ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

A Survey on Recommendation Systems based on Online Social Communities

Kashmira Nagwekar

M. E Student, Dept. of Computer Engineering, Vidyalankar Institute of Technology, Wadala, Mumbai, India

ABSTRACT: During the last two decades we have witnessed the tremendous amount of growth in e-commerce industry. People all over the world buy articles just by a click of mouse. Today recommendation system is an important part of almost every website. A user might not be able to find out all the desired articles and items from the endless information pool available on the internet. Recommender system suggests those items to the user which are most suitable to the user based on his data of items purchased and his ratings collected over a period of time, which helps to predict the buying behaviour of the user.Recommendation systems play an important role in suggesting relevant information to users. Community-wise social interactions as a new dimension for recommendations and a social recommendation system using trust-based collaborative filtering and community detection approaches is proposed. I will be using (i) community detection algorithm to extract friendship relations among users by analysing user-user social graph and (ii) trust-based collaborative filtering for rating prediction.

KEYWORDS: social recommendation system; trust-based collaborative filtering; community detection; cold-start

I. INTRODUCTION

E-Commerce sites are gaining popularity across the world. People visit them not just to shop products but also to know the opinion of other buyers and users of products. Online customer reviews are helping consumers to decide which products to buy and also companies to understand the buying behaviour of consumers.

Recommendation systems play an important role in suggesting relevant information to users. Community-wise social interactions form a new dimension for recommendations. A social recommendation system using collaborative filtering and community detection approaches is proposed. I will be using (i) community detection algorithm to extract friendship relations among users by analysing user-user social graph and (ii) trust-based collaborative filtering for rating prediction. I will be developing the approach using map-reduce framework. This approach will improve scalability, coverage and cold start issue of collaborative filtering based recommendation system.

II. RELATED WORK

Collaboration, interaction and information sharing are the main driving forces of the current generation of web applications referred to as 'Web 2.0'. Well-known examples of this emerging trend include weblogs (online diaries or journals for sharing ideas instantly), Friend-Of-A-Friend (FOAF) files (machine-readable documentsdescribing basic properties of a person, including links between the person and objects / people they interact with), wikis (web applications such as Wikipedia that allow people to add and edit content collectively) and social networking sites (virtual communities where people with common interests can interact, such as Facebook, dating sites, car addict forums, etc.). We focus on one specific set of Web 2.0 applications, namely *social recommender systems*. These recommender systems generate predictions (recommendations) that are based on information about users' profiles and relationships between users. Nowadays, such online relationships can be found virtually everywhere, think for instance of the very popular social networking sites Facebook, LinkedIn and MSN. Research has pointed out that people tend to rely more on recommendations from people they trust (friends) than on online recommender systems which generate recommendations based on anonymous people similar to them [2]. This observation, combined with the growing popularity of open social networks and the trend to integrate e-commerce applications with recommender systems, has generated a rising interest in *trust-enhanced recommendation systems*. The recommendations generated by these

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

systems are based on information coming from an (online) social network which expresses how much the members of the community trust each other.

Augmenting a recommender system by including trust relations can help solving the sparsity problem. Moreover, a trust-enhanced system also alleviates the cold start problem: it has been shown that by issuing a few trust statements, compared to a same amount of rating information, the system can generate more, and more accurate, recommendations.

Recommender Systems (RSs) are software tools and techniques providing suggestions for items to be of use to a user [2]. The suggestions relate to various decision-making processes, such as what items to buy, what music to listen to, or what online news to read.'Item' is the general term used to denote what the system recommends to users. A RS normally focuses on a specific type of item (e.g., CDs, or news) and accordingly its design, its graphical user interface, and the core recommendation technique used to generate the recommendations are all customized to provide useful and effective suggestions for that specific type of item.RSs are primarily directed towards individuals who lack sufficient personal experience or competence to evaluate the potentially overwhelming number of alternative items that a Web site, for example, may offer. A case in point is a book recommender system that assists users to select a book to read. In the popular Website, Amazon.com, the site employs a RS to personalize the online store for each customer. Since recommendations are usually personalized, different users or user groups receive diverse suggestions. In addition there are also non-personalized recommendations. These are much simpler to generate and are normally featured in magazines or newspapers. Typical examples include the top ten selections of books, CDs etc. While they may be useful and effective in certain situations, these types of non-personalized recommendations are not typically addressed by RS research.RS can play a range of possible roles. First of all, we must distinguish between the roles played by the RS on behalf of the service provider from that of the user of the RS. In fact, there are various reasons as to why service providers may want to exploit this technology:

- Increase the number of items sold.
- Sell more diverse items.
- Increase the user satisfaction.
- Increase user fidelity.
- Better understand what the user wants.

