A SURVEY ON MODERN RECOMMENDATION SYSTEM BASED ON BIG DATA

A PREPRINT

Yuanzhe Peng
Department of Electrical and Computer Engineering
University of Miami
Coral Gables, FL 33146
yxp412@miami.edu

ABSTRACT

Recommendation systems have become very popular in recent years and are used in various web applications. Modern recommendation systems aim at providing users with personalized recommendations of online products or services. Various recommendation techniques, such as content-based, collaborative filtering-based, knowledge-based, and hybrid-based recommendation systems, have been developed to fulfill the needs in different scenarios. This paper presents a comprehensive review of historical and recent state-of-the-art recommendation approaches, followed by an in-depth analysis of groundbreaking advances in modern recommendation systems based on big data. Furthermore, this paper reviews the issues faced in modern recommendation systems such as sparsity, scalability, and diversity and illustrates how these challenges can be transformed into prolific future research avenues.

1 Introduction

Recommendation systems have become very popular in recent years and are used in various web applications [1]. Recommendation systems are a specific type of information filtering system whose purpose is to predict a user's preference for an item. These recommenders relate to different decision-making processes, such as what items to buy and what music to listen to [2]. For example, Amazon's recommendation engine provides each user with a customized homepage. In addition, companies like Amazon, YouTube, and Netflix use recommendation systems to help users discover new and relevant videos, creating a better user experience while generating colossal income [3, 4]. Fig. 1 illustrates a modern recommendation system [5]. Besides, the recommendation system will also be involved in the field of human-computer interaction (HCI), and will further improve the performance of the interaction through the feedback mechanism [6, 7, 8, 9]. For example, the recommendation mechanism can help users obtain information about their exciting answers after the interactive query.

Recommendation systems are critical in some companies because when they are efficient, they can generate a lot of income or be a way to stand out from competitors [10, 11]. For example, Netflix launched the "Netflix Prize" challenge a few years ago. The target of this challenge was to train a recommender that could outperform Netflix's recommendation algorithm with a winning prize of 1 million dollars.

Furthermore, recommendation systems are one of the most common applications in the field of big data [12, 13]. It can predict a user's interest in purchasing items using a large amount of data in terms of purchase history, ratings, or online reviews. There are four types of commonly used recommendation systems [14]: content-based, collaborative filtering-based, knowledge-based and hybrid-based. Each type of recommendation system has its advantages and disadvantages [15]. For example, collaborative filtering-based recommenders may suffer from sparsity, scalability [16], and cold-start problems, while content-based recommenders may have limited ability to expand users' existing interests [17, 18].

The rest of this paper is organized as follows. Section II presents a comprehensive review of historical and recent state-of-the-art approaches in recommendation systems, followed by an in-depth analysis of groundbreaking advances

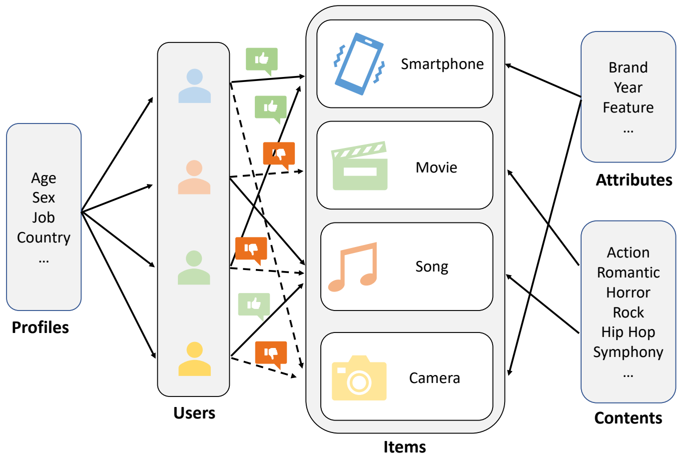

Figure 1: The logical process of a modern recommendation system.

in recommendation systems. Section III describes the issues faced in recommendation systems based on big data such as sparsity, scalability, diversity, and how to address these issues in modern recommendation systems. Finally, the summary is given in Section IV.

2 Recommendation Systems

Recommendation systems aim to predict users' preferences for a certain item and provide personalized services [19]. This section will discuss several commonly used recommender methods, such as content-based method, collaborative filtering-based method, knowledge-based method, and hybrid-based method.

