

# Survey Paper on Recommendation System

Mukta kohar, Chhavi Rana  
 Department of computer science and Engg  
 U.I.E.T, MDU, Rohtak

**Abstract-**Today there is a big variety of different approaches and algorithms of data filtering and recommendation . In this paper we describe the recommendation system related research and then introduces various techniques and approaches used by the recommender system User-based approach, Item-based approach, Hybrid recommendation approaches and related research in the recommender system. In the end we will show the main challenges and issues recommender systems come across.

**Keywords-**Recommender system,, Content based algorithm, Collaborative filtering algorithm, Hybrid approach

## INTRODUCTION

The First recommender system[1] was developed by Goldberg, Nichols, Oki & Terry in 1992. Tapestry was an electronic messaging system that allowed users to either rate messages ("good" or "bad") Recommender system[1,2,3] as defined by M. Deshpande and G. Karypis: A personalized information filtering technology used to either predict whether a particular user will like a particular item (prediction problem) or to identify a set of N items that will be of interest to a certain user. Recommender systems form or work from a specific type of information filtering system technique that attempts to recommend information items (movies, TV program/show/episode, video on demand, music, books, news, images, web pages, scientific literature etc.) or social elements (e.g. people, events or groups) that are likely to be of interest to the user. Typically, a recommender system compares a user profile to some reference characteristics, and seeks to predict the 'rating' or 'preference' that a user would give to an item they had not yet considered. These characteristics may be from the information item (the content-based approach) or the user's social environment (the collaborative filtering). The recommender system apply data mining[4,5] techniques and prediction algorithms to predict user's interest on information , product and services user .

Recommender systems apply techniques and methodologies from another neighboring areas-such as Human computer interaction (HCI) or Information Retrieval (IR). However, most of these systems bear in their core an algorithm that can be understood as a particular instance of a data mining (DM) technique. The process of data mining consists of 3 steps, carried out in succession: Data Preprocessing, Data Analysis and Result Interpretation. Examples of recommender system are

amazon.com[2], Reel.com[2], CDNOW[2], eBay[2], Levis[2], Moviefinder.com[2]

## BACKGROUND

Recommender systems typically produce a list of recommendations in one of two ways - through collaborative or content-based filtering.

Content based algorithm recommender system[8] are the recommender system which work with profiles of users that are created at the beginning. A profile has information about a user and his taste. Taste is based on how user rated items. In the recommendation process, the engine compares the items that were already positively rated by the user with the items he didn't rate and looks for similarities. Those items that are mostly similar to the positively rated ones, will be recommended to the user.

Collaborative filtering Algorithm[8,6] recommender system became one of the most researched techniques of recommender systems since this approach was mentioned and described by Paul Resnick and Hal Varian in 1997. [1] The idea of collaborative filtering is in finding users in a community that share appreciations [23]. If two users have same or almost same rated items in common, then they have similar tastes[11]. Such users build a group or a so called neighborhood. A user gets recommendations to those items that he/she hasn't rated before, but that were already positively rated by users in his/her neighborhood

Various approaches of Collaborative Filtering are

- 1) User-based approach: This approach was proposed in the end of 1990s by the professor of University of Minnesota Jonathan L. Herlocker. In the user-based approach, the users perform the main role. If certain majority of the customers has the same taste then they join into one group. Recommendations are given to user based on evaluation of items by other users from the same group, with whom he/she shares common preferences. If the item was positively rated by the community, it will be recommended to the user[12,13].
- 2) Item-based approach: This approach was proposed by the researchers of University of Minnesota in 2001 [14]. Referring to the fact that the taste of users remains constant or change very slightly similar items build neighborhoods based on appreciations of users. Afterwards the system generates recommendations with items in the neighborhood that a user would prefer [14] [23]

- 3) Hybrid recommendation approaches  
For better results some recommender systems combine different techniques of collaborative approaches and contentbased approaches. The combination of approaches can proceed in different ways [23]:

