

# Modeling the Temporal Dynamics of Social Rating Networks using Bidirectional Effects of Social Relations and Rating Patterns

Mohsen Jamali  
School of Computing Science  
Simon Fraser University  
Burnaby, BC, Canada  
mohsen\_jamali@cs.sfu.ca

Gholamreza Haffari  
School of Computing Science  
Simon Fraser University  
Burnaby, BC, Canada  
ghaffar1@cs.sfu.ca

Martin Ester  
School of Computing Science  
Simon Fraser University  
Burnaby, BC, Canada  
ester@cs.sfu.ca

## ABSTRACT

A social rating network (SRN) is a social network in which edges represent social relationships and users (nodes) express ratings on some of the given items. Such networks play an increasingly important role in reviewing websites such as Epinions.com or online sharing websites like Flickr.com. In this paper, we first observe and analyze the temporal behavior of users in a social rating network, who express ratings and create social relations. Then, we model the temporal dynamics of an SRN based on our observations, using the bidirectional effects of ratings and social relations. While existing models for other types of social networks have captured some of the effects, our model is the first one to represent all four effects, i.e. social relations-on-ratings (social influence), social relations-on-social relations (transitivity), ratings-on-social relations (selection), and ratings-on-ratings (correlational influence). Existing works consider these effects as static and constant throughout the evolution of an SRN, however our observations reveal that these effects are actually dynamic. We propose a probabilistic generative model for SRNs, which models the strength and dynamics of each effect throughout the network evolution. This model can serve for the prediction of future links, ratings or community structures. Due to the sensitive nature of SRNs, another motivation for our work is the generation of synthetic SRN data sets for research purposes. Our experimental studies on two real life datasets (Epinions and Flickr) demonstrate that the proposed model produces social rating networks that agree with real world data on a comprehensive set of evaluation criteria.

## Categories and Subject Descriptors

H.2.8.d [Information Technology and Systems]: Database Applications - Data Mining

## General Terms

Algorithms, Measurement, Experimentation

## Keywords

User Behavioral Modeling, Social Rating Networks, Temporal Dynamics, Generative Models

Copyright is held by the International World Wide Web Conference Committee (IW3C2). Distribution of these papers is limited to classroom use, and personal use by others.  
WWW 2011, March 28–April 1, 2011, Hyderabad, India.  
ACM 978-1-4503-0632-4/11/03.

## 1. INTRODUCTION

Social networks have attracted much attention these days. Social networking services such as Facebook, Orkut, Flickr and LiveJournal, have millions of users. Users are allowed to interact with other people besides having a profile with some attributes. There have been many works to analyze the behavior of social networks and their growth patterns [7][15][9][10]. Most existing models consider the social networks only as a graph with nodes and edges. Recently, some researchers have investigated social networks which include attributes for nodes [18][11]. In this paper we investigate Social Rating Networks (SRN). A Social Rating Network is a social network in which each user expresses ratings on some items besides creating social relations to other users. Users in a social rating network can perform two types of actions: creating a link to another user (social action) and creating a rating for an item (rating action). Note that the "rating action" is a general term and includes the real valued item ratings in Epinions and Flixster and binary rating values such as joining a community in Facebook and LiveJournal or adding a photo to your favorite list in Flickr. In other words different types of user behaviors can be formulated as a rating action. In our experiments, we use one data set with real valued ratings (Epinions) and another data set with binary ratings (Flickr).

In this paper, we analyze and model the behavior of users while performing different actions. We also take into account the timestamp to analyze the temporal dynamics of user behavior. Our model assume real valued ratings and binary ratings such as joining a community can be considered as special cases of real valued ratings.

Many effects influence the behavior of users in the evolution of a social rating network. A fundamental property of social networks is that people tend to behave similarly to their friends. The process of *social influence* [4] in SRNs leads to people adopting the rating behavior of their friends. Also people may adopt the rating behavior from users having similar rating patterns. We call this *correlational influence*. While this effect is the implicit foundation of the successful collaborative filtering recommenders [5], it has not yet been explicitly considered in social network models. The social behavior of users is also investigated. As discussed in [3], people tend to form relationships with others who are already similar to them, so-called *selection* [12]. However, we argue that social selection is not the only mechanism influencing the social relation creation process in SRNs. Implied by the well-known trust transitivity [6], people tend to create social relations to friends of their friends. We call this effect *transitivity*.

There are also some other effects that influence the behavior of users. For example users may create social relations to other users who live in the same location, or they may befriend their co-

workers. Users may rate items they face in their daily life independent of their friends. These effects are called environmental or external effects [1]. Since we do not have access to these kinds of information, we make a simplifying assumption about these kinds of behaviors: when creating a social relation, we assume that there is a chance that users create a link to a random existing user (representing unknown effects) or to a new user. Similarly, there is a chance that users rate a random existing item or a new item.

