A Trust-Based Matrix Factorization Method for Recommendations

Manjula H. Badiger
Dept. of CSE, SDM College of Engineering & Technology,(Affiliated to VTU, Belgavi), Dharwad-580002, Karnataka, India

Govind G. Negalur
Dept. of CSE,SDM College of Engineering & Technology,(Affiliated to VTU, Belgavi), Dharwad-580002, Karnataka, India

ABSTRACT
A trust-based matrix factorization method for recommendations merge several information sources into the recommendation model in order to diminish the data sparsity and cold start problems and their abasement of recommendation performance. An analysis of social trust data propose that not only the explicit trust influence the ratings but also the implicit influence should be taken into consideration in a recommendation model. The method therefore builds on top of the futuristic recommendation algorithm, SVD++ by further incorporating both the explicit and implicit influence of trusted and trusting users on the forecast of items for a current user. The proposed method extends SVD++ with social trust information.

Keywords
Recommender systems, social trust, matrix factorization, implicit trust, collaborative filtering

1. INTRODUCTION
The largest E-commerce sites offer millions of products for sale. Choosing among so many options is challenging for consumers. Recommender systems have emerged in response to this problem. A recommender system for an Ecommerce site receives information from a consumer about which products she is interested in, and recommends products that are likely to fit her needs. Today, recommender systems are deployed on hundreds of different sites, serving millions of consumers.

Recommender systems have been widely used to suggest items (movies, books, music, news, Web pages, images, etc.) that are likely to interest the users from a large volume of choices. Robust and accurate recommendations are important in e-commerce operations and in marketing. Collaborative filtering (CF) is one of the most popular techniques to implement a recommender system [1]. Collaborative filtering is a technique that forecasts the interest of a current user by gathering rating details from other indistinguishable users or items. The idea of CF is that active users will prefer those items which the identical users will prefer. CF has also been applied to tasks besides item recommendations, in domains such as image processing [2] and bioinformatics [3]. However, due to the nature of collaborative filtering, recommender systems based on this method suffers from two well-known issues: data sparsity and cold start [4]. The data sparsity arises due to the fact that users usually rate only a small portion of items, while the cold start indicates the dilemma that accurate recommendations are expected from new users but they only give few ratings and rate only few items that are difficult to reveal their desires. Both issues severely decreases the efficiency of a recommender system in modeling user desires and thus the correctness of fore casting a user's rating for an unrevealed item.

Traditional recommender systems assume that users are independent and identically distributed; this assumption ignores the social interactions or connections among users. But in our real life, when we are asking our friends for the recommendations we are essentially soliciting a verbal social recommendation. For example, when you ask a friend for a recommendation of a movie to see or a good restaurant, you are actually appealing for verbal social recommendations. Social recommendation is a daily occurrence, and we always turn out to our friends for recommendations. If a choice is between recommendations from friends and those from recommender systems, in terms of quality and usefulness, friends' recommendations are preferred, even though the recommendations given by the recommender systems have high novelty factor. Friends are seen as more qualified to make good and useful recommendations compared to traditional recommender systems. From this point of view, the traditional recommender systems that ignore the social network structure of users may no longer be suitable.

To help resolve these issues, many researchers [5], [6], [7], [8], [9] attempt to incorporate social trust information into their recommendation models, given that model-based CF approaches outperform memory-based approaches [10]. To investigate this phenomenon, we conduct an empirical trust analysis through which two important observations are concluded. First, trust information is also very sparse, yet complementary to rating information. Hence, focusing too much on either one kind of information may achieve only marginal gains in predictive accuracy. Second, users are strongly correlated with their trust neighbors whereas they have a weakly positive correlation with their trust-alike neighbors (e.g., friends). Given that very few trust networks exist, it is better to have a more general trust-based model that can well operate on both trust and trust-alike relationships.

These observations motivate us to consider both explicit and implicit influence of ratings and of trust in a trust-based model. The influence can be explicit (real values of ratings and trust) or implicit (who rates what (for ratings) and who trusts whom (for trust)). The implicit influence of ratings has been demonstrated useful in providing accurate recommendations and later will show how implicit trust can also provide added value over explicit trust.

Our approach builds on top of a state-of-the- art model SVD++ where both the explicit and implicit influence of user-item ratings are involved to generate prediction. The work is the first to extend SVD++ with social trust information. Specifically, on one hand the implicit influence of trust (who trusts whom) can be naturally added to the SVD++ model by extending the user modeling. On the other hand, the explicit influence of trust (trust values) is used to constrain that user specific vectors should conform to their social trust relationships.

This ensures that user-specific vectors can be learned from their trust information even if a few or no ratings are given. In this way, the data sparsity and cold start issues can be better alleviated. Our novel model thus incorporates both explicit and implicit influence of item ratings as well as user trust. In addition, a weighted-regularization technique is used to further avoid over-fitting for model learning.

