A Revisit to Social Network-Based Recommender Systems

Hui Li, Dingming Wu, Nikos Mamoulis

Department of Computer Science, The University of Hong Kong

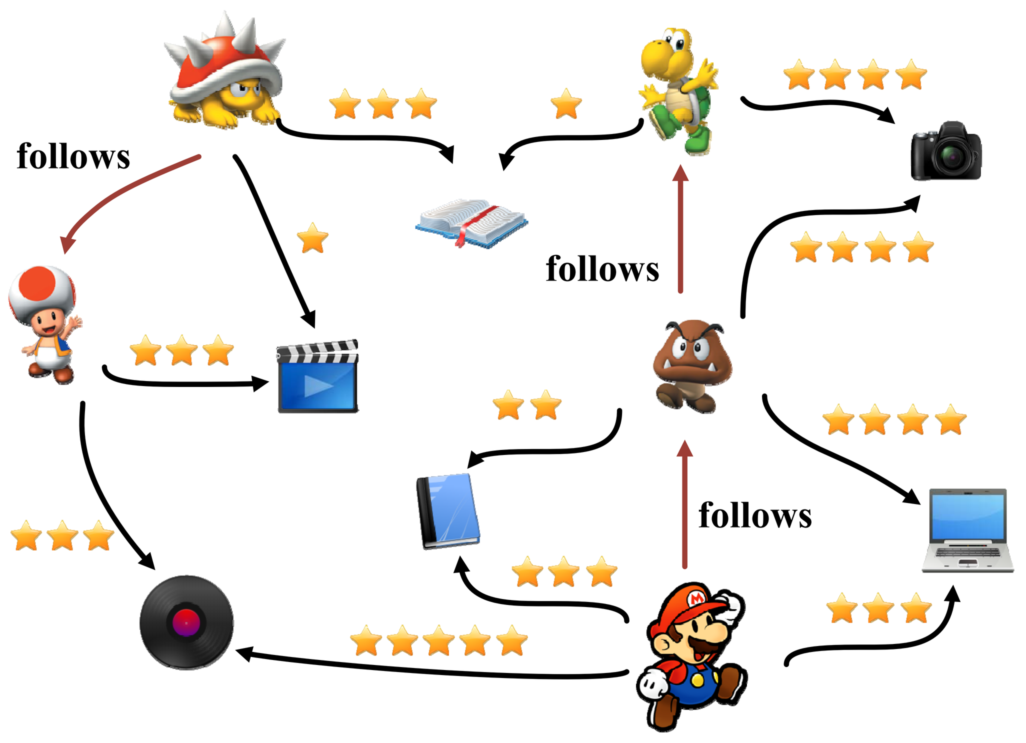
{hli2, dmwu, nikos}@cs.hku.hk

THE UNIVERSITY OF HONG KONG

DEPARTMENT OF COMPUTER SCIENCE

Social Recommender

Unlike traditional recommender systems which only utilize user-item rating matrix, a social recommender incorporates the information from social network to ameliorate the problems of data sparsity and cold-start users.



In this paper, we extended a social network-based recommender system (SNRS) [1]. Two methods were proposed to improve performance. One is classifying the correlations between pairs of users ratings to improve the accuracy of the system. The other is making the system robust to sparse data, i.e., when there are few immediate friends with few common ratings with regard to the target user.

Basic Approaches

CF User-based collaborative filtering algorithm using Pearson correlation to measure user similarity.

SNRS The social network-based recommender system [1] which utilizes a simplified homogeneous social network as a Bayesian network to improve the performance of traditional recommender systems. SNRS can be divided into three parts: user preference, item acceptance and influence from friends. The probability distribution of the target user U's rating on the target item I is calculated as:

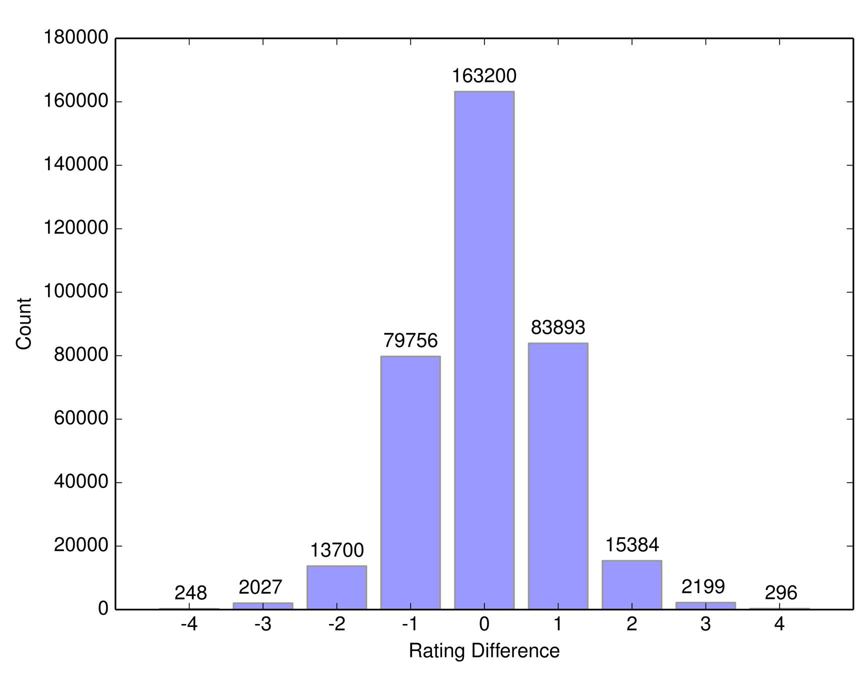
$$Pr(R_{UI}=k)=\frac{1}{Z}Pr(R_{UJ}=k|a_J=a_I)\times Pr(R_I=k)\times \prod_{V\in N(U)}\frac{1}{Z_V}H_{UV}(R_{UI}-R_{VI}), \tag{1}$$

where H_{UV} is a histogram recording the differences in ratings between users U and V and $N(U)$ denotes the immediate friends of U. SNRS considers the influences of all users in $N(U)$ to U.

[1] J. He and W. W. Chu. A social network-based recommender system (snrs). In Data Mining for Social Network Data, pages 47-74. 2010.

Dataset

We crawled a dataset from a real social network-based recommender system called Dianping (http://www.dianping.com) where a 5-scale rating system is used. The following figure shows the rating difference among friends in this dataset.



Classifying User Correlations

The correlations between the target user and his/her immediate friends learned from their common rated items are based on the assumption that each pair of friends behave consistently on reviewing the items. However, disagreement exists among friends in practice which is also proved in our analysis of Dianping dataset. We propose to use a **Bipolar Distribution** to classify correlations between friends into two categories: **LIKE** and **DISLIKE**, such that conflict ratings (the target user likes item I, while one immediate friend dislikes it, or vice versa) can be ignored when doing recommendation. Specifically, the existing rating of each user U on each item I is firstly converted into the deviation ΔR_{UI} from the average rating value $\overline{R_U}$ for U and then the rated items for each user U are classified into S_U^{LIKE} and $S_U^{DISLIKE}$ categories. Note that ΔR_{UI} is rounded to the nearest integer. Formally,

$$S_U^{LIKE}=\{I_i|\Delta R_{UI_i}=R_{UI_i}-\overline{R_U}\geq 0\}$$
$$S_U^{DISLIKE}=\{I_i|\Delta R_{UI_i}=R_{UI_i}-\overline{R_U}< 0\}$$

The correlations between the target user U and all the immediate friends V are calculated as

$$\begin{cases}\prod_{V\in N(U)}\frac{1}{Z_V}H_{UV}^{LIKE}(\Delta R_{UI}-\Delta R_{VI}), & \Delta R_{UI}\geq 0\\ \prod_{V\in N(U)}\frac{1}{Z_V}H_{UV}^{DISLIKE}(\Delta R_{UI}-\Delta R_{VI}), & \Delta R_{UI}< 0\end{cases}$$

Then, we use the above correlations as a substitute for original correlation (i.e., $\prod_{V\in N(U)}\frac{1}{Z_V}H_{UV}(R_{UI}-R_{VI})$) in Equation 1.

