

Optimizing Recommender System: Literature Review

Ghelani Vaibhavi Subhash^a, and Prof (Dr.) Purvi Ramanuj^b

^aM.E. Information Technology, L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India. 19itvai004@ldce.ac.in

^bAssistant Professor, L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India purviramanuj@gmail.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Recommending suitable product items to the target user is essential for the continued success of eCommerce. Today, many E-commerce systems adopt numerous recommendation techniques. In this review paper, we analyse various hybrid recommender systems. Existing recommendation methods have challenges of data sparsity and efficiency, as the numbers of users, items, and interactions between the two in real-world applications increase fast. In this paper, we review relevant problems about a recommendation system and describe the relevant recommendation techniques used to overcome them. We also explore the evaluation process and proposed future research directions. We also identify newer challenges such as recommendations based on user profile and providing cross-domain recommendations. Network algorithm-based hybrid recommendation system is a good base with which we can respond best by discovering innovative options such as domain-specific recommendations, processing larger datasets, etc.

Keywords: Data Analysis, Recommender System, Network-based Algorithm, Hybrid Recommender System, E-commerce

1. Introduction

With the rapid expansion of Internet technology and global computing, a variety of channels and methods to access information have brought great convenience for users. For this different domain experts take the help of a recommendation system that matches users with items. Recommendation system eases information overload and provides sales assistance through user guidance, advisory, and sometimes persuasion via various offers.

In this review paper, we summarize the state of the art of Recommendation Systems in the last 10 years. We follow a systematic methodology to analyze and interpret the available facts. This methodology provides an unbiased and reproducible way for undertaking a review works. Many primary studies were retrieved from digital libraries and the most relevant papers were selected for more detailed processing. We hope this work will help anyone working in the field of Recommendation Systems, especially by providing insights about future trends or opportunities.

2. Types Of Recommender System

There are 5 types of Recommendation System:

A. Personalized Recommendation System

This type of recommendation system makes use of user profile and some contextual parameters of users and provide personalized recommendations to users.

B. Collaborative Recommendation System

This type of recommendation system makes use of the user profile, some contextual parameters, and data of the community to which the user belongs. It recommends a similar product to a user which other users of his community are buying.

C. Content-based Recommendation System

This type of recommendation system makes use of user profile, contextual parameters as well as features of the product. Based on this it recommends the product to the user which has the same feature as the product he has already purchased before.

D. Knowledge-based Recommendation System

This type of recommendation system makes use of user profile, contextual parameters, product features and knowledge models which keeps track of certain event in user's demographics and accordingly do the recommendations. For example, in the case of the birthday of the user, it recommends a certain product. In the case of festivals related to the user's religion like Diwali, it recommends a certain product to the user.

E. Hybrid Recommendation

This type of recommendation system makes use of a combination of all four recommendation system methods/parameter and try to recommend the best suitable product for the user.

3. Research Problems

We summarize the most important Recommendation System problems the studies try to solve. Studies may address more than one problem. Below we describe each of the problems:

A. Cold-start

This problem has to do with recommendations for new users or items. The system has no information about new user preferences and thus fails to recommend anything to them. In the case of new items, the system has no ratings for these items and doesn't know whom to recommend. Cold-start was found in many studies.

B. Data sparsity

This problem arises from the fact that users usually rate a very limited number of the available items, especially when the catalogue is very large. Data sparsity was a research problem of many studies.

C. Scalability

This characteristic is difficult to achieve. It consists of several users and products. Nowadays many eCommerce systems consist of millions of customers. It is difficult to recommend them correct sets of products from their hundreds of products. Improving scalability was addressed in some studies.

D. Diversity

Our system must understand diversity. Our system should have a recommendation list that consists of all products which are similar to each other. If one customer is not interested in any of the products in that list, he might not be interested in any of them and gets no value from that recommendation list.

4. Recommendation Techniques

A. N2VSCDNNR [14].

In conventional recommendation methods, we face the problems of data sparsity and efficiency as the numbers of users, items, and interactions between the two increases fast. In this paper, the author has implemented an innovative clustering recommender system based on node2vec technology and a rich information network, namely, N2VSCDNNR, to solve these problems.

N2VSCDNNR Framework [14]:

There are 5 parts of this framework:

Phase I: Construction of Bipartite Networks

The first phase is to construct a bipartite network. In recommender systems, constructing a user-item bipartite network based on their relationships is abundant. However, many times we face the problem of data sparsity, for example, some of the customers have very few records; making the constructed bipartite network is not sufficient to capture the real relationship between users and items. Therefore, the author has introduced item categories to effectively solve the sparsity problem.

