlinear damping as well as the application of robust adaptive state observer and Lyapunov functions (LF). Similar results are obtained in [17, 18]. In [19], a method for identifying the ICS in rational form using input-output data is presented along with a Rasser-shaped model.

The adaptive control over a class of large-scale systems consisting of an arbitrary number of interacting subsystems with unknown parameters, nonlinearities, and bounded disturbances is considered in [20] using the reference model method. In [21], topological structural identification of large-scale subsystems having sparse flows of interconnected dynamical systems due to a small amount of data is considered. In [22], an approach to blind identification of a two-channel system (TCS) with finite impulse response from a limited number of output measurements in the presence of additive white noise is described. The proposed approach, which is based on data in the frequency domain, allows frequency-based estimation. In [23], fundamental problems of blind multichannel identification are considered on the basis of an analysis of some modern adaptive algorithms.

Review [24] covers some modern approaches based on decomposition of the problem of identifying systems with multiple inputs/single output. Proposed identification procedures combine low-dimensional solutions with the iterative version of the Wiener filter. The identification of stationary linear systems of this class are considered in [13, 25, 26]. In [25], a method for identifying multidimensional systems in the frequency domain using the correlation approach is proposed. A parametric model in the form of a vector-difference equation was further transformed into a corresponding frequency domain model. The identifiability of multidimensional discrete dynamical systems is considered in [26]. A priori estimates for the identifiability of coefficients are proposed on the basis of numerical characteristics of the asymptotic variance lower bound of coefficient estimates.

Thus, models in the form of transfer functions and in state space are used to study processes in ICS. Adaptive procedures are used to synthesize control algorithms. Adaptive control algorithms have been developed in the presence of unmodulated dynamics and disturbances. This condition can be explained by insufficient information about the state and parameters of the system or object, as well as the difficulty of accounting for interconnections in the system. The identification of multidimensional system parameters is based on applying statistical procedures, frequency methods, and neural network technologies. A number of studies deal with the application of adaptive methods, which are mainly used for adjusting (identifying) control device parameters. The resulting parametric uncertainties are compensated by selecting appropriate control algorithms. However, few publications cover the problem of TCS identification.

The present paper proposes a measurement-based approach to the ICS adaptive identification. The proposed approach to the identifiability of TCS having cross-connections involves various assumptions regarding the parameters of cross-connections. The stability of the adaptive system is demonstrated on the basis of obtained identifiability estimates to inform the considered approach to ICS identification.

1. TCSs

1.1. Problem statement

We consider the TCS with cross-connection. The links in the channels are considered identical. Transfer functions are used to analyze these systems [27]. For identification tasks, describing TCSs in the state space is more convenient. Let the system contain n sequentially interconnected vertical layers:

1) first layer

$$\begin{cases} \dot{\mathbf{X}}_{11} = \mathbf{A}_1\mathbf{X}_{11} + \mathbf{B}_1 v_{11}, \\ y_{11} = \mathbf{C}_1^{\mathrm{T}}\mathbf{X}_{11}, \\ \dot{\mathbf{X}}_{21} = \mathbf{A}_1\mathbf{X}_{21} + \mathbf{B}_1 v_{21}, \\ y_{21} = \mathbf{C}_1^{\mathrm{T}}\mathbf{X}_{21}, \end{cases} \quad (1)$$

2) kth layer ($1 < k \le n$)

$$\begin{cases} \dot{\mathbf{X}}_{1,k} = \mathbf{A}_k\mathbf{X}_{1,k} + \mathbf{B}_k u_{1,k}, \\ y_{1,k} = \mathbf{C}_k^{\mathrm{T}}\mathbf{X}_{1,k}, \\ u_{1,k} = y_{1,k-1} + v_{1,k-1}, \\ v_{1,k-1} = d_{k-1}(y_{2,k-1} + d_{k-2}v_{2,k-1}), \\ \dot{\mathbf{X}}_{2,k} = \mathbf{A}_k\mathbf{X}_{2,k} + \mathbf{B}_k u_{2,k}, \\ y_{2,k} = \mathbf{C}_k^{\mathrm{T}}\mathbf{X}_{2,k}, \\ u_{2,k} = y_{2,k-1} + v_{2,k-1}, \\ v_{2,k-1} = d_{k-1}(y_{1,k-1} + d_{k-2}v_{1,k-1}), \end{cases} \quad (2)$$

where $v_{11} = g_1 - y_{1,n}$, $v_{21} = g_2 - y_{2,n}$, $v_{i,k-1}$ is the output of cross-connection of the ith channel, i = 1, 2; $\mathbf{X}_{i,k} \in \mathbb{R}^q$ is the state vector of the kth layer of the ith channel of the TCS; $\mathbf{A}_k \in \mathbb{R}^{q_k \times q_k}$; $\mathbf{C}_k \in \mathbb{R}^{q_k}$; $\mathbf{B}_k \in \mathbb{R}^{q_k}$; $y_{i,k} \in \mathbb{R}$ is the output of kth layer of ith channel; $v_{i,k-1}$ is the output of cross-connection, and $g_i(t) \in \mathbb{R}$ is the system input (master control). Matrix $\mathbf{A}_k$ is a Hurwitz matrix.

Parameter d_{k-1} represents an operator. Depending on the problems solved by the TCS, d_{k-1} can be a constant, nonlinear function, or differential operator.

The information structure for system (1), (2) has the following form:

$$\mathbb{I}_{\mathrm{o}} = \left\{ g_i(t), y_{i,k}(t) \forall \left(i = 1,2\right) \& \left(k = \overline{1,n}\right), t \in J = \left[t_0, t_{\mathrm{e}}\right] \right\}, \quad (3)$$

where J is the data capture interval; t is time; and t_0 and t_{e} are the beginning and the end of time interval.

The system is identified using a model with a structure similar to (1) and (2) and with outputs $\hat{y}_{i,k} \in \mathbb{R}$, where $i = 1,2$; $k = \overline{1,n}$.

The problem is reduced to selecting algorithms for adjusting model parameters in such a way that

$$\lim_{t\to\infty} | \hat{y}_{i,k} - y_{i,k} | \le \delta_{i,k}, \quad (4)$$

where $\delta_{i,k} \ge 0$ is the specified value.

1.2. On structural aspects of the system

The identification of systems (1) and (2) depends on the possibility of estimating its parameters significantly.

We introduce a model for TCS (1):

$$\begin{cases} \dot{\hat{\mathbf{X}}}_{1,1} = \mathbf{K}_1(\hat{\mathbf{X}}_{1,1} - \mathbf{X}_{1,1}) + \hat{\mathbf{A}}_1\mathbf{X}_{1,1} + \hat{\mathbf{B}}_1 v_{1,1}, \\ \hat{y}_{1,1} = \mathbf{C}_1^{\mathrm{T}}\hat{\mathbf{X}}_{1,1}, \\ \dot{\hat{\mathbf{X}}}_{2,1} = \mathbf{K}_1(\hat{\mathbf{X}}_{2,1} - \mathbf{X}_{2,1}) + \hat{\mathbf{A}}_1\mathbf{X}_{2,1} + \hat{\mathbf{B}}_1 v_{2,1}, \\ \hat{y}_{2,1} = \mathbf{C}_1^{\mathrm{T}}\mathbf{X}_{2,1}, \end{cases} \quad (5)$$

where $\mathbf{K}_1 \in \mathbb{R}^{q_1 \times q_1}$ is the known stable matrix (reference model); $\hat{\mathbf{A}}_1 \in \mathbb{R}^{q_1 \times q_1}$, $\hat{\mathbf{B}}_1 \in \mathbb{R}^{q_1}$ is the matrices of model (5), and $\hat{\mathbf{X}}_{i,1}$ is the state vector.

We denote $\mathbf{E}_{1,1} \triangleq \hat{\mathbf{X}}_{1,1} - \mathbf{X}_{1,1}$, $\mathbf{E}_{2,1} \triangleq \hat{\mathbf{X}}_{2,1} - \mathbf{X}_{2,1}$. Then for the first layer

$$\begin{aligned} \dot{\mathbf{E}}_{1,1} &= \mathbf{K}_1\mathbf{E}_{1,1} + \Delta\mathbf{A}_1\mathbf{X}_{1,1} + \Delta\mathbf{B}_1 v_{1,1}, \\ \dot{\mathbf{E}}_{2,1} &= \mathbf{K}_1\mathbf{E}_{2,1} + \Delta\hat{\mathbf{A}}_1\mathbf{X}_{2,1} + \Delta\mathbf{B}_1 v_{2,1}, \end{aligned} \quad (6)$$

where $\Delta\mathbf{A}_1 \triangleq \hat{\mathbf{A}}_1 - \mathbf{A}_1$, $\Delta\mathbf{B}_1 \triangleq \hat{\mathbf{B}}_1 - \mathbf{B}_1$ are matrices of parametric mismatches.

Similarly, error equations for the remaining layers of TCSs are obtained:

$$\begin{aligned} \dot{\mathbf{E}}_{1,k} &= \mathbf{K}_k\mathbf{E}_{1,k} + \Delta\mathbf{A}_k\mathbf{X}_{1,k} + \Delta\mathbf{B}_k u_{1,k}, \\ \dot{\mathbf{E}}_{2k} &= \mathbf{K}_k\mathbf{E}_{2,k} + \Delta\mathbf{A}_k\mathbf{X}_{2,k} + \Delta\mathbf{B}_k u_{2,k}. \end{aligned} \quad (7)$$

Let input $g_i(t)$, i = 1, 2 satisfy the excitation constancy (EC) condition:

$$\mathscr{EC}_{\underline{\alpha}_i,\overline{\alpha}_i} : \underline{\alpha}_i \le g_i^2(t) \le \overline{\alpha}_i \quad \forall t \in [t_0, t_0 + T], \quad (8)$$

where $\underline{\alpha}_i$ and $\overline{\alpha}_i$ are positive numbers, $T > 0$. Further, condition (8) is written as $g_i(t) \in \mathscr{EC}_{\underline{\alpha}_i,\overline{\alpha}_i}$. If $g_i(t)$ does not have an EC property, then it is written as $g_i(t) \notin \mathscr{EC}_{\underline{\alpha}_i,\overline{\alpha}_i}$.

Theorem 1. Let the following conditions be satisfied: 1) $g_i(t) \in \mathscr{EC}_{\underline{\alpha}_i,\overline{\alpha}_i}$, where $(\underline{\alpha}_i, \overline{\alpha}_i) > 0$; 2) system (5) is stable and detectable; 3) matrix $\mathbf{K}_1 \in \mathbb{R}^{q_1 \times q_1}$ is a Hurwitz matrix; 3) outputs $v_{i1}(t) \in \mathscr{EC}_{\underline{\sigma}_{i1},\overline{\sigma}_{i1}}$, where $\underline{\sigma}_{i1} > 0$, $\overline{\sigma}_{i1} > 0$, $i = 1,2$; 4) $\mathbf{X}_{i,1}(t) \in \mathscr{EC}_{\underline{\alpha}_{\mathbf{X}_{i,1}},\overline{\alpha}_{\mathbf{X}_{i,1}}}$, where $(\underline{\alpha}_{\mathbf{X}_{i,1}}, \overline{\alpha}_{\mathbf{X}_{i,1}}) > 0$. Then system (5) is identifiable if

$$\pi_{12}\left\|\Delta\mathbf{A}_1\right\|^2 + 0.5\upsilon_{12}\left\|\Delta\mathbf{B}_1\right\|^2 \le (\lambda - 0.5)V_1, \quad (9)$$

where $\lambda_1 > 0.5$, $\pi_{12} = 2\max(\bar{\alpha}_{X_{11}}, \bar{\alpha}_{X_{12}})$, $\upsilon_{12} = \bar{\sigma}_{11} + \bar{\sigma}_{21}$, $\|\Delta\mathbf{A}_1\|^2 = \mathrm{tr}(\Delta\mathbf{A}_1^{\mathrm{T}}\Delta\mathbf{A}_1)$, tr is a matrix trace, and $V_1(t) = 0.5\mathbf{E}_{11}^{\mathrm{T}}(t)\mathbf{R}_1\mathbf{E}_{11}(t) + 0.5\mathbf{E}_{21}^{\mathrm{T}}(t)\mathbf{R}_1\mathbf{E}_{21}(t)$, $\mathbf{R}_1 = \mathbf{R}_1^{\mathrm{T}} > 0$ is a symmetric matrix.

When condition (9) is satisfied, system (5) is called parametrically $\mathscr{EC}_{1,X}$-identifiable on the set of state variables. The identifiability of subsystem layer (1) depends on the properties of the TCS output.

We consider system (7) and introduce LF:

$$V_k(t) = 0.5\mathbf{E}_{1,k}^{\mathrm{T}}(t)\mathbf{R}_k\mathbf{E}_{1,k}(t) + 0.5\mathbf{E}_{2,k}^{\mathrm{T}}(t)\mathbf{R}_k\mathbf{E}_{2,k}(t). \quad (10)$$

Theorem 2. Let the following conditions be satisfied: 1) matrix $\mathbf{K}_k \in \mathbb{R}^{q_k \times q_k}$ is a Hurwitz matrix; 2) $\|\mathbf{X}_{1,k}\|^2 \in \mathscr{EC}_{\underline{\alpha}_{\mathbf{X}_{1,k}}, \bar{\alpha}_{\mathbf{X}_{1,k}}}$, $\|\mathbf{X}_{2,k}\|^2 \in \mathscr{EC}_{\underline{\alpha}_{\mathbf{X}_{2,k}}, \bar{\alpha}_{\mathbf{X}_{2,k}}}$; $\pi_k = \max(\bar{\alpha}_{\mathbf{X}_{1,k}}, \bar{\alpha}_{\mathbf{X}_{2,k}})$, $y_{2k-1} \in \mathscr{EC}_{\underline{\alpha}_{y_{2,k-1}}, \bar{\alpha}_{y_{2,k-1}}}$; 3) system (5) is $\mathscr{EC}_{1,X}$-identifiable; 4) system (7) is stable and detectable; 5) $v_{2,k-1}^2 \in \mathscr{EC}_{\underline{\alpha}_{v_{2,k-1}}, \bar{\alpha}_{v_{2,k-1}}}$; 6) operator d_{k-1} is constant: $d_k \le \omega_k \le \omega$, where ω is some number; 7) $\lambda_k \ge 0.5$. Then system (7) is $\mathscr{EC}_{k,X}$-identifiable, if

$$0.5\pi_{k,i}\|\Delta\mathbf{A}_k\|^2 + 0.5\|\Delta\mathbf{B}_k\|^2 \times \times \left(\tilde{\bar{\alpha}}_k + 2\omega^2(\tilde{\bar{\alpha}}_k + 2\omega^2\beta_k)\right) \le (\lambda_k - 0.5)V_k, \quad (11)$$

where $\pi_{k,i} = 2\max(\bar{\alpha}_{X_{1,k}}, \bar{\alpha}_{X_{2,k}})$, $\tilde{\bar{\alpha}}_k = 2\max(\bar{\alpha}_{y_{2,k-1}}, \bar{\alpha}_{y_{1,k-1}})$, $\beta_k = \max_i(\bar{\alpha}_{y_{i,k-1}} + \omega^2\bar{\alpha}_{v_{i,k-1}})$.

Note 1: Conditions $\|\mathbf{X}_{i,k}\|^2 \in \mathscr{EC}_{\underline{\alpha}_{\mathbf{X}_{i,k}}, \bar{\alpha}_{\mathbf{X}_{i,k}}}$, $i = 1, 2$ follow from $v_{i1}(t) \in \mathscr{EC}_{\underline{\alpha}_{v_{i1}}, \bar{\alpha}_{v_{i1}}}$.

Note 2: The identifiability of kth layer (7) depends on the properties of previous layers of systems (1) and (2) and cross-connections. The parameters of cross-connections should be selected so that condition (11) is satisfied.

We consider the case when operator d_k is differentiable.

Theorem 3. Let the conditions of Theorem 2 be satisfied and: 1) operator d_{k-1} is differentiable, i.e., $v_{1,k-1} = d(y_{2,k-1} + d_{k-2}v_{2,k-1})/dt$; 2) systems (1) and (2) are stable, detectable, and recoverable. Then system (7) is $\mathscr{EC}_{k,X}$-identifiable, if

$$0.5\pi_{k,i}\|\Delta\mathbf{A}_k\|^2 + 0.5\|\Delta\mathbf{B}_k\|^2 \times \times \left(\tilde{\bar{\alpha}}_{k,\dot{y}} + 2(\bar{\alpha}_{\dot{y}} + \tilde{\bar{\alpha}}_{\ddot{v}})\right) \le (\lambda_k - 0.5)V_k, \quad (12)$$

where $\tilde{\bar{\alpha}}_{k,\dot{y}} = 2\max_i \bar{\alpha}_{\dot{y}_{2k-1}}$, $\tilde{\bar{\alpha}}_{k,\ddot{v}} = 2\max_i \bar{\alpha}_{\ddot{v}_{i,k-1}}$, $\pi_{k,i} = 2\max(\bar{\alpha}_{\mathbf{X}_{1,k}}, \bar{\alpha}_{\mathbf{X}_{2,k}})$, $V_k(t)$ is of the form (10).

Thus, the findings allow estimating the identifiability of systems (1) and (2) under measurability of the state vector of elements of all TCS layers. Most often, only set (3) is available for observation. In this case, it is necessary to operate with the available information to estimate the TCS identifiability.

We transform the TCS to a form convenient for using set $\mathbb{I}_o$ [28, 29] and consider the system (1). Let $\mathbf{A}_1$ be the Frobenius matrix with vector of parameters $\mathbf{A}_{1,s} \in \mathbb{R}^{q_1}$, $\mathbf{A}_{1,s} = [a_{1,s,1}, a_{1,s,2}, \ldots, a_{1,s,q_1}]^{\mathrm{T}}$, $\mathbf{C}_1 = [1, 0, \ldots, 0]^{\mathrm{T}}$, $\mathbf{B}_1 = [0, \ldots, 0, b_{1,s}]^{\mathrm{T}}$. In space (v_{11}, y_{11}), system (1) corresponds to the following representation:

$$\dot{y}_{1,1} = \bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}, \quad (13)$$

where $\bar{\mathbf{A}}_{1,1}^{\mathrm{T}} = [-a_{1,1,1}, a_{1,1,2}, \ldots, a_{1,1,q_1}; b_{1,s}, b_{1,2}, \ldots, b_{1,q_1}]$, $\bar{\mathbf{A}}_{1,1} \in \mathbb{R}^{2q_1}$, $\mathbf{P}_{1,1} \in \mathbb{R}^{2q_1}$.

Considering (13), we transform the system (1) in the input-output space to the following form:

$$\begin{cases} \dot{y}_{1,1} = \bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}, \\ \dot{y}_{2,1} = \bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{2,1}. \end{cases} \quad (14)$$

In order to evaluate the identifiability of system (14) by the output ($\mathscr{EC}_{1,y}$-identifiability), we consider the following model:

$$\begin{cases} \dot{\hat{y}}_{1,1} = -\chi e_{1,1} + \hat{\bar{\mathbf{A}}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}, \\ \dot{\hat{y}}_{2,1} = -\chi e_{2,1} + \hat{\bar{\mathbf{A}}}_{1,1}^{\mathrm{T}}\mathbf{P}_{2,1}, \end{cases} \quad (15)$$

where $\chi_1 > 0$, $e_{i,1} = \hat{y}_{i,1} - y_{i,1}$ is the prediction error of the $y_{i,1}$th output, i = 1, 2. The following equation is derived for $e_{i,1}$:

$$\dot{e}_{i,1} = -\chi_1 e_{i,1} + \Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{i,1}, \quad \Delta\bar{\mathbf{A}}_{1,1} = \hat{\bar{\mathbf{A}}}_{1,1} - \bar{\mathbf{A}}_{1,1}. \quad (16)$$

We consider LF $V_{1,2}(e_{1,1}, e_{2,1}) = 0.5(e_{1,1}^2 + e_{2,1}^2)$. For $\dot{V}_{1,2}$, we derive:

$$\dot{V}_{1,2} = -2\chi_1 V_{1,2} + \Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}(\mathbf{P}_{1,1}e_{1,1} + \mathbf{P}_{2,1}e_{2,1}) \le \le -\frac{\chi_1}{2}V_{1,2} + \frac{1}{2\chi_1}\Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_1\mathbf{P}_1^{\mathrm{T}}\Delta\bar{\mathbf{A}}_{1,1},$$

where $\mathbf{P}_1\mathbf{P}_1^{\mathrm{T}} = \mathbf{P}_{1,1}\mathbf{P}_{1,1}^{\mathrm{T}} + \mathbf{P}_{2,1}\mathbf{P}_{2,1}^{\mathrm{T}}$.

It follows from (16) that $\Delta\bar{\mathbf{A}}_{1,1}=0$, if vector $\mathbf{P}_i(t)$ is extremely nondegenerate, i.e., $\mathbf{P}_i(t)\in\mathscr{EC}_{\underline{\alpha}_{\mathbf{P}_i},\bar{\alpha}_{\mathbf{P}_i}}$ ($i=1,2$) and subsystem (1) is parametrically $\mathscr{EC}_{1,y}$-identifiable by the output. The $\mathscr{EC}_{1,y}$-identifiability of subsystem (2) is justified similarly.

For the kth layer (system (2)), the following representation similar to (13) can be obtained:

$$\begin{cases}\dot{y}_{1,k}=\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\mathbf{P}_{1,k},\\ \dot{y}_{2,k}=\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\mathbf{P}_{2,k}\end{cases}\tag{17}$$

where $\mathbf{P}_{i,k}\in\mathbb{R}^{2q_k}$ is the generalized input. For convenience, representations (13)–(15) and (17) are called noncanonical identification representations.

When identifying TCS and incomplete a priori information, the task of determining the type of cross-links may arise. In this case, the following approach can be used. We consider kth layer of system (2) with identical cross-connections and d_{k-1} = const and construct structure $\mathbf{S}_{u_{i,k},y_{i,k}}$ described by function $f_{u_{i,k},y_{i,k}}:u_{i,k}\to y_{i,k}$, i = 1, 2 for both channels of the kth layer. Structure $\mathbf{S}_{u_{i,k},y_{i,k}}$ represents the TCS input-output state. We define secants for $\mathbf{S}_{u_{i,k},y_{i,k}}$:

$$\xi_{u_{i,k},y_{i,k}}=a_{\xi_{0,i,k}}+a_{\xi_{1,i,k}}u_{i,k},\tag{18}$$

where $a_{\xi_{0,i,k}}$ and $a_{\xi_{1,i,k}}$ are parameters determined by the least squares method.

Since the cross-connection is rigid, the angle between secants $\xi_{u_{i,k},y_{i,k}}$ would not exceed a certain value: $|a_{\xi_{1,1,k}}-a_{\xi_{1,2,k}}|\le\delta_{\xi_{i,k}}$. Hence, cross-connections are positive. If cross-connections are asymmetric, then $|a_{\xi_{1,1,k}}-a_{\xi_{1,2,k}}|>\delta_{\xi_{i,k}}$. In this case, signal $v_{1,k-1}$ operates out of phase with output $y_{1,k-1}$ of the previous layer of the first channel.

Note 3: If the channels of each layer are nonidentical, the identifiability of the system layers can be easily derived from the above theorems.

1.3. Adaptive identification of TCSs

We consider representations (13)–(15) and (17) for TCS. For subsystem (17), the following model is introduced:

$$\begin{cases}\dot{\hat{y}}_{1,k}=-\chi_k e_{1,k}+\hat{\bar{\mathbf{A}}}_{1,k}^{\mathrm{T}}\mathbf{P}_{1,k},\\ \dot{\hat{y}}_{2,k}=-\chi_k e_{2,k}+\hat{\bar{\mathbf{A}}}_{1,k}^{\mathrm{T}}\mathbf{P}_{2,k},\end{cases}\tag{19}$$

where $\chi_k>0$, $e_{i,k}=\hat{y}_{i,k}-y_{i,k}$ is the prediction error for output $y_{i,k}$ (k = 1, 2, i is the ith element in the kth layer). For $e_{i,k}$, the following equation is derived:

$$\dot{e}_{i,k}=-\chi_k e_{i,k}+\Delta\bar{\mathbf{A}}_{i,k}^{\mathrm{T}}\mathbf{P}_{i,k},\ \Delta\bar{\mathbf{A}}_{1,k}=\hat{\bar{\mathbf{A}}}_{1,k}-\bar{\mathbf{A}}_{1,k}.\tag{20}$$

The adaptive algorithm for adjusting parameters of model (19) has the following form:

$$\Delta\dot{\bar{\mathbf{A}}}_{1,k}=\dot{\hat{\bar{\mathbf{A}}}}_{1,k}=-\boldsymbol{\Gamma}_k e_{i,k}\mathbf{P}_{1,k},\tag{21}$$

where $\boldsymbol{\Gamma}_k\in\Re^{q_k\times q_k}$ is the diagonal matrix with $\gamma_{k,j}>0$.

We consider subsystem (14)–(16) and LF $\tilde{V}_1(e_{1,1})=0.5e_{1,1}^2$. From condition $\dot{\tilde{V}}_1\le 0$, the adaptive algorithm for adjusting vector $\hat{\bar{\mathbf{A}}}_{1,1}$ is derived, as follows:

$$\Delta\dot{\bar{\mathbf{A}}}_{1,1}=\dot{\hat{\bar{\mathbf{A}}}}_{1,1}=-\boldsymbol{\Gamma}_1 e_{1,1}\mathbf{P}_{1,1},\tag{22}$$

where $\boldsymbol{\Gamma}_1\in\Re^{q_1\times q_1}$ is the diagonal matrix with $\gamma_{1,j}>0$, j is the diagonal element number.

Note 4: Since TCS has identical channels, only one of them is adjusted.

