An Extensive Review on Collaborative Filtering-Based Recommendation of Online Social Voting

Shubham Gangrade[1], Prof. Manish Rai[2], Dr. Mohit Gangwar[3]

[1]*M.Tech. Scholar* , [2]*Guide*, [3]*Co-Guide*

[1,2]*Department of Computer Science and Engineering, RKDF College of Engineering, Bhopal*

[3]*Principal, Bhabha Engineering Research Institute, Bhopal*

Abstract - The WWW has become an enormous source of information and it continues to increase in size and use. People are relying more and more on the Web not only for information sourcing, but also for other usages such as communicating, banking, investing, shopping, as well as for education and entertainment purposes. One of the popular usages of the WWW is for online shopping, where the buying and selling of products and services are conducted electronically. Recommender systems are the sub-class of the information filtering systems that are used to predict the user's rating or preference for a particular item. Recommender systems are applied in a wide variety of applications such as movies, news, music, books, search queries, research articles and products. Recommender systems use two ways for producing a set of recommendations. One is Collaborative Filtering and other is Content-based Filtering. This work present An Extensive Review on Collaborative Filtering-Based Recommendation system for Online Social Voting to attract more online customers.

Keywords- online social networks, recommender systems, social voting, Collaborative Filtering, Content-based Filtering.

I. INTRODUCTION

The growth of the internet and the cloud sources has made it difficult to extract the required useful information from all the available information. This large size of data requires techniques for efficient extraction of necessary information. This is called information filtering. An information filtering system is a system that removes redundant and unwanted information from an information stream using some automated or computerized methods before presenting it to the users.

Recommender systems or recommendation systems are a sub class of information filtering systems that are used to predict the rating or the preference given by the user to an item. There are different kinds of approaches for implementing the recommender systems among them collaborative filtering is one such approach.

Collaborative filtering is a technique used by some of the recommender systems in which pre-dictions are made automatically about the interest of the user by the information collected from users with similar interests.

Example: A group of users' rate different items (like videos, images, games). Generally, users do not give ratings for all the items. System predicts the ratings of items which are not yet rated. The motivation for CF is that people often get the best recommendation from the persons of similar interests. CF explores techniques of matching interests of different people and developing patterns in order to get recommendations.

Work flow of CF algorithms

Typically the work flow of CF system is

A user gives some rating to an item. This rating represents user's interest in that item.

The system searches for other users who gave similar ratings to this item.

Now, using the similarity information system predicts the ratings of items for which this user has not given any rating yet.

Types

A. Memory Based:

This computes the similarity between users or items based on user rating data which is used for making recommendations.

Advantages of this approach are implementation is easy and effective, results are explainable, new data can be added easily and scales well with co-rated items.

Disadvantages are its performance decreases when sparsity in data increases and even though it handles new users efficiently adding new items becomes complicated because it relies on a data structure.

Examples: User-based, Item-based recommendations.

B. Model based:

In this approach several algorithms are used to find patterns based on training data to develop models which are used to make predictions of real-time data.

Advantages of this approach are it handles sparsity better than memory-based algorithms, improves prediction performance and also gives an intuitive rationale for the recommendations. There are some disadvantages with this approach which are expensive model building and loss of information.

Examples: Singular value Decomposition and Probabilistic latent semantic analysis.

C. Hybrid:

This approach is the combination of memory-based and model-based ones. Advantages with this approach are this overcomes the limitations of native CF approaches, problems such as sparsity and loss of information. Disadvantage with this approach is that it is highly complex and expensive to implement.