- 1) Separate implementation of algorithms and joining the results.
- 2) Utilize some rules of content-based filtering in collaborative approach.
- 3) Utilize some rules of collaborative filtering in contentbased approach.
- 4) Create a unified recommender system, that brings together both approaches

We also have modern recommendation approaches Context –aware approaches, Semantic based approaches, Cross-domain based approaches, Cross-lingual approaches. separates the data in such a way that the margin is maximized

We also have various modern recommendation approaches such as context-aware approaches, Semantic-based approaches[20], cross-domain based approaches, peer-to-peer approaches and cross-lingual approaches.[17]

#### RELATED RESEARCH

In practice, research paper recommender systems do not exist. However, concepts have been published and partly implemented that could be used for their realisation. Some authors suggest using collaborative filtering and ratings. Ratings could be directly obtained by considering citations as ratings [15] or implicitly generated by monitoring readers' actions such as bookmarking or downloading a paper [16], [17]. Citation databases such as CiteSeer apply citation analysis (e.g. bibliographic coupling [18] or co-citation analysis [19], [20]), in order to identify papers that are similar to an input paper [21]. Scholarly search engines such as Google Scholar focus on classic text mining and citation counts.

Each concept does have disadvantages, which limits its suitability for generating recommendations.

For example[7], citation analysis cannot identify homographs<sup>2</sup>, and not all research papers are listed in citation databases. Likewise, reference lists can contain irrelevant entries caused by the Matthew Effect[16], self citations[17], citation circles[18] and ceremonial citations<sup>6</sup>. recommender systems cannot identify related papers if different terms are used.

Collaborative filtering in the domain of research paper recommendation is criticised for various reasons. Some authors claim that collaborative filtering would be ineffective in domains where more items than users exist [22]. Others believe that users would be unwilling to spend time for explicitly rating research papers [15]. Problematic with implicit ratings is that for obtaining the

required data, continuous monitoring of the researcher's work is necessary, which raises privacy issues[20]. In general, collaborative filtering has to cope with the possibility of manipulation. Another drawback is that a critical mass of ratings and users is required to receive useful recommendations.

#### CHALLENGES AND ISSUES

##### A. Cold-start[8]

It's difficult to give recommendations to new users as his profile is almost empty and he hasn't rated any items yet so his taste is unknown to the system. This is called the coldstart problem. In some recommender systems this problem is solved with survey when creating a profile. Items can also have a cold-start when they are new in the system and haven't been rated before. Both of these problems can be also solved with hybrid approaches.

##### B. Trust[8]

The voices of people with a short history may not be that relevant as the voices of those who have rich history in their profiles. The issue of trust arises towards evaluations of a certain customer. The problem could be solved by distribution of priorities to the users.

##### C. Scalability[8]

With the growth of numbers of users and items, the system needs more resources for processing information and forming recommendations. Majority of resources is consumed with the purpose of determining users with similar tastes, and goods with similar descriptions. This problem is also solved by the combination of various types of filters and physical improvement of systems. Parts of numerous computations may also be implemented offline in order to accelerate issuance of recommendations online.

##### D. Sparsity[8]

In online shops that have a huge amount of users and items there are almost always users that have rated just a few items. Using collaborative and other approaches recommender systems generally create neighborhoods of users using their profiles. If a user has evaluated just few items then it's pretty difficult to determine his taste and he/she could be related to the wrong neighborhood. Sparsity is the problem of lack of information. [14]

##### E. Privacy[8]

Privacy has been the most important problem. In order to receive the most accurate and correct recommendation, the system must acquire the most amount of information possible about the user, including demographic data, and data about the location of a particular user. Naturally, the question of reliability, security and confidentiality of the given information arises. Many online shops offer effective protection of privacy of the users by utilizing specialized algorithms and programs.

### CONCLUSION AND FUTURE WORK

This paper presented the various techniques and algorithm to build the recommender system

We also introduce various modern recommendation approaches such as context-aware approaches, Semantic-based approaches, cross-domain based approaches, peer-to-peer approaches and cross-lingual approaches. We also studied cold start problem [8]. A recent research topic in the context of the recommender system is recommendations to group. We have also uncovered areas that are open to many further improvements, and where there is still much exciting and relevant research to be done in coming years.