Herlocker et al. [17], in a paper that has become a classical reference in this field, define eleven popular tasks that a RS can assist in implementing. Some may be considered as the main or core tasks that are normally associated with a RS, i.e., to offer suggestions for items that may be useful to a user. Others might be considered as more "opportunistic" ways to exploit a RS. As a matter of fact, this task differentiation is very similar to what happens with a search engine, Its primary function is to locate documents that are relevant to the user's information need, but it can also be used to check the importance of a Web page (looking at the position of the page in the result list of a query) or to discover the various usages of a word in a collection of documents.

- Find Some Good Items
- Find all good items
- Annotation in context
- Recommend a sequence
- Recommend a bundle
- Just browsing
- Find credible recommender
- Improve the profile
- Express self
- Help others
- Influence others

As these various points indicate, the role of a RS within an information system can be quite diverse. This diversity calls for the exploitation of a range of different knowledge sources and techniques used to identify the right recommendations.

Several different types of recommender systems that vary in terms of the addressed domain, the knowledge used, but especially in regard to the recommendation algorithm, i.e., how the prediction of the utility of a recommendation is made. Other differences relate to how the recommendations are finally assembled and presented to the user in response

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

to user requests. A taxonomy provided by [17] that has become a classical way of distinguishing between recommender systems and referring to them. [17] distinguishes between six different classes of recommendation approaches:

I. Content-based Recommender Systems:

The system learns to recommend items that are similar to the ones that the user liked in the past. The similarity of items is calculated based on the features associated with the compared items. For example, if a user has positively rated a movie that belongs to the comedy genre, then the system can learn to recommend other movies from this genre.

II. Collaborative filtering based Recommender Systems:

The simplest and original implementation of this approach recommends to the active user the items that other users with similar tastes liked in the past. The similarity in taste of two users is calculated based on the similarity in the rating history of the users. This is the reason why collaborative filtering is referred to as as "people-to-people correlation." Collaborative filtering is considered to be the most popular and widely implemented technique in RS.

III. Demographic Recommender Systems:

This type of system recommends items based on the demographic profile of the user. The assumption is that different recommendations should be generated for different demographic niches. Many Web sites adopt simple and effective personalization solutions based on demographics. For example, users are dispatched to particular Web sites based on their language or country. Or suggestions may be customized according to the age of the user. While these approaches have been quite popular in the marketing literature, there hasbeen relatively little proper RS research into demographic systems.

IV. Knowledge-based Recommender Systems:

Knowledge-based systems recommend items based on specific domain knowledge about how certain item features meet users' needs and preferences and, ultimately, how the item is useful for the user. Notable knowledge based recommender systems are case-based [2]. In these systems a similarity function estimates how much the user needs (problem description) match the recommendations (solutions of the problem). Here the similarity score can be directly interpreted as the utility of the recommendation for the user.

Knowledge-based systems tend to work better than others at the beginning of their deployment but if they are not equipped with learning components they may be surpassed by other shallow methods that can exploit the logs of the human/computer interaction (as in CF).

V. Community-based Recommender Systems:

This type of system recommends items based on the preferences of the users friends. This technique follows the epigram "Tell me who your friends are, and I will tell you who you are" [2]. Evidence suggests that people tend to rely more on recommendations from their friends than on recommendations from similar but anonymous individuals. This observation, combined with the growing popularity of open social networks, is generating a rising interest in community-based systems or, as or as they usually referred to, social recommender systems. This type of RSs models and acquires information about the social relations of the users and the preferences of the user's friends. The recommendation is based on ratings that were provided by the user's friends. In fact these RSs are following the rise of social-networks and enable a simple and comprehensive acquisition of data related to the social relations of the users.

The research in this area is still in its early phase and results about the systems performance are mixed. Some reports suggest that social-network based recommendations are no more accurate than those derived from traditional CF approaches, except in special cases, such as when user ratings of a specific item are highly varied (i.e. controversial items) or for cold-start situations, i.e., where the users did not provide enough ratings to compute similarity to other users. Others have showed that in some cases social-network data yields better recommendations than profile similarity data and that adding social network data to traditional CF improves recommendation results.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

III. PROBLEM STATEMENT

Social recommendation systems (SRS) consider social relationships among users as an additional input while deriving recommendations for specific users. There are many techniques to build recommender systems, which can be classified into Content based methods, Collaborative Filtering (CF) based methods and Hybrid methods.

