Recommendations Based on Different Aspects of Influences in Social Media

Chin-Hui Lai[1], Duen-Ren Liu[2], and Mei-Lan Liu[3]

[1] Department of Information Management, Chung Yuan Christian University, Taiwan
chlai@cycu.edu.tw
[2,3] Institute of Information Management, National Chiao Tung University, Taiwan
dliu@iim.nctu.edu.tw, meilan0913@gmail.com

Abstract. Among the applications of Web 2.0, social networking sites continue to proliferate and the volume of content keeps growing; as a result, information overload causes difficulty for users attempting to choose useful and relevant information. In this work, we propose a novel recommendation method based on different types of influences: social, interest and popularity, using personal tendencies in regard to these three decision factors to recommend photos in a photo-sharing website, *Flickr*. Because these influences have different degrees of impact on each user, the personal tendencies related to these three influences are regarded as personalized weights; combining influence scores enables predicting the scores of items. The experimental results show that our proposed methods can improve the quality of recommendations.

Keywords: Web 2.0, Social media, Social network, Recommender System, Social Influence, Collaborative filtering.

1 Introduction

As social networking sites continue to proliferate and the volumes of their content keep growing, information overload causes users to experience difficulty in choosing useful and relevant information. In order to overcome the information overload problem, a recommender system [1] plays an important role in providing users with personalized recommendations on items.

People's preferences for items may be affected by three decision factors: social friends, personal interest and item popularity. For example, some users may like an item because their friends also like it, or they are interested in such items, or it is generally popular. However, most researches only utilize users' preference, the content of items or social influence to make recommendations. Although these methods can recommend the proper items to users, they do not take advantage of different types of influences, including social, interest and popularity influences in making recommendations in social media such as *Flickr*. Moreover, these decision factors have different impacts on each user. Each user has his/her personal tendency in regard to the influences of the three decision-related factors.

C. Huemer and P. Lops (Eds.): EC-Web 2013, LNBIP 152, pp. 194–201, 2013.

In this work, we propose a novel recommendation method based on different types of influences, including: social, interest and popularity, as well as the personal tendency related to these three decision-related factors to recommend photos in a photo-sharing website. Social influence means that users are affected by their friends and friends' friends; interest influence means that people are influenced by users with similar interests; and popularity influence signifies that users are affected by popular items. Additionally, we consider the perspectives of both an influenced user and an influential user with different weights in computing social influence and interest influence. Because those influences have different degrees of impact on each user, we exploit personal tendency in regard to the influence of social, interest and popularity as personalized weights, to combine these 3 decision-related factors in recommendation methods. Those photos with high predicted scores will be recommended to the target users. The experimental results show that applying the personalized tendency towards different types of influences in a recommendation method can improve prediction accuracy and quality of recommendation in social media. Because of better recommendation performance, users can save time in looking for relevant photos, and social media can attract more people to share photos, resulting in greater business opportunities.

The remainder of the paper is organized as follows. Section 2 introduces the works related to social networks, and recommender systems. In Section 3, we introduce the proposed recommendation methods based on social, interest and popularity influences, respectively. Section 4 shows the experimental results and evaluations of our methods. Finally, in Section 5, we discuss our results and conclude the paper.

2 Related Work

Social networking sites, such as *Facebook*, *Twitter* and *Flickr*, are popular business models that allow users to share information, publish articles and post photos. With the emergence of social networks, social recommender systems became helpful for users to acquire relevant and useful information. The goal of social recommender systems (SRSs) is to mitigate the problem of information overload for users of social media. Unlike traditional CF recommender systems, social recommender systems take advantage of explicit and implicit relationships among users, items or social contexts in order to make recommendations.

Social influence means that a user's behavior is affected by his/her friends [2]. However, most of the research estimates the social influence in recommender systems by taking a single perspective instead of the perspectives of both the influenced and influential user. Additionally, information can be disseminated quickly through those influential users, and the items recommended by those influential users are easily accepted by others. Several researchers combine social influence with traditional recommendation methods to make recommendations [3]. Nevertheless, they still do not take other influence factors, such as interest influence and popularity influence, into account. In this work, we adopt personalized tendency towards three decision factors: social influence, interest influence and popularity influence in our proposed approach for improving recommendation performance in social media.