II. RECOMMENDER SYSTEM

Burke (2000) defined a Recommender System (RS) as a computer system that provides advice to users about items they might wish to purchase or examine. A recommender system provides individual personalization to each user by customizing its recommendations and presenting different items for each user according to her/his tastes. By selecting and providing a list of products that are likely to fit a user's needs from a large number of product choices offered by an e- commerce site, recommender systems help the user deal with information overload, reduce the user search time for interesting items, and enhance the effectiveness of the user decision making. Furthermore, recommender systems also benefit merchandisers as they can enhance sales on their e-commerce sites by converting browsers into buyers, increasing cross-selling and building consumer loyalty. Nowadays, recommendation techniques are widely used in practical applications by commercial e-commerce websites such as Amazon.com, Moviefinder.com, and Reel.com, Levis.com and eBay for recommending various products such as books, CDs, movies, news, and articles to target users.

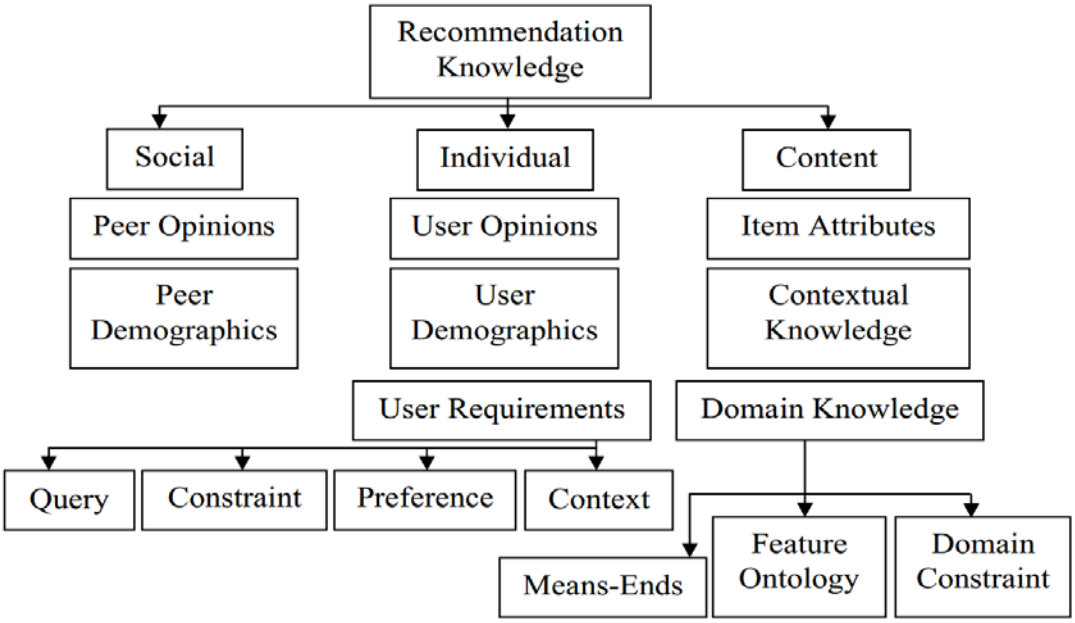

Figure 2.1 Knowledge Resources in Recommender System.

To provide a personalized set of recommendations, a recommender system incorporates a user's wishes into a user model and exploits suitable recommendation algorithms to map the user model into targeted product suggestions. There are three steps involved in a recommender system: acquiring preferences from a user's input data; computing the recommendation using proper techniques; and finally presenting the recommendation results to users. A recommender system suggests products by applying data analysis techniques to various pieces of knowledge gathered from different sources, that is, from the user, from peer users of the system, from data about the items being recommended, and also from the domain of recommendation itself, for example knowledge about what requirements recommended items satisfy. The items of knowledge can be acquired explicitly or implicitly from the sources. Examples of items of knowledge that can be acquired explicitly from the users are demographic data, ratings data, and product requirements as stated by the user in an online form. Knowledge about users' preferences can

also be acquired implicitly from the users' behaviour pattern data in the log data and also from transaction data in the database by using web mining and other web technologies. In addition, besides the knowledge generated by the user, knowledge can also be obtained from production data, for example, product attributes can be gathered from the product domain. Various knowledge sources used by recommender systems can be depicted in Figure 2.1.

