

Modified Funk SVD-Augmented Recommender Systems for Advancing Visual Data Processing in Industrial IoT

Olena Hordiichuk-Bublivska^{1,*†}, Halyna Beshley^{1,2,†}, Iryna Ivanochko^{2,†}, Mykola Beshley^{1,2,†} and Orest Kochan^{1,*†}

¹ Lviv Polytechnic National University, Bandera Str. 12, Lviv, 79013, Ukraine

² Comenius University in Bratislava, 82005 Bratislava 25, Slovakia

Abstract

Industrial recommender systems are currently an active area of research and development, with many businesses across various industries leveraging them to improve their customer experience and increase their revenue. The growth of 5G industrial networks has led to a massive increase in the volume of visual data (images, video information) generated from various sources, making it challenging to process and analyze the data. Additionally, traditional recommendation systems used in these networks may not be able to handle the massive amount of data and provide accurate recommendations. This article proposes a modified Funk Singular-Value Decomposition (SVD) approach for enhancing collaborative filtering in recommendation systems for 5G industrial networks. The proposed approach can effectively reduce the dimensionality of the data and capture the underlying patterns and relationships between users and items, thereby enhancing the performance of the recommendation system. According to the study results, it was determined that when using less data about users, the speed of providing recommendations increases. It was also established that the accuracy of calculations improves when additional item features are included. The simultaneous use of two modifications allows to improve the accuracy of providing recommendations by 2%, as well as to reduce the duration of calculations by an average of 10-15%. The proposed modifications of the FunkSVD algorithm can improve the level of providing user recommendations according to their requirements. The research results can be used to optimize the operation of industrial and 5G systems with dynamic parameter changes.

Keywords

recommender systems, collaborative filtering, big data, funk singular-value decomposition algorithm, 5G industrial network paper

BAIT'2024: The 1st International Workshop on "Bioinformatics and applied information technologies", October 02-04, 2024, Zboriv, Ukraine

* Corresponding author.

† These authors contributed equally.

✉ obublivska@gmail.com (O. Hordiichuk-Bublivska); halyna.v.beshlei@lpnu.ua (H. Beshley); irene.ivanochko@gmail.com (I. Ivanochko); mykola.i.beshlei@lpnu.ua (M. Beshley); orestvk@gmail.com (O. Kochan)

ORCID: 0000-0002-6439-549X (O. Hordiichuk-Bublivska); 0000-0001-5392-3499 (H. Beshley); 0000-0002-1936-968X (I. Ivanochko); 0000-0002-7122-2319 (M. Beshley); 0000-0002-3164-3821 (O. Kochan)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Introduction

The modernization of industrial systems has brought significant changes to the way production processes are managed. The Industry 4.0 concept has enabled the digitization of manufacturing, leading to the implementation of modern information systems that solve several problems to deliver the appropriate quality of service. In the industrial system, the problem of data collection from various end devices, storage, and analysis for further decision-making must be solved [1]. The quality of user service is one of the most important criteria for evaluating the efficiency of industrial systems. Defined characteristics of service provision for one user can be used to improve it for others. The more sources of information are analyzed, the easier it is to determine general patterns and possible problems in the operation of industrial systems. At the same time, the overabundance of data constantly arriving from various end devices is not always critical for computing. Data analysis and optimization methods should be used [2].

Storing and processing data received from various distributed sources is extremely important for modern industrial systems. Cloud technologies are used to improve the efficiency of big data analysis. Cloud manufacturing (CMfg) is an important component of the Industry 4.0 concept [3]. CMfg allows an industrial system to monitor its resources and quickly access them. Also, users can exchange data with other and global control nodes, saving them on cloud servers [4].

Recommender systems are widely used in various industries, including industrial networks, to provide personalized recommendations to users [5]. In industrial networks, recommender systems can be used to recommend products, services, and suppliers to businesses, based on their past purchases, interests, and behavior [6]. Overall, recommender systems have the potential to provide significant benefits to businesses in industrial networks by improving the efficiency of purchasing and reducing the time and effort required to find suppliers and products. In an industrial setting, recommender systems can be used in various ways, such as recommending maintenance schedules for machines, optimizing production processes, recommending suppliers for raw materials, and more. With the high bandwidth and low latency of 5G networks, these systems can provide real-time recommendations to users, allowing them to make decisions quickly and efficiently [7]. However, implementing recommender systems in this context requires careful consideration of the unique challenges and limitations of industrial networks. Since most endpoints in the Industrial Internet of Things (IIoT) are located remotely from each other, data is transmitted over wireless communication channels. 5G mobile communication systems allow serving users quickly and efficiently, providing them with a wide range of communication services [8, 9]. It often happens that the amount of information significantly exceeds the 5G bandwidth, which leads to delays or the need to reduce the quality of the transmitted content. Thus, it is necessary to solve the problem of data transmission to users, data optimization, resource allocation and system load.

It is also important to consider the specific needs of the industrial setting when developing a recommender system. This may involve identifying the most relevant data sources, selecting appropriate algorithms for data analysis, and incorporating domain-specific knowledge into the recommender system. In 5G industrial networks, there can be a large number of users and items, which can result in sparse data. Sparse data refers to situations where there are few ratings or interactions between users and items. This can

make it challenging to generate accurate recommendations using traditional collaborative filtering algorithms.

The collection of data on sales statistics and the use of services by different categories of users allows us to better understand their needs and adapt production processes to them [10-12]. The list of goods and services is constantly expanding, and the requirements for the quality of service are growing, so it is necessary to constantly look for new methods of data processing.

The novelty of this work lies in the proposal of modifications to the Funk SVD algorithm for processing big data in industrial settings, with the aim of developing effective 5G industrial recommender systems. The main contribution of the study is in two aspects. First, we propose to use only a part of the user's data to generate recommendations, which reduces the load on 5G channels and the duration of calculations, while maintaining a high level of accuracy in data processing. Second, we demonstrate the effectiveness of incorporating additional item features to improve the accuracy of recommender systems.

Overall, the proposed data optimization techniques have the potential to revolutionize the way 5G industrial networks operate, by enabling the development of more accurate and fast recommender systems. The study's results highlight the prospect of using modified FunkSVD algorithms to improve the quality of service for different types of 5G systems users, leading to increased efficiency, reduced costs, and optimized processes

The rest of the paper is organized as follows. In Section 2, we provide a literature review of collaborative filtering and its variants. In Section 3, we introduce the challenges and solutions for 5G industrial recommendation systems. In Section 4, we compare the SVD and Funk SVD algorithms for data optimization in industrial 5G recommender systems. In Section 5 and Section 6, we describe the proposed enhanced collaborative filtering approach using modifications of the Funk SVD. Section 7 contains the results and discussion, while the conclusion of the paper is presented in Section 8.

2. Related Works

We will analyze current research in this field. E.P. Xing et al. [13] defined the problem of big data processing in ML and proposed a framework for systematizing information. In [14], V. Vashishth et al. analyzed the interment of cloud technologies and IIoT and proposed a method for optimal access to remote resources. D. Tarek et al. proposed a distributed protocol for managing large traffic volumes in IoT systems [15]. In [16], S. Sennan et al. studied the problem of data clustering in IoT systems and proposed an optimization algorithm for Cluster Head (CH) selection. In [17], M.C. Adikari and T.P Amalan proposed a software model with optimal operating parameters for big data analytics and ML (ML) in FMCG (Fast Moving Consumer Goods).