3 Recommendation Based on Different Types of Influences

3.1 Overview

Because of the emergence of Web 2.0 technologies, users can post articles or share feelings in social networking sites such as *Facebook*, *Twitter* or *Flickr*. In *Flickr*, users can upload photos, mark photos as favorites, comment on photos and join specific groups to converse and share content. However, it is difficult for users to find relevant photos because of the volume of user-generated content and photos. Hence, more and more recommender systems are applied to social networks to filter out irrelevant information. Generally, recommendation methods in social networks use users' preferences, general acceptance of items and influence from social friends to provide recommendations [4]. However, because of differing personalities and behavior, users may be affected by different types of influences in making decisions, such as social influence, interest influence and popularity influence. Social influence refers to the ability of a user who follows his/her friends' footsteps to add their favorite photos into his/her favorite list. Interest influence means that a user is affected by users with similar interests; for example, a user may mark a photo as a favorite because similar users have added it to their lists. Popularity influence refers to the effect of popularity on a user's behavior.

In this work, we propose recommendation methods based on personalized tendencies towards different types of influence in the social media. Our proposed framework includes data collection, decision factor analyses, and recommendation phases. In the data collection phase, we collected a dataset from the *Flickr* website by using Flickr API to obtain information on photos and users. Then, the social, interest and popularity analyses are used to measure users' influences. Finally, according to a user's tendency to the influence of social, interest and popularity, we propose user-based collaborative filtering methods to recommend photos for users. Our methods not only provide personalized recommendations to users but also improve the performance of recommendations in social media.

3.2 Social Analysis

Social Influence from Direct Friends

We built social influence networks based on social influence links (SIL). If user u_c marks a photo as a favorite, which user u_f has marked, and u_c and u_f are friends, then we create a SIL from user u_c to u_f, which means user u_c is influenced by u_f or u_f influences u_c. The weight of SIL, the degree of social influence, indicates the degree of influence of user u_f over user u_c. Additionally, we take both perspectives, that of the influential and influenced users, into account, and then linearly combine the two values derived from these two perspectives.

User u_f's social influence on user u_c derived from the perspective of user u_c, i.e., the first part of Eq. (1), where $\left|FAV_{u_c}\right|$ is the number of favorite photos of user u_c; $FAV_{u_c F u_f} = \left\{ i \middle| i \in FAV_{u_c} \cap FAV_{u_f} \text{ and } t(u_c,i) > t(u_f,i) \right\}$ is the set of photos marked as favorites by user u_c after user u_f has marked the photos; and $t(u_c,i)$ is the time when u_c marks photo i. Additionally, user u_f's social influence on user u_c derived from the perspective of user u_f, i.e., the second part of Eq. (1), where $\left|FAV_{u_f}\right|$ is the number of favorite photos marked by user u_f. The parameter α ($0 \leq \alpha \leq 1$) is used to adjust the relative importance of the social influence derived from two different perspectives.

$$SI\left(u_c,u_f\right)=\alpha\times\frac{\left|FAV_{u_cFu_f}\right|}{\left|FAV_{u_c}\right|}+\left(1-\alpha\right)\times\frac{\left|FAV_{u_cFu_f}\right|}{\left|FAV_{u_f}\right|}, \tag{1}$$