2.1 Content-based Recommendation Systems

The main idea of content-based recommenders is to recommend items based on the similarity between different users or items [20]. This algorithm determines and differentiates the main common attributes of a particular user's favorite items by analyzing the descriptions of those items. Then, these preferences are stored in this user's profile. The algorithm then recommends items with a higher degree of similarity with the user's profile. Besides, content-based recommendation systems can capture the specific interests of the user and can recommend rare items that are of little interest to other users. However, since the feature representations of items are designed manually to a certain extent, this method requires a lot of domain knowledge. In addition, content-based recommendation systems can only recommend based on users' existing interests, so the ability to expand users' existing interests is limited.

2.2 Collaborative Filtering-based Recommendation Systems

Collaborative Filtering-based (CF) methods are primarily used in big data processing platforms due to their parallelization characteristics [21]. The basic principle of the recommendation system based on collaborative filtering is shown in Fig. 2 [22]. CF recommendation systems use the behavior of a group of users to recommend to other users [23]. There are mainly two types of collaborative filtering techniques, which are user-based and item-based.

- User-based CF: In the user-based CF recommendation system, users will receive recommendations of products that similar users like [24]. Many similarity metrics can calculate the similarity between users or items, such as Constrained Pearson Correlation coefficient (CPC), cosine similarity, adjusted cosine similarity, etc. For example, cosine similarity is a measure of similarity between two vectors. Let x and y denote two vectors, cosine similarity between x and y can be represented by

$$\cos(\theta) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}} \tag{1}$$

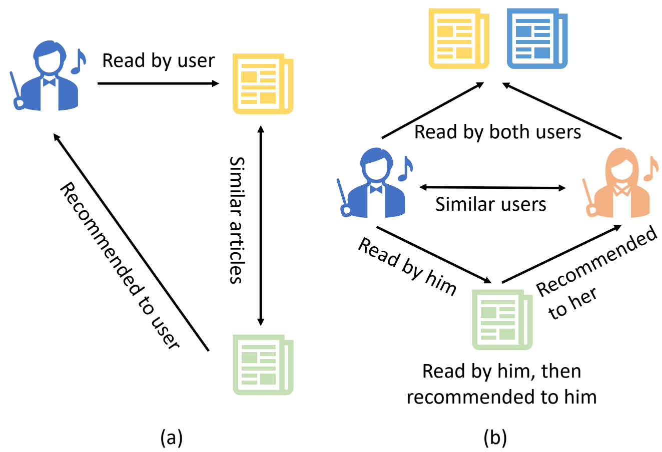

Figure 2: (a) Content-based Recommendation System (b) Collaborative Filtering-based Recommendation System.

Table 1: Summary of the Modern Recommendation Systems Methods.

Recommendation Systems	Descriptive Key Points	Papers
Content-based	Recommend items based on the similarity between different items.	Musto et al. [26] Volkovs et al.[27] Mittal et al.[28] Almaguer et al.[29]
Collaborative Filtering-based	Recommend items to some users based on the other users behavior.	Zhang et al.[30] Bobadilla et al.[31] Bobadilla et al.[32] Rezaimehr et al.[24]
Knowledge-based	Recommend items to users based on basic knowledge of users, items, and relationships between items.	Dong et al.[33] Gazdar et al.[34] Alamdari et al.[35] Cena et al.[36]
Hybrid-based	Recommend items to users based on more than one filtering approach.	Hrnjica et al.[37] Shokeen et al.[38] Zagra et al.[39] Ibrahim et al.[40]

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{2}$$

- Item-based CF: Item-based CF algorithm predicts user ratings for items based on item similarity. Generally, item-based CF yields better results than user-based CF because user-based CF suffers from sparsity and scalability issues. However, both user-based CF and item-based CF may suffer from cold-start problems [25].

2.3 Knowledge-based Recommendation Systems

The main idea of knowledge-based recommendation systems is to recommend items to users based on basic knowledge of users, items, and relationships between items [41, 42]. Since knowledge-based recommendation systems do not require user ratings or purchase history, there is no cold start problem for this type of recommendation [43]. Knowledge-based recommendation systems are well suited for complex domains where items are not frequently purchased, such as cars and apartments [44]. But the acquisition of required domain knowledge can become a bottleneck for this recommendation technique [33].

2.4 Hybrid-based Recommendation Systems

Hybrid-based recommendation systems combine the advantages of multiple recommendation techniques and aim to overcome the potential weaknesses in traditional recommendation systems [45]. There are seven basic hybrid recommendation techniques [40]: weighted, mixed, switching, feature combination, feature augmentation, cascade, and meta-level methods [46, 47]. Among all of these methods, the most commonly used is the combination of the CF recommendation methods with other recommendation methods (such as content-based or knowledge-based) to avoid sparsity, scalability, and cold-start problems [37, 39, 48].

