International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 06 | June-2016 www.irjet.net p-ISSN: 2395-0072

A Survey On Semantic Based Social Recommendation

Anjumol M[1], Ancy K Sunny[2]

[1]*Computer Science and Engineering, Vimal Jyothi Engineering College, Chemperi, Kannur, Kerala*
[2]*Asst. Professor, Computer Science and Engineering, Vimal Jyothi Engineering College, Chemperi, Kannur, Kerala*

Abstract - *Recommender systems or recommendation systems are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item. Recommender systems have become extremely common in recent years, and are applied in a variety of applications. Collaborative Filtering is one of the most successful approaches to building recommender system, which uses the known preferences of a group of users to make recommendations or predictions of the unknown preferences for other users. The most popular ones are probably movies, music, news, books, research articles, search queries, social tags, and products in general. However, there are also recommender systems for experts, jokes, restaurants, financial services, life insurance, and Twitter followers. The Matrix Factorization methods have been proposed for social recommendation due to their efficiency in dealing with large data sets. This paper provides an overview of recommender systems that include Collaborative, Content and Matrix factorization methods.*

1.INTRODUCTION

Social network users generate large amounts of information, which will be available on the web. But the main problem facing is difficult to find the information useful for us, among big amounts of useless one. The information on web is unstructured. As a result it can be very hard to find relevant elements of a specific category of item. The usual way to get around this problem is to use different retrieval and filtering techniques that are able to adapt to the user automatically or manually. To clarify the search of relevant items between all the available information, it is necessary to using the opinion of other people to know which of the items are really relevant. This procedure is known as Recommendation. The Recommender system is a part of Information Filtering Systems. Information filtering (IF) is one of the methods that is rapidly developing to manage large information flows. The aim of IF is to expose users to only information that is relevant or important to them. Many IF systems have been developed in recent years for various application spaces. Information filtering has recently begun to attract attention as a method for delivery of relevant information. IF system cover a wide scope of spaces, advancements and methods involved in the process of exposing users to the information they require. IF systems are applicable for unstructured or semi-structured data, handle large amounts of data, deal primarily with textual data, are based on user profiles, and their objective is to remove irrelevant data from incoming streams of data items.

There are many methods to get the opinion of the other people in order to retrieve relevant items for a specific person. The most of these methods work around the idea of finding similarities in the taste of the people, using historical information of them about preferences. Then the prediction for a specific person is based in the opinion of the most similar user to the person. This is called Collaborative filtering. Traditional Collaborative filtering techniques do not consider social relations, making them difficult to provide accurate recommendations[1]. Recently, Ma, Lyu and King [2], [3] proposed framework of Social recommender systems that made use of social relation data, can be exploited to improve accuracy of recommendations. Online social networks (OSN)[4] presents new opportunities as to further improve the accuracy of Recommendation Systems. Due to the stable and long-lasting social bindings, people are more willing to share their personal opinions with their friends, and typically trust recommendations from their friends more than those from strangers and vendors.

A novel framework for social recommendation is shown in fig 1. It shows the entire social contextual information which can be derived from links on social networks. Users typically examine item's content and information on senders. In this case the user cares about who the sender is and whether the sender is a close friend or authoritative. If more than one friend send the same content, the user may read it more attentively. This knowledge can be acquire from social relation and user-user interaction information. Both of these aspects are very important for the user to decide whether to

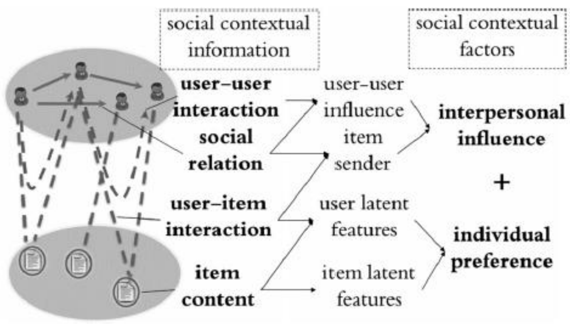

Fig- 1. A novel framework for social recommendation :it understands mechanism of user behavior on social networks, fully utilizes contextual information, and summarizes the knowledge as two social contextual factors.

adopt the particular item. The above can be summarized as two contextual factors. They are individual preference and interpersonal influence. In Bond's work in 1996, it is indicated that individuals are to some extent influenced by other's behavior, rather than making decisions independently. Therefore, when the individual preference and interpersonal influence are properly incorporated into recommendation, can the uncertainty be reduced and quality improved.