In [18], B. Raviteja et al. investigated the relevance of using ML algorithms to improve the efficiency of Supply Chain Management (SCM) systems. G.K. Singh et al. in [19] considered the importance of Big Data analysis for solving business and logistics management problems. H. Fan et al. analyze the operation of IIoT systems as a component of the Industry 4.0 concept and propose a federated learning-based privacy-preserving data aggregation scheme [20]. In [21] W. Gao et al. determine the importance of applying federated learning in IIoT systems and propose a resource allocation scheme that allows for choosing the best learning devices. Liang et al. in [22] also investigate the problem of big data processing in IIoT systems and the selection of optimal end devices for training.

The authors also propose a Deep Q-Network (DQN)-based scheme for optimal resource allocation.

In [23], K. Shah et al. analyze the features of the application of recommender systems for intelligent user selection of goods and services. Z. Guo et al. in [24] investigate the use of recommender systems for IIoT. The authors offer their own implicit feedback-based group recommender system. B. Wu et al. consider the problem of determining the relationships between users and goods in IoT systems and provide a framework for improving item recommendation [25]. In [26], S.M. Kasongo investigated the importance of intrusion detection in IIoT systems and proposed their Intrusion Detection Systems (IDSs), which are more efficient and reliable than existing ones. A. Simeone et al. in [27] analyze the service of users of intelligent production and proposed the provision of personal recommendations using an intelligent decision-making recommender system.

In [28], T. Vafeiadis et al. offer a smart recommender system that analyzes data from various IoT devices and enhances decision support. Y. Fu et al. investigate the problem of intelligent resource management in 5G networks [29]. In [30], A. Vulpe et al. consider performance indicators of the LTE RAN network to detect possible malfunctions of the terminal equipment. The authors also offer their analytics framework that can be used for 5G networks. Yongchang Wang and L. Zhu [31] investigate and compare different big data optimization methods. The authors note the advantages of the SVD algorithm for determining the most important information for processing. C. Zhang et al. [32] investigate the methods of automated processing of big data and propose a feature extraction method that uses SVD and PCA (Principal component analysis) algorithms.

K. Birul offers a modified Funk SVD that uses data on all products and only one user to form recommendations [33]. S. Guo et al. consider the Funk SVD algorithm for improving the performance of recommender systems [34]. As is clear from the analyzed sources, processing big data in industrial and 5G systems is still an actual problem. However, user data analysis using recommender systems, considering dynamic changes in service quality requirements, system performance indicators, etc., have not been sufficiently researched.

In our previous work [35], we investigated improving the performance of recommender systems and proposed a modified federated SVD, which demonstrated the accuracy and reliability of the results. Extending the research, we used the FunkSVD algorithm to process sparse data in this work. We also proposed modifying the Funk SVD algorithm for its more efficient use under different system parameters and loads in 5G systems. The proposed modifications improve both the accuracy and speed of providing recommendations to users and the analytics of production processes.

3. Challenges and Solutions for 5G Industrial Recommendation Systems

5G industrial recommendation systems have the potential to revolutionize the way factories and industrial plants operate, by enabling real-time data collection and analysis, predictive maintenance, and optimized production processes. Overall, the successful implementation of 5G industrial recommendation systems requires a combination of advanced technologies, secure data transmission protocols, and efficient data processing and analysis algorithms.

For the efficient operation of ML algorithms, it is necessary to use large amounts of information. The development of information technologies and rapid digitalization contribute to the daily generation of significant data sets from various sources. Thanks to

such data, it is possible to better identify features characteristic of individual tasks to accurately determine the result. However, the excess of information still complicates the work of computational algorithms. The processing and storage of large data arrays lead to the consumption of computing resources and the slowing down of calculations. The diversity of information sources also requires the adaptation of ML algorithms and distributed data processing [36-38].

Federated learning is one of the methods of effective organization of distributed data processing using ML [39]. Instead of sending all the data from different distributed nodes to each other, they are processed on local nodes. Only the results of the calculations are sent to the controller, which updates the entire machine-learning system. Federated learning is effectively used in industrial systems, as it solves the problem of processing data collected from different end devices. This improves the reliability and privacy of users' private information, as it does not need to be transmitted over the network [40-42].

Statistical analysis is extremely important for operating commercial systems that provide users with certain goods or services. Based on the previous actions of customers, it is possible to determine how satisfied they are with the services they use. Recommender systems (RS), are used to establish relationships between users and goods or services, which are generally called items. We can often see examples of recommender systems in everyday life. When visiting websites or applications, users are offered advertisements for items that may interest them in some way. For example, when searching for a certain object on the site, the recommender system also identifies similar ones according to certain parameters. Similar products or services are designated as "you may like this," "other products you may like," etc. There are different approaches to determining the most appropriate items for users. Recommendations can be personalized, that is, determined by taking into account the characteristics of a specific user. Non-personalized recommendations common to a certain group of users can also be calculated.

For the work of recommender systems, special methods are used for processing data from users. The result of RS is suggestions to users about new products or services they are most probably to like. The most common approaches to the work of recommender systems are:

Content-based, which analyzes the similarity of different products to each other. If a user liked a certain product or service before, there is a high probability that he will highly rate similar ones in the future. The advantage of this approach is the simplicity of implementation. The disadvantage is that new and unlike other products have little chance of getting into the recommendations;

Collaborative filtering forms recommendations by analyzing user profiles. In case of finding similarities between two or more users, they are recommended products or services previously chosen by other group members. This approach demonstrates high efficiency. At the same time, preparing a recommender system in advance is necessary, providing information about users and their preferences. In the absence of information, problems arise in the formation of recommendations;

A hybrid approach to the formation of recommendations, combining the above [43].

Intelligent production systems focus not only on producing certain products but also on their sale to end users. Determining the level of interest in certain goods or services is extremely important for the system's effective functioning [44, 45]. A situation often arises when a product is not in demand among users due to its features. As a result of monitoring feedback and sales statistics, possible modifications and improvements of services and goods can be identified and immediately implemented. Automating the management of

the intelligent production system contributes to the prompt resolution of existing problems and the prevention of new ones. 5G Wireless Network is an important component of smart manufacturing. Since many final IoT devices are located at a considerable distance from each other, wireless communication allows to combine them into a single data exchange system. End users can connect to the network when they need to send or receive information. 5G technology makes it possible to provide various communication services at high speed and ensure the required quality of service. However, large amounts of data exchanged between users and the system create a significant load on communication channels and processing devices. As a result, some of the services may be of lower quality or may not be available at all.

The architecture of the industrial system using the 5G network for visual data communication from IIoT devices is shown in Fig.1.

