A Modern Recommendation System Survey in the Big Data Era

Sony Peng, Doo-Soon Park, Dae-Young Kim, Yixuan Yang, Sophort Siet, Sadriddinov Ilkhomjon Rovshan Ugli, and HyeJung Lee

Abstract Recommendation Systems are a prominent field that is widely used and popular for assisting people in making appropriate automated selections. It is a mechanism that assists the users in identifying relevant information from a range of available data. In this paper, we create a survey of state-of-the-art recommendation models and focuses primarily on providing a concise summary of the many strategies for recommendation, ranging from traditional to modern approaches, to open the door for further investigation into recommendation systems. In addition, the specific domain of the recommendation models was also illustrated in this study.

Keywords Recommendation systems · State-of-the-art · Modern approaches

1 Introduction

We live in the era of flood information, with data significantly increasing from various sites. With the massive data information, recommendation systems are the well-known technique to filter the knowledge data (e.g., news, movies, songs, point-of-interests, books, journals, products, etc.) information and then return suitable lists to the user. From the traditional data mining approach to the deep learning model, the model of the recommendation is a promising technique at the root of state-of-art technology [1]. In the realm of eCommerce, users have several options in finding items of interest, where recommendation systems assist people in locating and analyzing the preferred items. In the Healthcare area, recommendations have been used to suggest

S. Peng · D.-S. Park (✉) · D.-Y. Kim · Y. Yang · S. Siet · S. I. R. Ugli
Department of Software Convergence, Soonchunhyang University, Asan, South Korea
e-mail: parkds@sch.ac.kr

D.-Y. Kim
e-mail: dyoung.kim@sch.ac.kr

H. Lee
Institute for Artificial Intelligence and Software, Soonchunhyang University, Asan, South Korea

J. S. Park et al. (eds.), *Advances in Computer Science and Ubiquitous Computing*,
Lecture Notes in Electrical Engineering 1028,
https://doi.org/10.1007/978-981-99-1252-0_77

suitable items (e.g., food, drug, physical activity, healthcare professional, hospital place). Moreover, location-based social networks (LBSNs), including Foursquare, Facebook, Instagram, LinkedIn, and Yelp are the centric sites where users can share their physical location, user tag, comment/feedback, like/dislike, rating, real-time check-ins, and user friend connections. In addition, users are closely connected and certainly know about their real-time locations.

The existing studies have created recommendations in various domains by applying simple recommendation techniques to complex models. At first, the recommendation models are categorized into two methods. These two models are collaborative filtering (CF), and content-based filtering (CBF). According to the demand for state-of-art applications, more models are introduced including knowledge-based, demographic-based, and hybrid method, etc.

Due to the popularity of the recommendation model, researchers are exploring the recommendation model to enhance its performance and the model's complexity.

To the best knowledge of our study, Sect. 2 introduces the detail of modern recommendation approaches. Section 3 provides simple explanations about recommendation-specific domains. Finally, Sect. 4 is described conclusion of the work.

2 Modern Recommendation Systems in the Big Data Era

2.1 How Do We Filter the Related Journals?

Within the scope of this investigation, we spot the journals from various well-known sites. Google Scholar [2] is used as a primary search engine, and in addition to this, we have also embraced Web of Science [3], which we refer to as an essential tool to discover related published papers.

Moreover, several keywords are used to filter the journal papers. These keywords include recommendation system, collaborative filtering, content-based filtering, hybrid model, traditional recommendation system, modern recommendation system, product recommendation, movie recommendation, product recommendation, news recommendation, job recommendation, location recommendation, point-of-interest recommendation, healthcare recommendation system, and recommendation challenges.

2.2 Overview of Recommendation Systems

With the massive amount of data, the recommendation system aims to predict the close items to the user. Recommendation systems collect information on their users' preferences for items (e.g., books, journals, songs, movies, products, location, travel,

etc.), which then return the personalized items to the active/target users. Data is a necessary part of the model of the recommendation. Refer to Fig. 1, there are two types of the data, namely explicit data, and implicit data. Explicit data refer to the data that users are actively provide to the system through their interaction such as like/dislike, rating and especially it's directed collection. This type of data is powerful and beneficial to the recommendation model since the data is given the clear purpose of the user's mind in terms of analysis. The disadvantage of this data is difficult to obtain. Another type is Implicit data which is collected indirectly indicates preference and taste. The data is mainly stored in the session system like browsing history, landing page, purchase history, and even mouse movement.

Figures 2 and 3 are shown the recommendation model from collaborative filtering, content-based filtering, knowledge-based, demographic-based, and finally, the hybrid model. More precisely, the recommendation system's general process is displayed in Fig. 2.