According to our observations from real life data, different effects influencing the behavior of users are dynamic. In other words, the strength of each effect is not constant throughout the evolution of an SRN. For example, when very few users have created social relations and the average number of social relations per user is very low, the strength of social influence and transitivity is low, but these effects become more important during the growth of a social rating network in which users rate more items, create more social relations and are exposed to more people and items. Effects influencing the behavior of users are not the only dynamic components governing the evolution of an SRN. At the beginning of the evolution when there are only few nodes in the system, the rate of new users joining the system is high. Also, the rate of creating a social relation to a new user decreases with the increase in the number of existing users in the system.

In this paper, based on our observations and analysis, we propose a probabilistic generative model to model the behavior of users and capture the temporal dynamics of different effects influencing the behavior of users. To capture the dynamism of effects, we model each of them as a function of some evolving features of the SRN such as the number of users, the number of neighbors for a users, etc.

This generative model can serve various purposes. First, there are very few SRN datasets publicly available. This is mainly due to privacy concerns in social networks, especially when some behaviors of the user should also get published. The lack of publicly available datasets is seriously restricting research on social network based recommendation. A generative model for social rating networks can be used to generate synthetic datasets with properties similar to those of real-life datasets, which can be used for research purposes. Moreover, the study of the evolution of social rating networks and the growth patterns for social networks and rating profiles is an important application of a generative model, providing insights into models developed in the social sciences. Also such a model can be used for predicting future links, ratings or community structures.

The main contributions of this paper are as follows:

- We analyze the temporal dynamics of SRNs and provide insights on bidirectional effects of social relations and rating patterns. We analyze and model the dynamism of effects influencing the behavior of users in an SRN. The dynamic nature of these effects has been neglected in the literature.
- We present a generative model for the evolution of SRNs, modeling different effects influencing the behavior of users. To the best of our knowledge, our model is the first one to comprehensively capture all the four effects between social relations and ratings: social-on-rating (social influence), social-on-social (transitivity), ratings-on-social (selection), ratings-on-rating (correlational influence).
- We perform an experimental study on two real life data sets (Epinions and Flickr), showing that the proposed model can indeed produce realistic SRNs. To compare different generative models we introduce several formal evaluation metrics.

- Our experimental study provides interesting insights into the factors that drive SRNs. For example, we find that *transitivity* is much more important than *selection* in the creation of social relations for both Epinions and Flickr.

The rest of the paper is organized as follows. We present our observations and analysis of the temporal dynamics of SRNs in section 2. Based on our observations and analysis, we present a model for the behavior of users and the evolution of SRNs in section 3. Related work is discussed in section 4. Evaluation metrics and experimental results are presented in section 5. Section 6 concludes the paper and presents some directions for future work.

## 2. ANALYZING THE TEMPORAL DYNAMICS OF SRNS

To observe the behavior of users in an SRN and to analyze the evolution of SRNs, we performed experimental studies on two real life data sets: Epinions and Flickr. Epinions.com is an online reviewing website. In Epinions, two types of actions are provided: users rating items with some rating values, and users creating social relations to other users. In Flickr, we consider two types of actions: users creating social relations to other users, and users adding photos to their favorite list. We consider adding a photo to favorites as a binary rating action (with rating value 1).

We represent the data set as a sequence of actions  $\mathcal{A}$ . Each action can be either a *rating action* creating a rating or a *social action* creating a social relation. Every action  $a$  in  $\mathcal{A}$  is also associated with a timestamp. We order the actions with respect to their timestamps and observe the behavior of users in chronological order.

For every action  $a$ , we observe and analyze the behavior of users as follows (see figure 1):

- Who is the user  $u$  performing the action? One of the existing users or a new user joining the SRN and performing his first action.
- What type of action is being performed? A rating action or a social action?
- If it is a social action, who is the target user  $v$  to whom the social relation is being created?
  - If transitivity is playing role in this action, then  $v$  should be one of the friends of friends of  $u$ .
  - Selection leads to select  $v$  among top similar users to  $u$ .
  - If unknown external effects are influencing this action, then  $v$  could be any existing user.
  - Finally, there is the possibility of  $v$  being a new user just joining the SRN.
- If it is a rating action, what item  $i$  is being rated and by what value?
  - If social influence is affecting the rating behavior of  $u$  in this action, then  $i$  should be one the items rated by friends of  $u$ . The value of the rating should also be affected by ratings expressed by direct neighbors of  $u$  on  $i$ .
  - If correlational influence is affecting the behavior of  $u$ , then  $i$  should be one of the items rated by top similar users to  $u$ . The rating of  $u$  on  $i$  should also be influenced by the ratings of similar users on  $i$ .

