Improving Recommendation using Implicit Trust Relationships from Tags

Hyunwoo Kim and Hyoung-Joo Kim,

Seoul National University, Seoul, Korea
hwkim@idb.snu.ac.kr, hjk@snu.ac.kr

Abstract. In this paper, we propose an implicit trust relationship extraction method for the alleviation of the sparsity problem in recommender systems. This problem, which is a serious weakness of collaborative filtering based recommender systems, refers to situations where such a system cannot generate relevant items because a user-item matrix is sparse. In a social tagging system, tagging information is a useful data source for recommendation. We investigate the elicitation of implicit trust relationships from tagging information. The relationships are derived by Kullback-Leibler divergence of users' tagged items and tags. The experimental results show the proposed method provides relevant items precisely and performs well in practice.

Keywords: Recommendation, Trust relationships, Tagging information, Recommender systems.

1 Introduction

Item recommendation is the task that of identifying and recommending items of high relevance to a given user. Correct item recommendation is increasingly important because of information overload: it is impossible for a user to search all items to discover items that match the user's preferences, because the number of existing items is too large.

Recommender systems have incorporated the methods of item content analysis and user ratings of items. Collaborative filtering is one of the most successful methods for recommendation. A collaborative filtering method is based on user similarity. Given a user, the system searches for similar users and recommends the items that the similar users are interested in. The system utilizes rating information associated with the users similar to the given user. User similarity is a subjective, personal, and symmetric relationship. The symmetry of user similarity means that if a user u_a is similar to a user u_b, then u_b is necessarily similar to u_{ci}. The sparsity problem is a serious weakness of the collaborative filtering method [1]. A user cannot rate all items in the recommender system because the number of items is too large. The problem occurs when the user-item matrix has high sparsity and no overlap between users [2]. It makes it difficult to search for similar users.

Trust is "a subjective expectation an agent has about another's future behavior based on the history of their encounters" [3]. In recommender systems, some users may trust a certain user but others may not trust that user. It depends on the users' personal interests or preferences. Like user similarity, trust is a subjective and personal relationship. Unlike user similarity, however, trust is an asymmetric relationship: if u_a trusts u_b, it does not necessarily mean that u_b trusts u_a. This asymmetry, which is a significant feature of trust relationships, facilitates the propagation of trust relationships. A trust relationship has direction and is inherently transitive. Trust relationships are also dynamic: they gradually develop and keep changing over time. By utilizing trust information, a recommender system can alleviate the sparsity problem [4]. Nevertheless, most recommender systems provide no means of representing explicit trust relationships between users.

Social tagging represents a new opportunity for researchers who study recommendation. In a social tagging environment, a user annotates an item with keywords, called tags, relevant to the item. Annotated tags form a folksonomy. Tags provide more information about an item than a rating because tags also reflect user preferences. Users generate a folksonomy with their tags. Many studies on recommendation have already taken advantage of these aspects of social tagging [5].

In this paper, to improve the performance of recommender systems, we propose a recommendation method based on the use of implicit trust relationships that are derived from tagged items and tags in social tagging system.

2 Related Work

There are two ways to incorporate the use of trust information in recommender systems. In the first approach, a recommender system utilizes explicit information about trust relationships provided by users. In the second approach, a recommender system elicits implicit trust relationships from user data. Users do not provide their trust information. Instead, the recommender system derives trust relationships from the user information, e.g., their profiles, rated items, ratings, and tags. Papagelis et al. have proposed a method of trust inference for the alleviation of the sparsity problem of collaborative filtering [6]. O'Donovan and Smyth have also proposed trust-based recommender systems [7]. Bhuiyan et al. have proposed a method of building trust networks from user tagging information [1]. Although this method uses tags in the building of a trust network, it also users item descriptions, and so is in applicable to recommender systems that do not provide item descriptions.

3 Recommendation Model

We propose a method of recommending items to a user with using implicit trust relationships of the user derived from user tagging information. A user tagging information consists of triples that include users, items, and tags. A *<user, item, tag>* triple is a basic building block of the tagging information. If a user u annotates a tag t to an item i, then the triple $<u, i, t>$ is stored in the recommender systems dataset. We utilize these triples to elicit implicit trust relationships. One naïve approach is to conduct the extraction of implicit trust relationships using conditional probability between two users. To do so, our proposed approach uses Kullback-Leibler divergence [8] instead.