2.5 Challenges in Modern Recommendation Systems

- Sparsity. As we know, the usage of recommendation systems is growing rapidly. Many commercial recommendation systems use large datasets, and the user-item matrix used for filtering may be very large and sparse. Therefore, the performance of the recommendation process may be degraded due to the cold start problems caused by data sparsity [49].
- Scalability. Traditional algorithms will face scalability issues as the number of users and items increases. Assuming there are millions of customers and millions of items, the algorithm's complexity will be too large. However, recommendation systems must respond to the user's needs immediately, regardless of the user's rating history and purchase situation, which requires high scalability. For example, Twitter is a large web company that uses clusters of machines to scale recommendations for its millions of users [38].
- Diversity. Recommendation systems also need to increase diversity to help users discover new items. Unfortunately, some traditional algorithms may accidentally do the opposite because they always recommend popular and highly-rated items that some specific users love. Therefore, new hybrid methods need to be developed to improve the performance of the recommendation systems [50].

3 Recommendation System based on Big Data

Big data refers to the massive, high growth rate and diversified information [51, 52]. It requires new processing models to have stronger decision-making and process optimization capabilities [53]. Big data has its unique "4V" characteristics, as shown in Fig. 3 [54]: Volume, Variety, Velocity, and Veracity.

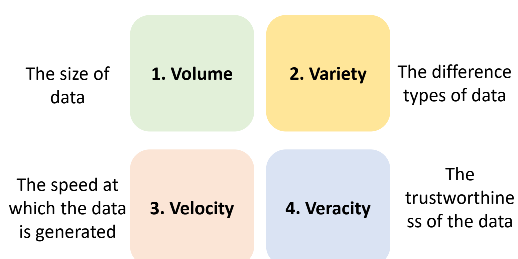

Figure 3: The 4V of big data.

3.1 Big Data Processing Flow

Big data comes from many sources, and there are many methods to process it [55]. However, the primary processing of big data can be divided into four steps [56]. Besides, Fig. 4 presents the basic flow of big data processing.

- Data Collection.
- Data Processing and Integration. The collection terminal itself already has a data repository, but it cannot accurately analyze the data. The received information needs to be pre-processed [57].
- Data Analysis. In this process, these initial data are always deeply analyzed using cloud computing technology [58].
- Data Interpretation.

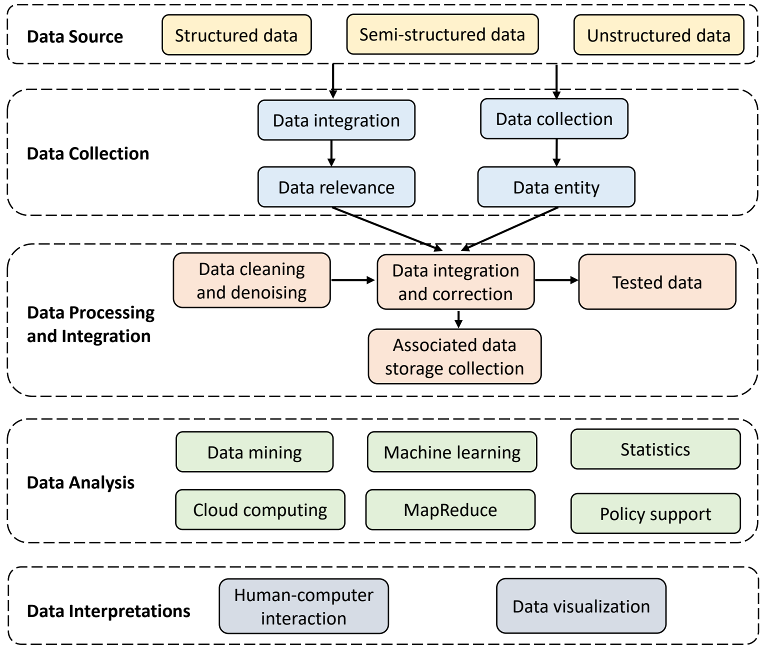

Figure 4: The basic flow of big data processing.

3.2 Modern Recommendation Systems based on the Big Data

The shortcomings of traditional recommendation systems mainly focus on insufficient scalability and parallelism [59]. For small-scale recommendation tasks, a single desktop computer is sufficient for data mining goals, and many techniques are designed for this type of problems [60].