Social Influence from Propagation

Because a target user and his/her direct friends may have no co-favorite photos, there is no direct social influence from friends on a target user. Social influence propagation can be used to infer the social influence on a target user through indirect social influence. Let user u_c denote the target (source) user; $SIP(u_c, u_f)$ denotes the social influence of user u_f on target user u_c derived from the influence propagation on a social network. If there is no direct social influence from user u_f on u_c, the propagation score of user u_f on user u_c based on the social influence is the average of the direct social influence of user u_f on user u_k, i.e., $SI(u_k, u_f)$ (Eq. (1)), weighted by the social influence propagation $SIP(u_c, u_k)$, as shown in Eq. (2) where $u_k : u_k \rightarrow u_f$ means that u_f has direct social influence on u_k, and $SIP(u_c, u_k)$ is the influence of user u_k on user u_c derived from the influence propagation of the social influence network:

$$SIP\left(u_c,u_f\right)=\frac{\sum_{u_k:u_k\rightarrow u_f} SIP\left(u_c,u_k\right)\times SI\left(u_k,u_f\right)}{\sum_{u_k:u_k\rightarrow u_f} SIP\left(u_c,u_k\right)}, \tag{2}$$

The Weight of Social Influence

Because each user may be affected by his/her friends to differing degrees, we use a weight to represent the personalized tendency towards social influence of each user. The weight of social influence is based on the proportion of the number of favorite photos which have been marked by both user u_c and user u_c's friends. Let $W_{u_c,SI}$ denote the weight of social influence for target user u_c; FRI_{u_c} be a set of friends of target user u_c; $\left|\bigcup_{u_f\in FRI_{u_c}} FAV_{u_cFu_f}\right|$ be the number of photos that user u_c marks as favorite photos after u_c's friends have marked these photos as favorites.

$$W_{u_c,SI}=\frac{\left|\bigcup_{u_f\in FRI_{u_c}} FAV_{u_cFu_f}\right|}{\left|FAV_{u_c}\right|}, \tag{3}$$

3.3 Interest Analysis

Interest Influence

Interest influence means that users who have similar or common interests may affect the behavior of one another; it is derived from the influential user, who is similar to the influenced user, on the influenced user. Again, as in the social influence discussed in Section 3.2, interest influence is also derived from both the perspectives of the influential and influenced users.

The interest influence of user u_x on user u_c from user u_c perspective, i.e., the first part of Eq. (4), where $\left|FAV_{u_c}\right|$ is the number of favorite photos of user u_c, and $FAV_{u_cFu_x} = \left\{i \middle| i \in FAV_{u_c} \cap FAV_{u_x} \text{ and } t(u_c,i) > t(u_x,i)\right\}$ is the set of photos marked as favorite by user u_c after user u_x has marked them as favorite. Similarly, the Second part of Eq. (4) is used to obtain the interest influence derived from user u_x's perspective, where $\left|FAV_{u_x}\right|$ is the number of favorite photos of user u_x. Then, these two parts of interest influence are linearly combined by using a parameter β ($0 \le \beta \le 1$) to evaluate user u_x's total interest influence on user u_c,

$$II(u_c,u_x) = \beta \times \frac{\left|FAV_{u_cFu_x}\right|}{\left|FAV_{u_c}\right|} + (1-\beta) \times \frac{\left|FAV_{u_cFu_x}\right|}{\left|FAV_{u_x}\right|}, \tag{4}$$

The Weight of Interest Influence

Not all users express interest in the photos in the 'favorite' lists of similar users; i.e., every user has a personalized tendency towards interest influence. Given this, we used a weight to represent the personalized tendency towards interest influence for each user. The weight of interest influence is based on a proportion of the number of favorite photos marked by both user u_c and his/her similar users, that is, when user u_c and his/her similar users have common photos in their favorites lists. For those common photos, user u_c's similar users marked them before user u_c marked them. Eq. (5) is used to measure the weight of interest influence for user u_c, where $W_{u_c.II}$ is the weight of interest influence for user u_c; NBR_{u_c} is a set of Top-K similar users of target user u_c; and $\left|\bigcup_{u_x \in NBR_{u_c}} FAV_{u_cFu_x}\right|$ is the number of favorite photos that user u_c marked after user u_c's similar users marked them.

$$W_{u_c.II} = \frac{\left|\bigcup_{u_x \in NBR_{u_c}} FAV_{u_cFu_x}\right|}{\left|FAV_{u_c}\right|}, \tag{5}$$