This paper is organized as follows. In section 2 give an introduction to Literature survey. In section 3, which describes the techniques of collaborative filtering. In section 4 describes the matrix factorization models. In section 5, which shows about the semantic based recommendations. Section 6 comes to the conclusion.

2. LITERATURE SURVEY

The techniques which is mainly used for Recommendations are Collaborative filtering and Content based filtering. The combination of these two technique is termed the Hybrid filtering approach.

One of the most researched techniques of recommender systems is Collaborative filtering technique, which was developed by Goldberg, Nichols, Oki, and Terry in 1992. This is also called Social filtering. This technique based on collecting and analyzing a large amount of information on users behaviors, activities or preferences and predicting what the users will like based on their similarity to other users. Collaborative filtering technique uses two systems, they are Memory based systems and Model based systems. One can add new data easily by using Memory based CF technique and Model based technique improves performance of prediction. There are certain limitations, cold-start, scalability and sparsity.

Another common approach when designing the recommendation system is the Content-based filtering. This methods are based on a description of a profile of the user's preference and the item. Basically this method uses an item profile characterizing the item within the system. This system helps the active user to build their own profile through the exclusive ratings provided by other users. This provide transparency to their active user by giving explanation how recommender system works. But it is harder to acquire feedback from users because they typically rank the items and therefore, not possible to determine whether the recommendation is correct.

Hybrid approach, combining collaborative and content-based filtering could be more effective in some cases. Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; it is by adding content-based capabilities to a collaborative-based approach and vice versa; or by unifying the approaches into one model. Several studies empirically compare the performance of the hybrid with the other two methods and demonstrate that the hybrid methods can provide more accurate recommendations than pure approaches. These methods can also be used to overcome some of the common problems in recommender systems such as cold start and the sparsity problem.

Among all the three filtering methods Jong Seo Lee found that Collaborative filtering is the most successful recommendation technique to date. Collaborative filtering decides recommendable items that a user might be interested in based on the information about many user's preference.

3. COLLABORATIVE FILTERING TECHNIQUES

Main categories of collaborative filtering techniques are Memory-based and Model-based CF techniques.

3.1. Memory-based CF Techniques

Memory-based CF[6] algorithm generate a prediction by using the entire or a sample of the user-item database. Every user is a part of a group of people with similar interests. By identifying the so called neighbors of a new user, a prediction of preferences on new items for the users can be produced. This algorithm is relatively simple to implement for any situation. The updation of the database is very easy, since it uses entire database every time it makes prediction. The memory-based technique uses the entire database every time it makes a prediction, so it needs to be in memory it is very slow. Sometimes the technique not make a prediction for certain active user or items. This can occur if the user has no items in common with all people who have rated the target items.

3.1.1. Neighborhood-based Collaborative Filtering

The neighborhood-based collaborative filtering also called heuristic-based[1] method. In this neighborhood-based filtering, the user-item ratings stored in the system are directly used to predict rating for new items.

3.1.1.1. User based Nearest Neighbor

User based algorithms[17] are CF algorithms that work on the assumption that each user belongs to a group of users who have similar behavior. The basic idea for the recommendation is composed by items that are liked by users. Items are recommended based on user's preference on items(in term of user's tastes). The algorithm considers that users who are similar (have similar attributes) will be interested on same items. User based algorithms are three steps algorithm; the first step is that in order to find which ones are similar to the target user, every user profiles will be created, the second step is to compute the union of the items selected by these users and based on the importance of the items in the set associate a weight with each of them and the third and final step is to choose and recommend items that have the highest weight and have not been already chosen by the active user. The most important step is the first one; creating the union of items liked by others or selecting the most important of them is easily done when the set of similar users is known. Thus the overall performance of the algorithm will depend on the method used to find users that are similar to the target user.

3.1.1.2. Item based Nearest Neighbor

The item-based approach investigates the set of items rated by target user and calculates their similarity with the target item i and then chooses *k* most similar items $\{i_1, i_2, \ldots, i_k\}$. Their representing similarities $\{s_{i1}, s_{i2}, \ldots, s_{ik}\}$ are also computed at the same time. Formerly the most similar items are discovered, after that the prediction is calculated in such a way that by taking a weighted mean of the target user's ratings on these similar items. Similarity computation and the prediction generation are two important factors which make item-based recommendation more powerful. For similarity computation basically different types of similarity measures are used and weighted sum and regression used for prediction computation. The main idea of this CF is that the user will like an item that is similar to items the user liked earlier, and will dislike an item that is similar to items the user disliked earlier.