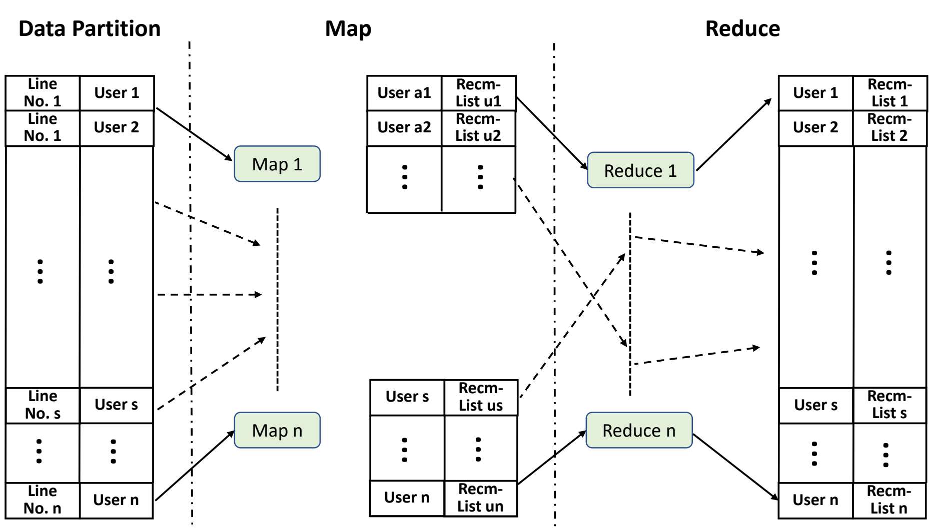

Figure 5: Mapreduce in the Recommendation Systems.

However, the rating data is usually so large for medium-scale recommendation systems that it is impossible to load all the data into memory at once [61]. Common solutions are based on parallel computing or collective mining, sampling and aggregating data from different sources, and using parallel computing programming to perform the mining process [62]. The big data processing framework will rely on cluster computers with high-performance computing platforms [63]. At the same time, data mining tasks will be deployed on a large number of computing nodes (i.e., clusters) by running some parallel programming tools [64], such as MapReduce [52, 65]. For example, Fig. 5 is the MapReduce in the Recommendation Systems.

In recent years, various big data platforms have emerged [66]. For example, Hadoop and Spark [52], both developed by the Apache Software Foundation, are widely used open-source frameworks for big data architectures [52, 67]. Each framework contains an extensive ecosystem of open-source technologies that prepare, process, manage and analyze big data sets [68]. For example, Fig. 6 is the ecosystem of Apache Hadoop [69].

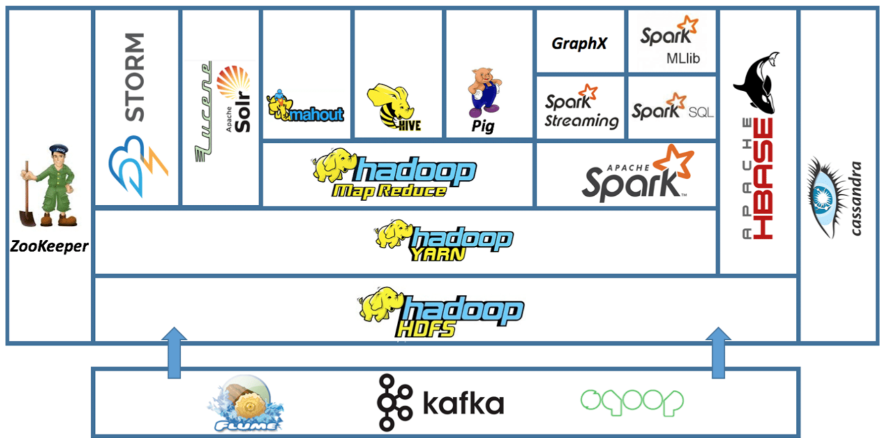

Figure 6: The ecosystem of Apach Hadoop.

Hadoop allows users to manage big data sets by enabling a network of computers (or "nodes") to solve vast and intricate data problems. It is a highly scalable, cost-effective solution that stores and processes structured, semi-structured and unstructured data.