### REFERENCES

- [1] Witten I. H. and Frank I. *Data Mining*, Morgan Kaufman Publishers, San Francisco, 2000.
- [2] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, John Riedl, GroupLens: an open architecture for collaborative filtering of netnews, Proceedings of the 1994 ACM conference on Computer supported cooperative work, p.175-186, October 22-26, 1994, Chapel Hill, North Carolina, United States
- [3] John S. Breese, David Heckerman and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, pages 43-52, July 1998
- [4] Deshpande, M., and Karypis, G. Item-based top- $n$  recommendation algorithms. *ACM Trans. Inf. Syst.* 22, 1 (2004), 143-177.
- [5] Breese, J., Heckerman, D., and Kadie, C., Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, page 4352, 1998.
- [6] [http://en.wikipedia.org/wiki/Recommender\\_system](http://en.wikipedia.org/wiki/Recommender_system)
- [7] Goldberg, D., Nichols, D., Oki, B.M. and Terry, D. Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*, 35, 12 (December 1992), 51-60.
- [8] Cover, T., and Hart, P., Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1):21-27, 1967
- [9] van Meteren, R., and van Someren, M. Using content-based filtering for recommendation.
- [10] Cristianini, N., and J. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, March 2000.
- [11] K. O. et al. Context-aware svm for context-dependent information recommendation. In *International Conference On Mobile Data Management*, 2006.
- [12] Ghani, R., and Fano, A., Building recommender systems using a knowledge base of product semantics. In *In 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems*, 2002
- [13] Miyahara, K., and Pazzani, M.J., Collaborative filtering with the simple bayesian classifier. In *Pacific Rim International Conference on Artificial Intelligence*, 2000.
- [14] Pronk, V., Verhaegh, W., Proidl, A., and Tiemann, M., Incorporating user control into recommender systems based on naive bayesian classification. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*, pages 73-80, 2007
- [15] Xia, Z., Dong, Y., and Xing, G., Support vector machines for collaborative filtering. In *ACMSE 44: Proceedings of the 44th annual Southeast regional conference*, pages 169-174, New York, NY, USA, 2006. ACM.
- [16] Xu, J., and Araki, K., A svm-based personal recommendation system for tv programs. In *Multi-Media Modelling Conference Proceedings*, 2006.
- [17] Algorithms and Methods in Recommender Systems, Daniar Asanov Berlin Institute of Technology Berlin, Germany.
- [18] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. on Knowl. and Data Eng.*, vol. 17, pp. 734-749, June 2005.
- [19] R. Burke, "The adaptive web," P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin, Heidelberg: Springer-Verlag, 2007, ch. Hybrid web recommender systems, pp. 377-408.
- [20] A. Elgohary, H. Nomir, I. Sabek, M. Samir, M. Badawy, and N. A. Yousri, "Wiki-rec: A semantic-based recommendation system using wikipedia as an ontology," in *Intelligent Systems Design and Applications (ISDA)*, 2010 10th International Conference on, 29 2010.
- [21] S. Berkovsky, T. Kuflik, and F. Ricci, "Cross-domain mediation in collaborative filtering," in *Proceedings of the 11th international conference on User Modeling*, ser. UM '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 355-359.
- [22] R. Kanawati and H. Karoui, "A p2p collaborative bibliography recommender system," in *Proceedings of the 2009 Fourth International Conference on Internet and Web Applications and Services*. Washington DC, USA: IEEE Computer Society, 2009, pp. 90-96.
- [23] A. Elgohary, H. Nomir, I. Sabek, M. Samir, M. Badawy, and N. A. Yousri, "Wiki-rec: A semantic-based recommendation system using wikipedia as an ontology," in *Intelligent Systems Design and Applications (ISDA)*, 2010 10th International Conference on, 29 2010.