In particular, the construction of bipartite networks is performed in two steps:

1. Build two bipartite networks, i.e., the user-item bipartite network and the item-category bipartite network.
2. Build the user-category bipartite network by integrating the user-item bipartite network and the item-category bipartite network, where the author has a weight between user and category which is the total number of times that the user checks the items in this category.

Phase II: One-Mode Projection of Bipartite Networks

This phase again is performed by following two steps:

1. Project the user-item network into two separate networks, i.e., a user-user network and an item-item network, where the weight is the number of the common neighbors between the user (or item) and user (or item) in the corresponding bipartite networks with category, respectively.
2. For either users or items, the two projection networks are integrated as one network.

Phase III: Network Representation Using Node2vec

Although in the first two-phase the network structure is enriched according to item categories, it is difficult to capture appropriate network features using traditional network analysis methods. So, in this paper, the author has adopted node2vec to learn continuous feature representations of nodes in a network. Because of its neighborhood sampling strategy, node2vec can learn fruitful information from the network. This helps to decrease the effect of data sparsity. Node2Vec is used to automatically capture network features of the generated projection networks to transform each user (or item) into a vector.

Phase IV: Clustering Users and Items by SCDNN

In this method, we cluster the users and items. For this, the SCDNN method is used. This method is based on dynamic nearest-neighbors (DNNs) and automatically cluster number determination algorithm, to cluster users (items) into several clusters based on the related user (item) vectors.

Phase V: Two-Phase Personalized Recommendation

In this step, the first author recommends item clusters to user clusters, and then further realize the personalized recommendation is being done in the following 2 phases:

1. First, this method uses the number of user-item relationships between each item cluster and the target user cluster to quantify the item cluster. Then it used a basic clustering method, the K-means method, to divide all item clusters into two classes, and recommend the item clusters in the class with large average weights to the user cluster.
2. Based on the k-means clustering recommendation results, some traditional recommendation algorithms are adopted to recommend items in the related item clusters to the users in each user cluster, based on the related rating records.

B. A new method to find neighbor users that improves the performance of Collaborative Filtering [15]

One of the most successful techniques used in the Recommender System is Collaborative Filtering (CF). This method checks purchases made by other users to find products that are most similar to the target user. The collaborative Filtering efficiency is largely affected by Data sparsity and high dimensionality. In this paper, the author has tried to solve the mentioned problems through a neighbour user finding method.

This method consists of the following 5 phases:

Phase I: Data representation

Phase II: Constructing items subspaces

Phase III: Removing redundancy

Phase IV: Neighbour user finding by using neighbour trees

Phase V: Estimating rating value and give recommendation

C. Collaborative filtering using ontology and dimensionality reduction techniques [16]

In this research, the author has used dimensionality reduction and ontology techniques to solve the problem of sparsity and scalability. Then, the author has used ontology to improve the accuracy of recommendations in Collaborative Filtering. In the Collaborative Filtering part, the author has also used a dimensionality reduction technique, Singular Value Decomposition (SVD), to find the most similar items and users which can significantly improve the scalability of the recommendation method. Then the method is evaluated on two real-world datasets to show its effectiveness and compare the results with the results of methods in the literature. The results showed that the proposed method is effective in improving the sparsity and scalability problems in Collaborative Filtering.

D. Structural Balance Theory-based E-commerce Recommendation over Big Rating Data [17]

In the simple user-based collaborative Filtering method, this paper recommends the target user all the products which have been purchased or rated good by their friend. But it may happen that sometimes we can find a smaller number of a friend or no friends of the target user. For this, the given paper has approached SBT-Rec: Structural Balance Theory-based Product Item Recommendation Approach. SBT-Rec is introduced to recommend appropriate product items to the target user if the target user does not have similar friends, and product items preferred by the target user have no similar product items. Generally, SBT-Rec is based on “enemy’s enemy is a friend” and “enemy’s friend is an enemy” rules of Structural Balance Theory.

E. Adaptive Deep Modelling of Users and Items Using Side Information for Recommendation [18]

Matrix factorization (MF) is widely applied in traditional recommendation systems to model user preferences and item features by mapping the user-item ratings into a low-dimension latent vector space. However, Matrix Factorisation has ignored the individual diversity where the user’s preference for different unrated items is usually different. In this paper, the author has proposed a novel latent factor model called adaptive deep latent factor model (ADLFM), which learns the preference factor of users adaptively by the specific items under consideration. The author has also proposed a novel user representation method that is resulting from their rated item descriptions instead of original user-item ratings. Based on this, a deep neural network framework with an attention factor is further proposed to learn the adaptive representation of users.