Due to feedback, some parameters of system (14)–(16), (22) may be unidentifiable. Considering this, Eq. (16) may be written in the following form:

$$\dot{e}_{1,1}=-\chi_1 e_{1,1}+\Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}+f(\bar{\mathbf{A}}_{1,1},p_{v_{11}}),\tag{23}$$

where $f(\cdot)$ is uncertainty resulting from the nonfulfillment of the EC condition.

We consider LF $\tilde{V}_{1,e}(e_{1,1})=0.5e_{1,1}^2$, $\tilde{V}_{1,\Delta}(\Delta\bar{\mathbf{A}}_{1,1})=0.5\Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\boldsymbol{\Gamma}_1^{-1}\Delta\bar{\mathbf{A}}_{1,1}$.

Theorem 4: Let the following conditions be satisfied: 1) $g_i(t)\in\mathscr{EC}_{\underline{\alpha}_i,\bar{\alpha}_i}$, $\mathbf{P}_{1,1}\notin\mathscr{EC}_{\underline{\alpha}_{\mathbf{P}_{1,1}},\bar{\alpha}_{\mathbf{P}_{1,1}}}$; 2) $y_{i,1}\notin\mathscr{EC}_{\underline{\alpha}_{y_{i,1}},\bar{\alpha}_{y_{i,1}}}$, $i=1,2$; 3) $|f|^2\le\alpha_f$, where $\alpha_f\ge 0$; 4) there exists such $\upsilon>0$ that $e_{1,1}\Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}=\upsilon(e_{1,1}^2+\Delta\bar{\mathbf{A}}_{1,1}^{\mathrm{T}}\mathbf{P}_{1,1}\mathbf{P}_{1,1}^{\mathrm{T}}\Delta\bar{\mathbf{A}}_{1,1})$ is valid at $t>>t_0$; 5) $\lambda_{\max}^{-1}(\boldsymbol{\Gamma}_1)\left\|\Delta\bar{\mathbf{A}}_{1,1}\right\|^2\le\tilde{V}_{1,\Delta}\le\lambda_{\min}^{-1}(\boldsymbol{\Gamma}_1)\left\|\Delta\bar{\mathbf{A}}_{1,1}\right\|^2$; 6) the following system of differential inequalities is valid for system (16), (23):

$$\begin{bmatrix}\dot{\tilde{V}}_{1,e}\\ \dot{\tilde{V}}_{1,\Delta}\end{bmatrix}\le\underbrace{\begin{bmatrix}-\frac{\chi_1}{2} & 2\frac{\pi_{P_{1,1}}}{\chi_1}\\ 4\upsilon & -\frac{\upsilon\eta}{2}\end{bmatrix}}_{\mathbf{A}_{S_1}}\begin{bmatrix}\tilde{V}_{1,e}\\ \tilde{V}_{1,\Delta}\end{bmatrix}+\underbrace{\begin{bmatrix}\alpha_f\\ \chi_1\\ 0\end{bmatrix}}_{\mathbf{B}_{S_1}}\tag{24}$$

while comparison system $\dot{\mathbf{S}}_1(t)=\mathbf{A}_{S_1}\mathbf{S}_1(t)+\mathbf{B}_{S_1}$ for (24), where $\mathbf{S}_1(t)=[s_{1,e}(t),s_{1,\Delta}(t)]^{\mathrm{T}}$, $s_{1,w}(t)$

$(w=e,\Delta)$ is majorant for $\tilde{V}_{1,w}(t)$ and $s_{1,w}(t_0)\ge\tilde{V}_{1,w}(t_0)$. Then system (14)–(16), (23) is exponentially dissipative with estimate

$$[\tilde{V}_{1,e}(t)\tilde{V}_{1,\Delta}(t)]^{\mathrm{T}}\le e^{\mathbf{A}_{S_1}(t-t_0)}\mathbf{S}_1(t_0)+\int_{t_0}^{t}e^{\mathbf{A}_{S_1}(t-\tau)}\mathbf{B}_{S_1}d\tau,$$

if $\chi_1^2\eta>8\vartheta_{P_{1,1}}$, where $\eta=\underline{\pi}_{P_{1,1}}\lambda_{\min}^2(\boldsymbol{\Gamma}_1)$, $\underline{\pi}_{P_{1,1}}\ge 0$, $\left\|\mathbf{P}_{1,1}\mathbf{P}_{1,1}^{\mathrm{T}}\right\|\le\vartheta_{\mathbf{P}_{1,1}}(\bar{\alpha}_{\mathbf{P}_{1,1}})$, $\lambda_{\min}(\boldsymbol{\Gamma}_1)$ is minimum eigenvalue of matrix $\boldsymbol{\Gamma}_1$.

It follows from Theorem 4 that the adaptive identification system allows obtaining biased estimates of system (14)–(16) parameters.

We consider (17), (19), (21). Let d_{k-1} in (2) is constant.

We represent $\mathbf{P}_{1,k}$ and $\Delta\bar{\mathbf{A}}_{1,k}$ as: $\mathbf{P}_{1,k}=[\tilde{\mathbf{P}}_{1,k}^{\mathrm{T}},p_{j,v_{1,2,k-1}}]^{\mathrm{T}}$, $\Delta\bar{\mathbf{A}}_{1,k}=[\Delta\tilde{\bar{\mathbf{A}}}_{1,k}^{\mathrm{T}},\Delta d_{j,k}]^{\mathrm{T}}$, where $p_{j,v_{i,2,k-1}}$ is transformation $v_{1,2,k-1}$, j is element $\Delta\bar{\mathbf{A}}_{1,k}$ that is the transformation of d_{k-1}. Then (20) has the following form:

$$\dot{e}_{1,k}=-\div_k e_{1,k}+\Delta\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\mathbf{P}_{1,k}+f_{1,k}(\bar{\mathbf{A}}_{1,k}\mathbf{P}_{1,k}),\quad(25)$$

where $f_{1,k}(\cdot)\in\mathbb{R}$ is uncertainty resulting from the non-fulfillment of the EC condition.

We consider system (19), (20), (25), and LF $\tilde{V}_{k,e}(e_{1,k})=0.5e_{1,k}^2$, $\tilde{V}_{k,\Delta}(\Delta\bar{\mathbf{A}}_{1,k})=0.5\Delta\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\boldsymbol{\Gamma}_k^{-1}\Delta\bar{\mathbf{A}}_{1,k}$.

Theorem 5. Let the Theorem 4 conditions be satisfied and: 1) $y_{1,k}(t)\notin\mathscr{C}\mathscr{C}_{\underline{\alpha}_{y_{i,k}},\bar{\alpha}_{y_{i,k}}}$, $\mathbf{P}_{1,k}\notin\mathscr{C}\mathscr{C}_{\underline{\alpha}_{\mathbf{P}_{1,k}},\bar{\alpha}_{\mathbf{P}_{1,k}}}$; 2) $u_{1,k}=\upsilon_{u_{1,k}}(y_{1,k-1},v_{1,k-1})$, $;u_{1,k}\notin\mathscr{C}\mathscr{C}$; 3) $f_{1,k}^2\le\alpha_{f_{1,k}}$, where $\alpha_{f_{1,k}}\ge 0$; 4) estimate $\lambda_{\max}^{-1}(\boldsymbol{\Gamma}_k)\left\|\Delta\bar{\mathbf{A}}_{1,k}\right\|^2\le\tilde{V}_{k,\Delta}\le\lambda_{\min}^{-1}(\boldsymbol{\Gamma}_k)\left\|\Delta\bar{A}_{1,k}\right\|^2$ is valid, where $\lambda_k(\boldsymbol{\Gamma}_k)$ is minimum eigenvalue of matrix $\boldsymbol{\Gamma}_k$; 5) $\left\|P_{1,k}P_{1,k}^{\mathrm{T}}\right\|\le\pi_{P_{1,k}}<\bar{\alpha}_{P_{1,k}}$; 6) there exists such $\upsilon>0$ that at $t>>t_0$ the following is valid:

$$e_{1,k}\Delta\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\mathbf{P}_{1,k}=\upsilon(e_{1,k}^2+\Delta\bar{\mathbf{A}}_{1,k}^{\mathrm{T}}\mathbf{P}_{1,k}\mathbf{P}_{1,k}^{\mathrm{T}}\Delta\bar{\mathbf{A}}_{1,k});\quad(26)$$

7) the following system of differential inequalities is valid for system (21), (25):

$$\underbrace{\begin{bmatrix}\dot{\tilde{V}}_{k,e}\\ \dot{\tilde{V}}_{k,\Delta}\end{bmatrix}}_{\dot{\tilde{\mathbf{V}}}_k}\le\underbrace{\begin{bmatrix}-\frac{\chi_k}{2} & 2\frac{\vartheta_{P_{1,k}}}{\chi_k}\\ 4\upsilon & -\frac{\upsilon\eta}{2}\end{bmatrix}}_{\mathbf{A}_{S_k}}\underbrace{\begin{bmatrix}\tilde{V}_{k,e}\\ \tilde{V}_{k,\Delta}\end{bmatrix}}_{\tilde{\mathbf{V}}_k}+\underbrace{\begin{bmatrix}\alpha_{f_{1,k}}\\ \chi_k\\ 0\end{bmatrix}}_{\mathbf{B}_{S_k}},\quad(27)$$

while comparison system $\dot{\mathbf{S}}_k(t)=\mathbf{A}_{S_k}\mathbf{S}_k(t)+\mathbf{B}_{S_k}$ for (27), where $\mathbf{S}_k(t)=[s_{k,e}(t),s_{k,\Delta}(t)]^{\mathrm{T}}$, $s_{k,w}(t)$ $(w=e,\Delta)$ is a majorant for $\tilde{V}_{k,w}(t)$ and $s_{k,w}(t_0)\ge\tilde{V}_{k,w}(t_0)$. Then system (19), (20), (25) is exponentially dissipative with estimate

$$\tilde{\mathbf{V}}_k(t)\le e^{\mathbf{A}_{S_k}(t-t_0)}\mathbf{S}_k(t_0)+\int_{t_0}^{t}e^{\mathbf{A}_{S_k}(t-\tau)}\mathbf{B}_{S_k}d\tau,$$

if condition $\chi_k^2\eta>32\vartheta_{\mathbf{P}_{1,k}}$ is satisfied, where $\underline{\pi}_{\mathbf{P}_{1,k}}\ge 0$, $\eta=\underline{\pi}_{\mathbf{P}_{1,k}}\lambda_{\min}(\boldsymbol{\Gamma}_k)$.

It follows from (25) that the properties of the kth layer adaptive identification system depend on cross-connections.

Thus, the TCS identifiability is proved by the state and output of the TCS kth layer. The results confirming the convergence of the obtained estimates for system parameters are obtained. The adaptive identification system properties depend on parameters of the system cross-connections and information properties of signals in the TCS.

Note 5: If TCS contains nonidentical channels, models (17) should be applied to each kth layer. The same is true for cross-connections.

2. ICS

We consider system S_{ICS}

$$\begin{aligned}&\dot{\mathbf{X}}(t)=\mathbf{A}\mathbf{X}(t)+\mathbf{D}\mathbf{F}_1(\mathbf{X},t)+\mathbf{B}\mathbf{U}(t),\\ &\mathrm{L}\mathbf{Y}(t)=\mathbf{C}\mathbf{X}(t)+\mathbf{F}_2(\mathbf{X},t),\end{aligned}\quad(28)$$

where $\mathbf{X}\in\mathbb{R}^m$ is the state vector, $\mathbf{A}\in\mathbb{R}^{m\times m}$ is the state matrix, $\mathbf{D}\in\mathbb{R}^{m\times q}$, $\mathbf{F}_1(\mathbf{X},t):\mathbb{R}^m\to\mathbb{R}^q$ is the nonlinear vector function, $\mathbf{U}\in\mathbb{R}^k$ is the input (control) vector, $\mathbf{B}\in\mathbb{R}^{m\times k}$, $\mathbf{Y}\in\mathbb{R}^n$ is the output vector, $\mathbf{C}\in\mathbb{R}^{n\times m}$, $\mathbf{F}_2(\mathbf{X},t):\mathbb{R}^m\to\mathbb{R}^n$ is the disturbance (measurement errors) vector, and L is the operator determining the way vector $\mathbf{Y}$ is formed. In some cases, L can be the differential operator representing dynamic properties of the measuring system or the way of mutual interaction between subsystems. Matrices **A**, **D**, **B** are block matrices representing the state of certain subsystems. Vector $\mathbf{F}_2(\mathbf{X},t)$ can be either a disturbance (measurement error) or a variable representing the influence of certain subsystems.

The data set has the following form:

$$\mathbb{I}_{\mathrm{o}}=\left\{\mathbf{Y}(t),\mathbf{U}(t),t\in\left[t_0,t_N\right]\right\},\ t_N<\infty.\quad(29)$$

Assumption 1: Elements $\varphi_{1,i}(x_j)\in\mathbf{F}_1$, $\varphi_{2,i}(x_j)\in\mathbf{F}_2$ are smooth single-valued functions.

For estimating parameters of matrices **A**, **D**, **B**, **C**, the following model is used:

$$\begin{cases} \dot{\hat{\mathbf{X}}}(t) = \hat{\mathbf{A}}(t)\hat{\mathbf{X}}(t) + \hat{\mathbf{D}}(t)\mathbf{F}_1(\mathbf{X},t) + \hat{\mathbf{B}}(t)\mathbf{U}(t), \\ \mathbf{L}\hat{\mathbf{Y}}(t) = \mathbf{C}\hat{\mathbf{X}}(t) + \mathbf{F}_2(\hat{\mathbf{X}},t), \end{cases} \quad (30)$$

where $\hat{\mathbf{A}}(t)$, $\hat{\mathbf{D}}(t)$, $\hat{\mathbf{B}}(t)$ are matrices with adjustable parameters.

Problem. For system (27) satisfying Assumption 1, to develop model (30) on the basis of analysis $\mathbb{I}_o$ and find such rules for adjusting parameters of matrices $\hat{\mathbf{A}}(t)$, $\hat{\mathbf{D}}(t)$, and $\hat{\mathbf{B}}(t)$ so that

$$\lim_{t\to\infty} \left\| \hat{\mathbf{Y}}(t) - \mathbf{Y}(t) \right\| \le \delta_y, \quad \delta_y \ge 0,$$

where $\|\cdot\|$ is the Euclidean norm.

Note 6: For some class of systems, Eq. (28) can be considered as an equation describing connections between subsystems. In this case, the type of vector $\mathbf{F}_2(\mathbf{X},t)$ should be estimated.

For the synthesis of adaptive identification algorithms, the approach described in Section 1.3 can be used.

Note 7: If system S_{ICS} contains nonlinear subsystems, then in order to decide on the class of nonlinearity under uncertainty, the interconnection graph should be plotted [30] from which the type of nonlinearity can be decided. The same is valid when the system contains several nonlinearities in one subsystem.

The adaptation models and algorithms match the equations derived in the previous section. For evaluating the quality of the adaptive identification subsystem, the theorems from Section 1.2 can be applied.

3. EXAMPLES

Example 1. We consider the two-channel target angle tracking system with identical azimuth and elevation channels and asymmetric cross-connections [31]:

$$\dot{x} = -a_x x + b_x(g - y), \quad (31)$$

$$\ddot{y} = -a_y \dot{y} + b_y \left(x - (g_1 - y_1) - k(g_1 - y_1)' \right), \quad (32)$$

$$\dot{x}_1 = -a_x x_1 + b_x(g_1 - y_1), \quad (33)$$

$$\ddot{y}_1 = -a_y \dot{y}_1 + b_y \left(x_1 + (g - y) + k(g - y)' \right), \quad (34)$$

where $a_x = T_x^{-1}$, $b_x = T_x^{-1}k_x$; T_x, k_x are the time constant and amplifier gain, respectively; k is the cross-connection parameter; $a_y = T_y^{-1}$, $b_y = T_y^{-1}k_y$ are servomotor parameters; g, g_1 are input influences. The cross-connection is shown in parentheses as a differentiating link with parameter k, $(g_1 - y_1)' = d(g_1 - y_1)/dt$.

The information about link outputs $x_i(t)$, $y_i(t)$, and outputs $g_i(t)$ at some time interval is available for measurement. The parameters of systems (31) and (32) should be estimated.

We use transformation [28] to obtain a model for $y(t)$ and consider the system of filters (transformation):

$$\begin{aligned} \dot{p}_y &= -\mu_1 p_y + y, \\ \dot{p}_\vartheta &= -\mu_1 p_\vartheta + \vartheta, \\ \dot{p}_\upsilon &= -\mu_1 p_\upsilon + \upsilon, \end{aligned} \quad (35)$$

where $\vartheta \triangleq x - g_1 + y_1$, $\upsilon \triangleq d(g_1 - y_1)/dt$, $p_i(0) = 0$, $i = y, \vartheta, \upsilon$, $\mu_1 > 0$ is a number that does not match the roots of the characteristic equation for the second equation in (31). Then the model for (31) has the following form:

$$\dot{\hat{x}} = -\chi_x e_x + \hat{a}_x x + \hat{b}_x(g - y), \quad (36.1)$$

$$\dot{\hat{y}} = -\chi_y e_y + \hat{a}_y y + \hat{a}_{p_y} p_y + \hat{a}_\vartheta p_\vartheta + \hat{a}_\upsilon p_\upsilon, \quad (36.2)$$

where $e_x = \hat{x} - x$, $e_y = \hat{y} - y$, χ_x, χ_y are positive numbers (reference model). Algorithms for adjusting parameters of system (36) are the following:

$$\begin{aligned} \dot{\hat{a}}_x &= -\gamma_{a_x} e_x x, \quad \dot{\hat{b}}_x = -\gamma_{b_x} e_x (g - y), \\ \dot{\hat{a}}_y &= -\gamma_{a_y} e_y y, \quad \dot{\hat{a}}_{p_y} = -\gamma_{a_{p_y}} e_y p_y, \\ \dot{\hat{a}}_\vartheta &= -\gamma_{a_\vartheta} e_y p_\vartheta, \ \dot{\hat{a}}_\upsilon = -\gamma_{a_\upsilon} e_y p_\upsilon \end{aligned} \quad (37)$$

where $\gamma_{a_i}, \gamma_{b_i}$ $(i = x, y, a_y, p_y, \vartheta, \upsilon)$ are positive numbers ensuring convergence of (36).

System (31), (32) is modelled with parameters: $a_x = 1.2$, $b_x = 2$, $a_y = 5.95$, $b_y = 1$, $k = 5$. The inputs are $g(t) = 1.5\sin(0.1\pi t)$, $g_1(t) = 1.5\sin(0.1\pi t)$.

The structures confirming the presence of asymmetric connections in the system are presented in Fig. 1. In the figure, the impact of cross-connections on the channel output is shown. Analyzing $a_{\xi_{0,\varepsilon_1,y}} = 0.53$, $a_{\xi_{0,\varepsilon,y_1}} = -0.45$ of secants $\xi_{\varepsilon_1,y}, \xi_{\varepsilon,y_1}$ proves the fulfillment of condition $|a_{\xi_{0,\varepsilon_1,y}} - a_{\xi_{0,\varepsilon,y_1}}| > \delta_{\xi_0}$, where $\delta_{\xi_0} = 0.2$. Hence, cross-connections are antisymmetric.

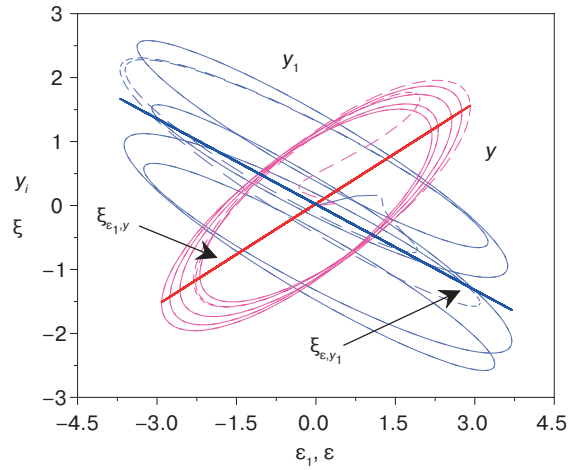


Fig. 1. Estimating the structure of cross-connections

The parameters of Eqs. (31) and (32) are estimated with respect to Note 4. The identification system parameters are: $\mu_1 = 1.5$, $\chi_x = 1.5$, $\chi_y = 2$.

The results of adaptive identification are presented in Figs. 2–6. The adaptive identification of parameters of model (36) is shown in Figs. 2 and 3, where t is hereafter the current time in relative units.

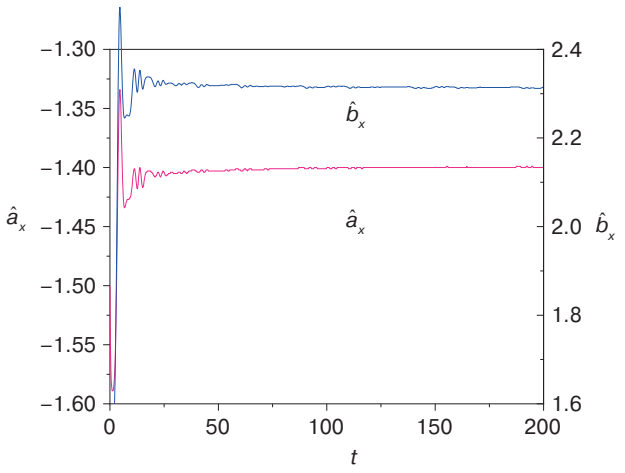


Fig. 2. Adjusting parameters of model (36.1)

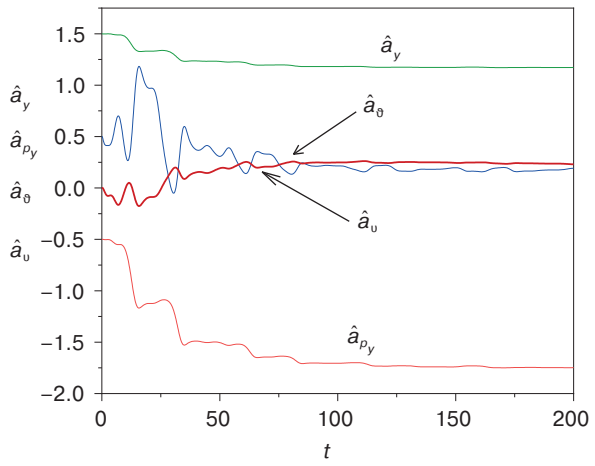


Fig. 3. Adjusting parameters of model (36.2)

The identification errors obtained using models (36) are shown in Fig. 4.

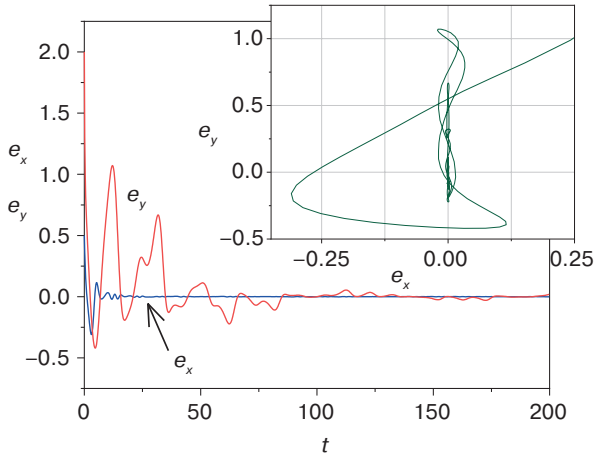


Fig. 4. A change in the prediction error

It follows from Figs. 3 and 4 that adjusting parameters of layer 2 depends on both the output of layer 1 (model (36.1)) and the properties of cross-connections. In addition, the quality of adjusting parameters of (36.2) is also dependent on the correlation between the outputs of elements (31) and (33). This correlation is transmitted through cross-connections. The conclusion drawn is supported by the structures shown in Fig. 5 as well as the statements of Theorem 5. Adjusting parameters of model (36.1) is smooth (Fig. 6).

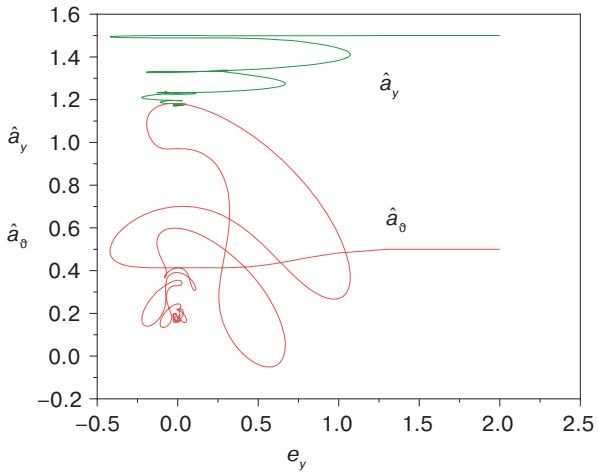


Fig. 5. Adjusting parameters of model (36.2)

Example 2. A pseudo-linear two-channel corrector is described by the system of equations in [32], as follows:

$$\begin{cases} \dot{x}_1 = -\alpha_1 x_1 + \beta_1 (g - x_6), \\ \dot{x}_2 = -\alpha_2 x_2 + \mu_2 (g - x_6)' + \beta_2 (g - x_6), \\ \dot{x}_6 = x_7, \\ \dot{x}_7 = -\alpha_{01} x_7 - \alpha_{02} x_6 + \beta_0 \left(x_1 \text{sign}(x_1)\right)\left(\text{sign}(x_2)\right), \end{cases} \quad (38)$$

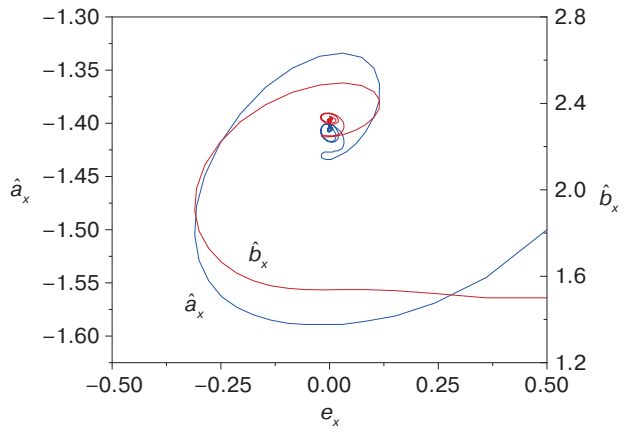


Fig. 6. Adjusting parameters of model (36.1)

where g is the master control, x_6 is the object output, $g - x_6$ is the mismatch error, x_1 is the amplitude channel output, x_2 is the phase channel output, $u = x_1\text{sign}(x_1) \cdot \text{sign}(x_2)$ is the regulator output, and $(g - x_6)' = d(g - x_6)/dt$, $\alpha_1, \beta_1, \alpha_2, \mu_2, \beta_2, \alpha_{01}, \alpha_{02}, \beta_0$ are system parameters.