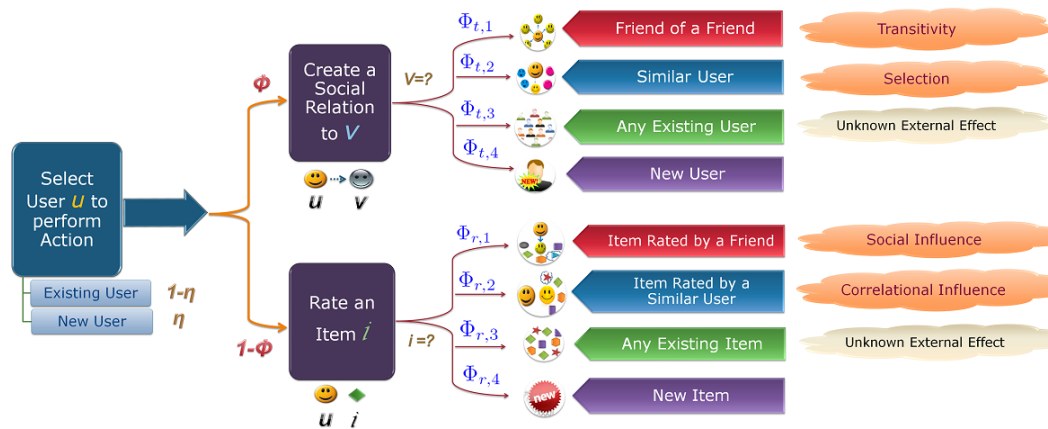


Figure 1: Illustration of different effects influencing the behavior of users in an action.

- If some unknown external effect is influencing the behavior of  $u$  in this action, then  $i$  could be any existing item.
- It could be also the case that  $i$  is a new item which has not been rated in the SRN before this action.

Note that the rating actions being affected by social influence, correlational influence and unknown external effects may overlap. Also the social actions influenced by transitivity, selection and unknown external effects may overlap. We compute the similarity among users using Pearson correlation [16] between their ratings. Similar to the idea used in [3] to construct a similarity network, the number of top similar users considered for each user  $u$  in rating actions is the same as the number of direct neighbors of  $u$  in the social network,  $N_u$ <sup>1</sup>. For social actions of user  $u$ , the number of similar users considered is the same as the number of users who are friends of friends of  $u$ .

First we analyze the role of new users and items in the evolution of SRNs. As shown in figure 1, the probability that an action is performed by a new user is  $\eta$ . However this probability is not a fixed number. Figure 2(a) presents our observation on the percentage of actions being performed by new users versus the total number of users existing in Epinions and Flickr. As shown in this figure, for both data sets  $\eta$  has a power law-like distribution over  $|N_t|$ , where  $N_t$  is the current set of existing users in the SRN at time  $t$ . At the very beginning of SRN's evolution, when there exists only a few users, the probability of an action being performed by a new user is very high, and this probability decreases very fast with the increase in the number of existing users. As depicted in figure 2(a), a similar behavior exists for the percentage of social actions in which a social relation is created to a new user, denoted by  $\Phi_{t,4}$ . If the action is a rating action, the percentage of rating actions in which a new item is rated (denoted by  $\Phi_{r,4}$ ) also exhibits a power law-like distribution over the number of existing items for both Epinions and Flickr, as shown in figure 2(b).

Our observations show that the percentage of social actions ( $\phi$ ) over the total number of actions (denoted by  $|A_t|$ ) remains fairly constant throughout the evolution of SRN for both Epinions and Flickr.

<sup>1</sup>Since there could be users with no direct neighbor but with some similar users, in our experiments, the number of similar users is set to be at least  $K=10$ .