Spark is a data processing engine for big data sets. Like Hadoop, Spark splits up large tasks across different nodes. However, it tends to perform faster than Hadoop, and it uses random access memory (RAM) to cache and process data instead of a file system. This enables Spark to handle use cases that Hadoop cannot. The following are some benefits of the Spark framework:

- It is a unified engine that supports SQL queries, streaming data, machine learning (ML), and graph processing.
- It can be 100x faster than Hadoop for smaller workloads via in-memory processing, disk data storage, etc.
- It has APIs designed for ease of use when manipulating semi-structured data and transforming data.

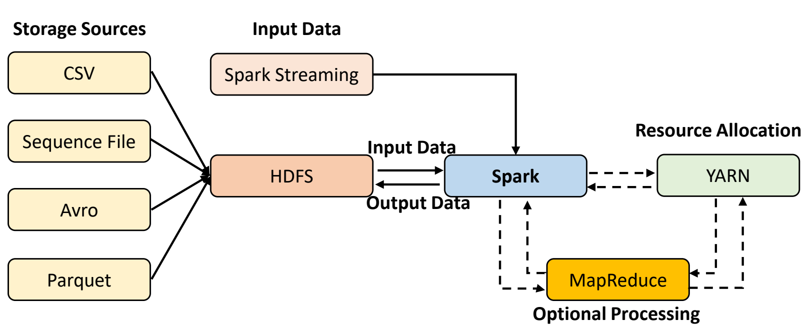

Figure 7: Spark uses the best parts of Hadoop through HDFS for reading and storing data, MapReduce for optional processing and YARN for resource allocation.

Furthermore, Spark is fully compatible with the Hadoop eco-system and works smoothly with Hadoop Distributed File System (HDFS), Apache Hive, and others. Thus, when the data size is too big for Spark to handle in-memory, Hadoop can help overcome that hurdle via its HDFS functionality. Fig. 7 is a visual example of how Spark and Hadoop can work together. Fig. 8 is the the architecture of the modern recommendation system based on Spark.

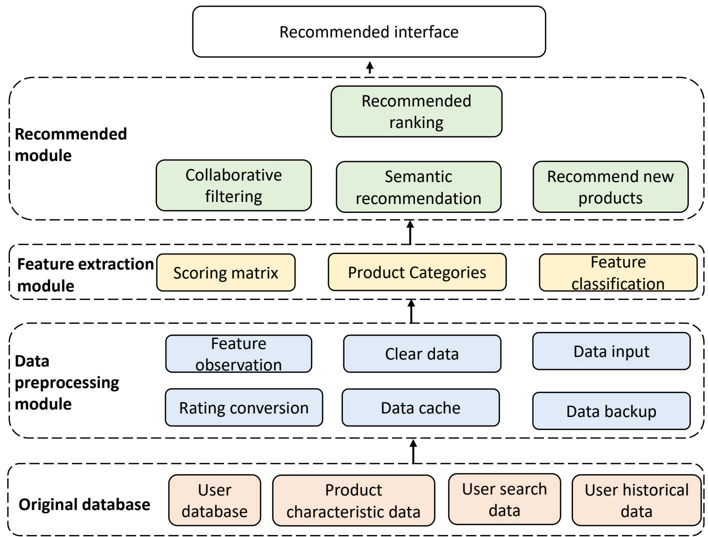

Figure 8: The architecture of the modern recommendation system based on Spark.

4 Summary

Recommendation systems have become very popular in recent years and are used in various web applications. Modern recommendation systems aim at providing users with personalized recommendations of online products or services. Various recommendation techniques, such as content-based, collaborative filtering-based, knowledge-based, and hybrid-based recommendation systems, have been developed to fulfill the needs in different scenarios.

This paper presents a comprehensive review of historical and recent state-of-the-art recommendation approaches, followed by an in-depth analysis of groundbreaking advances in modern recommendation systems based on big data. Furthermore, this paper reviews the issues faced in modern recommendation systems such as sparsity, scalability, and diversity and illustrates how these challenges can be transformed into prolific future research avenues.

References

[1] F. Zhou, B. Luo, T. Hu, Z. Chen, and Y. Wen, "A combinatorial recommendation system framework based on deep reinforcement learning," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 5733–5740.

[2] H. Wang, N. Lou, and Z. Chao, "A personalized movie recommendation system based on lstm-cnn," in *2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*. IEEE, 2020, pp. 485–490.

[3] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE transactions on knowledge and data engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[4] T. Omura, K. Suzuki, P. Siriaraya, M. Mittal, Y. Kawai, and S. Nakajima, "Ad recommendation utilizing user behavior in the physical space to represent their latent interest," in *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020, pp. 3143–3146.