3.4 Popularity Analysis

Popularity Influence

Popularity is also a factor that affects users' behavior; especially on social networking sites. The score of popularity influence is used to measure the degree of popularity of a photo in a period of time. If the popularity score of a photo is high in this period, the photo is popular. The score of popularity influence, as defined in Eq. (6), is a ratio of the total favorite count of each photo to the maximal favorite count of all photos, where PI_i is the score of popularity influence of photo i; FC_i is the total favorite count of photo i; and $\max_j(FC_j)$ is the maximal favorite count of all photos collected in the dataset.

$$PI_i = \frac{FC_i}{\max_j \left(FC_j \right)}, \tag{6}$$

The Weight of Popularity Influence

The popularity of items has a different impact on each user. Therefore, we define a weight to represent personalized tendency towards popularity influence for each user. The weight of popularity influence is based on the number of popular photos in a user's favorite list, as defined in Eq. (7). Let $W_{u_c,PI}$ denote the weight of popularity influence, which is a ratio of the number of photos that are included in the favorite count before u_c has marked them, with those that exceed the threshold of the number of photos marked as favorite by user u_c, and N_{u_c} denotes the number of photos that their favorite counts exceed a threshold before user u_c has marked the photos.

$$W_{u_c,PI} = \frac{N_{u_c}}{\left| FAV_{u_c} \right|}, \tag{7}$$

3.5 Recommendation

In this section, we propose a recommendation method based on our decision factor analyses, including social, interest and popularity influences. We combined these three types of influences with the personalized weights to predict the score of photo *i* for the target user u_c, i.e., $PS(u_c, i)$, as defined in Eq. (8). The influence scores of social, interest and popularity on a particular photo *i* are derived from Eqs. (1), (2), (3), (4), (5), (6) and (7). The predicted score of photo *i* is defined as follows where $W_{u_c,SI}$, $W_{u_c,II}$ and $W_{u_c,PI}$ are the weights of social, interest and popularity influences, respectively, for target user u_c; and *timefactor*(*i*) is a time factor of photo *i*, ranging from 0 to 1. A higher time weight is assigned to a photo marked in the recent past and, conversely, lower time weights are assigned to older photos. Finally, Top-*N* photos with the highest predicted scores, i.e. $PS(u_c, i)$, will be recommended to the target user.

$$PS(u_c,i) = \left(W_{u_c,SI} \times \sum_{u_f \in FRI_{u_c}, i \in FAV_{u_f}} SIP(u_c,u_f) + W_{u_c,II} \times \sum_{u_x \in NBR_{u_c}, i \in FAV_{u_x}} II(u_c,u_x) + W_{u_c,PI} \times PI_i \right) \times timefactor(i), \tag{8}$$

4 Experiment and Evaluations

In our experiment, we collected a data set from the famous social networking website *Flickr*. *Flickr* is a popular photo-sharing social media site where users can upload photos, add photos into their favorites list and make friends through the web platform. The data set consists of photos analyzed from Aug 15, 2011 to Nov 15, 2011. Our dataset was composed of 50,000 similar users, about 90,000 users and over 2 million photos. We then divided the data set: 75% for training, 10% for tuning and 15% for testing. The training set was used to implement our proposed method and generate recommendation lists, while the testing set was used to verify the quality of these recommendations. The tuning set was used to determine parameters for our methods. To compare the prediction accuracy of the proposed methods, we employed *F1*-metric [5, 6], which are widely used in recommender systems to evaluate the quality of recommendations.