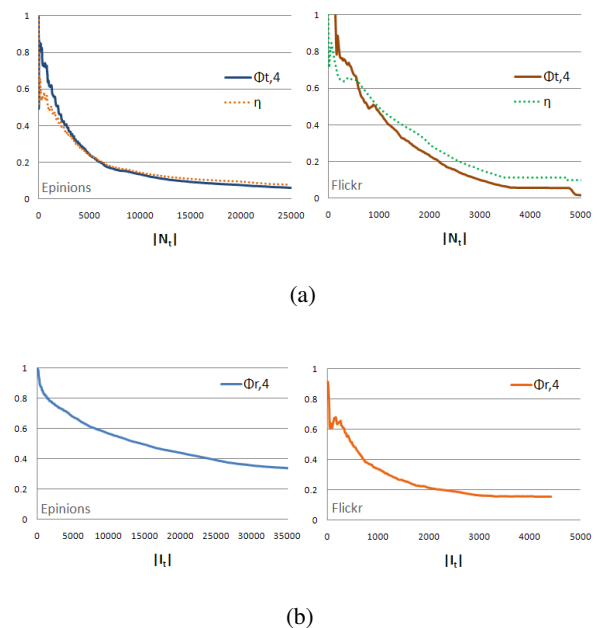
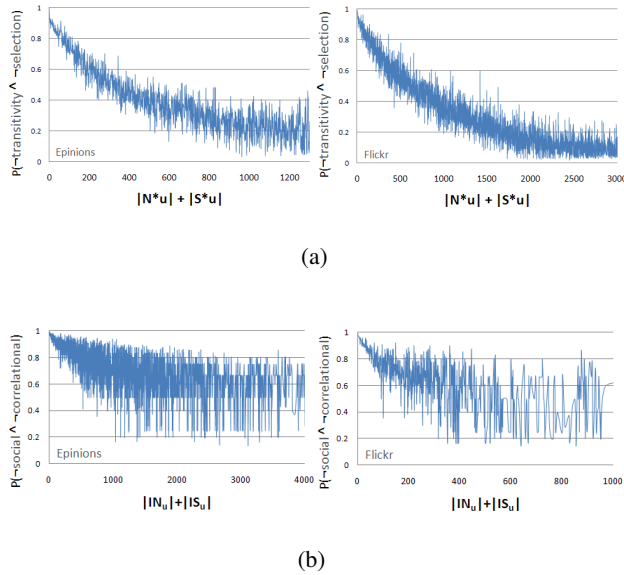


Figure 2: a) The distribution of  $\eta$  and  $\Phi_{t,4}$  versus the number of existing users in Epinions and Flickr.  $\eta$  is the percentage of actions that are performed by a new user.  $\Phi_{t,4}$  is the percentage of social actions that are created to a new user. b) Evolution of the percentage of ratings in which a new item is rated ( $\Phi_{r,4}$ ).

When a social relation is created by a user to some existing user, it could be either influenced by transitivity or selection, or by some unknown external effect. If unknown effects are playing a role, we assume that the social relation is being created to an arbitrary user. Our observations show that the percentage of social actions that are influenced by transitivity or selection increases almost linearly with growth of the number of social actions. We believe that in the course of time more social relations and more ratings are being expressed and hence users get more similar and also get to know friends of their friends. These events may lead to a stronger influence of similar users and direct neighbors on their behavior for cre-

ating social relations. Figure 3(a) depicts the percentage of social relations created to existing users that are not affected by transitivity or selection for different values of  $|N_u^*| + |S_u^*|$  for both data sets, where  $N_u^*$  is the set of friends of friends of  $u$  and  $S_u^*$  is the set of top similar users for  $u$ . As shown in this figure, the more users become available in  $N_u^*$  and  $S_u^*$ , the less likely it becomes that a social action can not be explained by the effect of transitivity or selection. Again, the diagram resembles a power law distribution.

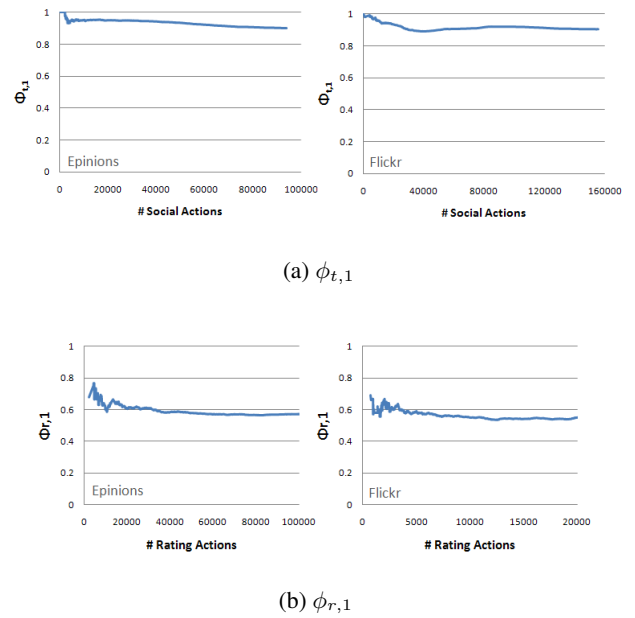


**Figure 3: a)  $P(\neg(\text{transitivity} \vee \text{selection}))$ : Percentage of social actions that are not affected by transitivity or selection for different values of  $|N_u^*| + |S_u^*|$  in Epinions and Flickr. b)  $P(\neg(\text{social} \vee \text{correlational}))$ : Percentage of rating actions that are not affected by social or correlational influence for different values of  $|IN_u| + |IS_u|$  in Epinions and Flickr.**