Note 8: In [32], system parameters have been preliminary selected.

It is assumed that elements of set $\mathbb{I}_o = \{x_1(t), x_2(t), x_6(t), t \in [0, t_k]\}$ are measured, where t_k is the known number. The parameters of system (38) are identified using the following model:

$$\dot{\hat{x}}_1 = -k_1 e_1 + \hat{\alpha}_1 x_1 + \hat{\beta}_1 (g - x_6), \quad (39)$$

$$\dot{\hat{x}}_2 = -k_2 e_2 + \hat{\alpha}_2 x_2 + \hat{\mu}_2 \, d(g - x_6)/dt + \hat{\beta}_2 (g - x_6), \quad (40)$$

$$\dot{\hat{x}}_6 = -k_6 e_6 + \hat{\alpha}_{01} x_6 + \hat{\alpha}_{p_{x_6}} p_{x_6} + \hat{\alpha}_{p_u} p_u, \quad (41)$$

where k_1, k_2, k_6 are known numbers (the reference model parameters); $e_i = \hat{x}_i - x_i$, $i = 1, 2, 6$; $\hat{\alpha}_i, \hat{\beta}_i$ $(i = 1, 2)$, $\hat{\alpha}_{p_{x_5}}, \hat{\alpha}_{p_{x_5}}$ are adjustable parameters; $\hat{x}_i$ $(i = 1, 2, 6)$ are model outputs; p_{x_6}, p_u are obtained by analogy with (33).

The adaptation algorithms are the following:

$$\dot{\hat{\alpha}}_1 = -\gamma_{\alpha_1} e_1 x_1, \qquad \dot{\hat{\beta}}_1 = -\gamma_{\beta_1} e_1 (g - x_6),$$

$$\dot{\hat{\alpha}}_2 = -\gamma_{\alpha_2} e_2 x_2, \dot{\hat{\beta}}_2 = -\gamma_{\beta_2} e_2 (g - x_6),$$

$$\dot{\hat{\mu}}_2 = -\gamma_{\mu_2} e_2 (g - x_6)', \quad (42)$$

$$\dot{\hat{\alpha}}_{01} = -\gamma_{\alpha_{01}} e_6 x_6, \dot{\hat{\alpha}}_{p_{x_6}} = -\gamma_{p_{x_6}} e_6 p_{x_6}, \dot{\hat{\alpha}}_{p_u} = -\gamma_{p_u} e_6 p_u,$$

where $\gamma_i > 0$ is the gain in the adjustment loop of the corresponding parameter.

System (38) is modelled with the following parameters: $\alpha_1 = 1.05$, $\beta_1 = 3.5$, $\alpha_2 = 2.2$, $\beta_2 = 2.2$, $\alpha_{01} = 3$, $\alpha_{02} = 5.03$, $\beta_0 = 5.2$, $g(t) = \sin(0.05\pi t)$. The results are shown in Fig. 7–12.

Figure 7 shows the structures (transient excluded) representing phase processes in system (38). Obviously, the system is nonlinear. The use of results [30] shows that the system is structurally identifiable. Therefore, the input is S-synchronizing and allows accounting for TCS nonlinear properties.

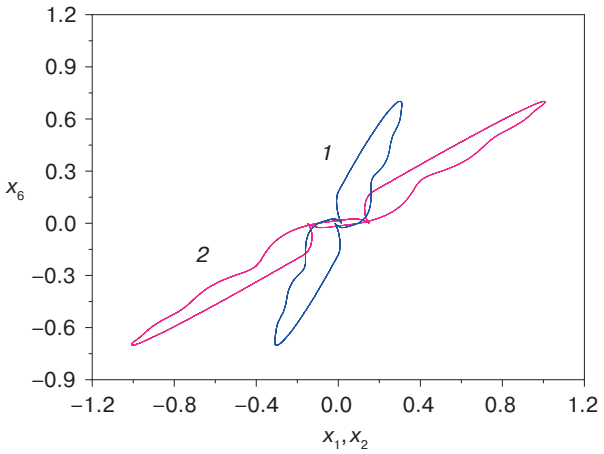


Fig. 7. Phase portraits of the system: (*1*) $x_6(x_2)$, (*2*) $x_6(x_1)$

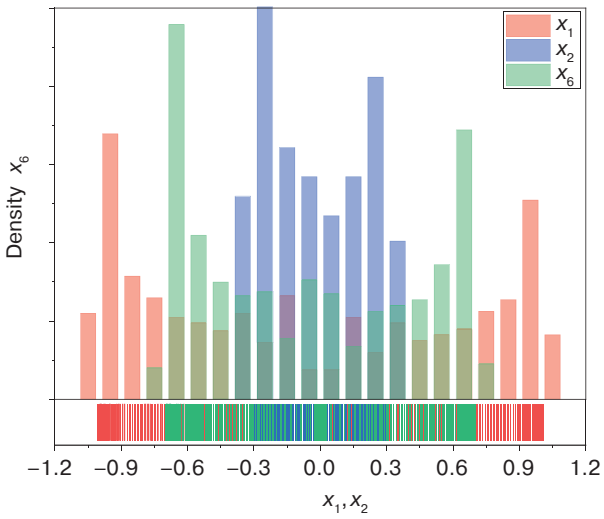


Fig. 8. Estimating the mutual influence of variables x_1, x_2 on x_6

The analysis shows the relationship between variables x_1 and x_2 (correlation coefficient is 0.94). Their impact on the TCS output is represented by the density diagram (Fig. 8). The density diagram characterizes the impact of x_1 and x_2 on the change in variable x_6. The intervals of the impact of variables on x_6 are clearly visible. The mutual influence affects the convergence of adaptive algorithms.

The adjustment of the amplitude channel model is shown in Fig. 9 while the phase channel model is shown in Fig. 10.

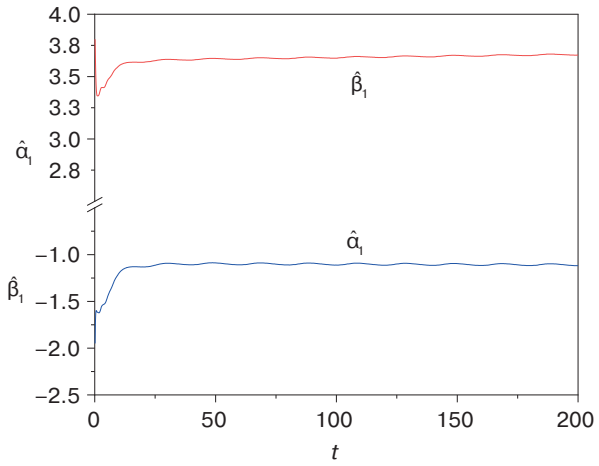


Fig. 9. Adjusting parameters of model (39)

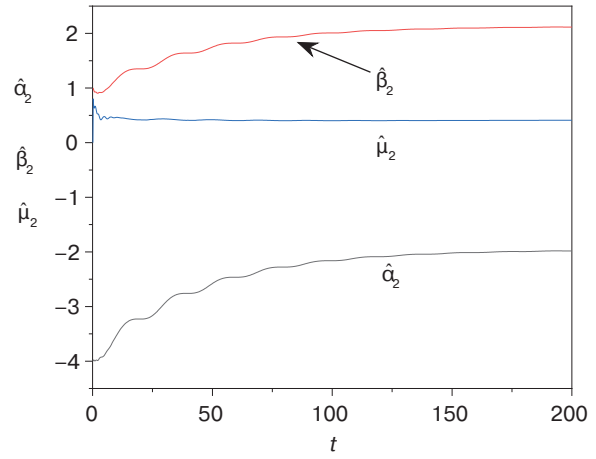


Fig. 10. Adjusting parameters of model (40)

Figure 11 shows the adjustment of the object model parameters. The adjustment is affected by the master control.

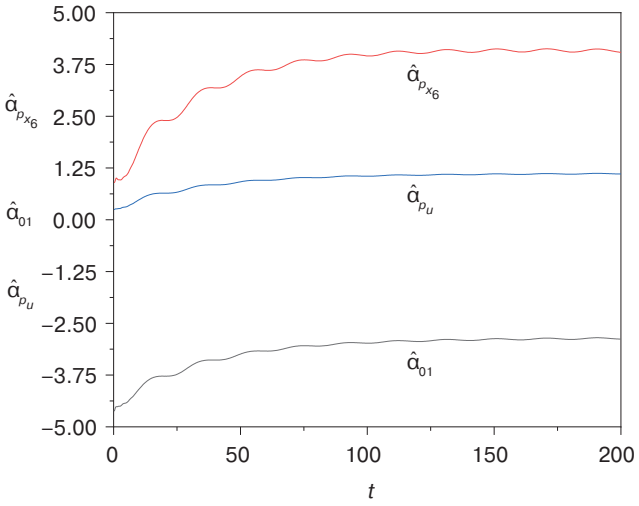


Fig. 11. Adjusting parameters of model (41)

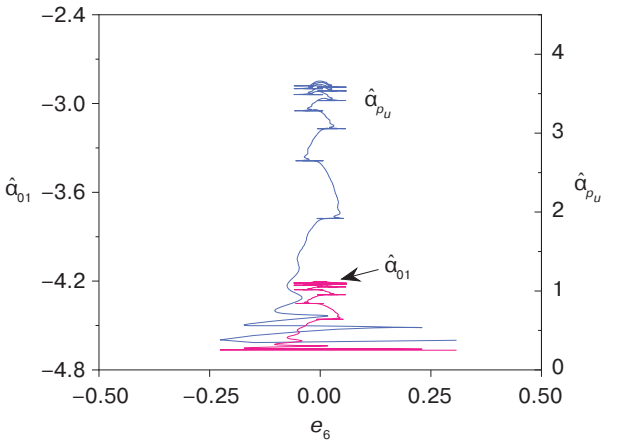


Fig. 12. Dynamics of the adjustment loop for model (41)

The dynamics of processes in adjusting model (40) is more complicated. Here, the input influence is encountered. For TCS, the statement of Theorem 5 is valid.

CONCLUSIONS

The proposed approach to identifying ICS is based on features of adaptive identification of a TCS with cross-connections. Structural aspects of TCS identification are based on obtained conditions of the system identifiability in the state space and output space, as well as the influence of input properties on the possibility of estimating system parameters. Adaptive algorithms for identifying TCS with identical channels are obtained. The identifiability of a TCS in terms of state and output is demonstrated. The results confirm the convergence of estimates for system parameters. The properties of the adaptive identification system depend on the system cross-connection parameters and informative properties of signals in the TCS. When considering the case of ICS, the role of a priori information while accounting for existing interconnections is emphasized. At appropriate ICS splitting, it would be possible to apply the approaches proposed for a TCS. The examples of adaptive identification of real systems are given. The case of TCS with nonlinear control and complex relationships is considered.

Due to the multifaceted nature of the subject area, the paper covers only certain aspects. The proposed approach can be used to estimate the dynamics of processes in an adaptive system taking into account the impact of system elements on the quality of the parameter adjustment process.

REFERENCES

1. Morozovskii V.T. *Mnogosvyaznye sistemy avtomaticheskogo regulirovaniya* (*Multi-Connected Automatic Control Systems*). Moscow: Energiya; 1970. 288 p. (in Russ.).
2. Zyryanov G.V. *Sistemy upravleniya mnogosvyaznymi ob"ektami* (*Control Systems for Multi-Connected Objects*): textbook. Chelyabinsk: South Ural State University Publishing Center; 2010. 112 p. (in Russ.).
3. Meerov M.V., Litvak B.L. *Optimizatsiya sistem mnogosvyaznogo upravleniya* (*Optimization of Multi-Connection Control Systems*). Moscow: Nauka; 1972. 344 p. (in Russ.).
4. Bukov V.N., Maksimenko I.M., Ryabchenko V.N. Control of multivariable systems. *Autom. Remote Control.* 1998;59(6): 832–842.
   [Original Russian Text: Bukov V.N., Maksimenko I.M., Ryabchenko V.N. Control of multivariable systems. *Avtomatika i Telemekhanika*. 1998;6:97–110 (in Russ.).]
5. Voronov A.A. *Vvedenie v dinamiku slozhnykh upravlyaemykh system* (*Introduction in Dynamics of Complex Controlled Systems*). Moscow: Nauka; 1985. 352 p. (in Russ.).
6. Egorov I.N., Umnov V.P. *Sistemy upravleniya elektroprivodov tekhnologicheskikh robotov i manipulyatorov* (*Control Systems for Electric Drives of Technological Robots and Manipulators*): textbook. Vladimir: Vladimir State University Publ.; 2022. 314 p. (in Russ.).
7. Egorov I.N. *Pozitsionno-silovoe upravlenie robototekhnicheskimi i mekhatronnymi ustroistvami* (*Positional Force Control of Robotic and Mechatronic Devices*). Vladimir: Vladimir State University Publ.; 2010. 192 p. (in Russ.).
8. Gupta N., Chopra N. Stability analysis of a two-channel feedback networked control system. In: *2016 Indian Control Conference* (*ICC*). 2016. https://doi.org/10.1109/INDIANCC.2016.7441129
9. Pawlak A., Hasiewicz Z. Non-parametric identification of multi-channel systems by multiscale expansions. In: *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2011. https://doi.org/10.1109/ICASSP.2002.5744953
10. Kholmatov U. The possibility of applying the theory of adaptive identification to automate multi-connected objects. *Am. J. Eng. Technol*. 2022;4(03):31–38. Available from URL: https://inlibrary.uz/index.php/tajet/article/view/5789
11. Aliyeva A.S. Identification of multiconnected dynamic objects with uncertainty based on neural technology and reference converters. *Informatics and Control Problems.* 2019;39(2):93–102. Available from URL: https://icp.az/2019/2-11.pdf
12. Hua C., Guan X., Shi P. Decentralized robust model reference adaptive control for interconnected time-delay systems. In: *Proceeding of the 2004 American Control Conference.* Boston, Massachusetts: June 30 – July 2, 2004. P. 4285–4289. https://doi.org/10.23919/ACC.2004.1383981
13. Vorchik B.G. Identifiability of a multivariable closed-loop stochastic system. Decomposition of a closed-loop system in identification. *Autom. Remote Control*. 1977;38(2):172–183.
   [Original Russian Text: Vorchik B.G. Identifiability of a multivariable closed-loop stochastic system. Decomposition of a closed-loop system in identification. *Avtomatika i Telemekhanika*. 1977;2:14–28 (in Russ.).]
14. Glentis G.-O., Slump C.H. A highly modular normalized adaptive lattice algorithm for multichannel least squares filtering. In: *1995 International Conference on Acoustics, Speech, and Signal Processing.* 1995;2:1420–1423. https://doi.ieeecomputersociety.org/10.1109/ICASSP.1995.480508
15. Ali M., Abbas H., Chughtai S.S., Werner H. Identification of spatially interconnected systems using neural network. In: *49th IEEE Conference on Decision and Control* (*CDC*). 2011. https://doi.org/10.1109/CDC.2010.5717080
16. Yang Q., Zhu M., Jiang T., He J., Yuan J., Han J. Decentralized Robust Adaptive output feedback stabilization for interconnected nonlinear systems with uncertainties. *J. Control Sci. Eng*. 2016;2016:article ID 3656578. https://doi.org/10.1155/2016/3656578
17. Wu H. Decentralized adaptive robust control of uncertain large-scale non-linear dynamical systems with time-varying delays. *IET Control Theory & Applications*. 2012;6(5):629–640. https://doi.org/10.1049/iet-cta.2011.0015
18. Fan H., Han L., Wen C., Xu L. Decentralized adaptive output-feedback controller design for stochastic nonlinear interconnected systems. *Automatica.* 2012;48(11):2866–2873. https://doi.org/10.1016%2Fj.automatica.2012.08.022
19. Ali M., Chughtai S.S., Werner H. Identification of spatially interconnected systems. In: *Proceedings of the 48h IEEE Conference on Decision and Control* (*CDC) held jointly with 2009 28th Chinese Control Conference*. 2010. https://doi.org/10.1109/CDC.2009.5399748
20. Ioannou P.A. Decentralized adaptive control of interconnected systems. *IEEE Transactions on Automatic Control* 1986;31(4):291–298. https://doi.org/10.1109/TAC.1986.1104282
21. Sanandaji B.M., Vincent T.L., Wakin M.B. A review of sufficient conditions for structure identification in interconnected systems. *IFAC Proceedings Volumes*. 2012;45(16):1623–1628. https://doi.org/10.3182/20120711-3-BE-2027.00254
22. Soverini U., Söderström T. Blind identification of two-channel FIR systems: a frequency domain approach. *IFAC-PapersOnLine.* 2020;53(2):914–920. https://doi.org/10.1016/j.ifacol.2020.12.855
23. Huang Y., Benesty J., Chen J. Adaptive blind multichannel identification. In: Benesty J., Sondhi M.M., Huang Y.A. (Eds.). *Springer Handbook of Speech Processing.* Berlin, Heidelberg: Springer Handbooks; 2008. P. 259–280. https://doi.org/10.1007/978-3-540-49127-9_13
24. Benesty J., Paleologu C., Dogariu L.-M., Ciochină S. Identification of linear and bilinear systems: a unified study. *Electronics*. 2021;10(15):1790. https://doi.org/10.3390/electronics10151790

25. Bretthauer G., Gamaleja T., Wilfert H.-H. Identification of parametric and nonparametric models for MIMO closed loop systems by the correlation method. *IFAC Proceedings Volumes*. 1984;17(2):753–758. https://doi.org/10.1016/S1474-6670(17)61062-0
26. Lomov A.A. On quantitative a priori measures of identifiability of coefficients of linear dynamic systems. *J. Comput. Syst. Sci. Int.* 2011;50:1–13. https://doi.org/10.1134/S106423071101014X
27. Krasovskii A.A. Two-channel automatic regulation systems with antisymmetric cross connections. *Autom. Remote Control*. 1957;18(2):139–150.
[Original Russian Text: Krasovskii A.A. Two-channel automatic regulation systems with antisymmetric cross connections. *Avtomatika i Telemekhanika*. 1957;18(2):126–136 (in Russ.).]
28. Karabutov N.N. On adaptive identification of systems having multiple nonlinearities. *Russ. Technol. J.* 2023;11(5):94−105 (in Russ.). https://doi.org/10.32362/2500-316X-2023-11-5-94-10
29. Karabutov N.N. *Adaptivnaya identifikatsiya system* (*Adaptive Systems Identification*). Moscow: URSS; 2007. 384 p. (in Russ.).
30. Karabutov N. Structural identifiability of systems with multiple nonlinearities. *Contemp. Math*. 2021;2(2):140–161. https://doi.org/10.37256/cm.222021763
31. Barskii A.G. *K teorii dvumernykh i trekhmernykh sistem avtomaticheskoyu regulirovaniya* (*On the Theory of Two-Dimensional and Three-Dimensional Automatic Control Systems*). Moscow: Logos; 2015. 192 p. (in Russ.).
32. Skorospeshkin M.V. Adaptive two-channel correction device for automatic control systems. *Izvestiya Tomskogo politekhnicheskogo universiteta = Bulletin of the Tomsk Polytechnic University*. 2008;312(5):52–57 (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Морозовский В.Т. *Многосвязные системы автоматического регулирования*. М.: Энергия; 1970. 288 с.
2. Зырянов Г.В. *Системы управления многосвязными объектами*: учебное пособие. Челябинск: Издательский центр ЮУрГУ; 2010. 112 с.
3. Мееров М.В., Литвак Б.Л. *Оптимизация систем многосвязного управления*. М.: Наука; 1972. 344 с.
4. Буков В.Н., Максименко И.М., Рябченко В.Н. Регулирование многосвязных систем. *Автоматика и телемеханика*. 1998;6:97–110.
5. Воронов А.А. *Введение в динамику сложных управляемых систем*. М.: Наука; 1985. 352 с.
6. Егоров И.Н., Умнов В.П. *Системы управления электроприводов технологических роботов и манипуляторов:* учебное пособие. Владимир: Изд-во ВлГУ; 2022. 314 с.
7. Егоров И.Н. *Позиционно-силовое управление робототехническими и мехатронными устройствами*. Владимир: Изд-во ВлГУ; 2010. 192 с.
8. Gupta N., Chopra N. Stability analysis of a two-channel feedback networked control system. In: *2016 Indian Control Conference* (*ICC*). 2016. https://doi.org/10.1109/INDIANCC.2016.7441129
9. Pawlak A., Hasiewicz Z. Non-parametric identification of multi-channel systems by multiscale expansions. In: *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2011. https://doi.org/10.1109/ICASSP.2002.5744953
10. Kholmatov U. The possibility of applying the theory of adaptive identification to automate multi-connected objects. *Am. J. Eng. Technol*. 2022;4(03):31–38. URL: https://inlibrary.uz/index.php/tajet/article/view/5789
11. Aliyeva A.S. Identification of multiconnected dynamic objects with uncertainty based on neural technology and reference converters. *Informatics and Control Problems.* 2019;39(2):93–102. URL: https://icp.az/2019/2-11.pdf
12. Hua C., Guan X., Shi P. Decentralized robust model reference adaptive control for interconnected time-delay systems. In: *Proceeding of the 2004 American Control Conference.* Boston, Massachusetts June 30 – July 2, 2004. 2004. P. 4285–4289. https://doi.org/10.23919/ACC.2004.1383981
13. Ворчик Б.Г. Идентифицируемость многосвязной замкнутой стохастической системы. Декомпозиция замкнутой системы при идентификации. *Автоматика и телемеханика*. 1977;2:14–28.
14. Glentis G.-O., Slump C.H. A highly modular normalized adaptive lattice algorithm for multichannel least squares filtering. In: *1995 International Conference on Acoustics, Speech, and Signal Processing.* 1995;2:1420–1423. https://doi.ieeecomputersociety.org/10.1109/ICASSP.1995.480508
15. Ali M., Abbas H., Chughtai S.S., Werner H. Identification of spatially interconnected systems using neural network. In: *49th IEEE Conference on Decision and Control* (*CDC*). 2011. https://doi.org/10.1109/CDC.2010.5717080
16. Yang Q., Zhu M., Jiang T., He J., Yuan J., Han J. Decentralized Robust Adaptive output feedback stabilization for interconnected nonlinear systems with uncertainties. *J. Control Sci. Eng*. 2016;2016:article ID 3656578. https://doi.org/10.1155/2016/3656578
17. Wu H. Decentralized adaptive robust control of uncertain large-scale non-linear dynamical systems with time-varying delays. *IET Control Theory & Applications*. 2012;6(5):629–640. https://doi.org/10.1049/iet-cta.2011.0015
18. Fan H., Han L., Wen C., Xu L. Decentralized adaptive output-feedback controller design for stochastic nonlinear interconnected systems. *Automatica.* 2012;48(11):2866–2873. https://doi.org/10.1016%2Fj.automatica.2012.08.022
19. Ali M., Chughtai S.S., Werner H. Identification of spatially interconnected systems. In: *Proceedings of the 48h IEEE Conference on Decision and Control* (*CDC*) *held jointly with 2009 28th Chinese Control Conference*. 2010. https://doi.org/10.1109/CDC.2009.5399748

20. Ioannou P.A. Decentralized adaptive control of interconnected systems. *IEEE Transactions on Automatic Control* 1986;31(4):291–298. https://doi.org/10.1109/TAC.1986.1104282
21. Sanandaji B.M., Vincent T.L., Wakin M.B. A review of sufficient conditions for structure identification in interconnected systems. *IFAC Proceedings Volumes*. 2012;45(16):1623–1628. https://doi.org/10.3182/20120711-3-BE-2027.00254
22. Soverini U., Söderström T. Blind identification of two-channel FIR systems: a frequency domain approach. *IFAC-PapersOnLine.* 2020;53(2):914–920. https://doi.org/10.1016/j.ifacol.2020.12.855
23. Huang Y., Benesty J., Chen J. Adaptive blind multichannel identification. In: Benesty J., Sondhi M.M., Huang Y.A. (Eds.). *Springer Handbook of Speech Processing.* Berlin, Heidelberg: Springer Handbooks; 2008. P. 259–280. https://doi.org/10.1007/978-3-540-49127-9_13
24. Benesty J., Paleologu C., Dogariu L.-M., Ciochină S. Identification of linear and bilinear systems: a unified study. *Electronics*. 2021;10(15):1790. https://doi.org/10.3390/electronics10151790
25. Bretthauer G., Gamaleja T., Wilfert H.-H. Identification of parametric and nonparametric models for MIMO closed loop systems by the correlation method. *IFAC Proceedings Volumes*. 1984;17(2):753–758. https://doi.org/10.1016/S1474-6670(17)61062-0
26. Lomov A.A. On quantitative a priori measures of identifiability of coefficients of linear dynamic systems. *J. Comput. Syst. Sci. Int.* 2011;50:1–13. https://doi.org/10.1134/S106423071101014X
27. Красовский А.А. О двухканальных системах автоматического регулирования с антисимметричными связями. *Автоматика и телемеханика*. 1957;18(2):126–136.
28. Карабутов Н.Н. Об адаптивной идентификации систем с несколькими нелинейностями. *Russ. Technol. J.* 2023;11(5):94−105. https://doi.org/10.32362/2500-316X-2023-11-5-94-10
29. Карабутов Н.Н. *Адаптивная идентификация систем*. М.: УРСС; 2007. 384 с.
30. Karabutov N. Structural identifiability of systems with multiple nonlinearities. *Contemp. Math*. 2021;2(2):140–161. https://doi.org/10.37256/cm.222021763
31. Барский А.Г. *К теории двумерных и трехмерных систем автоматическою регулирования*. М.: Логос; 2015. 192 с.
32. Скороспешкин М.В. Адаптивное двухканальное корректирующее устройство для систем автоматического регулирования. *Известия Томского политехнического университета*. 2008;312(5):52–57.