5. Research Gap

We need to adapt the latest clustering algorithm to increase efficiency. We need to explore more networking algorithms for data sparsity. The model should be more robust and scalable. The threshold and another parameter of the model should be minimum and dynamic. Time complexity is still high in most of the models. Computational complexity is still high in most of the models.

6. Future Work

There is a need to propose a frame work for predicting the best-suited product to a range of users with different demographics with utmost accuracy and F1 Score (Planned matric for evaluation). The framework will use a combination of rich information networks, clustering and decision-making supervised algorithms. The framework will also make use of side information of users like demographics whenever available.

7. Conclusion

In the era of big data, a recommendation system helps users spend less time finding their favourite items. In this paper, we have surveyed the recent papers of reputed journals on solving data sparsity. Here we have summarized the approaches and techniques of the traditional and hybrid recommender systems, and discuss the major challenges and the advantages. Rich information network, clustering and decision-making model-based recommender system can learn the latent representation of users and items from massive data, and then construct a recommendation model which generates an effective recommendation list for the user.

TABLE I
SUMMARY OF ARTICLES ON RECOMMENDATION SYSTEM APPROACHES

Problem	Algorithms/Methods	Research Gap
---------	--------------------	--------------

To solve a problem of Data Sparsity and efficiency in a traditional recommendation system	One mode Projection of Bipartite Networks, Node2vec, Spectral Clustering, Dynamic Nearest Neighbor, K-Means	There is a still chance of improving Precision, Recall, and Time complexity.
To reduce the problem of Data sparsity and high dimensionality which are common in the traditional Recommendation system.	Subspace clustering approach, A new similarity method to compute similarity value, 54/321 rating method.	Less efficient in calculating similarity value as compared to another clustering algorithm.
To solve a problem of sparsity and scalability, using dimensionality reduction and ontology techniques.	Ontology, non-incremental Singular Value Decomposition (SVD).	There is still a chance of improving the Clustering method, scalability can also be improved.
To reduce the problem of the sparsity of Big Rating Data.	Structural Balance Theory	Similarity threshold assigned a fixed default value
Develop a system that learns the preference factor of users adaptively by the specific items under consideration.	Deep Neural Network	Cannot work properly for sparse datasets.
Hybrid recommender system which uses hypergraph topology.	Social network, hypergraph topology, matrix factorization	The computational complexity is high, and the number of the parameters is too high
Improving traditional recommendation systems for improving performance in predicted accuracy and recommended precision.	User correlation, evolutionary clustering	Need to join some other information of the network into the design for further performance enhancement.
To develop a more robust model as compared to the traditional model with penalty terms adapted automatically.	Gaussian-Gamma distribution, Gibbs samples	The model establishment is relatively complex when the distribution of the data is skewed.
To better use the intrinsic structure of the user-item rating matrix and user/item content information.	Dynamic single-element-based Collaborative Filtering integrating manifold regularization (DSMMF)	Need to Check the effectiveness of constraints of manifolds and graph regularization
To improve recommendation performance and alleviate the data sparsity.	Context-aware enhanced model based on Gaussian mixture model.	The establishment of the training model is complicated.

References

1. “Recommender Systems – An Introduction” by Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich Cambridge University Press.
2. R. Chen, Q. Hua, Y.-S. Chang, B. Wang, L. Zhang, and X. Kong, “A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks,” IEEE Access, vol. 6, pp. 64301– 64320, 2018.
3. Dietmar Jannach, Michael Jugovac, “Measuring the Business Value of Recommender Systems”. arXiv:1908.08328
4. “Recommender Systems – An Introduction” by Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich Cambridge University Press.
5. L. Wang, X. Meng, and Y. Zhang, “Context-aware recommender systems,” Journal of Software, vol. 23, no. 1, pp. 1-20, 2012.
6. F. Ricci, L. Rokach, B. Shapira, et al, “Recommender systems handbook: context-aware recommender systems,” New York: Springer, pp. 217-253, 2010.
7. Grover and J. Leskovec, “node2vec: Scalable Feature Learning for networks,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Jul. 2016, pp. 855–864.
8. J. Chen et al., “Improved spectral clustering collaborative filtering with Node2vec technology,” in proc. Int. Workshop Complex Syst. Netw. (IWCSN), Dec. 2017, pp. 330–334.