Similar behavior is observed for rating actions. The percentage of rating actions that are affected by social or correlational influence increases almost linearly with the growth of the number of rating actions. Figure 3(b) depicts the percentage of rating actions rating an existing item that are not affected by social or correlational influence for different values of  $|IN_u| + |IS_u|$  for both data sets, where  $IN_u$  denotes the set of items rated by friends of  $u$  and  $IS_u$  denotes the set of items rated by top similar users for  $u$ . Figures 3(a) and 3(b) show that the percentage of actions that can be explained only by unknown effects decreases in a power law-like distribution in both Epinions and Flickr.

Furthermore we analyze the percentage of social actions in which a social relation is created to a friend of a friend versus the social actions in which the social relation is created to a top similar user. Surprisingly, this percentage remains fairly constant throughout the evolution of the Epinions data set. Figure 4(a) shows the percentage of social actions influenced by transitivity or selection that are solely influenced by transitivity (denoted by  $\phi_{t,1}$ ) versus the total number of social actions for Epinions and Flickr. Note that social actions influenced by transitivity may overlap the social actions influenced by selection. As shown in this figure, transitivity plays a more important role in creating social relations than selection in both data sets.

The next observation is the percentage of ratings being affected



**Figure 4: a) Relative strength of transitivity and selection for Epinions and Flickr. Each diagram shows the percentage of social actions influenced by transitivity or selection that are solely influenced by transitivity (denoted by  $\phi_{t,1}$ ) versus total number of social actions in a data set. b) Relative strength of social influence and correlational influence for Epinions and Flickr. The diagrams show the percentage of rating actions affected by social or correlational influence that are solely influenced by social influence (denoted by  $\phi_{r,1}$ ) versus total number of rating actions.**

by social influence versus the actions affected by correlational influence. Figure 4(b) depicts the percentage of rating actions affected by social or correlational influence that are solely influenced by social influence (denoted by  $\phi_{r,1}$ ) versus total number of rating actions for Epinions and Flickr. Again, similar to the case for social actions, this percentage is fairly constant throughout the evolution of the SRN in both Epinions and Flickr.

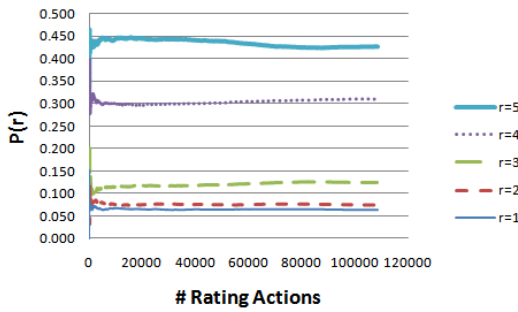
Our final observation is the distribution of rating values throughout the evolution of an SRN. Figure 5 depicts the evolution of the percentage of each rating value in Epinions<sup>2</sup>. Interestingly, we observe that the relative frequency of the rating values are fairly constant during the evolution of the SRN. The percentage of higher rating values are larger than those of low rating values, demonstrating that users in Epinions tend to be very generous in their ratings.

### 3. MODELING THE TEMPORAL DYNAMICS OF SRNS

In this section we exploit our observations and analysis to model the temporal dynamics of social rating networks. To do so, we represent different effects influencing the behavior of users in a probabilistic generative model.

A social rating network consists of a sequence of dated actions. The proposed model is illustrated in figure 1. In the following, we

<sup>2</sup>Note that since the ratings in Flickr are binary only, this observation does not apply to the Flickr data set.



**Figure 5: Distribution of rating values throughout the evolution of social network in Epinions. Possible rating values are integers in the range [1,5].**

explain the details of the proposed model for the generative process of performing an action.

In our model, first the user to perform the action is selected. With probability  $\eta$  this user is a new user. With probability  $1 - \eta$  one of the existing users performs the action.  $\eta$  is not a fixed value. Inspired by our observation in figure 2(a), we assume a power law function for  $\eta$  over the number of existing users as follows:

$$\eta = \text{power}(|N_t|) = a_1 \times (|N_t| - b_1)^{c_1} \quad (1)$$

where  $N_t$  is the current number of existing users. Now we introduce  $P(u|S_t)$ , the probability of user  $u$  performing the action.  $S_t$  denotes the current state of the SRN which summarizes the effects of all actions performed up to now. If the model decides that the next action is to be performed by a new user, then "selecting" a user  $u$  is deterministic: the model simply generates a new user  $u$  to perform the action. If it is decided that the next action is to be performed by an existing user, then the probability of a particular user  $u$  performing the action is proportional to the number of actions already performed by  $u$  taking into account a smoothing parameter  $\epsilon_0$ . This is similar to the "rich get richer" idea which is also the base of preferential attachment [15][14].