III. RELATED WORK

SR. NO.	TITLE	AUTHOR	YEAR	METHODOLOGY
1	Collaborative Filtering-Based Recommendation of Online Social Voting,	X. Yang et al.,	2017	Developed a set of matrix-factorization (MF) and nearest-neighbor (NN)-based recommender systems (RSs)
2	Little Is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds	Meng Jiang, Peng Cui, Nicholas Jing Yuan, Xing Xie, Shiqiang Yang	2016	A novel semi-supervised transfer learning method to address the problem of cross-platform behavior prediction, called XPTRANS
3	Personalized recommendation of human resources based on preferences and personality types a collaborative filtering-based approach,	H. Outman, A. Souad and D. Ali,	2016	A new collaborative approach for the generic company's information system
4	Point-of-Interest Recommendations via a Supervised Random Walk Algorithm	G. Xu, B. Fu and Y. Gu,	2016	A new POI recommendation framework by simultaneously incorporating user check-ins and reviews
5	Cross-Domain Recommendation via Tag Matrix Transfer,	Z. Fang, S. Gao, B. Li, J. Li and J. Liao,	2015	A novel way to exploit the rating patterns across multiple domains by transferring the tag co-occurrence matrix information
6	A community driven social recommendation system,	D. Lalwani, D. V. L. N. Somayajulu and P. R. Krishna,	2015	Community detection algorithm, user-item based collaborative filtering, using map-reduce framework
7	A neighbor selection method based on network community detection for collaborative filtering,	L. Guo and Q. Peng,	2014	Propose a method that employs the evolutionary algorithm to optimize neighbors for all target users

X. Yang et al. [1] Social voting is an emerging new feature in online social networks. It poses unique challenges and opportunities for recommendation. In this work, develop a set of matrix-factorization (MF) and nearest-neighbor (NN)-based recommender systems (RSs) that explore user social network and group affiliation information for social voting recommendation. Through experiments with real social voting traces, demonstrate that social network and group affiliation information can significantly improve the accuracy of popularity-based voting recommendation, and social network information dominates group affiliation information in NN-based approaches. Also observe that social and group information is much more valuable to cold users than to heavy users. In our experiments, simple metapath-based NN models outperform computation-intensive MF models in hot-voting recommendation, while users' interests for nonhot votings can be better mined by MF models. Further propose a hybrid RS, bagging different single approaches to achieve the best top-k hit rate.

Meng Jiang, Peng Cui et.al .[2] People often use multiple platforms to fulfill their different information needs. With the ultimate goal of serving people intelligently, a fundamental way is to get comprehensive understanding about user needs. How to organically integrate and bridge cross-platform information in a human-centric way is important. Existing transfer learning assumes either fullyoverlapped or non-overlapped among the users. However, the real case is the users of different platforms are partially overlapped. The number of overlapped users is often small and the explicitly known overlapped users is even less due to the lacking of unified ID for a user across

different platforms. In this work, Propose a novel semi-supervised transfer learning method to address the problem of cross-platform behavior prediction, called XPTRANS. To alleviate the sparsity issue, it fully exploits the small number of overlapped crowds to optimally bridge a user's behaviors in different platforms. Extensive experiments across two real social networks show that XPTRANS significantly outperforms the state-of-the-art. Demonstrate that by fully exploiting 26% overlapped users, XPTRANS can predict the behaviors of non-overlapped users with the same accuracy as overlapped users, which means the small overlapped crowds can successfully bridge the information across different platforms.

H. Outman, A. Souad and D. Ali,[3] In an open and competitive world, companies must be organized internally to gain external challenge. Several companies and organizations suffer from problems related to the management of knowledge and knowledge capitalization. Knowledge management has become essential to competitiveness and development for each organization, which explains the large number of research and studies conducted in the field, thus good development of knowledge management requires a good mastery of information process (Acquire, represent, manage and capitalize). Among the threats that affect the development and success of this debate in most cases: the marginalized role of each individual, the lack of appropriate and technological solutions to promote the contribution of each to the common capital information and Finally, the overload of information given to them. In this work, after the analysis of several studies in the fields of knowledge management and social sciences, propose a new collaborative approach for the generic company's information system which is able to develop the mind of collective participation in knowledge capital management.