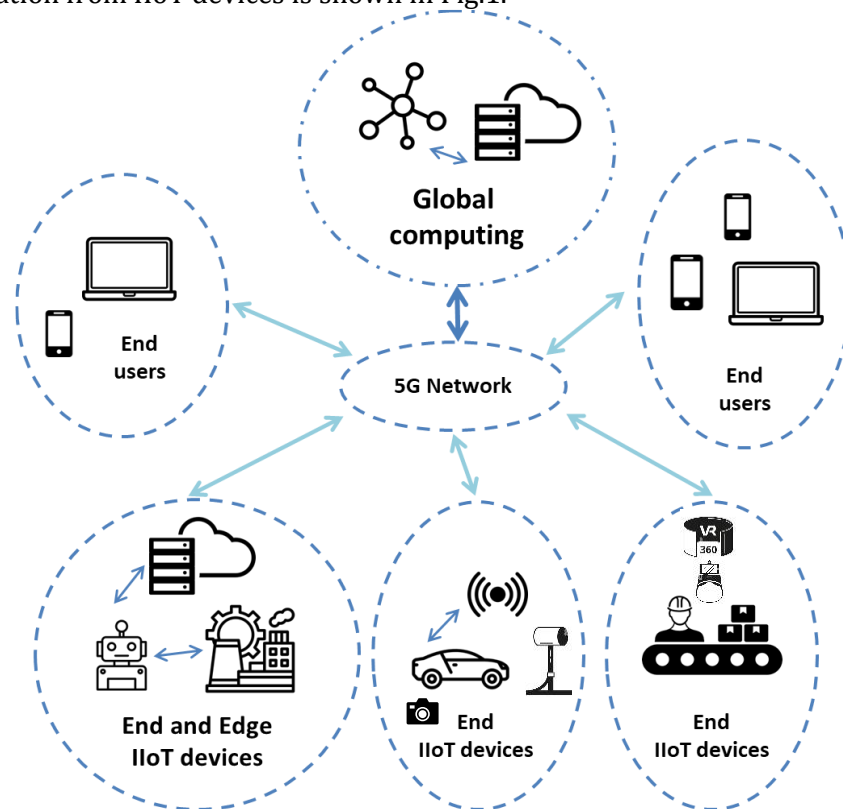


Figure 1: The architecture of the industrial system using the 5G network for visual data transmission

ML and artificial intelligence methods are used to optimize the operation of 5G networks. Analysis of the provided services and load changes depending on various parameters helps to better adjust the system parameters. ML methods make it possible to distribute network resources according to user needs and quickly solve problem situations. Thus, the mobile network can flexibly adapt to different operating conditions and determine priorities in user service. Having determined the areas of the mobile communication system that are the most loaded, it can be automatically attracted additional resources. For this purpose, constant monitoring of the load and performance indicators of the system is carried out. The process of processing user requests in 5G systems is shown in Fig.2.

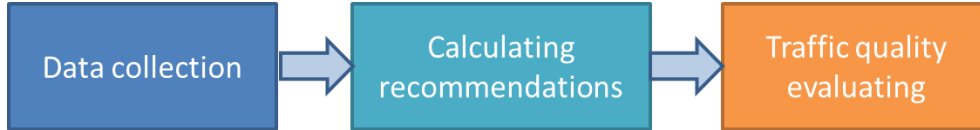


Figure 2: Processing user requests in 5G systems

The different types of traffic provided to users of 5G systems also require special attention to the provision of service priorities and allocating necessary communication resources [46]. Multimedia data transmission uses more power than voice messages. The reliability and confidentiality of user data should also be ensured when forwarding them for processing. ML methods solve the problems of big data optimization, analysis, and decision-making. Due to many end users, ML algorithms receive much information to train. However, since the data often contains redundancy and is too cumbersome to calculate, it should be pre-processed, and the most important ones should be selected.

4. Comprehensive Study of SVD and Funk SVD Algorithms for Visual Data Optimization in Industrial 5G Recommender Systems

The improvement of industrial systems made it necessary to process huge data sets. The information coming from different devices is often of various types, unstructured, and of large volume. Big data processing devices usually cannot process it efficiently, so users do not get timely results. For faster calculations, data should be pre-prepared and brought to a clear and concise form. Optimizing big data and discarding redundancy is one of the most relevant for modern industrial systems.

ML, Deep Learning, and Data Mining methods are used for intelligent information processing. There are many ways to transform data of a higher dimension into a smaller one while preserving its properties with a certain level of accuracy. For example, one of the most popular methods is Principal component analysis (PCA), which solves the problem of dimensionality reduction while preserving as many properties of the original data as possible. PCA represents data sequences as a set of interdependent variables, the principal components (PCs). Then, the most informative variables are searched, and unimportant ones are discarded. The Latent semantic analysis (LSA) algorithm is similar to PCA, but it is used more to analyze text sequences. Both methods can be extended to the Singular value decomposition (SVD) algorithm, an improved way to identify the most important data and discard redundancy.

SVD decomposes the original matrix into the product of three mutually orthogonal submatrices. This method is optimal for processing and optimizing matrices that are not rectangular. Compared to the PCA algorithm, SVD more effectively determines the fundamental properties in data arrays and considers their diversity. Singular Value Decomposition involves representing the initial data matrix A (m, n) as:

$$A = P \times \Delta \times Q, \quad (1)$$

where matrix P has dimension (m, m) , matrix Q has dimension (n, n) , matrix Δ has dimension (m, n) and demonstrates the relationship between P and Q .

SVD is widely used for recommender systems and has repeatedly proven its effectiveness [31]. However, in real systems where large data from various devices are processed, several problems arise. First, the information should be brought to a form

suitable for forming recommendations. Thus, all data on the interaction of users and items should be presented as a numerical score. Methods of pre-processing and optimization of information are used to solve this problem. It is also important to consider that only under ideal conditions we can get almost all the data on the evaluations of each item from all users and form the recommendations matrix. In such a table, the rows correspond to users, the columns correspond to items, and at the intersection, there are ratings, that is, evaluations.

In real recommender systems, we receive only part of the information about users' interest in certain products. Many products have not yet received customer reviews, and recently registered users do not have a purchase history. Recommender systems are designed to fill empty cells in recommendation tables but should also process data more compactly and efficiently. Recommendation tables usually have dimensions (m,n), where m is the number of users, and n is the items. Due to the sparseness of the data matrices, it does not make sense to process them all since they do not carry information for forming recommendations but use a lot of computing resources.

For determining the most relevant services for specific categories, recommender systems should be used that establish correspondences between users and products. Thus, the efficiency of service delivery is improved. 5G recommender systems have the problem of processing extremely large data, which are often not extremely important for forming recommendations. Recommendation matrices look like this:

$$A(n,m) = \begin{matrix} & a_{11} & N/A & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & N/A & \\ \dots & \dots & \dots & \dots & \\ a_{n1} & N/A & \dots & a_{nm} & \end{matrix}, \quad (2)$$

where a_{ij} is a value of interaction between user i and item j , N/A is a data Not Available or unknown.

An excess of information leads to a slowing down of calculations and an overloading of devices. Sparse data matrices are better optimized by discarding empty cells or those containing unimportant information. For sparse data processing, the paper proposes to improve the existing Funk SVD algorithm, which decomposes the initial matrix A into the product of two submatrices [34, 35]:

$$A(n,m) = M(n,k) \times N(k,m)^T, \quad (3)$$

where $k < n$ and $k < m$.

The FunkSVD algorithm uses the Stochastic Gradient Descent method to gradually reduce the error between the original matrix and the resulting one. Thus, it is possible to determine the Sum of the Squared Error for each decomposition:

$$\min \sum_{i \in n, j \in m} (a_{i,j} - m_i \times n_j^T)^2 + (|m_i|^2 + |n_j^T|^2), \quad (4)$$

where $a_{i,j}$ is the elements of the original recommendation matrix, n_i , m_j is the elements of matrices N and M accordingly.

The values of the element of the initial matrix $a'_{i,j}$ are obtained by multiplying the rows and columns with the corresponding indices:

$$a'_{i,j} = \sum_{i,j=k}(m_i \times n_j^T), \quad (5)$$

The calculation error can be defined as follows:

$$\text{Err} = a_{i,j} - a'_{i,j} \quad (6)$$

We can update the matrix M element, using learning coefficient ε and correction coefficient θ :

$$m'_i = m_i + \varepsilon(\text{Err} \times 2 \times n_j + \theta \times m_i), \quad (7)$$

Let's also calculate the element of the matrix N:

$$n'_j = n_j + \varepsilon(\text{Err} \times 2 \times m_i + \theta \times n_j). \quad (8)$$

Then the users recommendation is determined as:

$$a'_{i,j} = m'_i \times n_j'^T + d_{i,j}, \quad (9)$$

$d_{i,j}$ is the total coefficient of deviation of user and product indicators from the total:

$$d_{i,j} = d_i + d_j + \Delta, \quad (10)$$

where d_i is the deviation from the average value for the user, d_j the same for item, Δ is the regularization factor.