Next, the type of the action being performed is decided: social action or rating action. Based on the observations the probability of creating a social relation is considered to be a constant parameter  $\phi$ .

If the action is a social action, then there are four alternatives on how to create the social relation: creating a social relation to a new user, to a friend of a friend (transitivity), to a similar user (selection), or to a random existing user (unknown effect). With probability  $\Phi_{t,4}$  the action will create a social relation to a new user. According to our observations shown in figure 2(a),  $\Phi_{t,4}$  decreases with the increase of the number of existing users by a power function. So, we model this probability as follows:

$$\Phi_{t,4} = \text{power}(|N_t|) = a_2 \times (|N_t| - b_2)^{c_2} \quad (2)$$

With probability  $\Phi_{t,3} = (1 - \Phi_{t,4})\phi_{t,3}$ , the action creates a social relation to some random existing user influenced by unknown external effects. These are the social actions that can not be explained by transitivity or selection. To model  $\phi_{t,3}$  we exploit the observation shown in figure 3(a) and assign social actions that can not be explained by transitivity or selection to unknown external effects as follows:

$$\phi_{t,3} = \text{power}(|N_u^*| + |S_u^*|) = a_3 \times (|N_u^*| + |S_u^*| - b_3)^{c_3} \quad (3)$$

where  $N_u^*$  is the set of friends of friends of  $u$  and  $S_u^*$  is the set of top similar users to  $u$ . Note that as mentioned before,  $S_u^*$  is defined in a way such that  $|S_u^*|$  is the same as  $|N_u^*|$ . If we decide to create a social relation to an arbitrary existing user, we do so by assigning higher probabilities to users to whom more social relations have been created already. This is the same as in preferential attachment models [15].

$$P_e(v|u, S_t) = \frac{d_{in}(v) + \epsilon_3}{\sum_w (d_{in}(w) + \epsilon_3)} \quad (4)$$

In the above equation,  $d_{in}(v)$  is the in-degree of user  $v$  in the social network and  $\epsilon_3$  is a smoothing parameter.  $P_e(v|u, S_t)$  is the probability that  $u$  creates the social relation to  $v$  who is among existing users.

With probability  $(1 - \Phi_{t,4})(1 - \phi_{t,3})$ , the action creates a social relation to either a friend of a friend or a top similar user. Based on our observations in figure 4(a),  $\phi_{t,1}$  is constant throughout the evolution of the social rating network and can be learnt as model parameters. Hence, in our model a social action creates a social relation to a friend of a friend with probability  $\Phi_{t,1} = (1 - \phi_{t,4})(1 - \phi_{t,3})\phi_{t,1}$  and to a top similar user with probability  $\Phi_{t,2} = (1 - \phi_{t,4})(1 - \phi_{t,3})(1 - \phi_{t,1})$ .

If the social action is creating a social relation to a friend of a friend, then the probability of a friend of friend  $v$  being selected is proportional to the frequency of appearance of  $v$  in the set of friends of friends of  $u$ :

$$P_t(v|u, S_t) = \frac{f_t(v|u) + \epsilon_1}{\sum_{w \in N_u^*} (f_t(w|u) + \epsilon_1)} \quad (5)$$

where  $P_t(v|u, S_t)$  is the probability of creating a social relation from  $u$  to a friend of friend  $v$  when deciding to create the social relation to a friend of friend.  $f_t(v|u)$  is the frequency of appearances of  $v$  in the set of friends of friends of  $u$  and  $\epsilon_1$  is a smoothing parameter to be learnt.

Similarly, the probability of creating a social relation to a top similar user  $v$  is modeled as follows:

$$P_s(v|u, S_t) = \frac{\text{sim}_{u,v} + \epsilon_2}{\sum_{w \in S_u} (\text{sim}_{u,w} + \epsilon_2)} \quad (6)$$

where  $\epsilon_2$  is the smoothing parameter to be learnt for our model.  $P_s(v|u, S_t)$  is the probability that  $u$  creates a social relation to  $v$  who is among his top similar users.