About the author

Nikolay N. Karabutov, Dr. Sci. (Eng.), Professor, Department of Problems Control, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: karabutov@mirea.ru. Scopus Author ID 6603372930, ResearcherID P-5683-2015, RSCI SPIN-code 9646-9721, https://orcid.org/0000-0002-3706-7431

Об авторе

Карабутов Николай Николаевич, д.т.н., профессор, кафедра проблем управления, Институт искусственного интеллекта, ФГБОУ ВО МИРЭА – Российский технологический университет (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: karabutov@mirea.ru. Scopus Author ID 6603372930, ResearcherID P-5683-2015, SPIN-код РИНЦ 9646-9721, https://orcid.org/0000-0002-3706-7431

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 621.372.8
https://doi.org/10.32362/2500-316X-2024-12-5-77-89
EDN GTXQII

RESEARCH ARTICLE

Models of symmetric three-layer waveguide structures with graded-index core and nonlinear optical liners

Sergey E. Savotchenko @

MIREA – Russian Technological University, Moscow, 119454 Russia
@ *Corresponding author, e-mail: savotchenkose@mail.ru*

Abstract

Objectives. Determining the patterns of dispersion properties of waveguide modes of the optical range in layered media with distributed optical properties is a both a pressing and significant matter for study. It has fundamental and applied importance in nonlinear optics and optoelectronics. The combination of a nonlinear response and graded-index distributions of the optical properties of adjacent layers of a layered structure enables the desired values of the output characteristics using a wide range of control parameters to be selected easily. This renders such waveguides the most promising from the point of view of possible technical applications. The aim of this paper is to develop the theory of three-layer planar waveguide structures with a graded-index core and nonlinear optical liners with arbitrary profiles. By doing so it may be possible to find exact analytical solutions to nonlinear stationary wave equations describing explicitly the transverse electric field distribution of waveguide modes.

Methods. The analytical methods of mathematical physics and the theory of special functions applied to nonlinear and waveguide optics are used herein.

Results. The study provides a theoretical description of transverse stationary waves propagating along a symmetrical three-layer planar waveguide structure consisting of the inner graded-index layer sandwiched between nonlinear optical plates. It assumes an arbitrary spatial profile of the interlayer dielectric constant and the nature of the nonlinear response of the liner medium. The mathematical model of this waveguide structure formulated herein is based on nonlinear equations with distributed coefficients. The solutions obtained describe in general terms the transverse distribution of the amplitude of the electric field envelope. The transverse symmetry of the three-layer waveguide structure enables even and odd stationary modes corresponding to symmetric and antisymmetric transverse field profiles to be excited in it. A method was developed for constructing even (symmetric) and odd (antisymmetric) solutions which exist at certain discrete values of the effective refractive index/propagation constant. These discrete spectra were obtained in layers with graded-index linear, parabolic, and exponential profiles. The symmetrical three-layer waveguide structure with inner graded-index layer characterized by parabolic spatial profile and outer liners as Kerr nonlinear optical media is analyzed in detail, as an example of the application of the formulated theory. Analysis of the resulting exact analytical solution indicates that the electric field strength for the fundamental and first-order modes increases with increasing parabolic profile parameter, characterizing the relative change of the dielectric constant in the interlayer, while decreasing for higher order modes.

Conclusions. The theory developed in this paper supports the unambiguous description of the transverse distributions of the stationary electric field in planar symmetrical three-layer waveguides in an explicit analytical form. The results extend the understanding of the physical properties of nonlinear waves and the localization patterns of light beams in distributed media, and may be useful in the design of various optical waveguide devices.

Keywords: layered structure, layered waveguide, optical waveguide, nonlinear optics, optical nonlinearity, graded-index layer, nonlinear waves, Kerr nonlinear optical media, guided waves, waveguide mode

• **Submitted:** 01.03.2024 • **Revised:** 29.03.2024 • **Accepted:** 05.08.2024

For citation: Savotchenko S.E. Models of symmetric three-layer waveguide structures with graded-index core and nonlinear optical liners. *Russ. Technol. J*. 2024;12(5):77–89. https://doi.org/10.32362/2500-316X-2024-12-5-77-89

Financial disclosure: The author has no financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Модели симметричных трехслойных волноводных структур с градиентной сердцевиной и нелинейно-оптическими обкладками

С.Е. Савотченко @

МИРЭА – Российский технологический университет, Москва, 119454 Россия
@ *Автор для переписки, e-mail: savotchenkose@mail.ru*

Резюме

Цели. Выявление закономерностей дисперсионных свойств волноводных мод оптического диапазона в слоистых средах с распределенными оптическими характеристиками представляет собой актуальную и важную задачу, имеющую фундаментальное и прикладное значение в нелинейной оптике и оптоэлектронике. Сочетание нелинейного отклика и градиентных распределений оптических свойств соседних слоев слоистой структуры дает возможность легко подобрать требуемые значения выходных характеристик с помощью широкого ряда управляющих параметров, что делает такие волноводы наиболее перспективными с точки зрения возможных технических приложений. Цель работы – развитие теории трехслойных плоских волноводных структур с градиентной сердцевиной и нелинейно-оптическими обкладками с произвольными профилями, в рамках которой представляется возможным нахождение точных аналитических решений нелинейных стационарных волновых уравнений, описывающих в явном виде поперечное распределение электрического поля волноводных мод.

Методы. Использованы аналитические методы математической физики и теории специальных функций применительно к нелинейной и волноводной оптике.

Результаты. Проведено теоретическое описание поперечных стационарных волн, распространяющихся вдоль плоской симметричной трехслойной волноводной структуры, состоящей из внутреннего градиентного слоя, зажатого между нелинейно-оптическими обкладками, причем пространственный профиль диэлектрической проницаемости прослойки и вид нелинейного отклика среды обкладок предполагаются произвольными. Сформулирована математическая модель такой волноводной структуры на основе нелинейных уравнений с распределенными коэффициентами. Получены решения, описывающие в общем виде поперечное распределение амплитуды огибающей электрического поля. В силу поперечной симметрии трехслойной волноводной структуры в ней могут возбуждаться четные и нечетные стационарные моды, соответствующие симметричным и антисимметричным поперечным профилям поля. Разработан метод построения четных (симметричных) и нечетных (антисимметричных) решений, существующих при определенных дискретных значениях эффективного показателя преломления / константы распространения. Такие дискретные спектры получены в слоях с градиентными линейным, параболическим и экспоненциальным профилями. В качестве примера применения сформулированной теории детально проанализирован случай симметричной трехслойной волноводной структуры, внутренний градиентный слой которой характеризуется параболическим пространственным профилем, а внешние обкладки представляют собой керровские нелинейно-оптические среды. На основе анализа полученного точного аналитического решения установлено, что напряженность электрического поля для основной моды и моды первого порядка увеличивается с ростом параметра параболического профиля, характеризующего относительное изменение диэлектрической проницаемости в прослойке, однако уменьшается для мод более высоких порядков.

Выводы. Развитая в данной работе теория позволяет наглядно описать в явном аналитическом виде поперечные распределения стационарного электрического поля в плоских симметричных трехслойных волноводах. Полученные результаты расширяют представления о физических свойствах нелинейных волн и закономерностях локализации световых пучков в распределенных средах и могут быть полезными для разработки различных оптических волноводных устройств.

Ключевые слова: слоистая структура, слоистый волновод, оптический волновод, нелинейная оптика, оптическая нелинейность, градиентный слой, нелинейные волны, керровские нелинейно-оптические среды, управляемые волны, волноводная мода

• **Поступила:** 01.03.2024 • **Доработана:** 29.03.2024 • **Принята к опубликованию:** 05.08.2024

Для цитирования: Савотченко С.Е. Модели симметричных трехслойных волноводных структур с градиентной сердцевиной и нелинейно-оптическими обкладками. *Russ. Technol. J*. 2024;12(5):77–89. https://doi.org/10.32362/2500-316X-2024-12-5-77-89

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

The development of optical waveguides with desired characteristics is an important applied problem in the area of nonlinear optics [1–3]. The successful resolution of this issue requires theoretical modeling of the designed systems which will enable their properties to be described, characteristics to be predicted and the development process optimized. Thus, much attention is paid in scientific literature to the development of theoretical foundations for modeling the processes of excitation, propagation, and localization of electromagnetic waves in a variety of optical media [4, 5].

The requisite and often unique characteristics of fields in waveguide structures can be obtained most effectively in a combination of media with different optical properties [6, 7]. Classes of media where optical properties depend significantly on the spatial distribution of the refractive index (or dielectric permittivity) [8] or can be characterized by a nonlinear optical response in which the dielectric permittivity depends on the electric field intensity [9] are in particular considered promising and possessed of a wide variety of properties. The first group of media is called graded-index [10] while the second group is called nonlinear [11].

The dependence of optical characteristics on the quantities mentioned above may differ, and may be determined by the physical properties of real crystals. In particular, the most common form of nonlinear response is the linear dependence of dielectric permittivity on the square of amplitude (intensity) of the electric field (light), referred to as Kerr nonlinearity [12]. Waves and other localized disturbances in such media have been quite well studied in various modifications [13, 14]. This includes analytical methods [15, 16] which assume that exact solutions be obtained for the nonlinear wave equation used in various models [17, 18].

In order to describe the experimentally dependencies of the spatial distribution of optical characteristics as described above, a variety of functions (profiles) are used to model the change of the refractive index with distance from the optical media interface [19]. Certain profiles, such as linear [20], parabolic [21], exponential [22], and others [23, 24], allow exact analytical solutions to be established.

The theoretical study of waveguide properties of interfaces between graded-index and nonlinear media has in recent times intensified [25, 26]. In particular, solutions have been developed which describe the localization of light along the interface between the nonlinear Kerr medium and medium with linear [27, 28] and exponential refractive index profiles [29].

In terms of technical application, the studies on waveguide properties of multilayer media [30] including three-layer structures [31] are of great importance. Nonlinear waves in three-layer structures have been a focus of theoretical study for many years [32, 33], including in layered graded-index media [34]. In recent times, analytical solutions have been obtained for symmetric three-layer structures in which the inner layer is described by a symmetric linear graded-index profile. The outer layers are characterized by photorefractive nonlinear response [35], Kerr nonlinearity [36], and step nonlinearity [37]. The symmetric structure with a parabolic graded-index inner layer placed between media with Kerr nonlinearity has also been considered.

Due to the emerging variety of possible combinations of nonlinearities and graded-index layer profiles, it would be useful to construct a generalized model of a symmetric waveguide structure. This paper proposes a generalization of the model for the three-layer symmetric planar structure, in which the inner layer and the outer layers are characterized by an arbitrary graded-index profile and a nonlinear optical

response, respectively. Substituting the particular type of dielectric permittivity profiles and the shape of nonlinear response into model equations allows analytical solutions to be obtained which describe the amplitude spatial distribution of the envelope perpendicular to structure layers. The resulting analytical expressions, in turn, allow localization patterns of light beams to be determined in layered waveguide structures.

1. THEORETICAL MODELING OF THREE-LAYER WAVEGUIDE STRUCTURE

1.1. Model formulation

We consider a three-layer planar structure which is symmetric about the center. It is made of nonmagnetic materials with optically homogeneous properties in the longitudinal direction. The interfaces between layers are assumed to be planar. We place the origin of the coordinates in the middle of the inner layer (core or interlayer), in the *yz* plane; with the *x*-axis perpendicular to the planes of interfaces and the *z*-axis along the layers in the direction of wave propagation. Let layer interfaces be located in planes $x = \pm a$ (then the thickness of the layer is assumed to be $2a$). The media in all layers are considered with no allowance for dielectric losses.

In the model considered herein, the inner layer is characterized by spatial inhomogeneity of optical properties in the direction transverse to the plane of the layers (graded-index layer), while the outer adjacent layers (liners) are characterized by optical nonlinearity (nonlinear layers), i.e., by the dependence of the refractive index (or dielectric permittivity) on the light intensity. In this case, the interlayer thickness is considered to be much less than the thicknesses of the outer liners. Therefore, when studying the distribution of the electric field localized near the core, the liners can be considered to be semi-limited media, neglecting the influence of boundaries located at a further distance when compared to the value of a. This consideration is acceptable provided that the field rapidly decreases at a distance from interfaces and becomes negligible before reaching the outer boundaries of thick liners.

Let a transverse electric wave (TE wave) propagate along interfaces of a three-layer waveguide structure whose electric field strength component can be written in the following form:

$$E_y(x,z) = \psi(x)e^{i(\beta z-\omega t)}, \tag{1}$$

wherein $\psi(x)$ is spatial distribution of electric field strength in the transverse layer direction (envelope amplitude), ω is frequency, $\beta = kn$ is the propagation constant, $n = ck/\omega$ is effective refractive index, c is the speed of light in vacuum, $k = 2\pi/\lambda$ is wave number, λ is wavelength, and t is time.

It is known [4, 5] that field $\psi(x)$ is defined as the solution for the stationary equation (magnetic permeability is equal to unity):

$$\frac{d^2\psi(x)}{dx^2} + \{\varepsilon(x,I) - n^2\}k^2\psi(x) = 0, \tag{2}$$

wherein dielectric permittivity of the three-layer waveguide structure can be written as follows:

$$\varepsilon(x,I) = \begin{cases} \varepsilon_G(x), & |x| < a, \\ \varepsilon_N(I), & |x| > a, \end{cases} \tag{3}$$

wherein function $\varepsilon_G(x)$ defines the dependence of the dielectric permittivity on the spatial coordinate in the direction perpendicular to layers (dielectric permittivity of the graded-index layer), while function $\varepsilon_N(I)$ defines the dependence of the dielectric permittivity on the light intensity $I = |E|^2$, wherein E stands for the amplitude of the electric field strength (dielectric permittivity of the nonlinear layers).

We can represent the transverse field distribution in the following form:

$$\psi(x) = \begin{cases} \psi_N^{(-)}(x), & x < -a, \\ \psi_G(x), & |x| < a, \\ \psi_N^{(+)}(x), & x > a, \end{cases} \tag{4}$$

Then the following equations may be derived from (2):

$$\frac{d^2\psi_N^{(-)}(x)}{dx^2} + \{\varepsilon_N(I) - n^2\}k^2\psi_N^{(-)}(x) = 0, \quad x < -a, \tag{5}$$

$$\frac{d^2\psi_G(x)}{dx^2} + \{\varepsilon_G(x) - n^2\}k^2\psi_G(x) = 0, \quad |x| < a, \tag{6}$$

$$\frac{d^2\psi_N^{(+)}(x)}{dx^2} + \{\varepsilon_N(I) - n^2\}k^2\psi_N^{(+)}(x) = 0, \quad x > a, \tag{7}$$

which are supplemented by boundary conditions corresponding to the requirements of continuity of field components at layer interfaces:

$$\begin{gathered} \psi_N(\pm a \pm 0) = \psi_G^{(\pm)}(\pm a \pm 0), \\ \frac{d\psi_N(\pm a \pm 0)}{dx} = \frac{d\psi_G^{(\pm)}(\pm a \pm 0)}{dx}, \end{gathered} \tag{8}$$

as well as to the field vanishing condition at infinity: $|\psi_N^{(\pm)}(x)| \to 0, |x| \to \infty$.

In terms of physics, the requirement of limitation of the solution should be considered an obvious supplement. Thus, the set of Eqs. (5)–(7) and boundary conditions (8) represent a mathematical formulation of the proposed model for the three-layer waveguide structure with a dielectric permittivity profile described by means of the distributed expression (3).

1.2. The dispersion equation in the general case

In the inner layer, the solution to Eq. (6) can be represented as follows:

$$\psi_G(x) = C_1 F_1(x) + C_2 F_2(x), \quad (9)$$

wherein $C_{1,2}$ are the values depending on optical and geometrical parameters of the system and determined by boundary conditions (8). $F_{1,2}(x)$ are special functions which are linearly independent solutions to Eq. (6) at the given dielectric permittivity profile $\varepsilon_G(x)$. Since (6) is a linear homogeneous differential equation of the second order with a coefficient depending on variable x, its solutions are often expressed in the most general form through hypergeometric functions. In certain types of dielectric permittivity profiles, the solutions can be expressed through other pairs of linearly independent special functions, such as Bessel functions, Airy functions, and others. The main requirements for $F_{1,2}(x)$ are continuity and limitation on the interval $-a < x < a$ of these functions, as well as their derivatives $F'_{1,2}(x)$.

In the outer liners, the solutions to Eqs. (5) and (7) can be represented as follows:

$$\psi_N^{(\pm)}(x) = \psi_a^{(\pm)} \frac{\Psi_N^{(\pm)}(x)}{\Psi_N^{(\pm)}(\pm a)}, \quad (10)$$

wherein $\psi_a^{(\pm)}$ is field amplitudes at layer interfaces in planes $x = \pm a$. $\Psi_N^{(\pm)}(x)$ are solutions to nonlinear Eqs. (5) and (7) are limited in regions $x < -a$ and $x > a$, respectively, while satisfying requirement $|\Psi_N^{(\pm)}(x)| \to 0, |x| \to \infty$. The explicit form of functions $\Psi_N^{(\pm)}(x)$ is determined by the type of nonlinearity model of the outer liner medium. For example, for the most common Kerr nonlinearity, $\Psi_N^{(\pm)}(x)$ are expressed through hyperbolic functions depending on the nonlinear response sign.

Substituting solutions (9) and (10) into boundary conditions (8) results in the following system of algebraic equations for values $C_{1,2}$ and amplitudes $\psi_a^{(\pm)}$:

$$\begin{cases} \psi_a^{(\pm)} = C_1 F_1(\pm a) + C_2 F_2(\pm a), \\ \psi_a^{(\pm)} \varepsilon_{Neff}^{(\pm)} = C_1 F_1'(\pm a) + C_2 F_2'(\pm a), \end{cases} \quad (11)$$

where we denote

$$\varepsilon_{Neff}^{(\pm)} = \frac{1}{\Psi_N^{(\pm)}(\pm a)} \cdot \frac{d\Psi_N^{(\pm)}(\pm a)}{dx}. \quad (12)$$

The solvability condition of the system (11) allows the dispersion equation to be obtained which determined the values of the propagation constant for waveguide modes of the considered three-layer structure in the general case:

$$\Delta_1^{(+)} \Delta_2^{(-)} = \Delta_1^{(-)} \Delta_2^{(+)}, \quad (13)$$

wherein

$$\Delta_{1,2}^{(\pm)} = F_{1,2}(\pm a) \varepsilon_{Neff}^{(\pm)} - F'_{1,2}(\pm a). \quad (14)$$

This dispersion equation defines the relationship between propagation constant β, wave number k, the optical characteristics of the layers (unperturbed values of dialectic constants and parameters of dielectric permittivity dependencies in graded-index and nonlinear layers as determined by a certain type of model). The geometric parameter of the three-layer structure is considered the half-width of interlayer a.

The amplitude at one interface can be chosen as an independent characteristic, through which other parameters of solutions (9) and (10) can be expressed. In particular, the following ratio of amplitudes at the left and right interlayer boundaries may be derived from (11):

$$\frac{\psi_a^{(+)}}{\psi_a^{(-)}} = \frac{\Delta_1^{(-)} \Delta^{(+)}}{\Delta_1^{(+)} \Delta^{(-)}}, \quad (15)$$

wherein

$$\Delta^{(\pm)} = F_1(\pm a) F_2'(\pm a) - F_2(\pm a) F_1'(\pm a). \quad (16)$$

Then values $C_{1,2}$ can be written in the following form:

$$C_{1,2} = \psi_a^{(+)} \frac{\Delta_{1,2}^{(+)}}{\Delta^{(+)}} = \psi_a^{(-)} \frac{\Delta_{1,2}^{(-)}}{\Delta^{(-)}}. \quad (17)$$

Taking into account (15), field distribution in the inner layer can be rewritten in the following form:

$$\psi_G(x) = \frac{\psi_a^{(\pm)}}{\Delta^{(\pm)}}\{\Delta_1^{(\pm)}F_1(x) + \Delta_2^{(\pm)}F_2(x)\}. \quad (18)$$

Thus, the resulting expressions (10) and (18) determine the field distribution in the transverse layer direction. The parameters thereof are determined by expressions (12), (14), and (16) while the propagation constant is determined by dispersion Eq. (13).

1.3. Constructing waveguide modes of a given symmetry and discrete spectrum

Due to the symmetry of the considered three-layer waveguide structure, even and odd modes should clearly exist therein. They are described by symmetric and antisymmetric field distributions in the transverse direction, respectively. The solution to problem (5)–(8) can then be searched for on semiaxis $x > 0$. We continue it in the even or odd direction for symmetric or antisymmetric modes, respectively.

For symmetric distribution, the solutions should be even functions: $\psi_N^{(+)}(-x) = \psi_N^{(-)}(x)$ and $\psi_G(-x) = \psi_G(x)$, while for antisymmetric distribution, they should be odd: $\psi_N^{(+)}(-x) = -\psi_N^{(-)}(x)$ and $\psi_G(-x) = -\psi_G(x)$. The upper indices (±) can be omitted due to the given symmetry chosen.

The mode with given symmetry can be described by the following solution:

$$\psi_G(x) = \psi_a \frac{F_G(g(x))}{F_G(g(a))}, \quad (19)$$

where $F_G(g)$ is the special function selected in a certain manner to resolve Eq. (6) on semiaxis $x > 0$, with internal argument $g(x)$. The explicit form thereof is related to the spatial dependence profile type of the inner layer dielectric permittivity. The argument g contains propagation constant β, as well as optical and geometric parameters of the waveguide system.

For the symmetric mode, function F_G should have an extremum at the symmetry center of the three-layer waveguide structure (at $x = 0$). This implies that the derivative of function F_G should go to zero at $x = 0$. For the antisymmetric mode, function F_G should go to zero itself. Due to the necessity for symmetry of the desired field profile, these requirements result in the spectrum of propagation constant values (or effective refractive index) becoming discrete.

In particular, it should be as follows for the symmetric mode:

$$\left.\frac{dF_G(g(x))}{dx}\right|_{x=0} = F_G'(g(0))g'(0) = 0. \quad (20)$$

Thus, if $g'(0) \neq 0$, then

$$g(0) = \xi_j \ (j = 1, 2, \ldots), \quad (21)$$

wherein ξ_j are zeros of the derivative of special function F_G. Since argument g contains propagation constant β, resolving Eq. (21) with respect to it allows a discrete spectrum of its values $\beta = \beta(\xi_j)$ to be obtained. This is determined by the sequence of zeros of the derivative of special function F_G solving Eq. (6).

Similarly, it should be as follows for antisymmetric mode:

$$F_G(g(0)) = 0. \quad (22)$$

Hence,

$$g(0) = \zeta_j \ (j = 1, 2, \ldots), \quad (23)$$

wherein ζ_j are zeros of special function F_G. Solving Eq. (23) with respect to the propagation constant allows a discrete spectrum of its values $\beta = \beta(\zeta_j)$ to be obtained. This is determined by the sequence of zeros of special function F_G solving Eq. (6).

The solution to nonlinear Eq. (7) in the outer layer at $x > a$ can be represented in the following form:

$$\psi_N(x) = \psi_a \frac{\Psi_N(q_N(x - a - x_N))}{\Psi_N(q_N x_N)}, \quad (24)$$

where the dependence of coefficient q_N on the system parameters is known for a certain nonlinear response model, while value x_N is determined by the boundary conditions. Function Ψ_N should have parity coinciding with that of function F_G.

It should be noted that the solutions chosen in the form of (19) and (24) automatically satisfy the continuity conditions at interfaces of waveguide structure layers at $x = \pm a$.

In order to meet the continuity condition of the field derivative at the interface between the graded-index and nonlinear layers, (19) and (24) should be substituted into (8) at $x = a$ whence the following equation is derived:

$$\frac{F_G'(g(a))g'(a)}{F_G(g(a))} = \frac{q_N\Psi_N'(q_N x_N)}{\Psi_N(q_N x_N)}, \quad (25)$$

which allows value x_N to be determined depending on the optical and geometrical characteristics of the layered structure.

This shows the possible existence of waveguide modes of a given symmetry in the case when the propagation constant is given by a discrete spectrum of values.

2. RESULTS AND DISCUSSION

2.1. Some analytically solvable profiles of the graded-index layer

First, we consider the types of symmetric dielectric permittivity profiles in the graded-index layer for which exact analytical solutions are known. For comparison, we also consider the case of a step structure with the inner layer characterized by the constant dielectric permittivity value, independent of spatial coordinate (Fig. 1a): $\varepsilon_G = \varepsilon_0$ – const. Then the solution to Eq. (6) is determined by trigonometric functions and has the following form for even modes:

$$\psi_G(x) = \psi_a \frac{\cos(px)}{\cos(pa)}, \quad (26)$$

while for odd modes:

$$\psi_G(x) = \psi_a \frac{\sin(px)}{\sin(pa)}, \quad (27)$$

wherein $p^2 = k^2(\varepsilon_0 - n^2)$. These modes exist for values of effective refractive index $n^2 < \varepsilon_0$.