2. RELATED WORK

Trust-aware recommender systems have been widely studied, given that social trust provides an alternative view of user preferences other than item ratings (Guo, Zhang, and Yorke-Smith 2014). Specifically, Ma et al. (2008) propose a social regularization method (SoRec) by considering the constraint of social relationships. The idea is to share a common user-feature matrix factorized by ratings and by trust. Ma, King, and Lyu (2009) then propose a social trust ensemble method (RSTE) to linearly combine a basic matrix factorization model and a trust-based neighborhood model together. Ma et al. (2011) further propose that the active user's user-specific vector should be close to the average of her trusted neighbors, and use it as a regularization to form a new matrix factorization model (SoReg). Jamali and Ester (2010) build a new model (SocialMF) on top of SoRec by reformulating the contributions of trusted users to the formation of the active user's user-specific vector rather than to the predictions of items. Yang et al. (2013) propose a hybrid method (TrustMF) that combines both a truster model and a trustee model from the perspectives of trusters and trustees, that is, both the users who trust the active user and those who are trusted by the user will influence the user's ratings on unknown items. Tang et al. (2013) consider both global and local trust as the contextual information in their model, where the global trust is computed by a separate algorithm. Yao et al. (2014) take into consideration both the explicit and implicit interactions among trusters and trustees in a recommendation model. Fang, Bao, and Zhang (2014) stress the importance of multiple aspects of social trust. They decompose trust into four general factors and then integrate them into a matrix factorization model. All these works have shown that a matrix factorization model regularized by trust outperforms the one without trust. That is, trust is helpful in improving predictive accuracy. However, it is also noted that even the latest work (Fang, Bao, and Zhang 2014) can be inferior to other well-performing ratings-only models.

3. TRUST ANALYSIS

There are two main recommendation tasks in recommender systems, namely item recommendation and rating prediction. Many approaches have been proposed for rating prediction, including both memory- and model-based methods. Most algorithmic approaches are only (or best) designed for either one of the recommendations tasks, and our work focuses on the Rating prediction task. However, memory-based approaches have difficulty in adapting to large-scale data sets, and are often consume much time in searching candidate neighbors in a large user space. In contrast, model based approaches can be readily scaled up to large data sets and they generate rating predictions more efficiently. Most importantly, they have been demonstrated to achieve higher accuracy and better alleviate the data sparsity issue than memory-based approaches. Also shown that a matrix factorization model regularized by trust outperforms the same model without trust. That is, trust is helpful in improving predictive accuracy.

We conduct a trust analysis to investigate the value of trust in recommender systems. And we argue that the main reasons are in two-fold. First, the existing trust-based models consider only the explicit influence of ratings; the utility of ratings is not well exploited. We show that trust information could be even sparser than rating information. This motivates us to build a new trust-based model based on SVD++ that inherently and well considers both the explicit and implicit influence of ratings. Second, these trust-based models do not consider the explicit and implicit influence of trust simultaneously. These observations may lead to deteriorated performance when being applied to social relationships with smaller correlations with user similarity. Therefore, we incorporate into SVD++ both explicit and implicit influence of social trust, to enhance the generality of our proposed model. By doing so, a better way to utilize user–item ratings and user–user trust is proposed.

In social rating networks, a user can label (add) other users as trusted friends and thus form a social network. Trust is not symmetric; for example, users u1 trusts u3 but u3 does not specify user u1 as trustworthy. Besides, users can rate a set of items using a number of rating values, e.g., integers from 1 to 5. These items could be products, movies, music, etc. of interest. The recommendation problem is to predict the rating that a user will give to an unknown item, for example, the value that user u3 will give to item i3, based on both a user–item rating matrix and a user–user trust matrix.

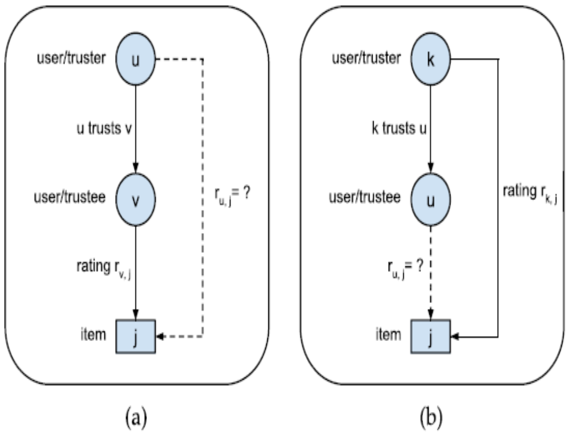

Figure 2. The influence of (a) trustees v and (b) trusters k on the rating prediction for the active user u and target item j.

Two important observations concluded are

Observation 1: Trust information is very sparse, yet is complementary to rating information. A trust-aware recommender system that focuses too much on trust (rather than rating) utility is likely to achieve only marginal gains in recommendation performance. In fact, the existing trust-based models consider only the explicit influence of ratings. That is, the utility of ratings is not well exploited. In addition, the sparsity of explicit trust also implies the importance of involving implicit trust in collaborative filtering. On the other hand, trust information is complementary to the rating information. Figure 1a shows that: (1) A portion of users have not rated any items but are socially connected with other users. (2) For the cold-start users who have rated few items (less than 5 in our case), trust information can provide a complementary part of source of information with ratio greater than 10% on average. (3) The warm-start users who have rated a lot of items (e.g., > 20) are not necessary to specify many other users as trustworthy (12% on the average). Although having differing distributions across the data sets, trust can be a complementary information source to item ratings for recommendations.