3.1.2. Dimensionality reduction techniques

Dimensionality reduction[18] approach, proposed by Sarwar et al. in 2000, suggests that sparsity problem should be removed in the data by filling the null entries in the ratings matrix with the average ratings for an item (or the average ratings for a user).They then use singular value decomposition to produce a low-dimensional representation of the original domain. A rating for a user is predicted by regenerating that user's properties from the reduced space, but with altered ratings due to the decomposition. However, they do not provide any theoretical foundation for why the average ratings for an item (over all users) would be a good representation of the missing item rating and, if it is, then why not simply present this value as the prediction. Their other approach for generating recommendations reduces the original matrix to a low-dimensional space, then computes a neighborhood in that space. A recommendation is made using the neighbor's opinions about products they purchased. However, the approach only considers user preference data as binary values by treating each non-zero entry as '1', which again does not reflect how much (or if) a user liked a product, but only if she consumed it or not. Furthermore, they only evaluated their work, empirically, based on the quality of recommendations, with no regards to performance. This work is similar to one of our proposed techniques (the basic model). However our approach has no restrictions on the user-preference data types, and enables to compute recommendations in roughly linear time.

3.2. Model-based CF Techniques

To achieve fast and scalable real time recommendations on the basis of very large datasets, we are make use of Model-based Filtering techniques. From the dataset we extract some information, and use that as a "model "to make recommendations without having to use complete dataset every time. This approach offers the benefits of both speed and scalability. The model based methods learn a model based on patterns recognized in the ratings of users. Liu, E. Chen, and Xiong in 2012, build a model-based collaborative filtering framework with three layers which is to help personalized ranking on recommender systems. The three layers are user, interests, item. In the case of collaborative filtering which only utilizes user-item interaction information, therefore it is not able to make full use of social relation and rich knowledge including user profiles and detailed item content.

One of the disadvantages of this model is inflexibility. The inflexibility is because building a model is a time and resource consuming process, it is more difficult to add data to this model-based system, making them inflexible. Another disadvantage is quality of prediction. Since we are not using all the information available to us, so that we do not get more accurate predictions. However, that the quality of predictions depends on the way the model is built.

3.2.1. Clustering Collaborative Filtering Model

Clustering Collaborative filtering model[5] work by identifying groups of users who appears to have similar preferences. Once the clusters created, prediction for an individual can be made by averaging the opinions of the other users in that cluster. Sometimes the Clustering method work well, but often work poorly because data are often highly sparse and partly because people often have different tastes which put them in multiple categories.

3.2.2. Bayesian Belief Nets Collaborative Filtering

One of the most successful machine learning algorithm in many classification domains is the Bayesian Belief Nets. The simple Bayesian classifier is fast because its learning time is linear in the number of examples in the training data. Based on the training set the Bayesian networks create a model with a decision tree. In the tree at each node and edges representing user information. Here the model can be built off-line over a matter of hours or days. The resulting model is very small, very fast, and accurate.

3.2.3. Matrix factorization Method

Matrix factorization in its basic form, characterizes both items and users by vectors of factors inferred from item rating patterns. High correspondence between item and user factors leads to a recommendation. These methods have become popular in recent years by combining good scalability with predictive accuracy. In addition, they offer much flexibility for modeling various real-life situations. More on this Matrix factorization method is specified in the next section.

4. MATRIX FACTORIZATION MODELS FOR SOCIAL RECOMMENDATION

The Matrix factorization methods[7][10] have been proposed for social recommendation[11] due to their efficiency in dealing with large data sets. The idea of Matrix factorization is to model the user-item interactions with factors representing latent characteristics of users and items in the system. This model is then trained using the available data, and later used to predict ratings of users for new items. High correspondence between item and user factors leads to a recommendation. These methods have become popular in recent years by combining good scalability with predictive accuracy. In addition, they offer much flexibility for modeling various real life situations. MF techniques have become a dominant methodology within collaborative filtering recommenders. Recently, matrix factorization models have gained popularity to their attractive accuracy and scalability. Matrix Factorization is also called Matrix decomposition, is a factorization of a matrix into a product of matrices.