Symmetric graded-index profiles:

1) Linear (Fig. 1b):

$$\varepsilon_G(x) = \varepsilon_0\left(1 - \Delta\frac{|x|}{a}\right), \quad (28)$$

wherein ε_0 is the dielectric permittivity at the center of the waveguide structure symmetry, and $\Delta = (\varepsilon_0 - \varepsilon_a)/\varepsilon_0$ is the change in dielectric permittivity from ε_0 to the value of dielectric permittivity at the interface of layers ε_a.

The solution to Eq. (6) on interval $0 < x < a$ with linear profile (28) can be written as [36]:

$$\psi_G(x) = \psi_a \frac{\mathrm{Ai}(x/x_G + \delta)}{\mathrm{Ai}(a/x_G + \delta)}, \quad (29)$$

wherein $F_G = \mathrm{Ai}(g)$ is the Airy function, $g(x) = x/x_G + \delta$,

$$\delta = -(\varepsilon_0 - n^2)(ak/\varepsilon_0\Delta)^{2/3}, \quad (30)$$

$$x_G = \left(\frac{a}{k^2\varepsilon_0\Delta}\right)^{1/3}. \quad (31)$$

When constructing the even solution (as noted in Section 1.3), there is a necessary requirement that function (29) has an extremum in the middle of the waveguide at $x = 0$. Therefore, $\delta = \xi_j$, $j = 1, 2, \ldots$, where for a linear profile, ξ_j are zeros of the derivative of the Airy function: $\mathrm{Ai}'(\xi_j) = 0$: $\xi_1 = -1.018792972$, $\xi_2 = -3.248197582$, $\xi_3 = -4.820099211$, … Then (30) yields the following discrete spectrum of effective refractive index values:

$$n_j^2 = \varepsilon_0 - |\xi_j|(\varepsilon_0\Delta/ak)^{2/3}. \quad (32)$$

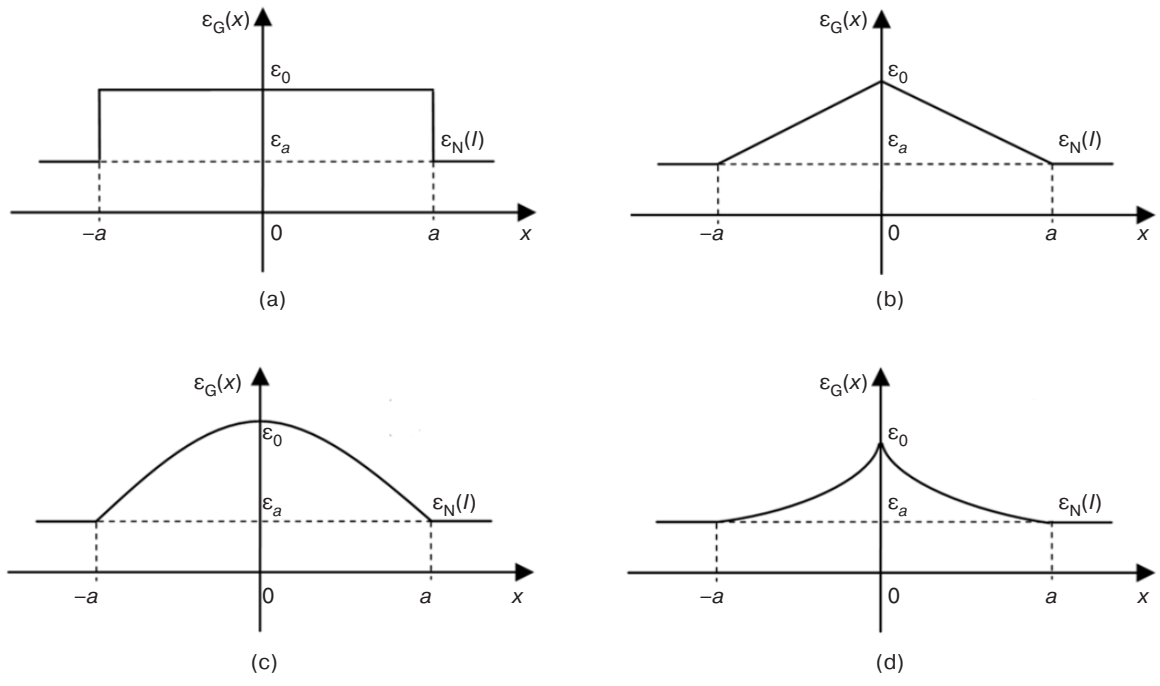


Fig. 1. Spatial symmetric profiles of dielectric permittivity of a three-layer waveguide structure: constant (a), linear (b), parabolic (c), and exponential (d)

When constructing an odd solution, there is a necessary requirement that function (29) turns to zero in the middle of the waveguide at $x = 0$. Therefore, $\delta = \zeta$ are zeros of the Airy function: $\mathrm{Ai}(\zeta_j) = 0$: $\zeta_1 = -2.338107410$, $\zeta_2 = -4.087949444$, $\zeta_3 = -5.520559828$, …

2) Parabolic (Fig. 1c):

$$\varepsilon_G(x) = \varepsilon_0\left(1 - \Delta\left(\frac{x}{a}\right)^2\right). \qquad (33)$$

For parabolic profile (33), the limited solution to Eq. (6) is known to be expressed through Hermite polynomials $H_j(x)$:

$$\psi_G(x) = \psi_a \frac{H_j(x/x_0)}{H_j(a/x_0)} e^{-(x^2-a^2)/2x_0^2}, \qquad (34)$$

wherein $j = 2m$ for even modes and $j = 2m + 1$ for odd modes, $m = 0, 1, 2, \ldots$, $x_0^2 = a / k\sqrt{\varepsilon_0\Delta}$, which exist at discrete values of effective refractive index [38]:

$$n_j^2 = \varepsilon_0 - (2j+1)\sqrt{\varepsilon_0\Delta}/ak. \qquad (35)$$

3) Exponential (Fig. 1d):

$$\varepsilon_G(x) = \varepsilon_0\{1 - \Delta(1 - e^{-|x|/a})\}. \qquad (36)$$

The solution to Eq. (6) with exponential profile (36) can be written as follows:

$$\psi_G(x) = \psi_a \frac{J_{2w}(2ve^{-|x|/2a})}{J_{2w}(2ve^{-1/2})}, \qquad (37)$$

wherein $F_G = J_{2w}(g)$ is the first order Bessel function of order $w = ak(n^2 - \varepsilon_a)^{1/2}$, $g = 2ve^{-|x|/2a}$, $v = ak(\Delta\varepsilon_0)^{1/2}$ [4].

Even modes are defined by the dispersion equation $J'_{2w}(2v) = 0$. With the roots of the equation $J'_{\xi j}(2v) = 0$ denoted by ξ_j, the following discrete spectrum of effective refractive index values is obtained:

$$n_j^2 = \varepsilon_a + (\xi_j / 2ak)^2. \qquad (38)$$

Odd modes are defined by the dispersion equation $J_{2w}(2v) = 0$. With the roots of the equation $J_{\zeta j}(2v) = 0$ denoted be ζ_j, the following discrete spectrum of effective refractive index values is obtained:

$$n_j^2 = \varepsilon_a + (\zeta_j / 2ak)^2. \qquad (39)$$

The analysis of roots of such equations and corresponding spectra is given in [39].

It should be noted that there are exact analytical Eqs. (6) for other spatial profiles of dielectric permittivity. These include a smooth step described by a hyperbolic tangent [40], symmetric Epstein profile (inverted symmetric Pöschl–Teller potential) described by a hyperbolic cosine [41]. Such rather complex solutions are expressed through a hypergeometric function. They cause difficulties for simple analysis, so they have not been considered here.

Thus, the solutions quoted above are the exact analytical solutions which describe the field distributions for three different dielectric permittivity spatial profiles.

2.2. Some analytically solvable models of outer liner nonlinearity

Now we consider some types of nonlinear medium models of the outer liners in which the dielectric constant depends on the electric field strength and for which exact analytical solutions are known. For comparison, we also consider the linear medium in which outer layers are characterized by a constant value of dielectric permittivity, independent of the field strength: $\varepsilon_N = \varepsilon_{0N}$ – const. Then the solution to Eq. (7) limited on semiaxis $x > a$ may be written as follows form:

$$\psi_N(x) = \psi_a e^{-q_N(x-a)}, \qquad (40)$$

wherein $q_N^2 = k^2(n^2 - \varepsilon_{0N})$. On the negative semiaxis, such a solution obviously continues in an even or odd way to describe symmetric and antisymmetric modes, respectively. Using conditions (8) for functions (19) and (40), the analog of dispersion Eq. (25) for a waveguide with linear liners $\gamma_G + q_N = 0$ may be derived. Here $\tilde{a}_G = F'_G(g(a))g'(a) / F_G(g(a))$, from which the spectrum of the effective refractive index values is obtained as follows:

$$n^2 = \varepsilon_{0N} + (\gamma_G / k)^2. \qquad (41)$$

By relating (41) to the discrete spectrum obtained for a particular graded-index profile of the inner layer, constraints on the mode orders excited in the interlayer of a given thickness can be obtained.

The simplest models of nonlinear media of the outer liners are the following:

1) Kerr nonlinearity:

$$\varepsilon_N(I) = \varepsilon_{0N} + \alpha I, \qquad (42)$$

where α is the Kerr nonlinearity coefficient, $I = |E|^2$ is field intensity. Then the even/odd solution to Eq. (7) with dielectric permittivity (42) at $|x| > a$ for self-focusing nonlinearity at $\alpha > 0$ has the following form:

$$\psi_N(x) = \pm\sqrt{\frac{2}{\alpha}} \cdot \frac{q}{k\mathrm{ch}(q(x \mp a \mp x_N))}. \quad (43)$$

Using conditions (8) for functions (19) and (43), we determine the following parameter value

$$x_N = -\frac{1}{q_N}\mathrm{Arth}\left(\frac{\gamma_G}{q_N}\right) \quad (44)$$

and the field amplitude at the boundary of layers

$$\psi_a = \sqrt{\frac{2}{\alpha}(n^2 - \varepsilon_{0N} - (\gamma/k)^2)}. \quad (45)$$

The discrete spectrum obtained for a particular graded-index profile of the inner layer should be substituted into expressions (43)–(45).

2) Step nonlinearity:

$$\varepsilon_N(I) = \begin{cases} \varepsilon_1, & I < I_s, \\ \varepsilon_2, & I > I_s, \end{cases} \quad (46)$$

wherein I_s is a threshold level of intensity (a known characteristic of the medium). When crossing this threshold, a sharp change from dielectric constant value ε_1 to ε_2 occurs [42, 43]. Thus, in the neighborhood of the layer interface in a nonlinear medium where $I > I_s$, a region (near-surface domain) of width x_s is formed with dielectric constant ε_2. Beyond it, further in the liners where $I < I_s$, the dielectric constant is ε_1. Such domains arise symmetrically on both sides of the inner layer [37]. The position of the boundaries of the near-surface domain is determined by x_s, found from additional field continuity requirements at domain boundaries, as follows:

$$\begin{gathered} \psi_N(\pm x_s + 0) = \psi_N(\pm x_s - 0) = I_s^{1/2}, \\ \psi'_N(\pm x_s + 0) = \psi'_N(\pm x_s - 0). \end{gathered} \quad (47)$$

In the step nonlinearity model, Eq. (7) with dielectric constant (46) decomposes into two:

$$\psi''_N(x) - (n^2 - \varepsilon_1)k^2\psi_N(x) = 0,\ I < I_s,\ |x| > x_s, \quad (48)$$

$$\psi''_N(x)(x) + (\varepsilon_2 - n^2)k^2\psi_N(x) = 0,\ I > I_s,\ a < |x| < x_s. \quad (49)$$

The solution to Eq. (48) at $n^2 > \varepsilon_1$ for even/odd modes may be written as follows:

$$\psi_N(x) = \pm I_s^{1/2} e^{\mp q_1(x \mp x_s)}, \quad (50)$$

wherein $q_1^2 = (n^2 - \varepsilon_1)k^2$, while the solution to Eq. (49) at $n^2 < \varepsilon_2$ has the following form:

$$\psi_N(x) = \pm\psi_a \cos(p_2(x \mp a) \mp \phi)/\cos(\phi), \quad (51)$$

wherein $p_2^2 = (\varepsilon_2 - n^2)k^2$ while values x_s, ϕ are determined from boundary conditions (8) and (47).

Substituting solutions (19), (50), and (51) into boundary conditions (8) and (47), parameters of even modes can be found (similar for odd modes), as follows:

$$\phi = \mathrm{arctg}\left(\frac{\gamma_G}{p_2}\right), \quad (52)$$

$$x_s = a + \frac{1}{p_2}\left\{\phi + \mathrm{arctg}\left(\frac{q_1}{p_2}\right)\right\}, \quad (53)$$

$$\psi_a = I_s^{1/2}\left(\frac{p_2^2 + q_1^2}{p_2^2 + \gamma_G^2}\right)^{1/2}. \quad (54)$$

The discrete values of effective refractive index obtained for a certain spatial profile of the inner graded-index layer should be substituted into expressions (50)–(54).

Thus, exact analytical solutions for two models of media nonlinearity are obtained.

2.3. Example of field distribution in a symmetric waveguide structure

The case when the inner graded-index layer is characterized by parabolic profile (33) and outer layers by Kerr self-focusing nonlinearity (42) may be considered as a particular example of the model of a symmetric three-layer waveguide structure.

In this structure, the spatial distribution of the electric field in the transverse layer direction is determined by expressions (34) and (43). The discrete spectrum of effective refractive index values is determined by expression (35) [38]. Limited by considering even modes for which $j = 2m$, $m = 0, 1, 2, \ldots$, the field distribution symmetric about the waveguide structure center can be written in the following form:

$$\psi(x) = \psi_0 \begin{cases} \sqrt{(n_{pm}^2 - n_{Geff}^2)}e^{(a^2 - x^2)/2x_0^2}\dfrac{H_{2m}(x/x_0)}{H_{2m}(a/x_0)}, \\ |x| < a, \\ \dfrac{n_{pm}}{\cosh(kn_{pm}(x \mp a \mp x_N))}, \\ |x| > a, \end{cases} \quad (55)$$

where $n_{pm}^2 = \varepsilon_0 - \varepsilon_{0N} - (4m+1)\sqrt{\varepsilon_0 \Delta}/ak$ and

$$n_{\text{Geff}} = \frac{1}{kx_0}\left\{\frac{H'_{2m}(a/x_0)}{H_{2m}(a/x_0)} - \frac{a}{x_0}\right\}. \qquad (56)$$

Thus, expression (55) represents an analytical solution to problem (5)–(8), when selecting symmetry, the parabolic profile of inner layer (33), and the self-focusing nonlinearity of liners (42).

The characteristic profiles of the solution (55) are shown in Fig. 2. It illustrates the impact of the relative change of dielectric constant in interlayer Δ on the electric field distribution at fixed values of other waveguide parameters for the main mode at $m = 0$ (Fig. 2a): even first-order mode $m = 1$ (Fig. 2b), and even second-order mode $m = 2$ (Fig. 2c).

The results show that the electric field strength increases with growing value of Δ for the main (Fig. 2a) and first (Fig. 2b) modes. However, the intensity decreases with increasing Δ for higher-order modes (Fig. 2c). Thus, the dependence of the field strength on Δ is not monotonic. Increasing the thickness of interlayer a gives the same effect.

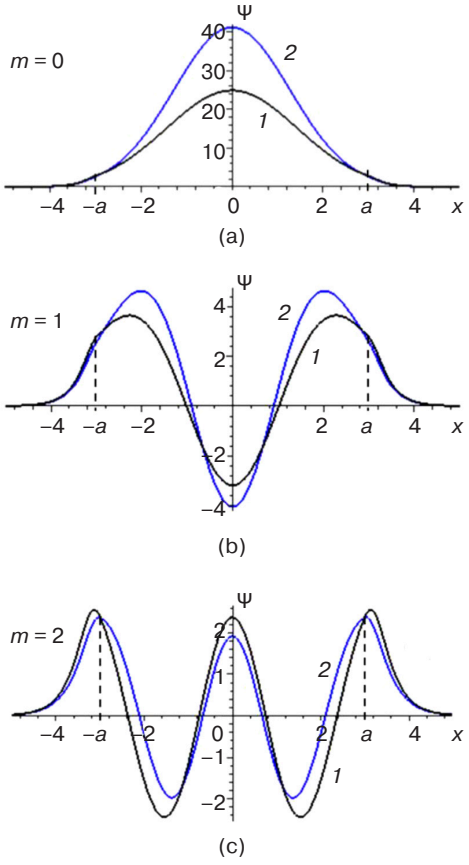


Fig. 2. Field distribution (55) at parameter values (in conventional units): $k = 0.65$, $\alpha = 6$, $\varepsilon_{0N} = 0.1$, $\varepsilon_0 = 30$, $a = 3$, and various $\Delta = 5$ (line 1), $\Delta = 8$ (line 2) for the first three even modes: $m = 0$ (a), $m = 1$ (b), $m = 2$ (c)

The existence of a discrete spectrum of effective refractive index values indicates the ability of waveguide modes to propagate at certain values of parabolic profile parameters. The propagation constant is related to the incidence angle of the beam exciting the waveguide mode. Thus, a discrete set of incidence angles should be taken into account since a waveguide mode of a certain order in the considered system can be excited only at a certain incidence angle. This depends on the dielectric permittivity at layer boundaries and the change in dielectric permittivity inside the graded-index layer.

Thus, the study obtained and analyzed the exact analytical solution which describes the field distribution in a symmetric three-layer waveguide structure where the inner graded-index layer and outer layers are characterized by parabolic profile and Kerr self-focusing nonlinearity.

CONCLUSIONS

This paper proposes a model of a symmetric three-layer planar waveguide structure with adjacent layers characterized by physically different optical properties. In particular, the inner layer is characterized by the dependence of dielectric permittivity on a spatial coordinate in the direction perpendicular to the interface plane. Outer layers are characterized by the dependence of dielectric permittivity on the electric field amplitude. In other words, the considered three-layer structure consists of the inner graded-index layer sandwiched between nonlinear optical liners. The spatial profile of the interlayer dielectric permittivity and the type of nonlinear response of the liner medium are assumed to be arbitrary.

The propagation of transverse electric waves with no losses taken into account is described from a theoretical point of view. The paper also formulates equations and boundary conditions for the transverse field distribution in a three-layer waveguide structure.

Solutions describing spatial distribution of the electric field transversely to layers are obtained in general form. The study shows that due to the transverse symmetry of the three-layer waveguide structure, even and odd stationary modes corresponding to symmetric and antisymmetric transverse field profiles can propagate along it. The paper also proposes a method of constructing even (symmetric) and odd (antisymmetric) solutions resulting in the existence of a discrete spectrum of the effective refractive index value.

Particular cases of specific spatial profiles of the inner layer dielectric permittivity with exact analytical solutions to the wave equation are considered. In particular, solutions for linear, parabolic, and exponential profiles described by corresponding special functions are given. The study also determined discrete spectra

of the effective refractive index values in layers with considered graded-index profiles.

In addition, specific models of media nonlinearity, such as Kerr and step nonlinearities, are also considered. For such nonlinearities, exact analytical solutions are given to the nonlinear wave equation describing the dependencies of the stationary electric field amplitude on the distance from layer interfaces in nonlinear optical media.

The study provides a detailed analysis of symmetric three-layer waveguide structure, where the inner graded-index layer is characterized by a parabolic spatial profile and the outer liners represent Kerr nonlinear optical media. The exact analytical solution to the formulated boundary value problem describing the transverse symmetric field distribution for self-focusing nonlinearity is obtained and analyzed. The main mode intensity significantly exceeds the intensity of higher-order modes. The electric field strength grows with the increasing relative change of dielectric permittivity in the interlayer for the main mode and first-order modes. However, it decreases with an increase of its value for higher-order modes.

The results obtained can be useful in developing various optical waveguide devices. The proposed theory also expands understanding of physical properties of nonlinear waves and localization patterns of light beams in distributed media.

REFERENCES

1. Zhao Y., Yang Y., Sun H.B. Nonlinear meta-optics towards applications. *PhotoniX.* 2021;2(1):3. http://doi.org/10.1186/s43074-021-00025-1
2. Bano R., Asghar M., Ayub K., Mahmood T., Iqbal J., Tabassum S., Zakaria R., Gilani M. A Theoretical Perspective on Strategies for Modeling High Performance Nonlinear Optical Materials. *Front. Mater*. 2021;8:783239. https://doi.org/10.3389/fmats.2021.783239
3. Dragoman D., Dragoman M. *Advanced Optoelectronic Devices.* Berlin: Springer; 1999. 436 p.
4. Adams M.J. *An Introduction to Optical Waveguides.* Chichester: Wiley; 1981. 432 p.
5. Chen C.-L. *Foundations for Guided-Wave Optics.* New York: John Wiley & Sons, Inc.; 2005. 462 p. https://doi.org/10.1002/0470042222
6. Malomed B.A., Mihalache D. Nonlinear waves in optical and matter-wave media: a topical survey of recent theoretical and experimental results. *Rom. J. Phys.* 2019;64(5–6):106. Available from URL: https://rjp.nipne.ro/2019_64_5-6/RomJPhys.64.106.pdf
7. Mihalache D. Localized structures in optical and matter-wave media: a selection of recent studies. *Rom. Rep. Phys*. 2021;73:403. Available from URL: https://rrp.nipne.ro/2021/AN73403.pdf
8. Agrawal G.P. *Physics and Engineering of Graded-Index Media*. Cambridge: Cambridge University Press; 2023. 348 p. https://doi.org/10.1017/9781009282086
9. Ablowitz M.J., Horikis T.P. Nonlinear waves in optical media. *J. Comp. Appl. Math.* 2010;234(6):1896–1903. https://doi.org/10.1016/j.cam.2009.08.039
10. Bednarik M., Cervenka M. Electromagnetic waves in graded-index planar waveguides. *J. Opt. Soc. Am. B*. 2020;37(12):3631–3643. https://doi.org/10.1364/JOSAB.408679
11. Kivshar Yu.S., Agrawa G.P. *Optical Solitons: From Fibers to Photonic Crystals*. San Diego: Academic Press; 2003. 540 p.
12. Čada M., Qasymeh M., Pištora J. Optical Wave Propagation in Kerr Media. In: *Wave Propagation. Theories and Applications*. London: IntechOpen; 2013. P. 175–192. http://doi.org/10.5772/51293
13. Kartashov Y.V., Malomed B.A., Torner L. Solitons in nonlinear lattices. *Rev. Mod. Phys*. 2011;83(1):247–305. http://doi.org/10.1103/RevModPhys.83.247
14. Laine T.A. *Electromagnetic Wave Propagation in Nonlinear Kerr Media*: Doctoral Thesis. Stockholm: Royal Institute of Technology (KTH), Department of Physics; 2000. 58 p. Available from URL: https://kth.diva-portal.org/smash/get/diva2:8732/FULLTEXT01.pdf
15. Mihalache D., Bertolotti M., Sibilia C. IV Nonlinear wave propagation in planar structures. *Prog. Opt*. 1989;27:227–313. https://doi.org/10.1016/S0079-6638(08)70087-8
16. Langbein U., Lederer F., Peschel T., Trutschel U., Mihalache D. Nonlinear transmission resonances at stratified dielectric media. *Phys. Rep*. 1990;194(5-6):325–342. https://doi.org/10.1016/0370-1573(90)90032-W
17. Mihalache D., Stegeman G.I., Seaton C.T., Wright E.M., Zanoni R., Boardman A.D., Twardowski T. Exact dispersion relations for transverse magnetic polarized guided waves at a nonlinear interface. *Opt. Lett*. 1987;12(3):187–189. https://doi.org/10.1364/OL.12.000187
18. Kadantsev V.N., Goltsov A.N., Kondakov M.A. Electrosoliton dynamics in a thermalized molecular chain. *Russ. Technol. J*. 2020;8(1):43–57 (in Russ.). https://doi.org/10.32362/2500-316X-2020-8-1-43-57
19. Shvartsburg A.B., Maradudin A. *Waves in Gradient Metamaterials*. Singapore: World Scientific; 2013. 339 p. https://doi.org/10.1142/8649