The accuracy and efficiency of the recommender system can be determined by presenting its confusion matrix (Table 1):

Table 1

Confusion matrix for recommendation systems

| | Actually Interested | Actually Not Interested |
|-----------------|------------------------------|------------------------------|
| Recommended | $K_{\text{true_positive}}$ | $K_{\text{false_positive}}$ |
| Not Recommended | $K_{\text{false_negative}}$ | $K_{\text{true_negative}}$ |

According to Table 1, $K_{\text{true_positive}}$ is the probability that the recommended item interested the user;

$K_{\text{false_positive}}$ is the corresponds to the probability of a positive recommendation for an uninteresting item;

$K_{\text{false_negative}}$ is the probability of the negative recommendation for the interesting item;

$K_{\text{true_negative}}$ is the probability of the negative recommendation for the uninteresting item.

The probability of giving a positive recommendation if the item is not interesting to the user can be represented as:

$$K_{\text{true_positive_rate}} = \frac{K_{\text{true_positive}}}{K_{\text{true_positive}} + K_{\text{false_negative}}}, \quad (11)$$

Probability of giving a negative recommendation if the product is not interesting to the user:

$$K_{\text{true_negative_rate}} = \frac{K_{\text{true_negative}}}{K_{\text{true_negative}} + K_{\text{false_positive}}}, \quad (12)$$

The accuracy of providing recommendations can be calculated:

$$K_{\text{positive_predictive_value}} = \frac{K_{\text{true_positive}}}{K_{\text{true_positive}} + K_{\text{false_positive}}}. \quad (13)$$

However, to evaluate the effectiveness of the FunkSVD algorithm, recommendations can be directly determined Mean Absolute Error (MAE) for matrix A with dimension N:

$$K_{\text{mean_absolute_error}} = \frac{\sum_{i,j=1}^N |a_{i,j} - a'_{i,j}|}{|N|}. \quad (14)$$

In the work, we use the Root Mean Square Error, (RMS Error, RMSE) to check the accuracy of the recommendation calculations:

$$K_{\text{rmse}} (\text{Accuracy}) = \sqrt{\frac{\sum_{i,j=1}^N (a_{i,j} - a'_{i,j})^2}{|N|}}. \quad (15)$$

For studying the effectiveness of data processing by the FunkSVD algorithm, a program model was created in the Python programming language, and data from an open recommender system was used. For a better understanding of the algorithm's operation, data of different volumes and with different levels of sparseness were extracted.

A comparison of data calculation durations by SVD and Funk SVD algorithms is shown in Fig. 3. Data matrices with different degrees of sparseness, i.e., cells filled with information, were studied.

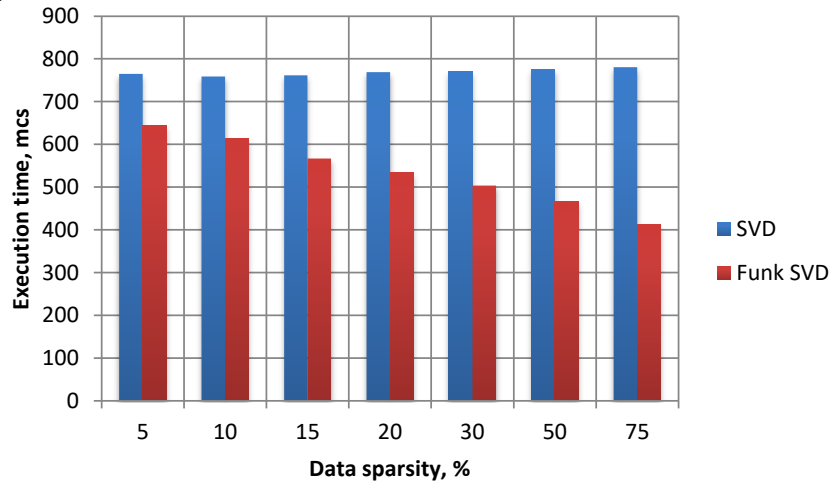


Figure 3: The comparison of execution time by SVD and Funk SVD algorithms

As we can see from Fig. 3, SVD works longer, while the difference with Funk SVD increases with the increase in the percentage of data sparsity. If the system processes a lot of redundant data, Funk SVD allows to discard part of it and speed up the calculation.

5. Modified Funk SVD Approach for Improving the execution time of Recommendation Systems

Although FunkSVD works quite well with sparse data, it can still be improved to speed up the calculations. According to eq. (5) the Funk SVD algorithm uses data about all users and products to determine the unknown cell value of the initial data matrix. In the paper, we proposed using different number of users δ in the calculations to modify the existing Funk SVD algorithm:

$$a'_{i,j} = \sum_{i=\delta, j=k} (m_i \times n_j^T), \quad (16)$$

where $\delta < k$.

The number of users whose data is considered in the calculations is chosen arbitrarily and can be modified to achieve a better result. Thus, it is possible to form recommendations by processing a smaller amount of data, which will undoubtedly speed up the work of recommender systems. The results of the comparison of calculation durations are shown in Fig.4.

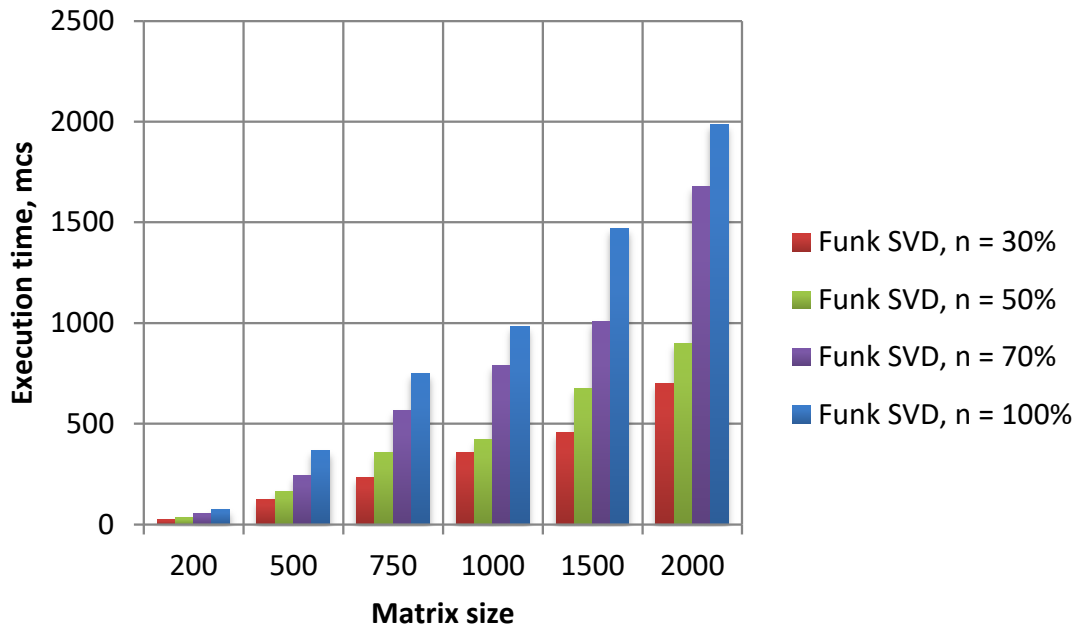


Figure 4: The comparison of execution time by Funk SVD algorithm using different number of users for providing recommendation.