This observation motivates us to consider both the explicit and implicit influence of ratings and trust, making better and more use of ratings and trust to resolve the concerned issues.

Observation 2: A user's ratings have a weakly positive correlation with the average of her social neighbors under the concept of trust-alike relationships, and a strongly positive correlation under the concept of trust relationships. Next, we consider the influence of trust in rating prediction, i.e., the influence of trusted neighbors on the active user's rating for a specific item, a.k.a. social influence. Specifically, we calculate the Pearson correlation coefficient (PCC) between a user's ratings and the average of her social neighbors. The observations made are (1) A weakly positive correlation is observed between a user's ratings and the average of the social neighbors the distributions of the two data sets are similar. Some adopts the symmetric friendship relationships whereas trust is directed.

Although few adopts the concept of trust (with values from 1 to 10), the publicly available data set contains only binary values (such degrading may cause much noise). We regard these relationships as trust alike, i.e., the social relationships that are similar with, but weaker (or more noisy) than social trust. (2) Under the concept of trust relationships, on the contrary, a user's ratings are strongly and positively correlated with the average of trusted neighbors.

In the social networks with relatively weak trust-alike relationships (e.g., friendship), implicit influence (i.e., binary relationships) may be more indicative than explicit (but noisy) values for recommendations. Hence, a trust-based model that ignores the implicit influence of ratings and trust may lead to deteriorated performance if being applied to such cases. The second observation suggests that incorporating both the explicit and implicit influence of ratings and trust may promote the generality of a trust-based model to both trust and trust-alike social relationships.

4. CONCLUSION

This paper proposed a trust-based matrix factorization model which incorporated both rating and trust information. Our analysis of trust in four real-world data sets indicated that trust and ratings are complementary to each other, and both pivotal for more accurate recommendations. Our approach takes into account both the explicit and implicit influence of ratings and trust information when predicting ratings of unknown items. A weighted-regularization technique was adapted and used to further regularize the user- and item-specific latent feature vectors. Comprehensive experimental results showed that our approach outperformed both trust- and ratings based methods in predictive accuracy across different testing views and across users with different trust degrees.

5. REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering (TKDE), vol. 17, no. 6, pp. 734–749, 2005.

[2] Y. Huang, X. Chen, J. Zhang, D. Zeng, D. Zhang, and X. Ding, "Single-trial fERPsg denoising via collaborative filtering on fERPsg images," Neurocomputing, vol. 149, Part B, pp. 914 – 923,2015.

[3] X. Luo, Z. Ming, Z. You, S. Li, Y. Xia, and H. Leung, "Improving network topology-based protein interactome mapping via collaborative filtering," Knowledge-Based Systems (KBS), vol. 90, pp. 23–32, 2015. IEEE Transactions on Knowledge and Data Engineering,Volume:28,Issue:7,Issue Date :July.1.2016 14

[4] G. Guo, J. Zhang, and D. Thalmann, "A simple but effective method to incorporate trusted neighbors in recommender systems," in Proceedings of the 20th International Conference on User Modeling, Adaptation and Personalization (UMAP), 2012, pp. 114– 125.

[5] H. Ma, H. Yang, M. Lyu, and I. King, "SoRec: social recommendation using probabilistic matrix factorization," in Proceedings of the 31st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2008, pp. 931–940.

[6] H. Ma, D. Zhou, C. Liu, M. Lyu, and I. King, "Recommender systems with social regularization," in Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM), 2011, pp. 287–296.

[7] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in Proceedings of the 4th ACM Conference on Recommender Systems (RecSys), 2010, pp. 135–142.

[8] B. Yang, Y. Lei, D. Liu, and J. Liu, "Social collaborative filtering by trust," in Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI), 2013, pp. 2747–2753.

[9] H. Fang, Y. Bao, and J. Zhang, "Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation," in Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI), 2014, pp. 30–36.

[10] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," Computer, vol. 42, no. 8, pp. 30–37, 2009.

[11] Y. Koren, "Factor in the neighbors: Scalable and accurate collaborative filtering," ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 4, no. 1, pp. 1:1–1:24, 2010.

[12] "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2008, pp. 426 434.

[13] G. Guo, J. Zhang, and N. Yorke-Smith, "TrustSVD: collaborative filtering with both the explicit and implicit influence of user trust and of item ratings," in Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI), 2015, pp. 123–129.

[14] R. Forsati, M. Mahdavi, M. Shamsfard, and M. Sarwat, "Matrix factorization with explicit trust and distrust side information for improved social recommendation," ACM Trans. Inform. Syst.,vol. 32, no. 4, pp. 17:1–17:38, 2014.