20. Touam T., Yergeau F. Analytical solution for a linearly graded-index-profile planar waveguide. *Appl. Opt.* 1993;32(3): 309–312. https://doi.org/10.1364/AO.32.000309
21. Lachance R.L., Belanger P.-A. Modes in divergent parabolic graded-index optical fibers. *J. Lightwave Technol*. 1991;9(11):1425–1430. https://doi.org/10.1109/50.97628
22. Taya S.A., Hussein A.J., Ramahi O.M., Colak I., Chaouche Y.B. Dispersion curves of a slab waveguide with a nonlinear covering medium and an exponential graded-index thin film (transverse magnetic case). *J. Opt. Soc. Am. B*. 2021;38(11): 3237–3243. https://doi.org/10.1364/JOSAB.439034
23. Shvartsburg A.B. Dispersion of electromagnetic waves in stratified and nonstationary media (exactly solvable models). *Phys. Usp.* 2000;43(12):1201–1228. https://doi.org/10.1070/pu2000v043n12abeh000827
24. Svendsen B.B., Söderström M., Carlens H., Dalarsson M. Analytical and Numerical Models for TE-Wave Absorption in a Graded-Index GNP-Treated Cell Substrate Inserted in a Waveguide. *Appl. Sci*. 2022;12(14):7097. https://doi.org/10.3390/app12147097
25. Almawgani A.H.M., Taya S.A., Hussein A.J., Colak I. Dispersion properties of a slab waveguide with a graded-index core layer and a nonlinear cladding using the WKB approximation method. *J. Opt. Soc. Am. B.* 2022;39(6):1606–1613. https://doi.org/10.1364/JOSAB.458569
26. Savotchenko S.E. Models of waveguides combining gradient and nonlinear optical layers. *Russ. Technol. J*. 2023;11(4): 84–93 (in Russ.). https://doi.org/10.32362/2500-316X-2023-11-4-84-93
27. Hussein A.J., Nassar Z.M., Taya S.A. Dispersion properties of slab waveguides with a linear graded-index film and a nonlinear substrate. *Microsyst. Technol*. 2021;27(7):2589–2594. https://doi.org/10.1007/s00542-020-05016-z
28. Taya S.A., Hussein A.J., Colak I. An exact solution of a slab waveguide dispersion relation with a linear graded-index guiding layer (TM case). *Microsyst Technol*. 2022;28(22):1213–1219. https://doi.org/10.1007/s00542-022-05281-0
29. Hussein A.J., Taya S.A., Vigneswaran D., Udiayakumar R., Upadhyay A., Anwa T., Amiri I.S. Universal dispersion curves of a planar waveguide with an exponential graded-index guiding layer and a nonlinear cladding. *Results in Physics*. 2021;20:103734. https://doi.org/10.1016/j.rinp.2020.103734
30. Panyaev I.S., Dadoenkova N.N., Dadoenkova Yu.S., Rozhleys I.A., Krawczyk M., Lyubchanckii I.L., Sannikov D.G. Four-layer nanocomposite structure as an effective optical waveguide switcher for near-IR regime. *J. Phys. D: Appl. Phys*. 2016;49(43):435103. http://doi.org/10.1088/0022-3727/49/43/435103
31. Zhong N., Wang Z., Chen M., Xin X., Wu R., Cen Y., Li Y. Three-layer-structure polymer optical fiber with a rough inter-layer surface as a highly sensitive evanescent wave sensor. *Sensors and Actuators B: Chem*. 2018;254:133–142. https://doi.org/10.1016/j.snb.2017.07.032
32. Akhmediev N.N. Novel class of nonlinear surface waves: asymmetric modes in a symmetric layered structure. *J. Exp. Theor. Phys*. 1982;56(2):299–303. Available from URL: http://jetp.ras.ru/cgi-bin/dn/e_056_02_0299.pdf
33. Fedyanin V.K., Mihalache D. *P*-Polarized nonlinear surface polaritons in layered structures. *Z. Phys. B.* 1982;47:167–173. https://doi.org/10.1007/BF01441299
34. Chatterjee S., Chaudhuri P.R. Some Unique Propagation Characteristics of Linearly Graded Multilayered Planar Optical Waveguides. *J. Basic Appl. Phys*. 2014;3(1):1–9.
35. Savotchenko S.E. Temperature controlled waveguide properties of the linearly graded-index film in semiconductor crystal with the photorefractive nonlinearity. *Appl. Phys. B: Lasers and Optics*. 2023;129(1):7. https://doi.org/10.1007/s00340-022-07950-4
36. Savotchenko S.E. New types of transverse electric nonlinear waves propagating along a linearly graded-index layer in a medium with Kerr nonlinearity. *Opt. Quant. Electron*. 2023;55(1):74. https://doi.org/10.1007/s11082-022-04323-1
37. Savotchenko S.E. Discrete spectrum of waveguide modes of a linearly graded-index film introduced into a medium with a stepwise nonlinearity. *Optik*. 2023;281(6):170835. https://doi.org/10.1016/j.ijleo.2023.170835
38. Savotchenko S.E. Guided waves propagating along a parabolic graded-index slab in Kerr nonlinear medium. *Opt. Quant. Electron*. 2023;55:898. https://doi.org/10.1007/s11082-023-05178-w
39. Savotchenko S.E. Surface waves propagating along an interface between media with an exponential spatial profile of the dielectric function and an abruptly appearance of a self-focusing nonlinear response in a near-surface layer at the strong light intensity. *Opt. Quant. Electron.* 2023;55(7):580. https://doi.org/10.1007/s11082-023-04871-0
40. Rana B., Svendsen B.B., Dalarsson M. TE-Wave Propagation Over an Impedance-Matched RHM to LHM Transition in a Hollow Waveguide. *Progress In Electromagnetics Research M*. 2022;110:1–10. http://doi.org/10.2528/PIERM22022505
41. Kaplan I.G. *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials*. Hoboken: John Wiley & Sons, Ltd; 2006. 384 p. https://doi.org/10.1002/047086334X
42. Khadzhi P.I., Fedorov L.V., Torstveit S. Nonlinear surface waves for the simplest model of nonlinear medium. *Phys. Tech. Lett*. 1991;61:110–113.
43. Lyakhomskaya K.D., Khadzhi P.I. Self-reflection effect in naïve model of nonlinear media. *Tech. Phys*. 2000;45(11): 1457–1461. https://doi.org/10.1134/1.1325030
[Original Russian Text: Lyakhomskaya K.D., Khadzhi P.I. Self-reflection effect in naïve model of nonlinear media. *Zhurnal Tekhnicheskoi Fiziki.* 2000;70(11):86–90 (in Russ.).]
44. Vigdorovich E.N. Radiation resistance of epitaxial structures based on GaAs. *Russ. Technol. J*. 2019;7(3):41–49 (in Russ.). https://doi.org/10.32362/2500-316X-2019-7-3-41-49

About the author

Sergey E. Savotchenko, Dr. Sci. (Phys.-Math.), Associate Professor, Professor of the High Mathematics Department, Institute for Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, ResearcherID N-9227-2018, RSCI SPIN-code 2552-4344, https://orcid.org/0000-0002-7158-9145

Об авторе

Савотченко Сергей Евгеньевич, д.ф.-м.н., доцент, профессор кафедры высшей математики, Институт кибербезопасности и цифровых технологий, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: savotchenkose@mail.ru. Scopus Author ID 6603577988, ResearcherID N-9227-2018, SPIN-код РИНЦ 2552-4344, https://orcid.org/0000-0002-7158-9145

Translated from Russian into English by K. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems

Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах

UDC 004.021
https://doi.org/10.32362/2500-316X-2024-12-5-90-97
EDN TAQCXC

RESEARCH ARTICLE

Service quality assessment in IT projects based on aggregate indicators

Andrey E. Krasnov [@], **Alexander A. Sapogov** [@]

Russian State Social University, Moscow, 129226 Russia
[@] *Corresponding authors, e-mail: krasnovmgutu@yandex.ru, sapogovmail@gmail.com*

Abstract

Objectives. Due to the need for prompt and rational assessment of service quality within the framework of complex IT projects, including infrastructure servicing and maintenance, which often involve a large number of identical or similar iterations, it becomes necessary to develop novel analysis methods based on nonlinear aggregation of indicators. As a result of changes in the structure of the process, territorial remoteness, automation, informatization, and the emergence of big data, the use of existing assessment methods often becomes impossible or labor-intensive. The purpose of the present work is to develop an approach to assessing the quality of work (services) in the framework of IT projects based on nonlinear aggregation of indicators.

Methods. The proposed approach to assessing service quality within IT projects is based on nonlinear aggregation of a number of indicators involving a preliminary decomposition of the system into private indicators. In order to meet the requirements of the decomposition process, service quality indicators must fully characterize the properties of the service as a whole at the different stages of its life cycle.

Results. The application of the proposed nonlinear aggregation methodology to quality indicators obtained by decomposing the system is described with the further calculation of a single indicator that takes all the essential initial parametric indicators into account. The decomposition of complex systems to the level of elementary relationship subsystems more adequately reflects interrelated phenomena in a complex system.

Conclusions. The practical application of the neural network parametric data aggregation model for assessing the quality of IT services is demonstrated. The use of an aggregated information and analytical indicator for assessing service quality increases the availability of analytical information for decision makers, reduces the dimension of analytical data, and improves the objectivity of the obtained generalized information.

Keywords: aggregation, assessment, quality, indicator, IT project, analytics

• **Submitted:** 05.02.2024 • **Revised:** 19.03.2024 • **Accepted:** 12.07.2024

For citation: Krasnov A.E., Sapogov A.A. Service quality assessment in IT projects based on aggregate indicators. *Russ. Technol. J*. 2024;12(5):90–97. https://doi.org/10.32362/2500-316X-2024-12-5-90-97

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Оценка качества услуг в рамках ИТ-проектов на основе агрегирования показателей

А.Е. Краснов [@], **А.А. Сапогов** [@]

Российский государственный социальный университет, Москва, 129226 Россия
[@] *Авторы для переписки, e-mail: krasnovmgutu@yandex.ru, sapogovmail@gmail.com*

Резюме

Цели. Необходимость оперативного и обоснованного оценивания качества услуг в рамках сложных ИТ-проектов, таких как сервисное, техническое обслуживание ИТ-инфраструктуры, включающей выполнение большого числа схожих или аналогичных итераций, предопределяет необходимость разработки новых методов оценки качества, основанных на нелинейном агрегировании показателей. Применение прежних методов контроля становится невозможным либо трудозатратным ввиду изменения структуры процесса, территориальной удаленности, автоматизации, информатизации и появления больших данных. Цель работы – разработка подхода к оцениванию качества работ (услуг) в рамках ИТ-проектов на основе нелинейного агрегирования показателей.

Методы. Предлагается подход к оцениванию качества работ (услуг) в рамках ИТ-проектов на основе нелинейного агрегирования ряда показателей с предварительной декомпозицией системы на частные индикаторы. Показатели качества услуги должны соответствовать требованиям процесса декомпозиции, т.е. полностью характеризовать свойства услуги как единого целого на стадиях ее жизненного цикла.

Результаты. Описано применение предложенной методологии нелинейного агрегирования к индикаторам качества, полученным путем декомпозиции системы, с дальнейшим расчетом единого показателя, учитывающего все существенные изначальные параметрические показатели индикаторов. Предложено производить декомпозицию сложных систем до уровня элементарных подсистем соотношений, описываемых этими индикаторами, которые изначально более адекватно отражают взаимосвязанные явления в сложной системе, нежели абсолютные показатели.

Выводы. Показано преимущество практического применения модели параметрического нелинейного агрегирования данных для оценки качества ИТ-услуг. Использование агрегированного информационно-аналитического показателя оценки качества услуг улучшает доступность аналитической информации для лиц, принимающих решения, снижает размерность аналитических данных, повышает объективность получаемой обобщенной информации.

Ключевые слова: агрегирование, оценка, качество, индикатор, ИТ-проект, аналитика

• **Поступила:** 05.02.2024 • **Доработана:** 19.03.2024 • **Принята к опубликованию:** 12.07.2024

Для цитирования: Краснов А.Е., Сапогов А.А. Оценка качества услуг в рамках ИТ-проектов на основе агрегирования показателей. *Russ. Technol. J*. 2024;12(5):90–97. https://doi.org/10.32362/2500-316X-2024-12-5-90-97

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

An important element in the management of a company's processes consists in assessing the quality of the services provided. Such an assessment provides a basis for feedback to support management decisions. However, a unified methodology for assessing all aspects of service quality is yet to be developed.

During the period of rendering services within the framework of complex IT projects, including service and maintenance works, it is necessary to ensure proper definitions that underly the information-analytical parameters of work performance. Such information-analytical parameters can be used in analytical reporting to support organizational and technical decisions as a means of improving work performance (service provision).

Quality assessment indicators reflecting parameters used for quality management can be measured in different ways. In any case, the result of the measurement comprises a set of numerical values for the parameters. However, it may not be sufficient to indicate that the quality of a project depends on the timing of its implementation. Thus, decisions must also be taken concerning whether all operations must necessarily begin and end at a certain time or only after a certain deadline. In addition, it should be stated which outcomes are to be monitored. The results of quality assessment are used in service delivery, monitoring and processes management.

The purpose of the present work is to propose a methodology for the nonlinear aggregation of heterogeneous quality indicators. In this connection, the main task is to apply the proposed methodology of nonlinear aggregation to the quality indicators obtained by decomposition of the system, with further calculation of a single indicator that takes into account all the significant initial parametric indicators.

PROBLEM STATEMENT

Depending on the specifics of a project, various indicators can be used for nonlinear calculation of a single quality indicator. These include defect density, failure probability, availability of service, reliability, etc. At present, many methods of service quality assessment have been developed, including critical-case methods such as INDSER, Kano, SERVQUAL, SERVPERF [1–4]. Various weighted average ranking methods are also applied, many of which are based on the ranking of essential attributes of the provided service quality, e.g., tangibility, reliability, assurance, trust, security, attention, communication, and customer understanding. However, traditional quality assessment schemes are often not adapted to the changed information realities. Often it becomes impossible to operate with big data; moreover, there are no methods of online monitoring in the process of service provision or there are no reliable mathematical methods for calculating the proposed characteristics due to their being based on expert assessments, whose results can be difficult to formalize. In [4], it is noted that "there is no single best method for assessing the state [...] since the state is a multidimensional characteristic, and different methods [...] reflect different aspects of its state […] so there is a need for research into methods […] that enable objectivizing the assessment of the subject's state on the basis of available heterogeneous data."

For example, the commonly used SERVQUAL model is based on the concept of customer service, which is based on the expectation-perception principle. The questionnaire used in the SERVQUAL model consists of five blocks of user questions, each of which responds to a specific request. 22 pairs of questions in each block are distinguished using the Likert scale, where each parameter has its own value. This methodology is based on the assessment of consumer attitudes towards the product or service provided to the customer. Despite the fact that SERVQUAL models are based on quantitative and qualitative indicators of customer satisfaction, they lack universality. Some of the shortcomings noted by researchers were taken into account in the development and supplementation of evaluation methodology concepts. At the same time, the majority of works do not contain conceptually new approaches to the methodology of service quality assessment [5–8].

Since the evaluation of services "has a qualitative character and is an object of nonnumerical nature," the essential evaluation of services cannot be correctly performed within a deterministic framework but becomes possible only within the context of a linguistic scale [9].

The widespread practice of service quality assessment focused on customer satisfaction cannot be recognized as successful due to the risk of substituting subjective feelings for objective concepts; moreover, there is a danger of manipulating consumer opinions. At the same time, when assessing the quality of services, it is impossible to completely abandon the evaluation of consumer opinion, i.e., it becomes necessary to achieve a balance between consumer opinion and objective evaluation [10].

Along with methods of comparative analysis based on approaches to solving problems of multi-criteria and multi-objective decision making in fuzzy conditions, clustering methods [11] attract attention due to the possibility of using the mathematical and conceptual apparatus formed on their basis for ranking research objects in quality assessment. Fuzzy

multi-criteria optimization methods include ELECTRE, PROMETHEE, VIKOR, TOPSIS, AHP, ANP, and DEMATEL approaches [12]. At the same time, the algorithms used in construction of decision matrices only achieve the goal of ranking within the group under consideration based on the available criteria. These are of little use in practical evaluation since they consider situational and relative rather than objective location problems. Such ranking algorithms evaluate an alternative solely within the presence of other alternatives. Clustering, on the other hand, represents a visualization method according to which the presence of relationships can be shown on a two- or three-dimensional projection. Meanwhile, the still-popular average value method can be supplemented with a number of weighting coefficients within the framework of metrics understandable to the researcher [13].

The issue of IT service quality assessment, which remains a priority for any organization, implies decision making with the help of one model or another [14, 15]. In this case, service quality can be defined as "a set of service characteristics that determine its ability to meet the established or anticipated needs of the consumer."[1]

The proposed use of a parametric nonlinear data aggregation model to assess the quality of IT services offers a number of advantages over traditional methods [16–19]. According to this methodology, the process of IT service quality assessment requires identification of the necessary quality indicators to inform the further assessment of the service quality level. Service quality indicators should meet the requirements of the decomposition process, i.e., fully characterize the properties of the service as a whole at the stages of its life cycle, which determine its ability to meet certain customer needs.

Thus, the identification and establishment of quality indicators in an IT project, especially in such works as service and maintenance of IT infrastructure, is a complex and fundamental procedure for the subsequent quality assessment according to the parametric data aggregation model, because it implies decomposition of the service delivery process in terms of quality into the corresponding aggregates. Since the process of decomposition has its own peculiarities and limitations, it should be performed according to the developed methodology rationally, expediently, and in accordance with the set goals [20, 21].

To this end, it is necessary to establish which initially qualitative characteristics are important in controlling the realization of the IT project, representing them in the form of vectors of values with criterion indicators.

[1] GOST R 50646-94. State Standard of the Russian Federation. *Service for people. Terms and definitions*. Moscow: Izdatelstvo standartov; 1994 (in Russ.).

RESULTS

It is proposed to use the following indicators for IT projects:

1. Inquiry Processing Timeliness Indicator (IPTI).

This indicator expresses the probability that a request sent to the IT service will be processed within the agreed timeframe:

$$\mathrm{IPTI} = \frac{D1}{D1 + D2}, \quad (1)$$

where $D1$ is the total number of processed requests for the reporting period, $D2$ is the total number of overdue requests for the reporting period.

2. Operational Readiness Indicator (ORI).

This indicator expresses the probability that the IT service is in an operational state at a given point in time:

$$\mathrm{ORI} = \frac{T1}{T1 + T2}, \quad (2)$$

where $T1$ is the time period of service operation (in hours) for the reporting period, $T2$ is the time of IT service unavailability for the reporting period.

3. Consumer Satisfaction Indicator (CSI).

This indicator expresses the probability that the IT service will meet the consumer's needs:

$$\mathrm{CSI} = \frac{Y1}{Y1 + Y2}, \quad (3)$$

where $Y1$ is the total number of users for the reporting period, $Y2$ is the total number of complaints for the reporting period.

As well as being methodologically justified by a single dimension and concept, the use of the introduced indicators having numerical values from 0 to 1 has practical significance, convenience of application, and visualizability.

It is proposed to decompose complex systems only to the level of elementary subsystems described by these indicators, which initially reflect interrelated phenomena in a complex system more adequately than absolute indicators.

Since a complex system is composed solely only inherent relations, representing interconnected and interdependent processes that reflect an internal balance, the description of elementary subsystems using absolute values will significantly complicate the calculation of aggregate indicators.

The next step in the parametric nonlinear aggregation model comprises the calculation of the aggregated indicator of the service quality assessment. A simpler interpretation of the methodological approach given in [17, 18] is proposed as an aggregation technique:

$$AIAI = \frac{1}{1 + \sum_{n=1}^{N} \alpha_n \frac{(i_n - 1)^2}{h^2}}, \quad 0 \le AIAI \le 1, \qquad (4)$$

where AIAI is the aggregated information-analytical indicator; α_n are weighting coefficients that verify the individual importance of the indicators of elementary aggregates; i_n are the values of the indicators of elementary aggregates, i.e., i_1 = IPTI, i_2 = ORI, i_3 = CSI, … Here N is the number of indicators which can be of any value; h^2 are the expert estimations of unknown interference intensities of the elementary aggregate indicators.

According to the conditions of formula (4), the AIAI monotonically increases to 1 when the values of all elementary aggregates approach to their upper values, i.e., to 1. Thus, the AIAI objectively reflects the qualitative characteristics of the process of rendering services (works performed). The AIAI can also be used to identify processes that require additional controls.

The presented methodology resolves many problems "associated with the use of indices, such as the choice of a base for calculation, the quality of source data and data aggregation" [22, 23].

The following table presents the calculation of AIAI for five organizations that are part of the group of companies providing IT support services (authors' data based on the materials of Digital Service[2]).

Calculated AIAI are presented in the diagrams (Figs. 1 and 2).

DISCUSSION OF RESULTS

As shown in the diagrams, the AIAI obtained by the model of parametric nonlinear aggregation correlates with the set of initial data represented by elementary aggregate indicators, as well as adequately summarizing the set of initial indicators within the framework of relations between them.

In the process of pilot application and operation taking into account expert opinion, the AIAI may be supplemented with the following conditions for best practical application:

- threshold value;
- target value;
- actual value.

The target direction of AIAI changes for all works over time: positive – increasing; negative – decreasing.

Target and threshold values can be used to set the AIAI scale of the work, including for comparison over time. The AIAI scale can also be divided into the following areas:

- works (services) do not require additional control,
- additional control of works (services) is required),
- preventive measures are required.

If the actual value of AIAI is not less than the target value, it corresponds to the wording "works (services) do not require additional control." If the actual value of AIAI is greater than or equal to the threshold value and less than the target value, this corresponds to the wording "additional control is required." In cases where the actual value of AIAI is less than the threshold value, it can be concluded that "preventive measures are required."

Table. Calculated data of indicators and AIAI

Pos. No.	IPTI	ORI	CSI	AIAI
1	0.85	0.75	0.88	0.66
2	0.79	0.76	0.74	0.58
3	0.93	0.95	0.98	0.88
4	0.65	0.55	0.64	0.46
5	0.58	0.5	0.45	0.40

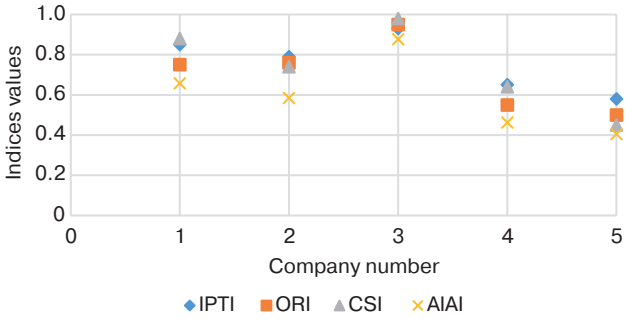


Fig. 1. AIAI visualization (dot diagram)

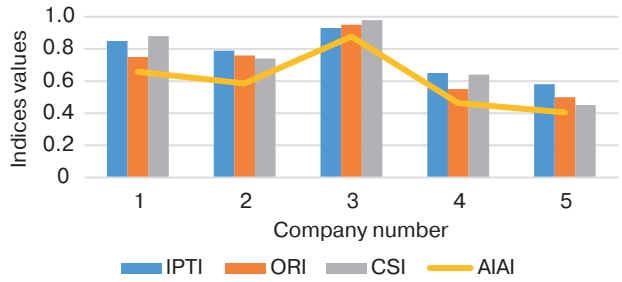


Fig. 2. AIAI visualization (combined diagram)

CONCLUSIONS

The parametric nonlinear data aggregation model for IT service quality assessment offers a number of advantages in terms of practical application. As well

[2] http://digitservice.ru (in Russ.). Accessed December 05, 2023.

as reducing the dimensionality of analytical data and increasing the objectivity of the obtained generalized information, the use of the aggregated information-analytical indicator for assessing the quality of services improves the availability of analytical information for decision makers.

Authors' contribution

The authors' contribution to the writing of the article, including the development of the concept (idea formation, formulation and development of key goals and objectives), conducting research, preparing and editing the text, and approval of the final version of the article is joint, balanced, and equivalent.

REFERENCES

1. Kurnosova O.A. Assessment of the quality of organization of the system of logistic service at industrial enterprises. *Uchenye zapiski Krymskogo federal'nogo universiteta imeni V.I. Vernadskogo. Ekonomika i upravlenie*. 2019;5(71):54–67 (in Russ.). Available from URL: https://sn-ecoman.cfuv.ru/wp-content/uploads/2019/05/54-67.pdf?ysclid=lx303wgj28650321072
2. Parasuraman A., Zeithaml V.A., Berry L.L. SERVQUAL: A multiple-item scale for measuring service quality. *J. Retail*. 1988;64(1):12–40.
3. Shatalova V.V., Likhachevskii D.V., Kazak T.V. Big Data: how Big Data technologies are changing our lives. *Big Data and Advanced Analytics*. 2021;7(1):188–192 (in Russ.).
4. Vainshtok A.P., Yurkov E.F. Ranking models of index estimation of the Russian Federation subjects in accordance with socio-economic indicators. *Informatsionnye protsessy = Information Processes*. 2023;23(1):138–147 (in Russ.). Available from URL: http://www.jip.ru/2023/138-147-2023.pdf
5. Limbourg S., Giangb H.Q., Coolsc M. Logistics Service Quality: The Case of Da Nang City. *Procedia Eng*. 2016;142: 124–130. https://doi.org/10.1016/j.proeng.2016.02.022
6. Rahman S. Quality management in logistics services: A comparison of practices between manufacturing companies and logistics firms in Australia. *Total Quality Management & Business Excellence*. 2008;19(5):535–550. https://doi.org/10.1080/14783360802018202
7. Franceschini F., Rafele C. Quality evaluation in logistic services. *Int. J. Agile Man. Syst.* 2000;2(1):49–54. http://doi.org/10.1108/14654650010312589
8. Gajewska T., Grigoroudis E. Importance of logistics services attributes influencing customer satisfaction. In: *4th IEEE International Conference on Advanced Logistics and Transport* (*ICALT*)*: Conference Paper*. IEEE; 2015. P. 53–58. https://doi.org/10.1109/ICAdLT.2015.7136590
9. Ershova T.B. General characteristics of quality IT services company. *Ekonomicheskie i gumanitarnye nauki = Economic Science and Humanities*. 2011;2(229):109–112 (in Russ.).
10. Borden L. *How to Measure and Improve the IT Service Desk Experience*. ISG White Paper; 2015. 12 p.
11. Lapko A.V., Lapko V.A., Tuboltsev V.P. Methodology for aggregating of the results of automatic classification of statistical data. In: *Reshetnev Readings: Materials of the 27th International Scientific and Practical Conference dedicated to the memory of the general designer of rocket and space systems, Academician M.F. Reshetnev*. Krasnoyarsk: 2023. P. 430–432 (in Russ.).
12. Rodzin S.I., Bozhenyuk A.V., Rodzina O.N. Methods of fuzzy multicriteria group decision-making for evacuation tasks in emergency situations. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2023;2(232):186–200 (in Russ.).
13. Shamasna Kh.A., Semashko A.V. Development of a data aggregation and visualization system in an intelligent poultry farm. *Nauchno-tekhnicheskii vestnik Povolzh'ya = Scientific and Technical Volga Region Bulletin*. 2023;9:135–139 (in Russ.).
14. Trainev V.A., Trainev O.V. *Parametricheskie modeli v ekspertnykh metodakh otsenki pri prinyatii reshenii* (*Parametric Models in Expert Assessment Methods for Decision Making*). Moscow: Prometei; 2003. 231 p. (in Russ.). ISBN 5-94798-023-1
15. Sakulin S.A., Alfimtsev A.N., Bobretsova A.G. An approach to decision support for choosing a telecommunications equipment supplier based on aggregation operators. *Vestnik komp'yuternykh i informatsionnykh tekhnologii = Herald of Computer and Information Technologies*. 2023;20(11):46–53 (in Russ.).
16. Krasnov A.E., Nadezhdin E.N., Nikol'skii D.N., Repin D.S., Kalachev A.A. Neural network approach to the estimation of the functioning efficiency of the organization based on the aggregation of its activities. *Informatizatsiya obrazovaniya i nauki = Informatization of Education and Science.* 2017;1(33):141–154 (in Russ.).
17. Krasnov A.E., Krasnikov S.A., Aniskin D.Yu., et al. Models of quantitative evaluation of objects4 quality of technologies, manufacture and business in IDFM standard. *Khranenie i pererabotka sel'khozsyr'ya = Storage and Processing of Farm Products*. 2006;3:53–56 (in Russ.).
18. Sapogov A.A. Existing methods for financial data aggregation. *Innovatsii i investitsii = Innovations and Investments Magazine.* 2023;8:247–250 (in Russ.).
19. Sapogov A.A., Krasnov A.E. Problem of selection of weight coefficients in aggregation operations. In: *Science. Production. Education: Collection of scientific papers of the All-Russian Scientific and Technical Conference*. Moscow: April 14, 2023. 2023. P. 132–139 (in Russ.).