3.1 Naïve approach

Collaborative filtering based recommender systems calculate the similarity between two users and then deploy the user similarity to recommend items. User similarity is a symmetric relationship. For example, when a user ti_c, is interested in an item i_n, and a user u_b is interested in both an item i_n, and an item i_n, the Jaccard similarity between u_n and u_b is 0.5. Let us examine the user similarity in depth. From u_n's point of view, u_b is interested in all the items ti_n is interested in (namely, i_{ni}). Thus, ti_c, might prefer the other items u_b is interested in but *It,* is not aware of (namely, i_n). However, from u_b's point of view, u_n is interested in some of the items u_b is interested in. u_b might not be interested in u_n's items, or u_b might have seen all of u_n's items already. We can calculate trust information from the asymmetric relationships.

Conditional probability is a measure that can be used to calculate an asymmetric relationship. Using tagged items, trust from u_n to u_b is defined as:

$$trust^{i}_{ua-nib} = \frac{P_i(u_a\ nub)}{p_i(u_a)}$$

$P_i(u_a)$ is the probability of the user $/l_a$'s items and $P_i(u_a\ n\ u_b)$ is the probability of items that both u_a and u_b are interested in. Trust is also calculated by tags. Using tags, trust from u_n to u_b is defined as:

$$trust^{t}_{ua->ub} \quad \frac{P_t(u_a\ n\ u_b)}{Pt(ua)}$$

$Pt(ua)$ is the probability of the user $/l_a$'s tags and $P_t(u_a\ n\ u_b)$ is the probability of tags annotated by both u_n and *ub.*

3.2 Kullback-Leibler divergence approach

Kullback-Leibler (KL) divergence is an asymmetric measure of the difference between two probability distributions. While the previously discussed naïve approach considers only the existence of tagged items or tags, the KL divergence approach takes into account the frequency of tagged items or tags. This enables the detailed analysis of the user's preference. The KL divergence approach is analogous to the weighted version of the naïve approach. Using tagged items or tags, trust from u_n to u_b is defined as:

$$Dia(ubilua) = \sum_{k\,E\,1} fi(ub, k) \log \frac{f_i(u_b, k)}{fi\,(u_a, k)\,(ua)}$$

$$Dia\,(ub = \sum_{so.\,(ua)} ft\,(u_b, s) \log \frac{ft\,(ub, s)}{ft\,(ua, s)}$$

I (u_a) is a set of tagged items in which u_a is interested, and T(u_a) is a set of tags annotated by u_a. $f_i(u_b, k)$ and $f_t(u_b, s)$ are probability mass functions defined as follows:

$$k) = \frac{n_i(u_b, k)}{{}_{E/E/(ub)} ni(ub, 1)}$$

$$f_t(u_b, s) = \frac{n_t(u_b, s)}{\overset{v,}{L} rET(ub) nt(nb'}$$

$n_i(u_b, /)$ is u_b*'s* distribution for the tagged item *1,* and $n_t(u_b$, r) is u_b*'s* distribution for the tag *r.* The proposed recommender system calculates a trust relationship between two users based on these equations.

3.3 Item recommendation

We propose a method that recommends items to a user using the user's tagging information. Given a user's tagging information, the system elicits implicit trust relationships from the user's dataset. To recommend appropriate items to the user, the system searches the user's trustful users, i.e., users who have high trust scores with the given user. Then, items of trustful users are aggregated to generate a final recommendation result. The system recommends items that match the user's preference in order of relevance.

4 Evaluation

Dataset. We conducted a series of experiments on *last fm'* dataset. *Last.fm* is a music website that provides a social tagging service to users. In *last.fm,* a user annotates tags to artists by according to the user's interests. Users manage and discover artists using tagging.

Using the *last fm* dataset, we evaluate the recommendation accuracy of derived implicit trust relationships of the different number of trustful users. Implicit trust relationships are extracted using two methods. Given a user, the first method searches for trustful users by using the conditional probability among other users. The second method searches for trustful users by using the KL divergence between two users.