G. Xu, B. Fu and Y. Gu, [4] Recently, location-based social networks (LBSNs) such as Foursquare and Whrrl have emerged as a new application for users to establish personal social networks and review various points of interest (POIs), triggering a new recommendation service aimed at helping users locate more preferred POIs. Although users' check-in activities could be explicitly considered as user ratings, in turn being utilized directly for collaborative filtering-based recommendations, such solutions don't differentiate the sentiment of reviews accompanying check-ins, resulting in unsatisfactory recommendations. This article proposes a new POI recommendation framework by simultaneously incorporating user check-ins and reviews along with side information into a tripartite graph and predicting personalized POI recommendations via a sentiment-supervised random walk algorithm. The experiments conducted on real data demonstrate the superiority of this approach in comparison with state-of-the-art techniques.

Z. Fang, S. Gao, B. Li, J. Li and J. Liao,[5] Data sparseness is one of the most challenging problems in collaborative filtering(CF) based recommendation systems. Exploiting social tag information is becoming a popular way to alleviate the problem and improve the performance. To this end, in recent recommendation methods the relationships between users/items and tags are often taken into consideration, however, the correlations among tags from different item domains are always ignored. For that, in this work Propose a novel way to exploit the rating patterns across multiple domains by transferring the tag co-occurrence matrix information, which could be used for revealing common user pattern. With extensive experiments Demonstrate the effectiveness of our approach for the cross-domain information recommendation.

D. Lalwani, D. V. L. N. Somayajulu and P. R. Krishna,[6] Recommendation systems play an important role in suggesting relevant information to users. In this work, Introduce community-wise social interactions as a new dimension for recommendations and present a social recommendation system using collaborative filtering and community detection approaches. Use (i) community detection algorithm to extract friendship relations among users by analyzing user-user social graph and (ii) user-item based collaborative filtering for rating prediction. Developed our approach using map-reduce framework. Our approach improves scalability, coverage and cold start issue of collaborative filtering based recommendation system. Carried out experiments on MovieLens and Facebook datasets, to predict the rating of the movie and produce top-k recommendations for new (cold start) user. The results are compared with traditional collaborative filtering based recommendation system.

L. Guo and Q. Peng,[7] The neighbor selection that determines which users are exploited to estimate a target user's ratings has an important influence on the accuracy of recommendations of collaborative filtering based recommender system. Two kinds of ways for neighbor selection: KNN and cluster-based, are lack of specificity which refers to selecting different appropriate neighbors for different given target users, and thus limit the accuracy of recommendation. Therefore, in this work, firstly, propose a method that employs the evolutionary algorithm to optimize neighbors for all target users. Secondly, overcoming the high time complexity of the first one, present another approach in which community detection algorithm is utilized as a preprocessing, and then the evolution algorithm is employed to optimize the neighborhood size for every community. present

experiments on a standard benchmark data-set, and the results show that the two methods both realize the specificity in neighbor selection, and accordingly lead to a higher accuracy of recommendations. Besides, the second one makes a good compromise between the specificity and time complexity.

IV. PROBLEM STATEMENT

During the early years of the internet, while shopping was moving online at a slow pace, the old paradigm of "selection by inspection" seemed to work at first. The physical browsing and window shopping was literally translated into jumping from a rudimentary online shop to another. Still, arguably, a manageable task. There were few products online to choose from, and yet fewer websites with online shops set up. Things changed significantly with the inception of the Web 2.0 around the year 2004. Everything and everybody moved to the internet, shops and shoppers alike, and radical changes to the GUI of web pages, particularly those enabled by the AJAX responsive technology, permitted a much richer and immersive experience of browsing.