According to Fig.4, a pattern of calculation acceleration can be observed when using less information about users. Now let's compare the calculations' accuracy, i.e., the difference between the original data matrix and the one restored after decomposition (Fig.5).

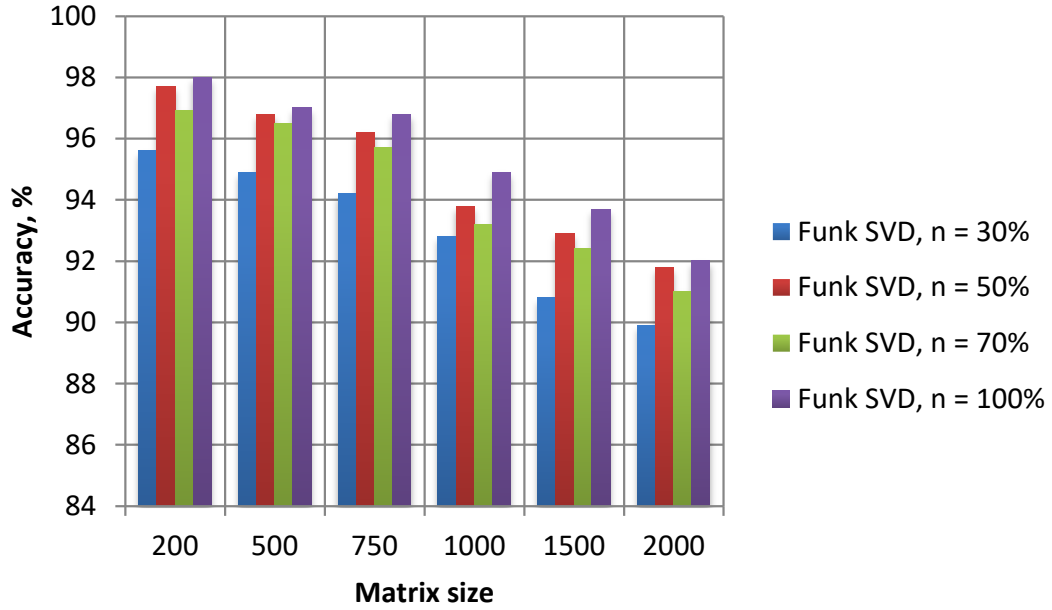


Figure 5: The comparison of accuracy by Funk SVD algorithm using different number of users for providing recommendation

Research results show a slight deterioration in the accuracy of calculations when using a smaller number of users. This approach allows to select different systems' optimal data processing parameters. We can attract fewer user data if it needs to process information quickly. More data can be used to make recommendations in newly created or dynamically updated systems.

6. Enhancing Accuracy of Recommendation Systems in 5G Industrial Networks using Modified Funk SVD Algorithm

The next modification of the FunkSVD algorithm proposed in this work is to improve the accuracy of recommendations. For improving the calculation accuracy of the Funk SVD algorithm, we suggest using more features about the items added to the matrix of goods after the layout by the standard algorithm. Such features are data on the dependencies of items and users, which affect the accuracy of the formation of recommendations. In 5G systems, additional data can be collected about groups of users, their interaction with goods, and features of interest in a certain service. There are no such features in the initial recommendation matrix, which contains only product ratings. Instead, we can use them to improve the quality of providing recommendations, if there is such a need (Fig. 6).

Thus, to calculate a recommendation to the user for a certain product, we form a new items matrix N_{mod} after FunkSVD decomposition containing additional features:

$$a'_{i,j} = \sum_{i,j=k} (m_i \times n_{j_mod}^T), \quad (17)$$

In 5G mobile communication systems, users are divided into groups. For example, special services are provided for business needs, collecting additional statistics. Such clients need high-quality recommendations according to individual characteristics. We

conducted performance studies for the proposed algorithm. A comparison of calculation durations for the non-modified and modified algorithms is shown in Fig.7.

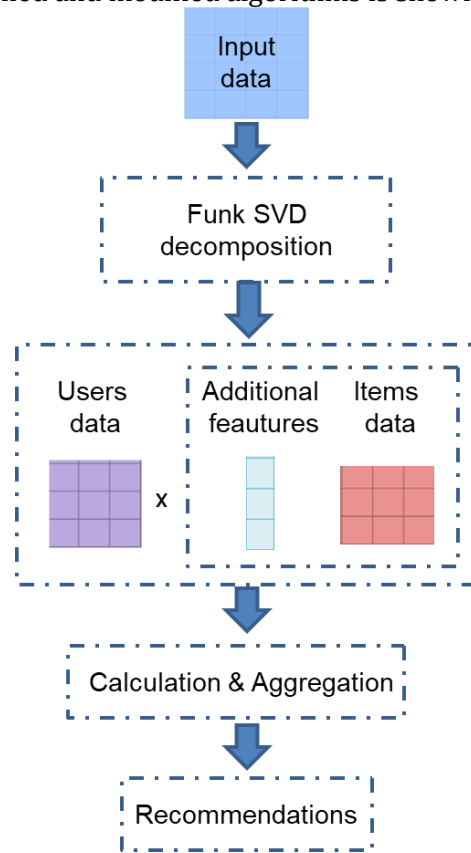


Figure 6: The scheme of using additional feautres about items for modified Funk SVD algorithm

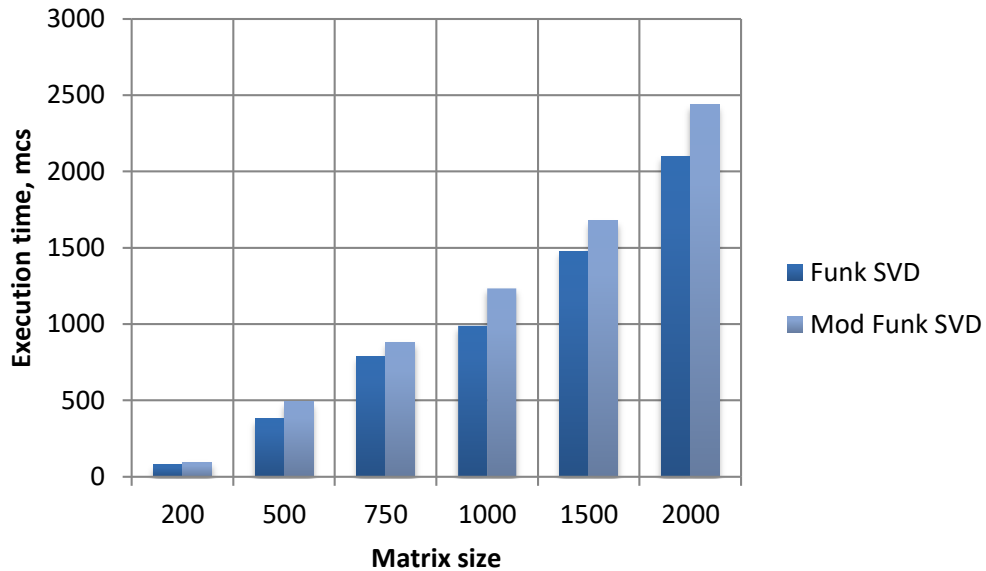


Figure 7: The comparison of execution time by non-modified and using additional feautres Funk SVD algorithms

From Fig.7, it can be concluded that the proposed modification requires the involvement of more computing power. Therefore the duration of data processing increases slightly. A comparison of calculation accuracy is shown in Fig.8.