9. Z.-D. Zhao and M.-S. Shang, "User-based collaborative-Filtering recommendation algorithms on Hadoop," in Proc. 3d Int. Conf. Knowl. Discovery Data Mining, Jan. 2010, pp. 478–481.
10. B. Sarwar, G. Karypis, G. Karypis, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in Proc. 10th Int. Conf. World Wide Web, vol. 1, pp. 285–295, May 2001.
11. C. H. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 1, pp. 45– 55, Jan. 2010.
12. D. Mayer and D. Butler, "Statistical Validation", Ecological Modelling, vol.68, no. 1, pp. 21-32, 1993.
13. Y. Koren, "Factorization Meets the Neighborhood: A Multifaceted Collaborative Filtering Model" Proc 14th ACM SIGKDD International conference on knowledge discovery and Data Mining (KDD'08), pp. 426- 434, August 2008.
14. Jinyin Chen, Yangyang Wu, Lu Fan, Xiang Lin, Haibin Zhang, Shanqing Yu, and Qi Xuan, Member, IEEE "N2VSCDNNR: A Local Recommender System Based on Node2vec and Rich Information Network" Published In: IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, VOL. 6, NO. 3, JUNE 2019.
16. Hamidreza Koochi, Kourosh Kiani "A new method to find neighbor users that improves the performance of Collaborative Filtering" Published in Expert Systems with Applications, vol.83, pp. 30-39, 2017 (Elsevier).
17. Mehrbakhsh Nilashi, Othman Ibrahim, Karamollah Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques" Published in Expert Systems with Applications, vol.92, pp.5-7-520,2018.
18. Lianyong Qi, Xiao long Xu, Xuyun Zhang, Wanchun Dou, Chunhua Hu, Yuming Zhou, Jianguo Yu "SBT-Rec: Structural Balance Theory-based Product Item Recommendation Approach" Published In: IEEE TRANSACTIONS ON BIG DATA 1, Volume: 4, Issue: 3, Sept. 1, 2018.
19. Jiayu Han, Lei Zheng, Yuanbo Xu, Bangzuo Zhang, Fuzhen Zhuang, Philip S. Yu, Fellow, IEEE, and Wanli Zuo "Adaptive Deep Modelling of Users and Items Using Side Information for Recommendation" Published in: IEEE Transactions on Neural Networks and Learning Systems (Volume: 31, Issue: 3, March 2020).
20. Y. Kim. (2014). "Convolutional neural networks for sentence classification." [Online] Available: <https://arxiv.org/abs/1408.5882>
21. Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," Computer, vol. 42, no. 8, pp. 30–37, Aug. 2009.
22. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," in Proc. Adv. Neural Inf. Process. Syst., 2008, pp. 1257–1264.
23. J. McAuley and J. Leskovec, "Hidden factors and hidden topics: understanding rating dimensions with review text," in Proc. 7th ACM Conf. Recommender Syst., Oct. 2013, pp. 165–172.
24. G. Ling, M. R. Lyu, and I. King, "Rating meet reviews, a combined approach to recommend," in Proc. 8th ACM Conf. Recommender Syst., Oct. 2014, pp. 105–112.
25. L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for a recommendation," in Proc. 10th ACM Int. Conf. Web Search Data Mining, Feb. 2017, pp. 425–434.
26. X. Zheng, Y. Luo, L. Sun, et al, "A novel social network hybrid recommender system based on hypergraph topologic structure," World Wide Web- internet & Web Information Systems, no.3, pp.1-29, 2017.
27. D. Mayer and D. Butler, "Statistical Validation", Ecological Modelling, vol. 68, no. 1, pp. 21-32, 1993.
28. Jianguo Chen, Chunxia Zhao, Uliji, Lifang Chen, "Collaborative filtering recommendation algorithm based on user correlation and evolutionary Clustering". Complex & Intelligent Systems, 2020-Springer.
29. ChengLuoa, BoZhangb, YangXiangc, ManQic, "Gaussian-Gamma collaborative filtering: A hierarchical Bayesian model for recommender systems", Complex & Intelligent Systems, 2020 – Springer.
30. Yangyang Li, Dong Wang, Haiyang He, Licheng Jiao, and Yu Xue, "Mining Intrinsic Information via Matrix-Factorization-Based Approaches for Collaborative Filtering in Recommender Systems", Neurocomputing - August 2017 (Elsevier).

31. Chaochao Chen, Kevin Chen-Chuan Chang, Xiaolin Zheng, "Towards Context-Aware Social Recommendation via Individual Trust", Knowledge-Based Systems, vol.127(C), pp.58-66, 2017 (Elsevier).
32. <https://www.igiglobal.com/dictionary/recommendersystem/>
33. http://ijcai-ijcai.org/files/tutorial_slides/td3.pdf
34. <https://towardsdatascience.com/recommender-systems-the-most-valuable-application-of-machine-learning-part-1>
<https://www.fi.muni.cz/~xpelane/PV254/slides/intro.pdf>
<http://www.movieLens.org>
<http://webscope.sandbox.yahoo.com>
<http://imdb.com>
<http://cs.cmu.edu/~rcm/websphinx/>