To improve the quality of recommendation, individual preference and interpersonal influence are properly incorporated. To address this problem, Meng Jiang, Peng and Fei in 2014 proposed a social contextual recommendation framework. This is based on a probabilistic matrix factorization method which is to incorporate individual preference and interpersonal influence to improve the accuracy of social recommendation. From psychological and sociological views, Bandura in 2001 gives a social cognitive theory of mass communication and argues that communication systems operate through two pathways. In the direct pathway, they promote changes by motivating and guiding participants to get what they prefer. In the socially-mediated pathway, participant's decisions are influenced by their friendship networks. Benjamin in 1974 shows the similar opinion that factors such as feeling, taste, interest and interpersonal relationship develop the structure of social behaviors and interactions. For social web, these two factors exactly represent individual preference and interpersonal influence. That motivates Meng Jiang, Peng Cui, Fei Wang in 2014[16] to propose a social contextual recommendation framework to incorporate them by analyzing both user motivation and application mechanism to recommender systems for social networks. Influence-based recommendation[12][13] involves interpersonal influence into social recommendation cases[14]. Trust-based approaches [15] exploit the trust network among users and make recommendation based on the ratings of users who are directly or indirectly trusted.

4.1. Singular Value Decomposition

Singular Value Decomposition (SVD) is a well established technique for identifying latent semantic factors in information retrieval. Applying SVD in the collaborative filtering domain requires factoring the user-item rating matrix. This often raises difficulties due to the high portion of missing values caused by sparseness in the user-item ratings matrix. Conventional SVD is undefined when knowledge about the matrix is incomplete. Moreover, carelessly addressing only the relatively few known entries is highly prone to over-fitting.

4.2. Non-negative Matrix factorization

Non-negative Matrix factorization, simply NMF, has successfully been used in many real world applications, such as information retrieval, computer vision, environmental study and computational social/network science. Different from the traditional decomposition methods such as Singular value decomposition, NMF (1) is usually additive, which results in a better interpretation ability; (2) does not require the factorized latent spaces to be orthogonal, which allows more flexibility to adapt the representation to the data set. NMF is factorize a nonnegative data matrix into the product of two nonnegative latent matrices. This will generally leads to sparse, part-based representation of the original data set, which is semantically much more meaningful compared to traditional factorization methods.

5. SEMANTIC BASED RECOMMENDATIONS

Exponential growth of information generated by social networks demands effective and scalable recommender systems to give useful results. Traditional techniques become unqualified because they ignore social relation data; existing social recommendation approaches consider social network structure, but social contextual information has not been fully considered. It is significant and challenging to fuse social contextual factors which are derived from user's motivation of social behaviors into social recommendation. So it is summarized as two contextual factors, individual preference and interpersonal influence. Besides the experiential assumptions, psychological and sociological studies have proved that individual both these affect user's decisions on information adoption. It demonstrates that the introduction of interpersonal influence into the preference-driven decision process makes user behaviors more complicated and thus increases the unpredictability of the item adoption. Therefore, only when individual preference and interpersonal influence are properly incorporated into recommendation, can the uncertainty be reduced and quality improved.

However, it is still an open issue about the relevance of the recommendation. Current recommendation systems have some common limitations: cold-start, sparsity, scalability, and diversity[8]. Although some particular combination of recommendation techniques can improve the recommendation's quality in some domains, but there is not a general solution to overcome these limitations. The use of semantics to formally represent data can provide several advantages in the context of personalized recommendation systems. We should think that the next generation of recommenders should focus on how their personalization processes can take advantage of semantics as well as social data to improve their recommendations. Here it describes how the accuracy of recommendation systems is higher when semantically enhanced methods are applied.

Semantic recommendation systems are characterized by the incorporation of semantic knowledge in their processes in order to improve the quality of recommendation. Most of them aim to improve the user-profile representation, employing a concept-based approach and using standard vocabularies and ontology languages like OWL(Web Ontology Language). Two different methods can be distinguished by Victor Codina and Luigi Ceccaroni. That are described below.

- Approaches employing spreading activation to maintain user interests and treating the user-profile as a semantic network. The interest scores of a set of concepts are propagated to other related concepts based on pre-computed weights of concepts relations.
- Approaches that apply domain-based inferences, which consist of making inferences about user's interests based on the hierarchical structure defined by the ontology. The most commonly used is the upward-propagation, whose main idea is to assume that the user is interested in a general concept if he is interested in a given percentage of its direct sub-concepts. This kind of mechanisms allows inferring new knowledge about the long-term user's interests and therefore modeling richer user-profiles.

Most of the recommendations make use of semantic similarity methods to enhance the performance of content-based approach, although there are some recommenders using semantics to enhance the user-profile matching of a collaborative filtering approach. The only recommender that makes use of semantic reasoning methods in both stages of the personalization process is AVATAR[9].