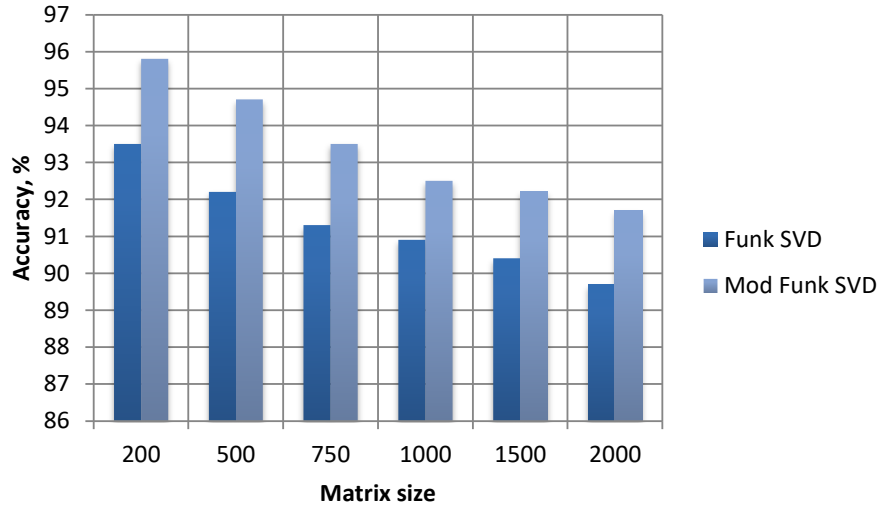


Figure 8: The comparison of accuracy by non-modified and using additional features Funk SVD algorithms

As shown in Fig.8, the modified Funk SVD demonstrates better calculation accuracy, so it can be used to form effective recommendations. As a result of the use of additional properties of products, it is possible to better determine their value for users.

7. Results and Discussion

Overall, the challenge of processing large amounts of data quickly and accurately is a key consideration in developing effective 5G industrial recommendation systems. As we can see from the obtained results, two modifications of the FunkSVD algorithm allow more flexible organization of the calculation of recommendations. If it is necessary to reduce the duration of the algorithm, we use the first modification. To improve the accuracy of calculations, we use the second. Now let's compare the effectiveness of simultaneously using two modifications of the FunkSVD algorithm in comparison with the non-modified one. Thus, we define the recommendations:

$$a'_{i,j} = \sum_{i=\delta, j=k} (m_i \times n_{j_mod}^T), \quad (18)$$

Let's calculate the error of the recommendations according to Eq.6. and update the matrices M and N:

$$m'_i = m_i + \varepsilon(\text{Err} \times 2 \times n_{j_mod} + \theta \times m_i), \quad (20)$$

$$n'_{j_mod} = n_{j_mod} + \varepsilon(\text{Err} \times 2 \times m_i + \theta \times n_{j_mod}). \quad (21)$$

Now we can calculate the updated recommendation:

$$a'_{i,j} = m'_i \times n'_{j_mod}^T + d_{i,j}, \quad (22)$$

The results of the calculation durations comparison for the twice modified FunkSVD compared to the unmodified one are shown in Fig. 9.

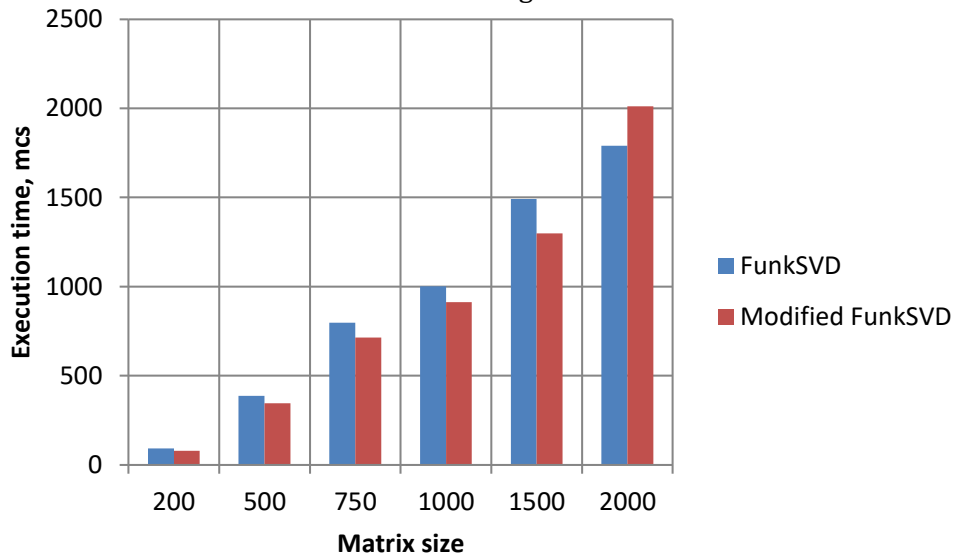


Figure 9: Execution time for twice modified FunkSVD algorithm compared to the non-modified

A comparison of the providing recommendations accuracy for two algorithms is shown in Fig.10.

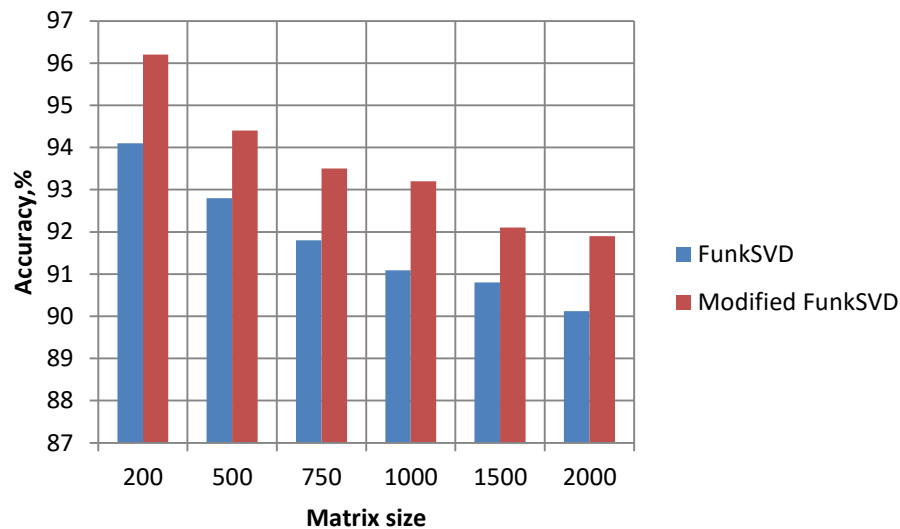


Figure 10: Accuracy of providing recommendation for twice modified FunkSVD algorithm compared to the non-modified

The conducted studies demonstrate the effectiveness of Funk SVD for processing large data. The proposed modifications allow using this algorithm adaptively to the work requirements and system load to achieve the best efficiency of user data calculations in 5G systems. For improving the reliability and privacy of data collection from end users, we recommend the modified federated SVD algorithm, which was considered in our previous

study [35]. Improving the reliability of the modified FunkSVD method proposed in this work will be the subject of our future research.

The speed and accuracy of data processing contribute to a better understanding of client needs and prompt provision of appropriate services. Thus, the proposed modifications contribute to the modernization of 5G systems, expanding the range of their services and improving interaction with users. Given the constant appearance of new services for users in 5G mobile communication networks, the problem of flexible resource allocation is already relevant. Depending on the needs and status of customers, the level of services they need, and the state of the mobile network, we can choose the appropriate recommendation algorithm. Also, the simultaneous use of two modifications improves both the accuracy and duration of calculations.