Despite significant improvements on recommendation approaches, some important problems still remain. Some of the weaknesses of collaborative filtering systems are:

- Due to data sparsity, a collaborative filtering algorithm experiences a lot of difficulties when trying to identify good neighbours in the system. Consequently, the quality of the generated recommendations might suffer from this.
- Moreover, it is also very challenging to generate good recommendations for users that are new to the system (i.e., cold start users), as they have not rated a significant number of items and hence cannot properly be linked with similar users.
- Thirdly, because recommender systems are widely used in the realm of e-commerce, there is a natural motivation for producers of items (manufacturers, publishers, etc.) to abuse them so that their items are recommended to users more often. For instance, a common 'copy-profile' attack consists in copying the ratings of the target user, which results in the system thinking that the adversary is most similar to the target.
- Finally, the users prefer more transparent systems, and that people tend to rely more on recommendations from people they trust ('friends') than on online recommender systems which generate recommendations based on anonymous people similar to them.

Social media and networks are prominently influencing people's social activities with their family, friends and colleagues, which results in rich social relations such as friendships in Facebook, follow relations in Twitter and trust relations in Epinions. Homophily indicates that users with similar preferences are more likely to be connected, and social influence reveals that users who are connected are more likely to have similar preferences. The state-of-the-art SRSs do not consider the role of connectedness among users (i.e., communities). The proposed approach use community driven social interactions (or trust) to improve the effectiveness of recommendation. Social recommendation systems (SRS) consider social relationships among users as an additional input while deriving recommendations for specific users.

IV. PROPOSED ALGORITHM

A SRS which is based on Trust based CF and interactions over the social network, is proposed. Unlike traditional recommendation systems that assume all users are independent, we exploit social interactions or connections among users to make recommendation. We will be using the Louvain's community detection algorithm to find communities in user-user social graph using friendship dimension.

Mining a Trust Network

The most common trust-enhanced recommender strategies ask their users to explicitly issue trust statements about other users. Take for instance the e-commerce site Epinions.com which orders reviews based on a trust network that it maintains by asking its users to indicate which members they trust (i.e., their personal web of trust) or distrust (block list).All these systems exploit the relations in the trust network to determine which opinions or ratings should weigh more or less in the recommendation process. In other words, this group of algorithms uses the trust estimates as weights in the decision process. This weighting can be done in several ways. The most commonly used strategies that are adaptations of the collaborative filtering mechanism.

A. Trust Based Collaborative Filtering

In collaborative filtering, a rating of target item i for target user a can be predicted using a combination of the ratings of the neighbours of a (similar users) that are already familiar with item i.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

The classical formula is given by the following equation:

$$p_{a,i} = \overline{r_a} + \frac{\sum_{u \in R^{+}} w_{a,u} (r_{u,i} - \overline{r_u})}{\sum_{u \in R^{+}} w_{a,u}} \qquad \text{eq. (1)}$$

The unknown rating pa,i for item i and target user a is predicted based on the mean ra of ratings by a for other items, as well as on the ratings ru,i by other users u for i. The formula also takes into account the similarity wa,u between users a and u, usually calculated as Pearson's Correlation Coefficient (PCC). In practice, most often only users with a positive correlation wa,u who have rated i are considered. We denote this set by R+. However, instead of a PCC-based computation of the weights, once can also infer the weights through the relations of the target user in the trust network (again through propagation and aggregation); see eq. (2) which adapts eq. (1) by replacing the PCC weights wa,u by the trust values ta,u. This strategy is also supported by the fact that trust and similarity are correlated.

$$p_{a,i} = \overline{r_a} + \frac{\sum_{u \in R} t_{a,u} (r_{u,i} - \overline{r_u})}{\sum_{u \in R^{T}} t_{a,u}} \qquad \text{eq. (2)}$$

This alternative is called trust-based collaborative filtering.

B. Louvain's Community Detection (CD)

Louvain's CD algorithm is based on the modularity optimization and can analyze significantly larger networks very accurately with considerably less computational time.The modularity of a partition is a scalar value between -1 and 1 that measures the density of links inside communities as compared to links between communities.

Description of the Proposed Algorithm:

In the proposed approach,

(1) Community structure is obtained from social network,

(2) This structure is used to separate the user-item rating matrix into a number of groups with the number being the same the number of communities,

(3) Both community level average rating and user-specific rating are calculated, and

(4) The two types of rating are used to predict the rating for both cold-start users and normal users.