20. Pivneva S.V., Blokhina M.V. Decomposition of the school management process in a unified information environment. In: *Modern Technologies and Automation in Technology, Management and Education: Collection of proceedings of the Second International Scientific and Practical Conference*. Balakovo: December 18, 2019. 2020. V. 1. P. 216–219 (in Russ.).
21. Elkin V.I. Aggregation and decomposition of systems of partial differential equations and control systems with distributed parameters. *Comput. Math. Math. Phys*. 2023;63(9):1741–1750.
[Original Russian Text: Elkin V.I. Aggregation and decomposition of systems of partial differential equations and control systems with distributed parameters. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki*. 2023;63(9):1575–1586 (in Russ.). https://doi.org/10.31857/S0044466923090089]
22. Lubenets K.A. Indices: prospects and problems. In: *Innovative Research: Problems of Implementation of Results and Directions of Development: Collection of Articles of the 17th International Scientific Conference*. Tyumen: November 18, 2023. St. Petersburg: Lomonosov International Institute for Advanced Studies; 2023. P. 43–45 (in Russ.).
23. Krasnov A., Pivneva S. Hierarchical quasi-neural network data aggregation to build a university research and innovation management system. In: Murgul V., Pukhkal V. (Eds.). *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies* (*EMMFT 2019*). *Advances in Intelligent Systems and Computing*. 2021. V. 1259. P. 12–25. https://doi.org/10.1007/978-3-030-57453-6_2

СПИСОК ЛИТЕРАТУРЫ

1. Курносова О.А. Оценка качества организации системы логистического сервиса на промышленных предприятиях. *Ученые записки Крымского федерального университета имени В.И. Вернадского. Экономика и управление*. 2019;5(71):54–67. URL: https://sn-ecoman.cfuv.ru/wp-content/uploads/2019/05/54-67.pdf?ysclid=lx303wgj28650321072
2. Parasuraman A., Zeithaml V.A., Berry L.L. SERVQUAL: A multiple-item scale for measuring service quality. *J. Retail*. 1988;64(1):12–40.
3. Шаталова В.В., Лихачевский Д.В., Казак Т.В. Большие данные: как технологии Big Data меняют нашу жизнь. *Big Data and Advanced Analytics*. 2021;7(1):188–192.
4. Вайнштok А.П., Юрков Е.Ф. Ранговые модели индексации субъектов РФ по социально-экономическим показателям. *Информационные процессы*. 2023;23(1):138–147. URL: http://www.jip.ru/2023/138-147-2023.pdf
5. Limbourg S., Giangb H.Q., Coolsc M. Logistics Service Quality: The Case of Da Nang City. *Procedia Eng*. 2016;142: 124–130. https://doi.org/10.1016/j.proeng.2016.02.022
6. Rahman S. Quality management in logistics services: A comparison of practices between manufacturing companies and logistics firms in Australia. *Total Quality Management & Business Excellence*. 2008;19(5):535–550. https://doi.org/10.1080/14783360802018202
7. Franceschini F., Rafele C. Quality evaluation in logistic services. *Int. J. Agile Man. Syst.* 2000;2(1):49–54. http://doi.org/10.1108/14654650010312589
8. Gajewska T., Grigoroudis E. Importance of logistics services attributes influencing customer satisfaction. In: *4th IEEE International Conference on Advanced Logistics and Transport* (*ICALT*)*: Conference Paper*. IEEE; 2015. P. 53–58. https://doi.org/10.1109/ICAdLT.2015.7136590
9. Ершова Т.Б. Общая характеристика качества ИТ-услуг предприятия. *Экономические и гуманитарные науки*. 2011;2(229):109–112.
10. Borden L. *How to Measure and Improve the IT Service Desk Experience*. ISG White Paper; 2015. 12 p.
11. Лапко А.В., Лапко В.А., Тубольцев В.П. Методика агрегирования результатов автоматической классификации статистических данных. В сб.: *Решетневские чтения: Материалы XXVII Международной научно-практической конференции, посвященной памяти генерального конструктора ракетно-космических систем академика М.Ф. Решетнева*. Красноярск: 2023. С. 430–432.
12. Родзин С.И., Боженюк А.В., Родзина О.Н. Методы нечеткого многокритериального группового принятия решений для задач эвакуации при чрезвычайных ситуациях. *Известия ЮФУ. Технические науки*. 2023;2(232):186–200.
13. Шамасна Х.А., Семашко А.В. Разработка системы агрегации и визуализации данных в интеллектуальной птицефабрике. *Научно-технический вестник Поволжья*. 2023;9:135–139.
14. Трайнев В.А., Трайнев О.В. *Параметрические модели в экспертных методах оценки при принятии решений*. М.: Прометей; 2003. 231 с. ISBN 5-94798-023-1
15. Сакулин С.А., Алфимцев А.Н., Бобрецова А.Г. Подход к поддержке принятия решений по выбору поставщика телекоммуникационного оборудования на основе операторов агрегирования. *Вестник компьютерных и информационных технологий*. 2023;20(11):46–53.
16. Краснов А.Е., Надеждин Е.Н., Никольский Д.Н., Репин Д.С., Калачев А.А. Нейросетевой подход к проблеме оценивания эффективности функционирования организации на основе агрегирования показателей ее деятельности. *Информатизация образования и науки*. 2017;1(33):141–154.
17. Краснов А.Е., Красников С.А., Анискин Д.Ю., Воробьева А.В., Кузнецова Ю.Г., Краснова Н.А., Сагинов Ю.Л. Модели количественного оценивания качества объектов технологий, производства и бизнеса в стандарте IDFM. *Хранение и переработка сельхозсырья*. 2006;3:53–56.
18. Сапогов А.А. Существующие методики агрегирования финансовых данных. *Инновации и инвестиции*. 2023;8:247–250.

19. Сапогов А.А., Краснов А.Е. Задачи определения весовых коэффициентов при операциях агрегирования. В сб.: *Наука. Производство. Образование: Сборник научных трудов Всероссийской научно-технической конференции*. Москва: 14 апреля 2023 г. 2023. С. 132–139.
20. Пивнева С.В., Блохина М.В. Декомпозиция процесса управления школой в единой информационной среде. В сб.: *Современные технологии и автоматизация в технике, управлении и образовании: Сборник трудов II Международной научно-практической конференции*. Балаково: 18 декабря 2019 г. 2020. Т. 1. С. 216–219.
21. Елкин В.И. Агрегирование и декомпозиция систем дифференциальных уравнений с частными производными и систем управления с распределенными параметрами. *Журнал вычислительной математики и математической физики*. 2023;63(9):1575–1586. https://doi.org/10.31857/S0044466923090089
22. Лубенец К.А. Индексы: перспективы и проблемы. В сб.: *Инновационные исследования: проблемы внедрения результатов и направления развития: сборник статей XVII международной научной конференции*. Тюмень: 18 ноября 2023 г. Санкт-Петербург: Международный институт перспективных исследований имени Ломоносова; 2023. С. 43–45.
23. Krasnov A., Pivneva S. Hierarchical quasi-neural network data aggregation to build a university research and innovation management system. In: Murgul V., Pukhkal V. (Eds.). *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies* (*EMMFT 2019*). *Advances in Intelligent Systems and Computing*. 2021. V. 1259. P. 12–25. https://doi.org/10.1007/978-3-030-57453-6_2

About the authors

Andrey E. Krasnov, Dr. Sci. (Phys.-Math.), Professor, Head of the Department of Information Security, Russian State Social University (4, Vil'gel'ma Pika ul., Moscow, 129226 Russia). E-mail: krasnovmgutu@yandex.ru. Scopus Author ID 57192947423, RSCI SPIN-code 5431-8134, https://orcid.org/0000-0002-4075-4427

Alexander A. Sapogov, Postgraduate Student, Russian State Social University (4, Vil'gel'ma Pika ul., Moscow, 129226 Russia). E-mail: sapogovmail@gmail.com. RSCI SPIN-code 2253-5104, https://orcid.org/0009-0007-8095-2771

Об авторах

Краснов Андрей Евгеньевич, д.ф.-м.н., профессор, заведующий кафедрой информационной безопасности, ФГБОУ ВО «Российский государственный социальный университет» (129226, Москва, ул. Вильгельма Пика, д. 4, стр. 1). E-mail: krasnovmgutu@yandex.ru. Scopus Author ID 57192947423, SPIN-код РИНЦ 5431-8134, https://orcid.org/0000-0002-4075-4427

Сапогов Александр Александрович, аспирант, ФГБОУ ВО «Российский государственный социальный университет» (129226, Москва, ул. Вильгельма Пика, д. 4, стр. 1). E-mail: sapogovmail@gmail.com. SPIN-код РИНЦ 2253-5104, https://orcid.org/0009-0007-8095-2771

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Philosophical foundations of technology and society

Мировоззренческие основы технологии и общества

UDC 378.1
https://doi.org/10.32362/2500-316X-2024-12-5-98-110
EDN WAZLGB

RESEARCH ARTICLE

A conceptual approach to digital transformation of the educational process at a higher education institution

Alexey A. Kytmanov [1, @], Yuliya N. Gorelova [2], Tatiana V. Zykova [3], Olga A. Pikhtilkova [1], Elena V. Pronina [1]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *Kazan (Volga Region) Federal University, Kazan, 420008 Russia*
[3] *Siberian Federal University, Krasnoyarsk, 660041 Russia*
[@] *Corresponding author, e-mail: kytmanov@mirea.ru*

Abstract

Objectives. The research aims to develop a conceptual approach to the digital transformation of university educational processes. The approach is based on a detailed analysis of the stages, participants, and components of the educational process at universities in order to develop a roadmap for digitalization and the development of a data-driven educational process management system. The main objectives of digital transformation are: (1) improve convenience for all groups of end users by providing access to data and operations with data related to the educational process; (2) increase the transparency of all components of the educational process; (3) release human and time resources by minimizing routine operations and improving the quality of decisions. The development of a data-driven educational process management system is based on digital culture principles of process management, which imply that the data collected in university systems are consistent, organized into a single structure. and stored in a form convenient for the development of new digital services. The development of tools for intelligent decision support and learning analytics is executed cooperatively by developers, analysts, and end users at all levels.

Methods. The research considers the work experience of the authors and their colleagues in Russian and international universities as users of information systems and services, developers of educational analytics services, and managers at various levels, as well as the stages of university digital transformation.

Results. The proposed conceptual approach increases comprehension by setting goals and organizing the planning of digital transformation processes in education. As well as providing a detailed description of the major participants and components of the educational process, comprising students, teachers and educational programs, the article discusses data selection criteria.

Conclusions. The development of a conceptual approach for creating a data-driven educational process management system at a university is becoming a priority task, whose successful execution will underpin further university advancement and competitiveness.

Keywords: digitalization, digital transformation, data-driven management, educational process, student, teaching staff, educational program, learning analytics

• **Submitted:** 09.02.2024 • **Revised:** 18.03.2024 • **Accepted:** 13.08.2024

For citation: Kytmanov A.A., Gorelova Yu.N., Zykova T.V., Pikhtilkova O.A., Pronina E.V. A conceptual approach to digital transformation of the educational process at a higher education institution. *Russ. Technol. J.* 2024;12(5):98–110. https://doi.org/10.32362/2500-316X-2024-12-5-98-110

Financial disclosure: The authors have no financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Концептуальный подход к цифровой трансформации образовательного процесса в вузе

А.А. Кытманов [1, @], Ю.Н. Горелова [2], Т.В. Зыкова [3], О.А. Пихтилькова [1], Е.В. Пронина [1]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *Казанский (Приволжский) федеральный университет, Казань, 420008 Россия*
[3] *Сибирский федеральный университет, Красноярск, 660041 Россия*
[@] *Автор для переписки, e-mail: kytmanov@mirea.ru*

Резюме

Цели. Целью работы является разработка концептуального подхода к цифровой трансформации образовательного процесса в вузе. В основе выбранного подхода лежит детальный анализ этапов, участников и компонентов образовательного процесса в вузе с целью выработки дорожной карты по его цифровизации и созданию системы управления образовательным процессом на основе данных. Основными задачами цифровой трансформации являются: повышение удобства доступа к данным и работы с данными, относящимися к образовательному процессу, для всех групп конечных пользователей; повышение прозрачности всех составляющих образовательного процесса; высвобождение человеко-временных ресурсов за счет минимизации рутинных операций и повышения качества принимаемых решений. В основе создания системы управления образовательным процессом на основе данных лежат принципы цифровой культуры управления процессами, которые подразумевают, что собираемые в университетских системах данные упорядочены в единую структуру, согласованы между собой, непротиворечивы и хранятся в виде, удобном для разработки новых цифровых сервисов. Разработка инструментов интеллектуальной поддержки принятия решений и учебной аналитики ведется в тесном взаимодействии разработчиков, аналитиков и конечных пользователей всех уровней.

Методы. В работе использован опыт работы авторов и их коллег в российских и зарубежных вузах в качестве пользователей информационных систем и сервисов, разработчиков сервисов учебной аналитики и руководителей разного уровня. Приведены этапы цифровой трансформации организации.

Результаты. Предложен концептуальный подход к пониманию, постановке целей и планированию процессов цифровой трансформации образовательного процесса. Подробно описаны данные основных участников и составляющих образовательного процесса: обучающихся, преподавателей и образовательных программ, необходимые для управления вузом на основе данных; аргументирован их отбор.

Выводы. Разработка концептуального подхода для создания системы управления образовательным процессом на основе данных в вузе становится приоритетной задачей, от качества решения которой во многом будут зависеть развитие и конкурентоспособность университета в будущем.

Ключевые слова: цифровизация, цифровая трансформация, управление на основе данных, образовательный процесс, обучающийся, профессорско-преподавательский состав, образовательная программа, учебная аналитика

• **Поступила:** 09.02.2024 • **Доработана:** 18.03.2024 • **Принята к опубликованию:** 13.08.2024

Для цитирования: Кытманов А.А., Горелова Ю.Н., Зыкова Т.В., Пихтилькова О.А., Пронина Е.В. Концептуальный подход к цифровой трансформации образовательного процесса в вузе. *Russ. Technol. J.* 2024;12(5):98–110. https://doi.org/10.32362/2500-316X-2024-12-5-98-110

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The digitalization of education is an important process aimed at solving problems related to the creation of technologies and the development of services for optimizing the educational process and making it more adequate to the needs of its key stakeholders: students, teachers, employers and graduates [1]. Over the past two decades, a large number of information systems and services have been implemented for collection, storage, and processing of data on participants and components of the educational process [2]. First of all, we mean here learning management systems (LMS), representing platforms for providing teaching and learning materials related to the corresponding courses. Depending on the level of development of teaching and learning materials, as well as the automation of assessment procedures, e-courses hosted on such platforms can serve both as a support for the educational process conducted in face-to-face format and as comprehensive resources within the framework of distance learning [3]. Examples of such systems include intra-university online learning platforms, many of which are Moodle-based[1], as well as Massive Open Online Course (MOOC) platforms, including such well-known platforms as Coursera[2], edX[3], and Udacity[4] (a more detailed list can be found on Wikipedia[5,6]). Such platforms allow for the collection of large amounts of data about learners and their process of mastering educational material. The emergence of such data collection tools has given rise to new areas of research such as Educational Data Mining and Learning Analytics. These fields, which were formalized as distinct areas of research in the early 2000s, have been developing much more rapidly over the past decade. A comparison of these concepts is presented in [4]. The ultimate goal of both processes is the ability to predict learning outcomes based on the analysis of educational platform data, while the focus of Learning Analytics is on the learning process itself and, accordingly, the performance (success) of the learner in mastering the course material. Educational Data Mining focuses directly on the process of extracting information from various sources [5].

A number of studies in the field of Learning Analytics have been related to the creation of systems for predicting learning outcomes based on the analysis of grades received, time spent on assignments, and overall course performance [6]. In addition, learning analytics tools are viewed as a source of real-time information for participants of the educational process (teachers and students). They allow learners to assess their own progress in the course compared to other participants, as well as to plan their time for completing assignments [7].

We note, however, that the potential of such systems is limited, as it is based on student activity data in e-learning courses, which can also be supplemented with attendance and activity data in face-to-face classes [8]. In particular, most existing systems do not take into account learners' individual characteristics such as cognitive style, motivational component, language and cultural aspects. In [9], the author emphasizes that in the digital educational environment, unlike traditional in-class learning, the stakeholders of the educational process have difficulties in determining the level of student engagement and motivation due to the lack of a conceptual approach to the process of modeling, forming, and maintaining student engagement in learning using digital educational resources. Moreover, such systems typically work with intra-disciplinary level data, i.e., data produced by student work activity within the studied course. At the same time, to create a decision support system that functions at the university level, it is necessary to use data from different hierarchical levels (see [10]), since aggregated data used for university statistics often do not reveal the underlying causes of problems arising in the educational process. In addition to students, educational processes involve teaching staff, who provide training and monitoring over the formation of knowledge, skills and abilities, as well as the degree programs (DP) themselves [11].

The present work proposes a conceptual approach to understanding, goal-setting and planning the processes of digital transformation of the educational process to enable the development of intelligent decision support

[1] https://moodle.org/. Accessed January 15, 2024.
[2] https://www.coursera.org/. Accessed January 15, 2024.
[3] https://www.edx.org/. Accessed January 15, 2024.
[4] https://www.udacity.com/. Accessed January 15, 2024.
[5] https://en.wikipedia.org/wiki/List_of_MOOC_providers. Accessed January 15, 2024.
[6] Roskomnadzor: the foreign owner of the resource violates the law of the Russian Federation.

tools. The described approach can serve as a basis for the development of a data-driven educational process management system.

1. METHODS

The study is based on the experience of the authors and their colleagues in Russian and international universities as users of information systems and services, developers of learning analytics services, and managers at different levels (teachers, e-course developers, academic heads of the educational program, department heads, school deans, heads and deputy heads of student affairs, developers of student success forecasting models, heads of IT departments, etc.). In addition, we study the global research experience in the fields of Digitalization in Education, Learning Analytics, and Educational Data Mining.

The following section describes the stages of digital transformation of the organization to provide a more informed vision of the roadmap and planning of activities for digitalization of the educational process. Its main objectives are to increase the transparency of all components of the educational process and subprocesses within it, minimize the routine burden on the participants of the educational process, and optimize the educational process by improving the quality of decisions taken at different levels. One of the main goals of digital transformation of an educational organization is to create a data-driven educational process management system as a set of Learning Analytics tools providing intelligent decision-making support.

2. STAGES OF THE DIGITAL TRANSFORMATION

The digital transformation of an organization is based on its level of digital maturity, representing an awareness of the need to transform its core processes related to data acquisition and information exchange. At the initial level of digital maturity, such transformations are spontaneous and typically initiated by individual departments as a means of optimizing their internal processes. A high level of digital maturity implies the consistent implementation of activities for the coordinated transformation and integration of all key processes of the organization in accordance with the developed transformation strategy and roadmap. As shown in the figure below, the digital transformation of an organization can be divided into three main stages.

2.1. Organization of work with data in digital form

The main disadvantage of document flow on physical media is the high labor intensity of their verification, statistical processing, analysis, and, as a consequence, decision-making based on the information extracted from them. Here, an acute problem arises in terms of accessing data from past periods, for example, for the preparation of reports or visualization of achievements over time.

When implementing electronic document or file management systems at this stage, a critical feature is the availability of data export tools in widely used formats, as well as options for presenting data in various formats, in particular, for the creation of summary forms and reports. In other words, one of the important criteria for the usability of the abovementioned systems and services is the availability of tools that provide flexibility in working with data. An example that illustrates the actual absence of document digitization is the storage of scanned versions of previously printed documents within the system.

2.2. Technological optimization of processes

The next step in digital transformation involves changing the methods of working with data and organizing interactions between departments based on

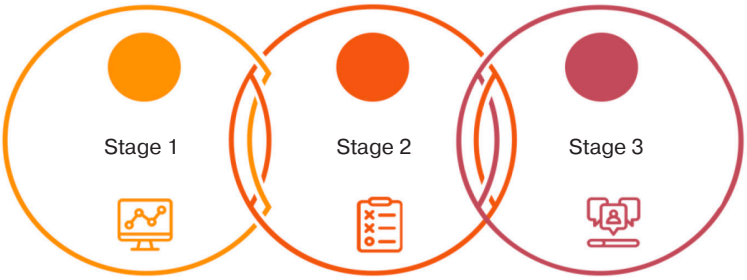


Figure. Stages of the organization's digital transformation

new working methods and tools to enhance process transparency and release human resources. An example of the transition to this stage is the optimization of the reporting process by automating the collection of data required for a specific type of reports throughout the reporting period. The results of an employee's or team's work during the reporting period should ideally be entered into the appropriate system as they arise. In this case, preparing a report does not require special efforts, since the process itself automates the generation of summarized data on outcomes achieved over a certain period. The approaches used at this stage not only save human resources but also significantly increase process transparency and reduce the likelihood of providing inaccurate data.

2.3. Transition to a digital culture of process management

At this stage, all key processes of the organization's functioning should comprise parts of a single whole. All collected data should be organized into a unified and coordinated structure to organically complement each other. As described in detail in [10], information systems function on the basis of database(s) designed to take into account the principles of student-centeredness, data continuity, and data consistency. Compliance with these principles when designing and developing the digital infrastructure of the university guarantees consistency and the absence of conflicts in data when working across different systems. It also guarantees completeness by storing all data with timestamps that allow for the accurate reconstruction of a learner's educational history, as well as providing ease of use. The information systems should allow for the connection of new modules and be updatable based on feedback from end users. The development of new modules and services is carried out in close cooperation between developers, analysts, and end users at all levels: top and middle managers, student affairs staff, as well as representatives of teaching staff and students. This digital infrastructure facilitates the development of data-driven learning analytics and decision support tools for more effectively identifying issues in the learning process and finding their possible solutions.

3. RESULTS

In this section, the main participants and components of the educational process (hereinafter subjects or objects of the educational process) are considered: students, teachers and DPs. The data structure for each subject (or object) is proposed and its detailed description is given in the context of the concept and objectives of data-driven management. The more information can be collected, the more clearly it can be structured and described in detail, resulting in higher system potential.

3.1. Student data

Since the student is a key participant in the educational process, it becomes especially important to initially classify his or her data in such a way that their subsequent use will simplify the process of data extraction and analysis [12]. When building a model of a student, two types of data should be taken into account: (1) the data that do not change or change very slowly over time (socio-demographics, gender, age, nationality and cultural characteristics, psychological characteristics and cognitive features, etc.); (2) accumulated data on the process and results of learning (accumulated digital footprints (entrance examination scores, results of participation in academic competitions, secondary education certificate grades, data on his or her activity on educational platforms) comprising a student's digital educational history, the concept of which was introduced in [13].

One of the important tasks of Learning Analytics is modeling the student, including the creation of his or her digital twin. It is clear that the more data about the student and his or her learning process can be collected, the more accurate the learner model will be. In particular, an important task is developing tools for collecting data on student's use of learning materials from external sources (educational content, educational forums, various reference materials).

When collecting data, it is important to adhere to the principle of data continuity [10] to ensure that no data that changes over time is lost, in particular, in the process of updating (overwriting). For example, in case a student does not pass an exam at the first attempt, all dates and results of attempts should be stored in the database. This gives a more comprehensive picture of academic performance, helping to identify problems with learning in a timely manner and adjust personal learning trajectories. In cases when a student switches from one major (degree program) to another, such changes in the student's interests or difficulties in mastering certain courses should also be recorded. In other words, such indicators can be used to identify and address problem areas with subsequent adjustment of the DP in order to improve its quality.

We will conditionally divide all data into three groups. The first group includes general data used in educational and administrative processes. The second group operates with data related to a specific DP and its learning outcome requirements. Finally, the third group contains information about the student's activity and performance in a particular course of the DP.

The first block includes *basic data on the student*:

- basic personal data,
- socio-demographics,
- health records,
- admission data,
- academic status of the student.

Basic personal data contains the student's surname, first name, patronymic, date and place of birth, gender, data on main identification documents (passport, social security number, taxpayer identification number).