[5] N. Entezari, E. E. Papalexakis, H. Wang, S. Rao, and S. K. Prasad, "Tensor-based complementary product recommendation," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 409–415.

[6] F. Ali, D. Kwak, P. Khan, S. H. A. Ei-Sappagh, S. M. R. Islam, D. Park, and K.-S. Kwak, "Merged ontology and svm-based information extraction and recommendation system for social robots," *IEEE Access*, vol. 5, pp. 12 364–12 379, 2017.

[7] Y. Peng, W. Han, and Y. Ou, "Semantic segmentation model for road scene based on encoder-decoder structure," in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2019, pp. 1927–1932.

[8] X. Ma, G. Jiang, Y. Peng, T. Ma, C. Liu, and Y.-s. Ou, “An intelligent speed-suggestion planner for coverage path with multiple constraints,” in *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2021, pp. 1213–1218.

[9] Y. Peng, Y. Ou, and W. Feng, “Learning stable control for a wheeled inverted pendulum with fast adaptive neural network,” in *2020 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2020, pp. 227–232.

[10] R. Rismanto, A. R. Syulistyo, and B. P. C. Agusta, “Research supervisor recommendation system based on topic conformity.” *International Journal of Modern Education & Computer Science*, vol. 12, no. 1, 2020.

[11] Z. Cui, X. Xu, X. Fei, X. Cai, Y. Cao, W. Zhang, and J. Chen, “Personalized recommendation system based on collaborative filtering for iot scenarios,” *IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685–695, 2020.

[12] B. Li, A. Maalla, and M. Liang, “Research on recommendation algorithm based on e-commerce user behavior sequence,” in *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, vol. 2. IEEE, 2021, pp. 914–918.

[13] X. Li and F. Sun, “Sports training recommendation method under the background of data analysis,” in *2021 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*. IEEE, 2021, pp. 12–16.

[14] T. Numnonda, “A real-time recommendation engine using lambda architecture,” *Artificial Life and Robotics*, vol. 23, no. 2, pp. 249–254, 2018.

[15] J. Xiao, M. Wang, B. Jiang, and J. Li, “A personalized recommendation system with combinational algorithm for online learning,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 3, pp. 667–677, 2018.

[16] Z. Huang, X. Xu, J. Ni, H. Zhu, and C. Wang, “Multimodal representation learning for recommendation in internet of things,” *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 675–10 685, 2019.

[17] H. Zhang, T. Huang, Z. Lv, S. Liu, and Z. Zhou, “Mcrs: A course recommendation system for moocs,” *Multimedia Tools and Applications*, vol. 77, no. 6, pp. 7051–7069, 2018.

[18] I. Benouaret and S. Amer-Yahia, “A comparative evaluation of top-n recommendation algorithms: Case study with total customers,” in *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020, pp. 4499–4508.

[19] B. Yi, X. Shen, H. Liu, Z. Zhang, W. Zhang, S. Liu, and N. Xiong, “Deep matrix factorization with implicit feedback embedding for recommendation system,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 8, pp. 4591–4601, 2019.

[20] P. Lops, M. d. Gemmis, and G. Semeraro, “Content-based recommender systems: State of the art and trends,” *Recommender systems handbook*, pp. 73–105, 2011.

[21] M. Elahi, F. Ricci, and N. Rubens, “A survey of active learning in collaborative filtering recommender systems,” *Computer Science Review*, vol. 20, pp. 29–50, 2016.

[22] B. Alhijawi and Y. Kilani, “Using genetic algorithms for measuring the similarity values between users in collaborative filtering recommender systems,” in *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*. IEEE, 2016, pp. 1–6.

[23] L. Al Hassanieh, C. Abou Jaoudeh, J. B. Abdo, and J. Demerjian, “Similarity measures for collaborative filtering recommender systems,” in *2018 IEEE Middle East and North Africa Communications Conference (MENACOMM)*. IEEE, 2018, pp. 1–5.

[24] F. Rezaimehr and C. Dadkhah, “A survey of attack detection approaches in collaborative filtering recommender systems,” *Artificial Intelligence Review*, vol. 54, no. 3, pp. 2011–2066, 2021.

[25] F. Zhang, T. Gong, V. E. Lee, G. Zhao, C. Rong, and G. Qu, “Fast algorithms to evaluate collaborative filtering recommender systems,” *Knowledge-Based Systems*, vol. 96, pp. 96–103, 2016.