Next, we model rating actions. There are four alternative ways to rate an item: rating a new item, rating an item rated by friends (social influence), rating an item rated by similar users (correlational influence), and rating a random existing item (unknown effect). With probability  $\Phi_{r,4}$  a new item is rated by action performer  $u$ . As we observed in figure 2(b), this probability exhibits power law behavior. So, we model it as follows:

$$\Phi_{r,4} = \text{power}(|I_t|) = a_4 \times (|I_t| - b_4)^{c_4} \quad (7)$$

where  $I_t$  is the current set of existing items. To select the rating value for this rating action, we apply the observation in figure 5 and model the probability of each rating value  $r$  by  $\mu_r$ , where  $\mu_r$  is a model parameter indicating the prior probability of rating value  $r$ .

With probability  $\Phi_{r,3} = (1 - \Phi_{r,4})\phi_{r,3}$ , user  $u$  rates an arbitrary existing item influenced by unknown external effects. These are the

rating actions that can not be explained by social or correlational influence. Similar to our model for social relations, we exploit the observation shown in figure 3(b) to model  $\phi_{r,3}$  and assign all rating actions that can not be explained by social or correlational influence to unknown external effects as follows:

$$\phi_{r,3} = \text{power}(|IN_u| + |IS_u|) = a_5 \times (|IN_u| + |IS_u| - b_5)^{c_5} \quad (8)$$

where  $IN_u$  is the set of items rated by direct neighbors of  $u$  and  $IS_u$  is the set of items rated by top similar users of  $u$ . In this case, the probability of selecting an existing item  $i$ ,  $P_e(i|u, S_t)$ , is modeled to be proportional to the number of ratings already expressed for  $i$ :

$$P_e(i|u, S_t) = \frac{|I_i| + \epsilon_8}{\sum_j (|I_j| + \epsilon_8)} \quad (9)$$

where  $I_i$  is the set of users who have rated  $i$ . The next step is to decide the rating value for this action. For each rating value  $r$ , we combine the prior probability  $\mu_r$  with the percentage of existing ratings for item  $i$  with value  $r$ :

$$P_e(r|i, S_t) \propto \frac{f_i(r) + \epsilon_9}{\sum_{r'} (f_i(r') + \epsilon_9)} + \mu_r \quad (10)$$

where  $f_i(r)$  is the frequency of rating value  $r$  for item  $i$  and  $\epsilon_9$  is the smoothing parameter.  $P_e(r|i, S_t)$  is the probability of assigning rating value  $r$  to  $i$  if  $i$  is a random existing user.

With probability  $\Phi_{r,1} = (1 - \Phi_{r,4})(1 - \phi_{r,3})\phi_{r,1}$ , the rating action selects the item from one of the items rated by direct neighbors of  $u$ . Based on our observations in figure 4(b),  $\phi_{r,1}$  remains fairly constant throughout the evolution of the SRN, and can be learnt as a model parameter. If an item already rated by direct neighbors is being rated, then the probability of item  $i$  being rated is as follows:

$$P_t(i|u, S_t) = \frac{f_t(i|u) + \epsilon_4}{\sum_{j \in IN_u} (f_t(j|u) + \epsilon_4)} \quad (11)$$

where  $f_t(i|u)$  is the frequency of item  $i$  appearing in the set of items rated by direct neighbors of  $u$ . To compute the probability of rating the item by value  $r$  (denoted by  $P_t(r|u, i, S_t)$ ), we use the prior knowledge of general probability of rating  $i$  by value  $r$  combined with the distribution of rating values on  $i$  among direct neighbors  $u$  as follows:

$$P_t(r|u, i, S_t) \propto \frac{f_t(r|i, u) + \epsilon_5}{\sum_{r'} (f_t(r'|i, u) + \epsilon_5)} + P_e(r|i, S_t) \quad (12)$$

In the above equation,  $P_e(r|i, S_t)$  is the prior probability of rating item  $i$  by value  $r$  as defined in equation 11.  $f_t(r|i, u)$  is the frequency of rating value  $r$  for item  $i$  among ratings expressed by friends of  $u$  for item  $i$ . It should be noted that since the ratings are binary in Flickr, models for rating values (equations 11 and 13) apply to Epinions only. In Flickr, all rating values are 1.

With probability  $\Phi_{r,1} = (1 - \Phi_{r,4})(1 - \phi_{r,3})(1 - \phi_{r,1})$ , the rating action selects the item from one of the items rated by top similar users to  $u$ . The process for rating an item in this case is similar to the process of rating an item rated by direct neighbors of  $u$  in the social network.  $P_s(i|u, S_t)$  and  $P_s(r|u, i, S_t)$  are defined in a way similar to the definition of  $P_t(i|u, S_t)$  and  $P_t(r|u, i, S_t)$ .