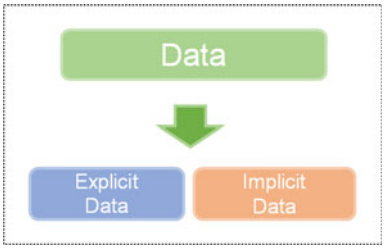

Fig. 1 An overview of the basic the data

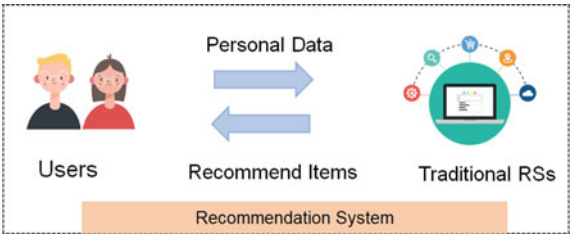

Fig. 2 Basic recommendation system model

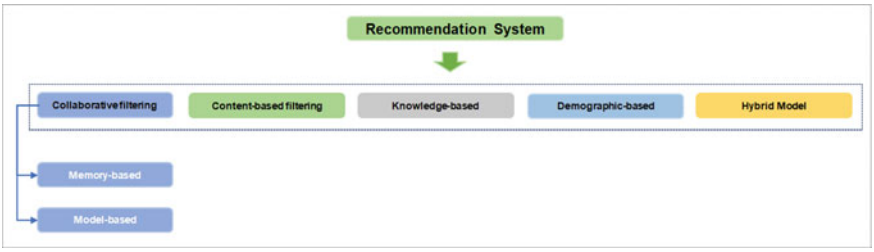

Fig. 3 Recommendation system types

2.3 Modern Recommendation Systems

Various recommendation models have been deployed in the study area's wide range. However, there are two main categories in the recommendation systems, including collaborative filtering and content-based filtering. In our current work, we are not focusing on only two models but aim to provide more studies with other models with the interaction of the big data.

Collaborative Filtering. Collaborative filtering is the most successful technique in the recommendation system and is widely used in the recommendation area. This technique uses the behavior of the group of users to suggest to other users [4]. Simply, the recommendation is based on nearest neighbor or similar users. There are two mainly types in the collaborative filtering including user-based and item-based.

User-based: In the mechanism, the system will suggest the items based on user's favorite and find the most similar item to the one user like [5]. Several similar metrics can be utilized to calculate the similarity between users/items, such as cosine similarity, adjusted cosine similarity, Pearson correlation coefficient (PCC).

Item-based: In collaborative filtering, item-based predicts user rating for the item similarity. For this model, their performance is better than the user-based since it suffers from several issues, such as sparsity and scalability [6]. Moreover, both user- and item-based suffer from cold-start problems (Item cold-start and User cold-start).

Content-based Filtering. The fundamental principle underlying content-based recommenders is to recommend products based on their content items' similarity [7]. By examining the descriptions of those items, the algorithms discover and separate the primary common attributes of a particular user's personal favorites. These choices are then saved in this user's profile. The system then recommends more similar things to the user's profile.

Furthermore, content-based recommendation systems can capture the user's individual preferences and propose products that would be of little interest to other people. However, because item feature representations are built manually to some extent, this technique necessitates a great deal of domain knowledge. Furthermore, content-based recommendation systems may only propose based on the user's existing interests, and in addition, it limits their capacity to broaden users' existing interests.

Knowledge-based model. The primary purpose of the knowledge-based recommendation model is to suggest to the users according to basic knowledge of the users, items, and the relationship between items or users [8]. More importantly, this model has no cold-start problem because the model doesn't depend on the user/item rating or purchase history. In addition, knowledge-based is particularly useful for complex domains where items are not frequently purchased [9]. However, this recommendation mechanism can run into trouble if people don't have the proper domain knowledge [10].

Demographic-based model. Demographic-based is the recommendation system type where the user's demographic data is utilized as an alternative input for making suggestions [11]. The data can be user age, gender, location, etc.

Hybrid model. Hybrid model is the method that merge multiple recommendation models to over-come the issue in the existing recommendation models.

3 Specific Domain of Recommendation Systems

Recommendation systems are utilized in various application perspectives. It is a state-of-art application where people use techniques because it assists the user to save time to filter the valuable item/information to reduce their finding times. Recommendation systems have been applied in many eCommerce companies. For instance, amazon deployed the product recommendation in their web sites to suggest the most suitable item to their target customers. Existing works [12–15] designed a recommendation model to suggest the products to the target customers.