Luckily, online shops faced the question of targeted recommendation rather early, perhaps realizing that information would only scale with time, and that people would need help to find the items they want. Their quest took them to devise algorithms that could either learn one's taste, or measure the "similarity" between our taste and that of others, and present us with an educated recommendation as to what might like to browse next. Social and group information is much more valuable to improve recommendation accuracy for cold users than for heavy users. This is due to the fact that cold users tend to participate in popular votings. Voting content information can be mined for recommendation, especially for cold votings. Developing voting RSs customized for individual users, given the availability of multichannel information about their social neighborhoods and activities is the need of modern online social networking experience.

V. CONCLUSION

In this work study of various on Collaborative Filtering-Based Recommendation methods for online recommendation and voting. There drawbacks has been studied, and scope for development has been discussed. The quality of the results offered by a RS greatly depends on the quality of the results provided by its CF phase; i.e. it is essential to be capable of adequately selecting the group of users most similar to a given individual. In Collaborative filtering recommendations for each user (active user) are obtained in line with the preference of other users who have rated the products (items) in a similar way to the active user. Traditional CF research approaches are based on directly improving CF metrics and similarity measures. Traditional metrics display a marked tendency to show high similarity between users based on the similarity of their votes on a very small set of items. The popularly used collaborative filtering (CF) recommendation approach requires a large amount of explicit ratings data for making meaningful recommendations. However, this data is not always available as it requires high involvement from the users to provide explicit ratings of the products they already know.

REFERENCES

[1] X. Yang et al., "Collaborative Filtering-Based Recommendation of Online Social Voting," in IEEE Transactions on Computational Social Systems, vol. 4, no. 1, pp. 1-13, March 2017.

[2] Meng Jiang, Peng Cui, Nicholas Jing Yuan, Xing Xie, Shiqiang Yang, "Little Is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds" Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16).

[3] H. Outman, A. Souad and D. Ali, "Personalized recommendation of human resources based on preferences and personality types a collaborative filtering-based approach," 2016 Third International Conference on Systems of Collaboration (SysCo), Casablanca, 2016, pp. 1-6.

[4] G. Xu, B. Fu and Y. Gu, "Point-of-Interest Recommendations via a Supervised Random Walk Algorithm," in IEEE Intelligent Systems, vol. 31, no. 1, pp. 15-23, Jan.-Feb. 2016.

[5] Z. Fang, S. Gao, B. Li, J. Li and J. Liao, "Cross-Domain Recommendation via Tag Matrix Transfer," 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, 2015, pp. 1235-1240.

[6] D. Lalwani, D. V. L. N. Somayajulu and P. R. Krishna, "A community driven social recommendation system," 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, 2015, pp. 821-826.

[7] L. Guo and Q. Peng, "A neighbor selection method based on network community detection for collaborative filtering," 2014 IEEE/ACIS 13th International Conference on Computer and Information Science (ICIS), Taiyuan, 2014, pp. 143-148.

[8] H. Steck, "Training and testing of recommender systems on data missing not at random," in Proc. ACM KDD, 2010, pp. 713-722.

[9] B. Marlin and R. Zemel, "Collaborative prediction and ranking with non-random missing data," in Proc. ACM RecSys, 2009, pp. 5-12.

[10] X. Yang, H. Steck, Y. Guo, and Y. Liu, "On top-k recommendation using social networks," in Proc. ACM RecSys, 2012, pp. 67-74.

[11] Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu, "PathSim: Meta path- based top-k similarity search in heterogeneous information networks," Proc. VLDB Endowment, vol. 4, no. 11, pp. 992-1003, 2011.

[12] Y. Sun, B. Norick, J. Han, X. Yan, P. S. Yu, and X. Yu, "Integrating meta-path selection with user-guided object clustering in heterogeneous information networks," in Proc. ACM KDD, 2012, pp. 1348-1356.
[13] R. Pan et al., "One-class collaborative filtering," in Proc. ICDM, 2008, pp. 502-511.