Solution to Cold-Start Problem

The problem of cold-start for a user happens in recommendation systems when sufficient information is not available or the user is new to the system. In such cases, if the user has a history in another system then we can use his/her external profile to recommend relevant items. For example, a user is new to movie system but has a profile in book system; we can find users similar to him/her in terms of their rating in the book system, to have recommendation in movie system.
In our case, if the user is new to movie system (Movie Lens), but having profile in Facebook then we find friends of his/her using CD algorithm.
We assume that a user preference is similar to or influence by their socially connected friends.
Let G(V, E) is user-user social undirected and un-weighted graph. Vertices V represent users in social network and edges E represent the friendship relation between users.
We apply Louvain's CD algorithm on graph G to find communities in user-user social network.
Next, we divide user-item rating data into p groups according to communities found in user-user social graph.
Then, we calculate two different prediction matrices:

1. For each item i belonging to community C, find the average rating of the item i in community C.
2. Second prediction matrix is calculated using Item based CF.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

1) Recommendation to a cold-start user

Consider a new user W with no rating history.

Our recommendation approach to the user W works as follows:

1. Find the community C to which the user W belongs to.
2. Compute the prediction matrix based on item average rating for user W in community C, i.e., Pavg(w,i,c) where i Є Ic.
3. Based on score Pavg(w,i,cw) obtained in learning, recommend the top-n items, where n is number of recommended items.

2) Recommendation to Normal User

For recommendation to normal user we combine two prediction matrices.

Let user u belongs to community C.

Prediction value for item i by user u in community C is calculated as P(u,i,c) = α Pavg (u,i,c) + β PCF(u,i,c), where α + β =1. In this work, we considered α = β = 0.5. Based on score P(u,i,c) obtained in learning, set of recommended items will be maxnicЄI P(u,i,c).

V. PSEUDO CODE

Implementation details of CCSRS algorithm using MapReduce is given as follows.

An assumption is made that process of recommendation in each community is independent of each other. The algorithm is such that it runs p same independent jobs, but with different inputs, one for each community C. Three important parts of computation are:
(1) Compute average rating of each item in community C,
(2) Compute similarity between items in community C, and
(3) Compute the rating prediction for item by user in community C.

For community C, the whole computation executes in five MapReduce phases:

- MR-1 calculates an average rating for each item in C.
- MR-2 phase prepares data for similarity calculation in C.
- MR-3 phase calculates similarity between items in C using cosine similarity matrix.
- MR-4 prepares data for rating prediction calculation.
- MR-5 computes rating prediction for item by user in C.

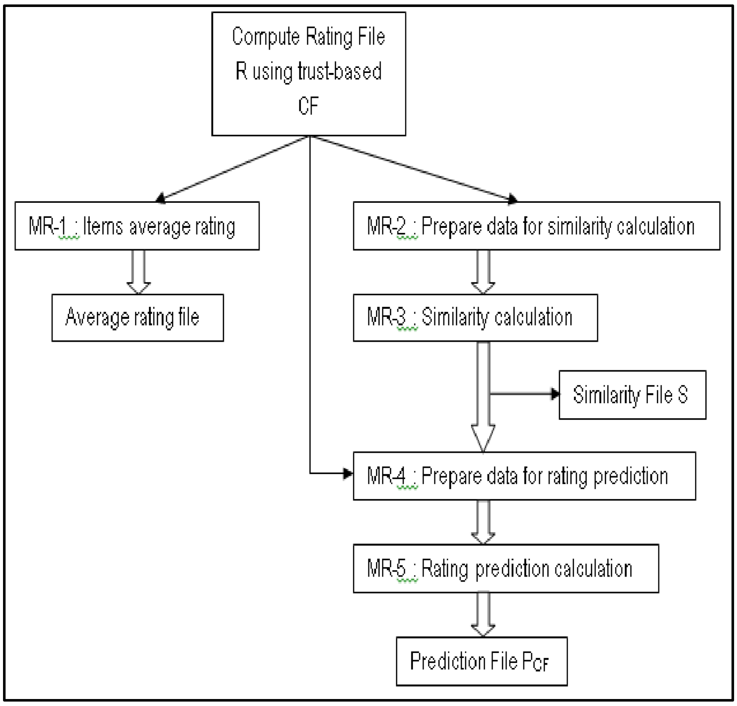

Fig. 1. Block Diagram for CCSRS approach.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