Other applications in movie domains proposed by Ref. [16–19] created a model to recommend movie entertainment based on various machine learning mechanisms. Not only in the range of entertainment or eCommerce, but recommendation systems are also widely used in healthcare. Ozsoy et al. [20] proposed a recommendation model to identify new targets for known drugs. This model aims to over-come the problem related to traditional drug discovery by using the existing drugs to treat newly discover diseases. Ref. [21] deploys the model of recommendation to predict a patient's health condition by analyzing the patient's lifestyle, physical health record, and social activities.

4 Conclusion

The simple survey focuses on modern recommendation systems in the big data era. In the study, we provide a simple explanation and survey of current recommendation systems such as collaborative, content-based, knowledge-based, demographic-based, and hybrid approaches. Furthermore, we illustrate the several issues which existed in the traditional recommendation models. Due to the demand of the big data era, recommendation models are built and researched more critically to adapt to real-world application and to find valuable knowledge filtering to figure out the pattern in the massive amount of data (big data).

Acknowledgements This research was supported by the National Research Foundation of Korea (No. NRF-2022R1A2C1005921) and BK21 FOUR (Fostering Outstanding Universities for Research) (No. 5199990914048).

References

1. Sahoo AK, Chakraverty S (2022) Machine intelligence in dynamical systems: a state-of-art review. Wiley Interdisc Rev Data Min Knowl Discov 12(4):e1461
2. Google Scholar Homepage. https://scholar.google.com. Last accessed 2022/10/31
3. Web of Science Homepage. https://www.webofscience.com. Last accessed 2022/10/31
4. Koren Y, Rendle S, Bell R (2022) Advances in collaborative filtering. In: Recommender systems handbook
5. Koohi H, Kiani K (2016) User-based collaborative filtering using fuzzy C-means. Measurement 91:134–139
6. Zhang Z, Zhang Y, Ren Y (2020) Employing neighborhood reduction for alleviating sparsity and cold start problems in user-based collaborative filtering. Inf Retrieval J 23(4):449–472
7. Mathew P, Kuriakose B, Hegde V (2016) Book recommendation system through content based and collaborative filtering method. In: 2016 international conference on data mining and advanced computing (SAPIENCE). IEEE, pp 47–52
8. Aggarwal CC (2016) Knowledge-based recommender systems. In: Recommender systems. Springer, Cham, pp 167–197
9. Tarus JK, Niu Z, Mustafa G (2018) Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. Artif Intell Rev 50(1):21–48
10. Dong M, Zeng X, Koehl L, Zhang J (2020) An interactive knowledge-based recommender system for fashion product design in the big data environment. Inf Sci 540:469–488
11. Safoury L, Salah A (2013) Exploiting user demographic attributes for solving cold-start problem in recommender system. Lect Notes Softw Eng 1(3):303–307
12. Cheng T (2019) Product recommendation system design. In: Proceedings of the 2019 2nd international conference on information management and management sciences, pp 71–74
13. Ge X, Zhang Y, Qian Y, Yuan H (2017) Effects of product characteristics on the bundling strategy implemented by recommendation systems. In: 2017 international conference on service systems and service management. IEEE, pp 1–6
14. Linden G, Smith B, York J (2003) Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput 7(1):76–80
15. Chiu MC, Huang JH, Gupta S, Akman G (2021) Developing a personalized recommendation system in a smart product service system based on unsupervised learning model. Comput Ind 128:103421
16. Vilakone P, Park DS, Xinchang K, Hao F (2018) An efficient movie recommendation algorithm based on improved k-clique. HCIS 8(1):1–15
17. Cui BB (2017) Design and implementation of movie recommendation system based on Knn collaborative filtering algorithm. In: ITM web of conferences, vol 12. EDP Sciences, p 04008
18. Ahuja R, Solanki A, Nayyar A (2019) Movie recommender system using K-means clustering and K-nearest neighbor. In: 2019 9th international conference on cloud computing, data science and engineering (confluence). IEEE, pp 263–268
19. Zhang J, Wang Y, Yuan Z, Jin Q (2019) Personalized real-time movie recommendation system: practical prototype and evaluation. Tsinghua Sci Technol 25(2):180–191
20. Ozsoy MG, Özyer T, Polat F, Alhajj R (2018) Realizing drug repositioning by adapting a recommendation system to handle the process. BMC Bioinf 19(1):1–14
21. Sahoo AK, Mallik S, Pradhan C, Mishra BSP, Barik RK, Das H (2019) Intelligence-based health recommendation system using big data analytics. In: Big data analytics for intelligent healthcare management, pp 227–246