8. Conclusions

We found that with the development of the IIoT and the use of a large number of visual sensors, the amount of visual data is growing rapidly. This creates the challenge of effectively processing, analyzing, and using this data to ensure productivity and optimize processes in industrial environments. We concluded that the 5G industrial recommender systems can help manufacturing companies optimize their processes and increase efficiency. Collaborative filtering is a popular technique used in recommendation systems, but there are several challenges associated with it that need to be addressed in order to develop effective 5G recommendation systems. In particular, collaborative filtering relies on user-item interactions to make recommendations. However, in many industrial settings, data can be sparse and incomplete, making it difficult to accurately capture user preferences and generate meaningful recommendations. That's why we analyzed big data optimization algorithms and recommendation systems. The advantages of the Funk SVD algorithm for processing big data in industrial systems are determined. For improving the efficiency of processing information from users, modifications of Funk SVD are proposed. First, not all data from users is used to form recommendations, but part of it. The study results showed that the duration of calculations with this approach decreases, and the accuracy of data processing remains relatively high, which can be effectively used for systems with a high load. Secondly, additional items' features are used for more accurate work of recommender systems. The study's results demonstrate the prospect of using the proposed methods to optimize the work of industrial recommender systems. Modified FunkSVD algorithms can be applied to improve the quality of service for different types of 5G systems users. Overall, the proposed data optimization is a critical aspect for developing effective recommender systems in future 5G industrial networks. By ensuring that the system is working with high-quality data that is relevant to the specific needs of the industrial setting, it is possible to generate accurate and useful recommendations that can improve productivity, reduce costs, and optimize processes.

Acknowledgements

This paper is supported by the National Research Foundation of Ukraine, project number 0123U103529 (2022.01/0009) "Assessing and forecasting threats to the reconstruction and sustainable operation of objects of critical infrastructure" from the contest "Science for reconstruction of Ukraine in the war and post war periods.

References

- [1] T. Steclik, R. Cupek, and M. Drewniak, "Automatic grouping of production data in Industry 4.0: The use case of internal logistics systems based on Automated Guided Vehicles," *J. Comput. Sci.*, vol. 62, no. 101693, p. 101693, 2022, doi: 10.1016/j.jocs.2022.101693.
- [2] L. Tang and Y. Meng, "Data analytics and optimization for smart industry," *Front. Eng. Manag.*, vol. 8, no. 2, pp. 157–171, 2021, doi: 10.1007/s42524-020-0126-0.
- [3] Y. Xin, D. Liu and X. Zhou, "Evolutionary Analysis of Cloud Manufacturing Platform Service Innovation Based on a Multiagent Game Perspective," in *IEEE Access*, vol. 10, pp. 104543-104554, 2022, doi: 10.1109/ACCESS.2022.3208915.
- [4] Y. Gao, B. Yang, S. Wang, G. Fu, and P. Zhou, "A multi-objective service composition method considering the interests of tri-stakeholders in cloud manufacturing based on an enhanced jellyfish search optimizer," *J. Comput. Sci.*, vol. 67, no. 101934, p. 101934, 2023, doi: 10.1016/j.jocs.2022.101934.
- [5] B. Zhu, F. Ortega, J. Bobadilla, and A. Gutiérrez, "Assigning reliability values to recommendations using matrix factorization," *J. Comput. Sci.*, vol. 26, pp. 165–177, 2018, doi: 10.1016/j.jocs.2018.04.009
- [6] M. Fu, L. Huang, A. Rao, A. A. Irissappane, J. Zhang and H. Qu, "A Deep Reinforcement Learning Recommender System With Multiple Policies for Recommendations," in *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 2049-2061, Feb. 2023, doi: 10.1109/TII.2022.3209290
- [7] W. Ruoxi, H. Beshley, Y. Lingyu, O. Urikova, M. Beshley and O. Kuzmin, "Industrial 5G Private Network: Architectures, Resource Management, Challenges, and Future Directions," 2022 IEEE 16th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 2022, pp. 780-784, doi: 10.1109/TCSET55632.2022.9766945.
- [8] R. Wang et al., "Radio Resource Management Methods for Ultra-Reliable Low-Latency Communications in 5G LTE Narrowband Industrial Internet of Things," 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT), Lviv, Ukraine, 2021, pp. 239-244, doi: 10.1109/AICT52120.2021.9628913.
- [9] Su, J., Beshley, M., Przystupa, K., et al. (2022). 5G multi-tier radio access network planning based on Voronoi diagram. *Measurement*, 192, 110814.
- [10] Xiong, G., Przystupa, K., Teng, Y., et al. (2021). Online measurement error detection for the electronic transformer in a smart grid. *Energies*, 14(12), 3551.
- [11] Chen, X., Przystupa, K., Ye, Z., et al. (2022). Forecasting short-term electric load using extreme learning machine with improved tree seed algorithm based on Lévy flight. *Eksploracja i Niezawodność – Maintenance and Reliability*, 24(1),153-162
- [12] Sun, L., Qin, H., Przystupa, K., Majka, M., & Kochan, O. (2022). Individualized short-term electric load forecasting using data-driven meta-heuristic method based on LSTM network. *Sensors*, 22(20), 7900.
- [13] E. P. Xing et al., "Petuum: A New Platform for Distributed ML on Big Data," in *IEEE Transactions on Big Data*, vol. 1, no. 2, pp. 49-67, 1 June 2015, doi: 10.1109/TBDATA.2015.2472014.

- [14] V. Vashishth, A. Chhabra and A. Sood, "A predictive approach to task scheduling for Big Data in cloud environments using classification algorithms," 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida, India, 2017, pp. 188-192, doi: 10.1109/CONFLUENCE.2017.7943147.
- [15] D. Tarek, A. Benslimane, M. Darwish and A. M. Kotb, "Distributed Packets Scheduling Technique for Cognitive Radio Internet of Things Based on Discrete Permutation Particle Swarm Optimization," 2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Rhodes, Greece, 2020, pp. 142-151, doi: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00040.
- [16] S. Sennan, S. Ramasubbareddy, S. Balasubramaniam, A. Nayyar, M. Abouhawwash and N. A. Hikal, "T2FL-PSO: Type-2 Fuzzy Logic-Based Particle Swarm Optimization Algorithm Used to Maximize the Lifetime of Internet of Things," in IEEE Access, vol. 9, pp. 63966-63979, 2021, doi: 10.1109/ACCESS.2021.3069455.
- [17] M. C. Adikari and T. P. Amalan, "Distribution cost optimization using Big Data Analytics, ML and Computer Simulation for FMCG Sector," 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE), Colombo, Sri Lanka, 2019, pp. 63-69, doi: 10.23919/SCSE.2019.8842697.
- [18] B. Raviteja, K. A. Pandya, F. Khan, Z. Tufail Khan, R. Prajwal and A. Kekatpure, "Smart Supply Chain Management using Big Data Analysis and ML," 2022 International Conference on Edge Computing and Applications (ICECAA), Tamilnadu, India, 2022, pp. 190-193, doi: 10.1109/ICECAA55415.2022.9936359.
- [19] G. K. Singh, M. Dadhich, V. Chouhan and A. Sharma, "Impact of Big Data Analytics & Capabilities on Supply Chain Management (SCM) - An Analysis of Indian Cement Industry," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2021, pp. 313-318, doi: 10.1109/ICAC3N53548.2021.9725531.
- [20] H. Fan, C. Huang and Y. Liu, "Federated Learning-Based Privacy-Preserving Data Aggregation Scheme for IIoT," in IEEE Access, vol. 11, pp. 6700-6707, 2023, doi: 10.1109/ACCESS.2022.3226245.
- [21] W. Gao, Z. Zhao, G. Min, Q. Ni and Y. Jiang, "Resource Allocation for Latency-Aware Federated Learning in Industrial Internet of Things," in IEEE Transactions on Industrial Informatics, vol. 17, no. 12, pp. 8505-8513, Dec. 2021, doi: 10.1109/TII.2021.3073642.
- [22] F. Liang, W. Yu, X. Liu, D. Griffith and N. Golmie, "Toward Deep Q-Network-Based Resource Allocation in Industrial Internet of Things," in IEEE Internet of Things Journal, vol. 9, no. 12, pp. 9138-9150, 15 June 2022, doi: 10.1109/JIOT.2021.3093346.
- [23] K. Shah, A. Salunke, S. Dongare and K. Antala, "Recommender systems: An overview of different approaches to recommendations," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-4, doi: 10.1109/ICIIECS.2017.8276172.