Social recommendations are based on social community of the user. It contains user's friends with trust values which expressing how much the active user trusts his friends. The user annotates his relationships with such information. Trust can be binary (trust or don't trust) or on some scale, 1-5 scale where 1 is low trust and 5 is high trust. Based on these trust values, user's social neighborhood can be inferred over the social network. In the Semantic view which represents user's interest about content of the item. For this purpose semantic content representation of the item is needed.

Recommendation systems can take advantage of semantic reasoning-capabilities to overcome common limitations of current systems and improve the recommendation's quality. The recommender uses domain ontologies to enhance the personalization: on the one hand, user's interests are modeled in a more effective and accurate way by applying a domain-based inference method; on the other hand, the matching algorithm used by our content-based filtering approach, which provides a measure of the affinity between an item and a user, is enhanced by applying a semantic similarity method. The new generation of recommenders should focus on how their personalization processes can take advantage of semantics as well as social data to improve their recommendations. The problem statement is that, by applying semantically-enhanced methods how the accuracy of recommendation systems is improved.

6. CONCLUSION

Recommender systems are turning out to be a useful tool that will provide suggestion to user according to their requirements. Recommender system typically produce a list of recommendations in one of two ways through collaborative or content based filtering. The matrix factorization method techniques have become a dominant methodology within collaborative filtering recommenders. The matrix factorization methods have been proposed for social recommendation due to their efficiency in dealing with large data sets. By using the matrix factorization methods the users will not get the exact recommendations. It will be more accurate by take advantage of semantics as well as social data to improve their recommendations.

REFERENCES

[1] G. Adomavicius, A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", IEEE Trans. Knowledge and Data Eng., June 2005.

[2] H. Ma, H. Yang, M.R. Lyu, I. King, "SoRec Recommendation Using Probabilistic Matrix Factorization", Proc. 17th ACM Conf. Information and Knowledge Management (CIKM 08), 2008.

[3] H. Ma, D. Zhou, C. Liu, M.R. Lyu, I. King, Recommender System with Social Regularization, Proc. Fourth ACM Intl Conf. Web Search and Data Mining (WSDM 11) ,2011.

[4] Yang Guo Xiwang Yang, A survey of collaborative filtering based social recommender systems, Preprint submitted to Computer Communications, 2013.

[5] Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl, Item-based collaborative filtering recommendation algorithms, Proceedings of the 10th International Conference on World Wide Web (New York, NY, USA), ACM, 2001.

[6] Y. Koren, Collaborative Filtering with Temporal Dynamics, Comm. ACM,2010.

[7] Y. Koren, Matrix Factorization Techniques for Recommender Systems, Computer,2009.

[8] Sunita Barve Poonam B. Thorat, R. M. Goudar, Survey on collaborative filtering, content-based filtering and hybrid recommendation system, International Journal of Computer Applications,2015.

[9] Blanco-Fernndez, Y. et al, A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems, Knowledge-Based Systems,2008.

[10] F. Wang, H. Tong, and C.Y. Lin, Towards Evolutionary Nonnegative Matrix Factorization, Proc. Assoc. Advancement of Artificial Intelligence (AAAI 11), 2011.

[11] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang, Social Contextual Recommendation, Proc. 21st ACM Intl Conf. Information and Knowledge Management (CIKM 12), 2012.

[12] J. Huang, X. Cheng, J. Guo, H. Shen, and K. Yang, Social Recommendation with Interpersonal Influence, Proc. 19th European Conf. Artificial Intelligence (ECAI 10), 2010.

[13] A. Goyal, F. Bonchi, and L.V.S. Lakshmanan, Learning Influence Probabilities in Social Networks, ACM Int'l Conf. Web Search and Data Mining, 2010.

[14] P. Cui, F. Wang, S. Yang, and L. Sun, Item-Level Social Influence Prediction with Probabilistic Hybrid Factor Matrix Factorization, Proc. Assoc. Advancement of Artificial Intelligence (AAAI 11) 2011.

[15] B. Carminati, E. Ferrari, and J. Girardi, Trust and Share: TrustedInformation Sharing in Online Social Networks, Proc. IEEE 28th Intl Conf. Data Eng. (ICDE 12), 2012.

[16] Meng Jiang, Peng Cui, Fei Wang, Wenwu Zhu, Fellow, ScalableRecommendation with Social Contextual Information, IEEE, November 2014.

[17] Dhoha Almazro, Ghadeer Shahatah,Lamia Albdulkarim, Mona Kherees, A Survey Paper on Recommender Systems, arXiv ,December 2010.

[18] A. M. Roumani, D. B. Skillicorn, A Dimensionality Reduction Technique for Collaborative Filtering, Queen's University ,November 2006.