4.1 Comparison of Different Variations of the Proposed Methods

Different variations of the proposed methods are compared in this experiment. The S-IF method predicts photos by only considering social influence with the time factor; the I-IF method makes recommendations by using interest influence with the time factor; and the P-IF method takes the popularity influence into account. In addition, the combination of any two decision factors is utilized in the recommendation methods, i.e., SI-IF, SP-IF and IP-IF, for improving the accuracy of prediction. The SI-IF method recommends photos by integrating the influence of social and interest; the SP-IF method recommends photos based on the combination of both social and popularity influences; and the IP-IF method makes recommendations based on the composite of both interest and popularity. Besides these 6 methods, the SIP-IF method, which makes predictions based on the combination of all three decision factors, is also compared and evaluated. All parameters in these recommendation methods are derived from the experimental results based on the pretesting data. That is, α is set as 0.6, β is set as 0.5, the number of neighbors is 40, τ equals to 1/10. The average F1 value, calculated over various top-N (top-5, top-10, top-20, top-40, top-60) recommendations, is used to measure the recommendation quality.

Fig. 1 shows the experimental results by averaging the F1 values of the S-IF, I-IF, P-IF, SI-IF, SP-IF, IF-IF and SIP-IF methods, respectively. For the methods which make recommendations based on one decision factor, the S-IF method outperforms the I-IF and P-IF methods. The performance of the SI-IF model is better than both the IP-IF and SP-IF methods. Combining two decision factors is better than using only

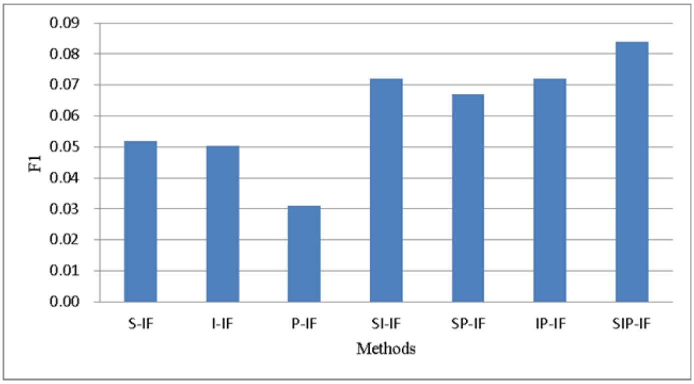

Fig. 1. The evaluation of different variations of the proposed methods

one decision factor in recommendation methods. Additionally, the SIP-IF method, which predicts photos based on the integration of social, interest and popularity influences and considers the personalized tendency towards such influences, has the best performance among these compared methods. In summary, the proposed decision influence types: social, interest and popularity, are useful and effective in making recommendations.

5 Conclusions

In this work, we proposed novel recommendation methods based on different types of influences: social, interest and popularity, and personalized tendency towards these three decision factors to recommend photos in a photo-sharing website *Flickr*. The perspectives of both the influenced user and the influential user were taken into account when computing the influence of social friends and interest. In addition, because these three decision factors have differing degrees of impact on each user, the personalized tendencies of these users towards these three decision factors were regarded as personalized weights to combine the influence scores for predicting the scores of items. The experimental results show that considering both of these perspectives when computing social and interest influences effectively enhances recommendation quality. Moreover, our proposed methods, which apply the personal tendency towards different types of influences as weights to combine the influence scores to make recommendations, indeed outperform other methods.

Acknowledgements. This research was supported by the National Science Council of the Taiwan under the grant NSC 101-2410-H-033-022-MY2.

References

1. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems: An Introduction. Cambridge University Press (2010)
2. Friedkin, N.E.: A structural theory of social influence. Cambridge University Press (1998)
3. Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S.: Feedback effects between similarity and social influence in online communities. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 160–168. ACM, Las Vegas (2008)
4. He, J., Chu, W.W.: A social network-based recommender system (SNRS). In: Memon, N., Xu, J.J.J., Hicks, D.L.L., Chen, H. (eds.) Data Mining for Social Network Data, pp. 47–74. Springer US (2010)
5. Salton, G., Harman, D.: Information retrieval. John Wiley and Sons Ltd. (2003)
6. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Analysis of recommendation algorithms for e-commerce. In: Proceedings of the 2nd ACM Conference on Electronic Commerce, pp. 158–167. ACM, Minneapolis (2000)