[26] C. Musto, G. Semeraro, M. d. Gemmis, and P. Lops, “Learning word embeddings from wikipedia for content-based recommender systems,” in *European conference on information retrieval*. Springer, 2016, pp. 729–734.

[27] M. Volkovs, G. W. Yu, and T. Poutanen, “Content-based neighbor models for cold start in recommender systems,” in *Proceedings of the Recommender Systems Challenge 2017*, 2017, pp. 1–6.

[28] D. Mittal, S. Shandilya, D. Khirwar, and A. Bhise, “Smart billing using content-based recommender systems based on fingerprint,” in *ICT Analysis and Applications*. Springer, 2020, pp. 85–93.

[29] Y. Pérez-Almaguer, R. Yera, A. A. Alzahrani, and L. Martínez, “Content-based group recommender systems: A general taxonomy and further improvements,” *Expert Systems with Applications*, vol. 184, p. 115444, 2021.

[30] F. Zhang, V. E. Lee, R. Jin, S. Garg, K.-K. R. Choo, M. Maasberg, L. Dong, and C. Cheng, "Privacy-aware smart city: A case study in collaborative filtering recommender systems," *Journal of Parallel and Distributed Computing*, vol. 127, pp. 145–159, 2019.

[31] J. Bobadilla, S. Alonso, and A. Hernando, "Deep learning architecture for collaborative filtering recommender systems," *Applied Sciences*, vol. 10, no. 7, p. 2441, 2020.

[32] J. Bobadilla, F. Ortega, A. Gutiérrez, and S. Alonso, "Classification-based deep neural network architecture for collaborative filtering recommender systems." *International Journal of Interactive Multimedia & Artificial Intelligence*, vol. 6, no. 1, 2020.

[33] M. Dong, X. Zeng, L. Koehl, and J. Zhang, "An interactive knowledge-based recommender system for fashion product design in the big data environment," *Information Sciences*, vol. 540, pp. 469–488, 2020.

[34] A. Gazdar and L. Hidri, "A new similarity measure for collaborative filtering based recommender systems," *Knowledge-Based Systems*, vol. 188, p. 105058, 2020.

[35] P. M. Alamdari, N. J. Navimipour, M. Hosseinzadeh, A. A. Safaei, and A. Darwesh, "A systematic study on the recommender systems in the e-commerce," *IEEE Access*, vol. 8, pp. 115 694–115 716, 2020.

[36] F. Cena, L. Console, and F. Vernero, "Logical foundations of knowledge-based recommender systems: A unifying spectrum of alternatives," *Information Sciences*, vol. 546, pp. 60–73, 2021.

[37] B. Hrnjica, D. Music, and S. Softic, "Model-based recommender systems," *Trends in Cloud-based IoT*, pp. 125–146, 2020.

[38] J. Shokeen and C. Rana, "A study on features of social recommender systems," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 965–988, 2020.

[39] A. Zagranovskaia and D. Mitura, "Designing hybrid recommender systems," in *IV International Scientific and Practical Conference*, 2021, pp. 1–5.

[40] A. J. Ibrahim, P. Zira, and N. Abdulganiyyi, "Hybrid recommender for research papers and articles," *International Journal of Intelligent Information Systems*, vol. 10, no. 2, p. 9, 2021.

[41] S. Shishehchi, S. Y. Banihashem, N. A. M. Zin, S. A. M. Noah, and K. Malaysia, "Ontological approach in knowledge based recommender system to develop the quality of e-learning system," *Australian Journal of Basic and Applied Sciences*, vol. 6, no. 2, pp. 115–123, 2012.

[42] C. C. Aggarwal, "Knowledge-based recommender systems," in *Recommender systems*. Springer, 2016, pp. 167–197.

[43] R. Cabezas, J. G. Ruizº, and M. Leyva, "A knowledge-based recommendation framework using svn," *Neutrosophic Sets and Systems*, vol. 16, p. 24, 2017.

[44] J. K. Tarus, Z. Niu, and G. Mustafa, "Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning," *Artificial intelligence review*, vol. 50, no. 1, pp. 21–48, 2018.

[45] M. T. Ribeiro, A. Lacerda, A. Veloso, and N. Ziviani, "Pareto-efficient hybridization for multi-objective recommender systems," in *Proceedings of the sixth ACM conference on Recommender systems*, 2012, pp. 19–26.