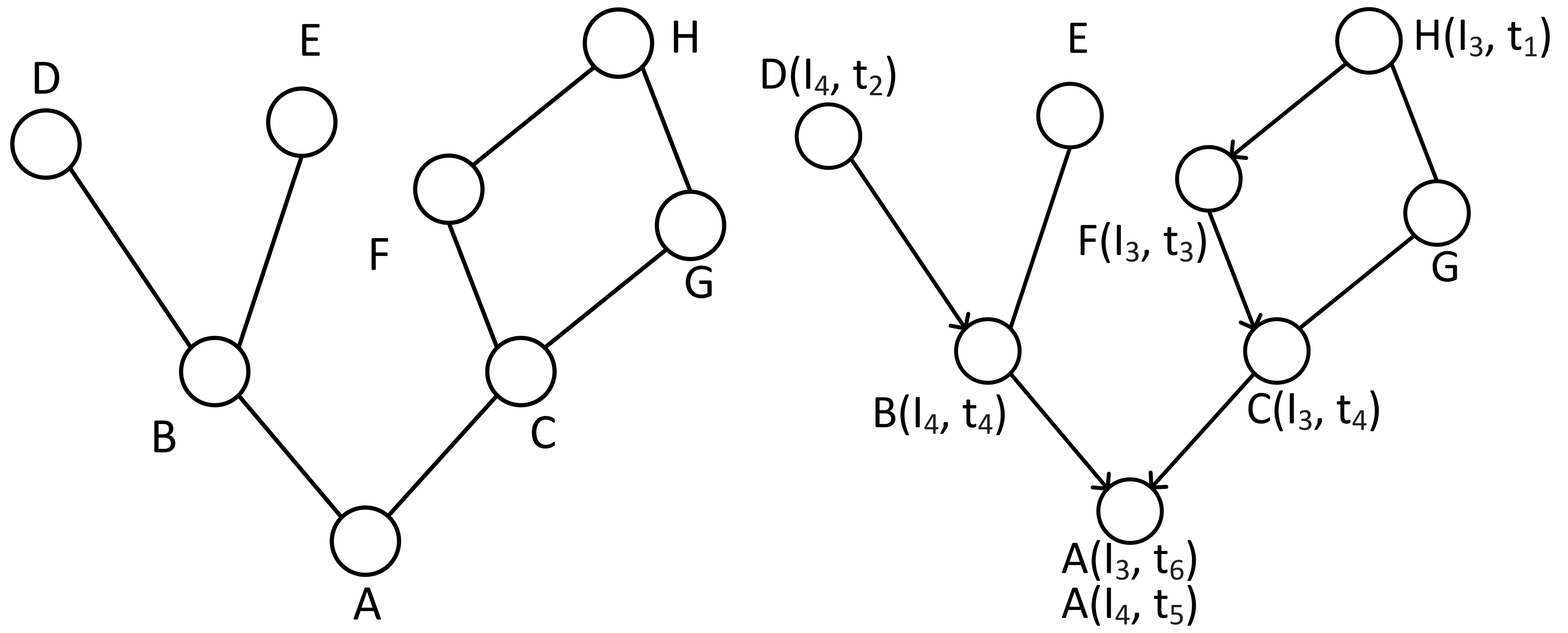
Temporal Influence Links

In a social network, some users may have few immediate friends, and friends may have few common ratings. The social influence may not be well considered in the recommendation if we only take the influence from immediate friends (i.e., $N(U)$) into account. To address this problem, we define temporal influence links in the social network.

Definition 1. A temporal influence link $E(V,U)$ is a directed edge from user V to user U if and only if (i) U and V are immediate friends, (ii) they both rated at least one common item I, and (iii) V rated I (at timestamp t) before U did (at timestamp $t+\Delta t$).

Definition 2. User V_0 can reach user V_n via temporal influence links if there exists a sequence of users $V_1,V_2,\cdots,V_{n-1}$, such that temporal influence links $E(V_i,V_{i+1}), i=0,1,\cdots,n-1$ exist.

The following examples show an example social network. The left figure shows the original social network, while the right figure illustrates the temporal influence links in this network. By considering the temporal influence links, the new user set $N'(U_A)$ includes users U_B,U_C,U_D,U_F,U_H and we utilize their history ratings for making recommendation for user U_A. Recall that in the SNRS system, only U_B, and U_C are considered in $N(U_A)$.



Results

We use Mean Absolute Error (MAE) and Coverage to evaluate performance. The experimental results on Dianping dataset are displayed in the following table. Approach SNRS* uses the classified correlations and approach SNRS** incorporates both enhancements from classified correlations and temporal influence links.

Training Data	Metrics	CF	SNRS	SNRS*	SNRS**
90%	MAE	0.8062	0.6941	0.5859	0.5910
	Coverage	0.7153	0.7033	0.7033	0.7416
80%	MAE	0.7900	0.6532	0.5584	0.5642
	Coverage	0.6204	0.7076	0.5917	0.6252

The experimental results demonstrate that our approach has a better performance compared to CF and the original SNRS.