http://www.last.fm

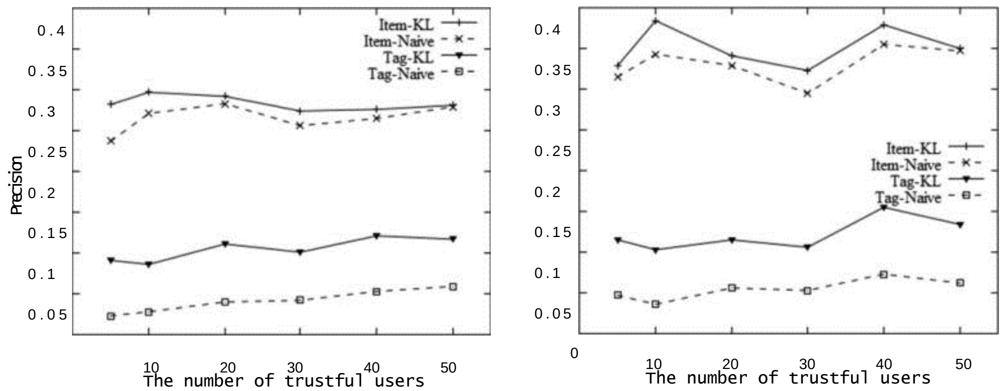

Fig. 1. Recommendation accuracy of the naïve and KL divergence approaches

Fig. 1 shows the average precision and recall values of the naïve and KL divergence approaches. The values are averaged over a randomly selected set of 100 users from the dataset. In the figure, solid lines represent our proposed method's results and dashed lines represent the naïve approach's results. The precision and recall values of the KL divergence method are higher than those of the naïve method. The naïve method captures the existence of tagged items or tags. The KL divergence method captures the number of tagged items or tags. This counting feature explains the superior performance of the KL divergence method. Independent of the extraction method used, using tagged items leads to better accuracy than using tags. Fundamentally, this is because the task under consideration is the recommendation of items, not tags.

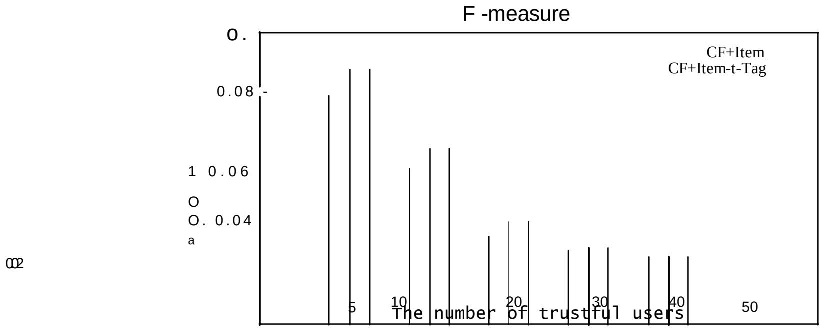

Fig. 2. Improvement over collaborative filtering

Fig. 2 shows our proposed recommender system's improvement over a collaborative filtering system. In the figure, *CF+Item* is the result of utilizing tagged item-based implicit trust relationships using the collaborative filtering method, while *CF+Item+Tag* is the result of utilizing both tagged item-based and tag-based implicit trust relationships using the collaborative filtering method. The use of implicit trust relationships improves the performance of the recommender system. As can be seen in Fig. 2, the improvement achieved is greater when the number of trustful users is small. Therefore, if the user has a small number of trustful users, the use of implicit trust relationships is a key means of improving the quality of the recommendation results. Overall, the results indicate that the proposed approach alleviates the sparsity problem of traditional recommender systems.

5 Conclusion

In this paper, we presented a recommendation method for use in social tagging systems. This method elicits implicit trust relationships to alleviate the sparsity problem. Under our approach, the system generates efficient and accurate recommendations with trust information. To derive trust information, we exploit asymmetric nature of user trust relationships. We utilize conditional probability as a naïve approach and Kullback-Leibler divergence as a more sophisticated approach. Experimental results show that the proposed approach performs well in practice. In future work, we plan to propose a hybrid recommendation method, and to investigate tag abstraction and trust propagation.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 20110030812).

References

1. Bhuiyan, T., Xu, Y., Josang, A., Liang, H., Cox, C.: Developing Trust Networks based on User Tagging Information for Recommendation Making. Web Information Systems Engineering 2010, pp. 357-364 (2010)
2. Massa, P., Avesani, P.: Trust-aware Recommender Systems. RecSys 2007, pp. 1724 (2007)
3. Mui, L., Mohtashemi, M., Halberstadt, A.: A Computational Model of Trust and Reputation. In: Proceedings of the 35th Hawaii International Conference on System Science, pp. 2431-2439 (2002)
4. Massa, P., Bhattacharjee, B.: Using Trust in Recommender Systems: An Experimental Analysis. Trust Management, pp. 221-235 (2004)
5. Milicevic, A.K., Nanopoulos, A., Ivanovic, M.: Social Tagging in Recommender Systems: A Survey of the State-of-the-art and Possible Extensions. Artificial Intelligence Review 33, pp. 187-209 (2010)
6. Papagelis, M., Plexousakis, D., Kutsuras, T.: Alleviating the Sparsity Problem of Collaborative Filtering using Trust Inferences. Trust Management, pp. 125-140 (2005)
7. O'Donovan, J., Smyth, B.: Trust in Recommender Systems. Intelligent User Interfaces, pp. 167-174 (2005)
8. Kullback, S., Leibler, R.A.: On Information and Sufficiency. The Annals of Mathematical Statistics 22, pp. 79-86 (1951)