[46] M. Hassan and M. Hamada, "Enhancing learning objects recommendation using multi-criteria recommender systems," in *2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE, 2016, pp. 62–64.

[47] Y. Zhang, X. Liu, W. Liu, and C. Zhu, "Hybrid recommender system using semi-supervised clustering based on gaussian mixture model," in *2016 international conference on cyberworlds (CW)*. IEEE, 2016, pp. 155–158.

[48] G. George and A. M. Lal, "Review of ontology-based recommender systems in e-learning," *Computers & Education*, vol. 142, p. 103642, 2019.

[49] J. D. West, I. Wesley-Smith, and C. T. Bergstrom, "A recommendation system based on hierarchical clustering of an article-level citation network," *IEEE Transactions on Big Data*, vol. 2, no. 2, pp. 113–123, 2016.

[50] X. He and X. Ke, "Research summary of recommendation system based on knowledge graph," in *The 2021 3rd International Conference on Big Data Engineering*, 2021, pp. 104–109.

[51] H. Chen, "A dqn-based recommender system for item-list recommendation," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 5699–5702.

[52] S. D. Kadam, D. Motwani, and S. A. Vaidya, "Big data analytics-recommendation system with hadoop framework," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, vol. 3. IEEE, 2016, pp. 1–5.

[53] D. P. Acharjya and K. Ahmed, “A survey on big data analytics: challenges, open research issues and tools,” *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 2, pp. 511–518, 2016.

[54] X. Zhou, W. Liang, I. Kevin, K. Wang, R. Huang, and Q. Jin, “Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data,” *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 1, pp. 246–257, 2018.

[55] P. Ram Mohan Rao, S. Murali Krishna, and A. Siva Kumar, “Privacy preservation techniques in big data analytics: a survey,” *Journal of Big Data*, vol. 5, no. 1, pp. 1–12, 2018.

[56] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, “Data mining with big data,” *IEEE transactions on knowledge and data engineering*, vol. 26, no. 1, pp. 97–107, 2013.

[57] C. K. Emani, N. Cullot, and C. Nicolle, “Understandable big data: a survey,” *Computer science review*, vol. 17, pp. 70–81, 2015.

[58] H.-Y. Lin and S.-Y. Yang, “A cloud-based energy data mining information agent system based on big data analysis technology,” *Microelectronics Reliability*, vol. 97, pp. 66–78, 2019.

[59] Y. Cheng and X. Bu, “Research on key technologies of personalized education resource recommendation system based on big data environment,” in *Journal of Physics: Conference Series*, vol. 1437, no. 1. IOP Publishing, 2020, p. 012024.

[60] K. Al Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, “Hybrid recommender system for tourism based on big data and ai: A conceptual framework,” *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 47–55, 2021.

[61] A. V. Dev and A. Mohan, “Recommendation system for big data applications based on set similarity of user preferences,” in *2016 International Conference on Next Generation Intelligent Systems (ICNGIS)*. IEEE, 2016, pp. 1–6.

[62] J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, “A disease diagnosis and treatment recommendation system based on big data mining and cloud computing,” *Information Sciences*, vol. 435, pp. 124–149, 2018.

[63] Z. Wan, “Research on e-commerce recommendation system based on big data technology,” in *Journal of Physics: Conference Series*, vol. 1883, no. 1. IOP Publishing, 2021, p. 012159.

[64] B. Asiya Banu and S. Banu, “Keyword based movie recommendation service using mapreduce.”

[65] J. P. Verma, B. Patel, and A. Patel, “Big data analysis: recommendation system with hadoop framework,” in *2015 IEEE International Conference on Computational Intelligence & Communication Technology*. IEEE, 2015, pp. 92–97.

[66] M. Uzun-Per, A. B. Can, A. V. Gürel, and M. S. Aktaş, “Big data testing framework for recommendation systems in e-science and e-commerce domains,” in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 2353–2361.

[67] Y.-w. Zhang, Y.-y. Zhou, F.-t. Wang, Z. Sun, and Q. He, “Service recommendation based on quotient space granularity analysis and covering algorithm on spark,” *Knowledge-Based Systems*, vol. 147, pp. 25–35, 2018.

[68] G. Chaithra *et al.*, “User preferences based recommendation system for services using mapreduce approach,” 2015.

[69] B. Ait Hammou, A. Ait Lahcen, and S. Mouline, “A distributed group recommendation system based on extreme gradient boosting and big data technologies,” *Applied Intelligence*, vol. 49, no. 12, pp. 4128–4149, 2019.