- [24] Z. Guo, K. Yu, T. Guo, A. K. Bashir, M. Imran and M. Guizani, "Implicit Feedback-based Group Recommender System for Internet of Things Applications," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan, 2020, pp. 1-6, doi: 10.1109/GLOBECOM42002.2020.9348091.
- [25] B. Wu, L. Zhong, L. Yao and Y. Ye, "EAGCN: An Efficient Adaptive Graph Convolutional Network for Item Recommendation in Social Internet of Things," in IEEE Internet of Things Journal, vol. 9, no. 17, pp. 16386-16401, 1 Sept.1, 2022, doi: 10.1109/JIOT.2022.3151400.
- [26] S. M. Kasongo, "An Advanced Intrusion Detection System for IIoT Based on GA and Tree Based Algorithms," in IEEE Access, vol. 9, pp. 113199-113212, 2021, doi: 10.1109/ACCESS.2021.3104113.
- [27] Simeone, A., Zeng, Y. & Caggiano, A. Intelligent decision-making support system for manufacturing solution recommendation in a cloud framework. Int J Adv Manuf Technol 112, 1035–1050 (2021), <https://doi.org/10.1007/s00170-020-06389-1>
- [28] T. Vafeiadis et al., "Intelligent Information Management System for Decision Support: Application in a Lift Manufacturer's Shop Floor," 2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), Sofia, Bulgaria, 2019, pp. 1-6, doi: 10.1109/INISTA.2019.8778290.
- [29] Y. Fu, S. Wang, C. -X. Wang, X. Hong and S. McLaughlin, "Artificial Intelligence to Manage Network Traffic of 5G Wireless Networks," in IEEE Network, vol. 32, no. 6, pp. 58-64, November/December 2018, doi: 10.1109/MNET.2018.1800115.
- [30] A. Vulpe, M. Idu, D. Gheorghe, A. Martian and O. Fratu, "ML-based Analytics Framework for beyond 5G Mobile Communication Systems," 2020 28th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2020, pp. 1-4, doi: 10.1109/TELFOR51502.2020.9306534.
- [31] YongchangWang and L. Zhu, "Research and implementation of SVD in ML," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), Wuhan, China, 2017, pp. 471-475, doi: 10.1109/ICIS.2017.7960038.
- [32] C. Zhang, J. Wang, J. Mou, X. Li and R. Wang, "Digital Text Feature Extraction Using Singular Value Decomposition and Principal Component Analysis," 2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China, 2019, pp. 10-13, doi: 10.1109/ICISCAE48440.2019.221578.
- [33] B. Kumar, "A novel latent factor model for recommender system," J. Inf. Syst. Technol. Manag., vol. 13, no. 3, 2016, <https://doi.org/10.4301/S1807-17752016000300008>.
- [34] S. Guo and C. Li, "Hybrid Recommendation Algorithm based on User Behavior," in 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), 2020, <https://doi.org/10.1109/ITAIC49862.2020.9339083>
- [35] Hordiichuk-Bublivska O., Beshley H., Kyryk M., Pyrih Y., Urikova O., Beshley M. A Modified Federated Singular Value Decomposition Method for Big Data and ML Optimization in IIoT Systems. In: Mikhailo Klymash, Andriy Luntovskyy, Mykola Beshley, Igor Melnyk, Alexander Schill. (eds) Emerging Networking in the Digital Transformation Age: Approaches, Protocols, Platforms, Best Practices, and Energy Efficiency. Lecture Notes in Electrical Engineering, 2023, Springer, Cham. 965, P.246-268.

- [36] P. Wu et al., "Fast data assimilation (FDA): Data assimilation by ML for faster optimize model state," *J. Comput. Sci.*, vol. 51, no. 101323, p. 101323, 2021, <https://doi.org/10.1016/j.jocs.2021.101323>.
- [37] T. Z. Emara and J. Z. Huang, "Distributed Data Strategies to Support Large-Scale Data Analysis Across Geo-Distributed Data Centers," in *IEEE Access*, vol. 8, pp. 178526-178538, 2020, doi: 10.1109/ACCESS.2020.3027675.
- [38] X. Zheng, L. Tian, B. Hui and X. Liu, "Distributed and Privacy Preserving Graph Data Collection in Internet of Thing Systems," in *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9301-9309, 15 June 15, 2022, doi: 10.1109/JIOT.2021.3112186.
- [39] F. Yin et al., "FedLoc: Federated Learning Framework for Data-Driven Cooperative Localization and Location Data Processing," in *IEEE Open Journal of Signal Processing*, vol. 1, pp. 187-215, 2020, doi: 10.1109/OJSP.2020.3036276.
- [40] K. Yin et al., "DLDP-FL: Dynamic local differential privacy federated learning method based on mesh network edge devices," *J. Comput. Sci.*, vol. 63, no. 101789, p. 101789, 2022, <https://doi.org/10.1016/j.jocs.2022.101789>.
- [41] K. Xu, W. Zhang, and Z. Yan, "A privacy-preserving mobile application recommender system based on trust evaluation," *J. Comput. Sci.*, vol. 26, pp. 87-107, 2018, <https://doi.org/10.1016/j.jocs.2018.04.001>.
- [42] Z. Wang and S. Ulukus, "Symmetric Private Information Retrieval at the Private Information Retrieval Rate," in *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 2, pp. 350-361, June 2022, doi: 10.1109/JSAIT.2022.3188610.
- [43] C. Das, A. K. Sahoo, and C. Pradhan, "Multicriteria recommender system using different approaches," in *Cognitive Big Data Intelligence with a Metaheuristic Approach*, S. Mishra, H. K. Tripathy, P. K. Mallick, A. K. Sangaiah, and G.-S. Chae, Eds. San Diego, CA: Elsevier, 2022, pp. 259-277, <https://doi.org/10.1016/B978-0-323-85117-6.00011-X>.
- [44] Xie, F., Zhang, Y., Przystupa, K., & Kochan, O. (2023). A Knowledge Graph Embedding Based Service Recommendation Method for Service-Based System Development. *Electronics*, 12(13), 2935.
- [45] Xu, X., Przystupa, K., & Kochan, O. (2023). Social Recommendation Algorithm Based on Self-Supervised Hypergraph Attention. *Electronics*, 12(4), 906.
- [46] M. Beshley, N. Kryvinska and H. Beshley, "Energy-Efficient QoE-Driven Radio Resource Management Method for 5G and Beyond Networks," in *IEEE Access*, vol. 10, pp. 131691-131710, 2022, doi: 10.1109/ACCESS.2022.3228758.