Socio-demographic data includes information on marital status, family members, as well as the number of dependents and average income. This information may prove useful when considering the possibility of applying for social scholarships and other assistance.

Health records form an important basis for determining whether inclusive education is necessary. This includes data on the presence of medical conditions and the opinions of medical experts.

Admission data traditionally include entrance examination scores or results of previous final attestation, as well as the results of participation in academic competitions, which confer certain privileges when applying for the next level of education. An applicant's portfolio data can give additional points at admission to some areas of training. At admission to a Master's program, information about the student's bachelor's diploma, results of entrance examinations, as well as publication record, will be required.

Academic status contains information about the current status of the student, e.g., studying / completed training / on academic leave / expelled / in the process of readmission. It can also indicate the status of the student in the educational context, namely, information about courses mastered within the framework of the DP, attained internships, internships, supervisors, the topic of a graduation thesis, etc.

These data can be supplemented with data on a student's psychological characteristics, cognitive features and styles. Taking into account such student characteristics can be useful in planning and organizing his educational and extracurricular activities, for example, when designing a personalized learning path.

The second group contains the following data blocks of interim and final attestation, as well as pre-professional training of student:

Data block for *intermediate and final assessment:*

- results of intermediate assessment for all years of study;
- term exams results;
- amount of academic debts at the present moment;
- number of attempts to pass the assessment for the course.

Pre-professional training data block*:*

- topics of completed research works and data on scientific advisors;
- research publications;
- conference presentations and talks;
- experience in professional activity in organizations relevant to the field of study;
- reviews and feedback on research and graduation theses;
- experience in team projects/startups participation with record on personal contribution.

The intermediate assessment section can provide useful information about the most challenging courses for students and help to select the appropriate educational materials, taking into account their background in the field of study, motivation and ambitions. Based on such data, it is possible to create more advanced services, such as services for predicting the student's success in completing an educational program or recommendation systems [14], which can build a personalized learning path depending on preferences, abilities and previously accumulated information about the user (digital footprints and digital educational history).

The research interests of the student and his or her readiness for research activity can be judged based on pre-professional training data. This is important when choosing a base for internships, or selection of potential job places for future graduates. Data on the scientific supervisor, reviews and feedback can provide students with guidance on choosing research topics, the relevance of these topics, and prospects for further development or career opportunities in the science and technology sector.

We move on to the third group of data comprising educational course data. This includes:

- number of accesses to the LMS and total time spent in it;
- number of transitions and number of clicks within the LMS;
- number of viewed educational videos and their viewing duration;
- participation in discussions on educational forums;
- number of references to external educational sources, total time spent on mastering educational material;
- success in completing assignments (accuracy, timeliness, independence level).

The data of the third group characterize the student's performance within a particular course. Currently, most educational courses are presented in electronic form in the university LMS (typically Moodle-based; a detailed list of LMS can be found on Wikipedia[7,8]). Course structure,

[7] https://en.wikipedia.org/wiki/List_of_learning_management_systems. Accessed January 15, 2024.

[8] Roskomnadzor: the foreign owner of the resource violates the law of the Russian Federation.

the content of assessment materials, evaluation criteria, and deadlines for assignments are generally determined by the course instructor on the basis of the approved course syllabus.

The course instructor, using access to the event log and gradebook, can obtain data on the student's activity within the course. It is also possible to get such data as study time, number of accesses to educational materials, educational video viewing time, and results of interim testing. In future, it will be possible to use these or other indicators to track the dynamics of academic performance within a course.

To summarize, the data in the first group usually remain unchanged or change slowly over time, while the data of the second group are updated on a regular basis, typically at least once per academic term. The data of the third group are updated most frequently, usually weekly.

3.2. Data on tertiary teachers (course instructors)

The next key subject of the educational process is the teacher, whose qualifications, experience and pedagogical skills largely determine the effectiveness of knowledge delivery and consequent effectiveness of the educational process. The author's interpretation of a particular course in the curriculum influences the motivation of students and their involvement, knowledge and skills gained as well as the learning outcomes of the course [15]. Unlike student data, on which most researchers of learning analytics are focused, the exploration of teacher data has drawn much less attention. However, interest in this topic has recently started to grow (see, for example, [16]).

Teachers undergo regular performance appraisals, for example, as part of employment or competitive election. For this procedure, the applicant prepares a list of his or her achievements for a certain period in the form approved by the organization where he/she plans (to continue) to work. However, data submitted in this way may contain inaccuracies; for example, there may be errors in their design. Most importantly, such data are usually used once and are not stored for future use. In addition, the preparation of such reports is time-consuming and routine. On the other hand, a properly organized data collection on the teacher's achievements in different areas will make it possible to receive such reports automatically without time and labor costs, while significantly reducing the risks of inaccurate or erroneous information. From a manager's point of view, a digital service for working with faculty data can provide opportunities to monitor the achievements not only of an individual employee, but also to generate summary reports on teams (employees of a department, institute, members of a scientific team) or groups (3rd-year students, PhDs under the age of 35, etc.), as well as the dynamics of employee- or team achievements over a certain period.

For the convenience of assessment, monitoring and timely adjustment of the educational process in the context of the teacher, it is proposed to consider the data on the teacher by analogy with that of students as multidimensional, structured, dynamically updated data on professional competence, communication skills, as well as digital competence and digital culture, including a group of personal data and groups of indicators. Each indicator is considered with respect to the information that can be obtained in the current moment—static data and dynamic data changing over time under the influence of some external factors, experience, or issued recommendations.

The first block of data contains *basic personal and professional information*:

- basic personal data (surname, name, patronymic, date and place of birth, gender, data of basic documents);
- socio-demographic data (marital status, family members, average income);
- health data (in relation to labor duties);
- data on education, academic degrees and titles;
- data on previous employment and positions held;
- data on mother-tongue and foreign language proficiency;
- data on teaching experience;
- profile data in scientometric databases and professionally-oriented social networks.

This block, which contains a standardized set of data that does not change or changes slowly over time, can be used to form a picture of the main stages of work history, e.g., relating to periods of employment.

The following blocks describe the teacher's competence as researcher, mentor, or practitioner.

The *teaching and learning competence* data block contains:

- data on courses taught (linked to academic years and semesters, indicating the type of classes and form of delivery);
- data on the number of students in the courses taught;
- data on developed teaching and learning materials and e-courses;
- data on external resources used in teaching practice;
- developed information content used in teaching activities;
- data on completed professional development / retraining / vocational training courses;
- data on student assessment of the teacher (if available);
- student attendance in the course during the semester;
- student academic performance within the course and the results of the term exam.

The *scientific research competency* data block includes:

- data on publications and intellectual property objects;
- data on participation in projects supported by research grants or carried out within the framework of contractual works, as a manager or contractor;
- data on membership in dissertation councils, editorial boards of scientific journals;
- data on work performance as an expert.

The *mentoring competency* data block contains:

- data on scientific supervision of graduation theses / dissertations of students / master's degree students;
- data on scientific supervision of postgraduate students / scientific advising of doctoral students;
- data on publications of students/postgraduate students co-authored with an academic staff member;
- data on graduation theses / candidate/doctoral dissertations, defended under the supervision of an academic staff member;
- data on students/postgraduate students who won prizes in competitions/contests/conferences (with indication of the event level).

The *practical competence* data block includes:

- data on work experience in organizations/enterprises in the field of training in which the teaching is provided (indicating places of work, positions held, main duties performed);
- data on developed cases drawn from practical experience;
- data on practice-oriented tasks developed for competitions/hackathons;
- data on expert experience in the professional field.

In addition to these blocks, it is also appropriate to introduce an additional data block, where an employee could provide other information indicating his or her professional qualifications, for example, awards received, competitions won, interuniversity and international team participation, consortia, organization of events, presentations of popular science lectures, experience in preparing and conducting trainings, business games, etc. These data can be subsequently classified into separate blocks to facilitate their effective application.

These blocks can be supplemented by a block of communicative and other soft skills, which can include the ability to deliver the course according to the audience preparation and proficiency (adaptability), motivation building and critical thinking skills, the ability to build productive interpersonal relationships within the team. However, it should be noted that the measurement of such skills requires special consideration. The level of communicative competence can be evidenced, for example, by the results of student evaluation of the instructor, the number of co-authors in publications, or participation in collective projects.

The block of teacher digital competence and digital culture requires a separate detailed exploration. It needs to consider the proficiency of using various digital tools in teaching and research, as well as the acceptance of changes brought by digitalization and the readiness to implement these changes. As faculty digital culture is an integral part of the corporate culture, its development becomes a significant issue of change management in the university.

A data collection and storage system should be capable of automatically retrieving data as they become available. An example is the retrieval of data from scientific metrics systems on published articles. Where automatic retrieval is impossible, the system should provide standardized data entry with supporting documents or references.

3.3. Degree program data

DP is one of the most significant components of the educational process and its quality directly affects its popularity among students and applicants. Insufficiently high indicators of the educational process (e.g., low academic performance, high percentages of students who change a DP over the course of training, low employment rates) can represent the evidence not only of insufficiently trained students, but also problems with the program itself. The most acute of these is the low demand for the program during the admission campaign and consequently low competition among the applicants. This could result in enrollment of students with lower exam scores who tend to be less motivated. This, in turn, entails further problems with their learning within the program. Low demand for DP typically has two reasons behind it: external (obsolescence, i.e., insufficient compliance with the changing demands of the labor market) and internal (inconsistency of the content and structure of the program with the stated learning outcomes). Despite the fact that the educational community—including DP developers themselves—recognizes the existence of the mentioned problems, there is currently no methodology to clearly describe an algorithm for DP analysis and evaluation, as well as identifying its weaknesses and correcting them.

In this regard, it is important to define a set of internal DP indicators for characterizing the features of its structure and content, as well as external DP indicators, which form a basis for judging DP quality and demand. Each indicator then needs to be associated with a set of data on which basis it will be calculated.

The main *internal characteristics* of the DP can be considered as follows:

- field of study, including level of education, form of education;

- requirements of the federal state educational standard or the educational standard of the educational organization, including learning outcomes expressed in term of competencies;
- professional standards, for which the DP trains students, including formed professional competencies;
- the structure of an DP, defined by its curriculum and including a set of disciplines and practices, as well as their labor intensity and mastering schedule;
- DP development team;
- Personnel implementing the DP;
- material and technical support for the DP;
- other characteristics, such as involvement of employer representatives in the learning process, availability of real-life tasks (cases) adapted to the learning process, geography range of internship sites, etc.

These characteristics, which determine the design, structure, and content of the DP, have a direct impact on its quality. However, in order to assess the quality of the DP, it is necessary to evaluate external indicators independent of its internal content, as they indicate how much external stakeholders (graduates, their employers, various administrative bodies of regional and federal levels) are satisfied with DP quality.

Among the main *external indicators* characterizing the quality of a DP, we note the following:

- distribution of Unified State Exam scores of applicants admitted to the DP, including their average and passing scores;
- proportion of entrants to the DP who have particular achievements (winners and prize-winners of academic olympiads and competitions at various levels);
- indicators related to the results of the intermediate certification: grade distribution records by course within a specific term / examination period, dynamics of the proportion of students who have a certain level of academic debt, etc.;
- student academic status updates: dynamics of dropout rates, academic leaves, changes of degree program (major) or higher education institution, dynamics of re-enrollments, transfer to DPs from other DPs within the university, from other universities, etc.;
- share of students (by year of study) who have achievements in research activity (presentations at scientific conferences, publications, intellectual property objects);
- share of students (by year of study) who have achievements in professional activities, noted by potential employers as a result of internships, participation in project work, case studies, etc.;
- employment rate of graduates in organizations relevant to their field of study;
- results of surveys of students regarding their satisfaction with the components of the educational process within the DP;
- results of surveys of graduates regarding their satisfaction with training within the DP, employment opportunities, competitiveness in the labor market, etc.;
- results of surveys of employers regarding their satisfaction with the qualification of graduates, interaction with the university, development team, course instructors implementing the DP.

Note that the development of a methodology for working with the above indicators requires special attention, since their absolute values (as they are) do not carry much meaning in terms of determining the quality of DP. The following approaches can prove useful

- consideration of these indicators in dynamics over a certain period of time;
- consideration of relative values of the indicators compared to similar indicators from other degree programs within one or similar groups of fields of study within the university;
- consideration of relative values of the indicators compared to similar indicators of degree programs in the same field of study at other universities, ranging from the nearest regional competitors to all universities in the country implementing programs in this field.

The development of methods for analyzing and assessing the quality of degree programs, as well as identifying their strengths and weaknesses, is a distinct and under-researched field of study. The very definition of the quality of a degree program is an important task in itself. We suggest that the quality of a degree program be defined in terms of achieving the desired learning outcomes, expressed in terms of competencies a graduate should possess upon successful completion of the program. However, measuring the level of possession of competencies, or professional skills of graduates are nontrivial tasks that do not have a straightforward solution. It should also be understood that a degree program is not an unchanging set of documents, but rather a dynamically evolving entity that connects students, faculty and teaching staff, administrative personnel, employer representatives, and other stakeholders within the educational process. As a key component of the educational process, DP is an important object of study within the framework of data-driven educational process management. In the context of a rapidly changing labor market, the tasks of creating tools for analyzing DPs, managing and optimizing their portfolio within the university are highly relevant.

4. DISCUSSION

Processes of digitalization in higher education institutions are currently implemented mainly through the creation of specialized systems and services which

function within particular areas and are aimed at solving a limited range of tasks [17]. One of the main disadvantages of this approach is that such services usually do not include protocols of interaction with each other, since their operation is often based on data stored in databases specially designed for these systems. This leads to the problems described in [10], namely, data duplication and redundancy, as well as absence of unified storage standards and consequent conflicts in data exchange between different information systems. The inability to integrate these systems as components of a coherent whole presents a challenge to addressing higher-level objectives, such as, for instance, determining the root causes of a decline in student retention rate in a given DP or the increase of student attrition from that program and their subsequent enrollment in other programs. We believe that the use of a single information platform, referred to as the "core", which instantiates the principles for data collection, storage, and processing for all university information systems, can help address the aforementioned challenges. Such systems would represent modules connected to the core. In addition, the sequential connection of new modules, modified according to end user requests, will significantly improve the transparency of internal processes, reduce labor and time expenditures, and increase the speed and accuracy of decision-making.

It should be noted that the core data elements of students, faculty, and DPs discussed in this work are interconnected in numerous ways. For example, the outcomes of a student's mid-term assessment in a subject are not solely determined by the student themselves, their motivation, level of preparation, and assignments completed during the term. Rather, they depend on the qualifications of the teachers who taught this subject, and the level of their expectations, methods, formats, teaching materials, as well as the structure of the course (in particular, the total workload and the distribution of workload between contact hours and independent study), as outlined in its syllabus. Additionally, if the course content is based on knowledge gained in other subjects of the educational program, then the structure of the course, as determined by the curriculum, also plays a significant role. In practice, there are instances where the skills and knowledge acquired by students in one subject are applied in another, with both subjects being taught concurrently in a single semester, which can create understandable challenges for students. Thus, the purpose of data-driven educational process management is to, among other things, simplify the process of accessing a diverse range of data related to the education process as much as possible, and to develop data analysis and visualization tools to facilitate the identification of areas for improvement in the educational process with a view to further optimizing it.

The next significant consideration is that the development, customization, and technical maintenance of such a system must be undertaken throughout its entire lifecycle. Both the development and subsequent refinement or modification of individual modules should occur in close collaboration with end users, based on the systematic collection of feedback. Having a team consisting of representatives from various roles, including developers, technical support personnel, senior management, and representatives from all groups of users and stakeholders, can help to make the process of evolving such a system more effective. In addition, it is essential that the process of gathering feedback and making subsequent changes to the system is regular and timely. Otherwise, the positive impact of the process may be significantly diminished, which could in turn negatively affect the overall effectiveness of the system. In this regard, using proprietary software developed by third-party vendors may offer less practical benefit to the organization compared to developing its own software in-house by a dedicated team of developers.

Finally, we would like to highlight another aspect that plays a crucial role in the successful implementation and development of a data-driven educational process management system. This aspect involves the adoption of the system by all relevant stakeholders and their willingness to contribute to its improvement and enhancement. The introduction of such a system has the potential to significantly enhance the transparency of various aspects of an employee's performance, which may lead to at least a mixed reaction towards the system and decisions made based on data generated by it. A very important point here is the interpretation of the data and indicators obtained. In view of the above, building a system of decision-making in relation to participants in the education process in the context of rewarding them for their achievements and supporting them in eliminating problem areas and realizing potential points of growth will be an important component of the successful implementation and further development of such a system.

CONCLUSIONS

In the present article, we propose a conceptual approach to the digitalization of the educational process in a higher education institution, describing its main stages, as well as providing data blocks for the key participants and components of the educational process: students, teaching staff and DPs. We highlight certain features and challenges associated with the development and implementation of information systems and digital services for building a data-driven educational process management system.

In subsequent works, the authors plan to study in detail each component of the proposed conceptual

approach, analyzing the specifics of collecting and assessing indicators related to students, teaching staff, and DPs.

Digitalization of the educational process represents an essential stage in the development of a university, without whose implementation it will be impossible to reach a significantly new level of educational activity. The continuously growing market of educational services, especially in terms of continuing professional education, puts universities in a position of catching-up as compared to online educational platforms, which have significantly excelled in the use of digital tools for Learning Analytics. Consequently, the development of a conceptual approach to create a digital system of educational process management in the university becomes a priority task, the quality of which will greatly influence the future development and competitiveness of the university.

Authors' contributions

A.A. Kytmanov—conceptualization, methodology, writing (original draft preparation, review, and editing), supervision.

Yu.N. Gorelova—formal analysis, writing (original draft preparation, review, and editing).

T.V. Zykova—methodology, formal analysis.

O.A. Pikhtilkova—formal analysis, writing (original draft preparation).

E.V. Pronina—formal analysis, writing (original draft preparation).

All authors have read and agreed to the published version of the manuscript.

REFERENCES

1. Selwyn N., Gašević D. The datafication of higher education: discussing the promises and problems. *Teach. High. Educ.* 2020;25(4):527–540. https://doi.org/10.1080/13562517.2019.1689388
2. Williamson B., Bayne S., Shay S. The datafication of teaching in Higher Education: critical issues and perspectives. *Teach. High. Educ*. 2020;25(4):351–365. https://doi.org/10.1080/13562517.2020.1748811
3. Taamneh A., Alsaad A., Elrehail H., Al-Okaily M., Lutfi A., Sergio R.P. University lecturers acceptance of moodle platform in the context of the COVID-19 pandemic. *Global Knowledge, Memory and Communication.* 2023;72(6/7):666–684. http://doi.org/10.1108/GKMC-05-2021-0087
4. Lemay D.J., Baek C., Doleck T. Comparison of learning analytics and educational data mining: A topic modeling approach. *Computers and Education: Artificial Intelligence.* 2021;2(1):100016. https://doi.org/10.1016/j.caeai.2021.100016
5. Papamitsiou Z., Economides A.A. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *J. Educ. Technol. Soc.* 2014;17(4):49–64. Available from URL: http://www.jstor.org/stable/jeductechsoci.17.4.49
6. Arnold K.E., Pistilli M.D. Course Signals at Purdue: Using learning analytics to increase student success. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (*LAK'12*). 2012. P. 267–270. Available from URL: https://dl.acm.org/doi/10.1145/2330601.2330666
7. Teasley S.D., Kay M., Elkins S., Hammond J. User-Centered Design for a Student-Facing Dashboard Grounded in Learning Theory. In: Sahin M., Ifenthaler D. (Eds.). *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*. Cham., Switzerland: Springer; 2021. P. 191–212. https://www.doi.org/10.1007/978-3-030-81222-5_9
8. Klein C., Lester J., Rangwala H., et al. Technological barriers and incentives to learning analytics adoption in higher education: insights from users. *J. Comput. High. Educ*. 2019;31:604–625. https://doi.org/10.1007/s12528-019-09210-5
9. Talbi O., Ouared A. Goal-oriented student motivation in learning analytics: How can a requirements-driven approach help? *Educ. Inf. Technol.* 2022;27(8):12083–12121. https://doi.org/10.1007/s10639-022-11091-8
10. Kustitskaya T.A., Esin R.V., Kytmanov A.A., Zykova T.V. Designing an Education Database in a Higher Education Institution for the Data-Driven Management of the Educational Process. *Educ. Sci.* 2023;13(9):947. https://doi.org/10.3390/educsci13090947
11. Datnow A., Hubbard L. Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *J. Educ. Change*. 2016;17(1):7–28. https://doi.org/10.1007/s10833-015-9264-2
12. Brown M. Seeing students at scale: how faculty in large lecture courses act upon learning analytics dashboard data. *Teach. High. Educ*. 2020;25(4):384–400. https://doi.org/10.1080/13562517.2019.1698540
13. Esin R.V., Zykova T.V., Kustitskaya T.A., Kytmanov A.A. Digital educational history as a component of the digital student's profile in the context of education transformation. *Perspektivy nauki i obrazovaniya = Perspectives of Science & Education.* 2022;5(59):566–584 (in Russ.). https://doi.org/10.32744/pse.2022.5.34
14. Permana A.A.J., Pradnyana G.A. Recommendation Systems for internship place using artificial intelligence based on competence. *J. Phys.: Conf. Ser.* 2019;1165:012007. https://doi.org/10.1088/1742-6596/1165/1/012007
15. Raffaghelli J.E., Stewart B. Centering complexity in 'educators' data literacy' to support future practices in faculty development: a systematic review of the literature. *Teach. High. Educ*. 2020;25(4):435–455. https://doi.org/10.1080/13562517.2019.1696301

16. Lewis S., Holloway J. Datafying the teaching 'profession': remaking the professional teacher in the image of data. *Cambridge J. Educ.* 2019;49(1):35–51. https://doi.org/10.1080/0305764X.2018.1441373
17. Williamson B. The hidden architecture of higher education: building a big data infrastructure for the 'smarter university'. *Int. J. Educ. Technol. High. Educ*. 2018;15:12. https://doi.org/10.1186/s41239-018-0094-1

About the authors

Alexey A. Kytmanov, Dr. Sci. (Phys.-Math.), Head of the Department of Higher Mathematics – 3, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kytmanov@mirea.ru. Scopus Author ID 6602129708, RSCI SPIN-code 6866-6079, https://orcid.org/0000-0003-3325-099X

Yuliya N. Gorelova, Cand. Sci. (Phil.), Head of the Master's Center, Institute of Management, Economics and Finance, Kazan (Volga Region) Federal University (18, Kremlevskaya ul., Kazan, 420008 Republic of Tatarstan, Russia). E-mail: JNGorelova@kpfu.ru. Scopus Author ID 56521686700, RSCI SPIN-code 9505-0346, http://orcid.org/0000-0003-1114-9751

Tatiana V. Zykova, Cand. Sci. (Phys.-Math.), Assistant Professor, Department of Applied Mathematics and Data Science, School of Space and Information Technology, Siberian Federal University (79, Svobodnyi pr., Krasnoyarsk, 660041 Russia). E-mail: tzykova@sfu-kras.ru. Scopus Author ID 57188699496, RSCI SPIN-code 1959-9769, https://orcid.org/0000-0002-7332-2372

Olga A. Pikhtilkova, Cand. Sci. (Phys.-Math.), Associate Professor, Department of Higher Mathematics – 3, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: pihtilkova@mirea.ru. RSCI SPIN-code 5589-7411, https://orcid.org/0009-0004-4632-5158

Elena V. Pronina, Cand. Sci. (Phys.-Math.), Assistant Professor, Department of Higher Mathematics – 3, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: pronina@mirea.ru. RSCI SPIN-code 3391-3440, https://orcid.org/0000-0002-2447-7175

Об авторах

Кытманов Алексей Александрович, д.ф.-м.н., доцент, заведующий кафедрой высшей математики – 3, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kytmanov@mirea.ru. Scopus Author ID 6602129708, SPIN-код РИНЦ 6866-6079, https://orcid.org/0000-0003-3325-099X

Горелова Юлия Николаевна, к.фил.н., доцент, заведующий центром магистратуры, Институт управления, экономики и финансов, ФГАОУ ВО «Казанский (Приволжский) федеральный университет» (420008, Россия, Республика Татарстан, Казань, ул. Кремлевская, д. 18, корп. 1). E-mail: JNGorelova@kpfu.ru. Scopus Author ID 56521686700, SPIN-код РИНЦ 9505-0346, http://orcid.org/0000-0003-1114-9751

Зыкова Татьяна Викторовна, к.ф.-м.н., доцент, кафедра прикладной математики и анализа данных, Институт космических и информационных технологий, ФГАОУ ВО «Сибирский федеральный университет» (660041, Красноярск, пр. Свободный, д. 79). E-mail: tzykova@sfu-kras.ru. Scopus Author ID 57188699496, SPIN-код РИНЦ 1959-9769, https://orcid.org/0000-0002-7332-2372

Пихтилькова Ольга Александровна, к.ф.-м.н., доцент, кафедра высшей математики – 3, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: pihtilkova@mirea.ru. SPIN-код РИНЦ 5589-7411, https://orcid.org/0009-0004-4632-5158

Пронина Елена Владиславовна, к.ф.-м.н., доцент, кафедра высшей математики – 3, Институт перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: pronina@mirea.ru. SPIN-код РИНЦ 3391-3440, https://orcid.org/0000-0002-2447-7175

Translated from Russian into English by Lyudmila O. Bychkova, Alexey A. Kytmanov, and Yuliya N. Gorelova
Edited for English language and spelling by Thomas A. Beavitt

MIREA – Russian Technological University.
78, Vernadskogo pr., Moscow, 119454 Russian Federation.
Publication date September 27, 2024.
Not for sale.

МИРЭА – Российский технологический университет.
119454, РФ, г. Москва, пр-т Вернадского, д. 78.
Дата опубликования 27.09.2024 г.
Не для продажи.