VI. CONCLUSION AND FUTURE WORK

The Social Recommendation System, SRS uses collaborative filtering to provide suggestions. Accuracy issues of recommendation system, in case of cold start and over-personalization are resolved by utilizing data from SRS. Augmenting a recommender system by including trust relations can help solving the sparsity problem. Moreover, a trust-enhanced system also alleviates the cold start problem: it has been shown that by issuing a few trust statements, compared to a same amount of rating information, the system can generate more, and more accurate, recommendations. Moreover, a web of trust can be used to produce an indication about the trustworthiness of users and as such make the system less vulnerable to malicious insiders: a simple copy-profile attack will only be possible when the target user, or someone who is trusted by the target user, has explicitly indicated that he trusts the adversary to a certain degree.

REFERENCES

1. Deepika Lalwani, D. V. L. N. Somayajulu, P. Radha Krishna, "A community driven social recommendation system", IEEE International Conference on Big Data (Big Data), Year: 2015, Pages: 821-826, DOI: 10.1109/BigData.2015.7363828.
2. Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor., "Recommender Systems Handbook", Vol. 1, New York: Springer, 2011, ISBN 978-0-387-85819-7, e-ISBN 978-0-387-85820-3, DOI 10.1007/978-0-387-85820-3.
3. Venkata Rajeev P., Smrithi Rekha V., "Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features", 2015 Int. Conf. on Circuit, Power and Computing Technologies [ICCPCT].
4. Aymen Elkhelifi, Firas Ben Kharrat, Rim Faiz, "Recommendation Systems Based on Online User's Action", DOI 10.1109/CIT/IUCC/DASC/PICOM 2015.
5. Akshay Chadha, Preeti Kaur, "Comparitive Analysis of Recommendation System", 2015 4th Int. Symposium on Emerging Trends and Technologies in Libraries and Information Services.
6. D. Blondel, J. Guillaume, R. Lambiotte and E. Lefebvre, "Fast Unfolding of Communities in Large Networks – Louvain Method", J. of Statistical Mechanics: Theory and Expt., vol. P1000 (10), 2008.
7. S. Chen, S. Owusu and L. Zhou, "Social Network Based Recommendation Systems: A Short Survey",2013 Int. Conf. on Social Computing (SocialCom), pp. 882-885, 2013.
8. M. Fatemi and L. Tokarchuk, "A Community Based Social Recommender System for Individuals & Groups", IEEE Int. Conf. on Social Computing, pp. 351–356, 2013.
9. M. Ge, C. Delgado-Battenfeld and D. Jannach, "Beyond Accuracy: Evaluating Recommender Systems by Coverage and Serendipity", ACM RecSys'10,2010.
10. J. Jiang, J. Lu, G. Zhang and G. Long, "Scaling-up Item-based Collaborative Filtering Recommendation Algorithm based on Hadoop", IEEE SERVICES, pp. 490-497, 2011.
11. J. Kamahara, T. Asakawa, S. Shimojo and H. Miyahara, "A Community-based Recommendation System to reveal Unexpected Interests", IEEE 11th Int.Multi.Model. Conf., 2005.
12. I. Konstas, V. Stathopoulos and J. M Jose, "On Social Networks and Collaborative Recommendation", ACM SIGIR'09, pp. 195-202, 2009.
13. H. Ma, H. Yang, M. R. Lyu and I. King, "Sorec: Social recommendation using probabilistic matrix factorization", CIKM 2008.
14. S. Sahebi and W. Cohen, "Community Based Recommendations: a Solution to the Cold Start Problem", RSWEB Workshop, 2011.
15. B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based Collaborative Filtering Recommendation Algorithms", 10th Int. Conf. on WWW, pp. 285-295, 2001.
16. J. Tang, X. Hu and H. Liu, "Social Recommendation: A Review", Social Network Analysis and Mining, vol. 3(4), pp. 1113-1133, 2013.
17. Burke, R.: Hybrid web recommender systems. In: The AdaptiveWeb, pp. 377–408. Springer Berlin / Heidelberg (2007).

BIOGRAPHY

Ms.KashmiraSandeepNagwekaris currentlypursuing Master's Degree in Computer Engineering from Vidyalankar Institute of Technology, Wadala, Mumbai (affiliated to University of Mumbai). She received Bachelor's Degree in Information Technology from Finolex Academy of Management and Technology (affiliated to University of Mumbai), Ratnagiri, Maharashtra, India.