### 3.1 Learning the model parameters

In this section, we present our approach to learn the parameters of our model. As discussed above, parameters of our model include

10 smoothing parameters ( $\epsilon_0 \dots \epsilon_9$ ), 15 parameters for power law functions ( $a_1, b_1, c_1 \dots a_5, b_5, c_5$ ),  $\phi$ , and  $\phi_{t,1}, \phi_{r,1}$ . We denote the set of all parameters by  $\Theta$ .

We compute the likelihood of the ratings and social relations observed in the data set under our model, and resort to maximum likelihood (ML) estimation to learn the values of the model parameters. We consider the network generation from the very beginning, and maximize the likelihood of the observed sequence of social and rating actions.

From the start time to the end time  $T$ , we observe  $K$  actions which transform the SRN from the state  $S_0$  to the state  $S_T$ .  $S_t$  denotes the state of the SRN at time  $t$  which summarizes the effects of all actions performed up to this time. Let  $\mathcal{A}$  denote the sequence of all actions performed between times 0 and  $T$ , in which the actions are sorted according to their timestamps. We denote each action in  $\mathcal{A}$  by  $a_k$  and its timestamp by  $t_k$ , where  $k \in [1..K]$ . Based on the chain rule, the likelihood of the model is the product of the probabilities of each individual action in  $\mathcal{A}$  given the previous actions:

$$P(\mathcal{A}|S_0, \Theta) = \prod_{k=1}^K P(a_k|S_{t_k}, \Theta) \quad (13)$$

For the two types of actions (social actions, rating actions), we compute the likelihood of the action in a different way. The probability of selecting a user  $u$ , and performing a social action  $T_{u,v}$  to create a social relation between user  $u$  and user  $v$  is:

$$P(T_{u,v}|S_t, \Theta) = P(u|S_t) \times \phi \times \left[ \Phi_{t,4} g_n(v|S_t) + (1 - \Phi_{t,4}) \phi_{t,3} P_e(v|u, S_t) + (1 - \Phi_{t,4})(1 - \phi_{t,3}) \phi_{t,1} P_t(v|u, S_t) + (1 - \Phi_{t,4})(1 - \phi_{t,3}) \phi_{t,2} P_s(v|u, S_t) \right] \quad (14)$$

where  $g_n(v|S_t)$  is an indicator of whether  $v$  is a new user. Likewise, the likelihood of an action  $R_{u,i,r}$  in which the user  $u$  gives the rating  $r$  for item  $i$  is:

$$P(R_{u,i,r}|S_t, \Theta) = P(u|S_t) \times (1 - \phi) \times \left[ \Phi_{r,4} g_n(i|S_t) \mu_r + (1 - \Phi_{r,4}) \phi_{r,3} P_e(i|u, S_t) P_e(r|i, u, S_t) + (1 - \Phi_{r,4})(1 - \phi_{r,3}) \phi_{r,1} P_t(i|u, S_t) P_t(r|i, u, S_t) + (1 - \Phi_{r,4})(1 - \phi_{r,3}) \phi_{r,2} P_s(i|u, S_t) P_s(r|i, u, S_t) \right] \quad (15)$$

where  $g_n(i|S_t)$  is an indicator of whether  $i$  is a new item. We use expectation maximization (EM) to estimate the maximum likelihood model parameters.

## 4. RELATED WORK

There has been a lot of research on modeling the evolution of social networks [7][9] [10][13]. However, these works only model the creation of social relations and do not consider attributes for nodes. In the following, we review some recent works which address the problem of modeling the evolution of social networks together with node attributes.

[3] introduces an evolution model for social networks with user activities. In the Wikipedia dataset used in their experiments, activities are the editing of a certain Wikipedia article, and network edges represent interaction, i.e. participation in the discussions on another user's profile. The generative model of [3] considers only two factors: social influence and selection.

The co-evolution of social and affiliation networks has been explored by [18]. In this scenario, nodes are associated with group labels. The proposed model creates edges and group labels. For each user, groups to join are selected from those of his friends (social influence), and new friends are selected from his friends of friends (transitivity). However, the effects of selection and correlational influence are not modeled.

Like our work, [11] investigates social networks with user ratings. Similar to our work the bidirectional effects between trust and ratings are explored, but only *selection* and *correlational influence* are considered. The effect of social relations on the behavior of users is not discussed in this paper.

Evaluating generative models and the fit of synthetic networks with actual networks is an important issue in designing models for the evolution of social networks. In the following we review two recent works which investigate evaluation metrics for generative models.

[1] addresses the problem of distinguishing social influence (causation) from correlation. A shuffle test is proposed based on the intuition that if social influence is not a likely source of correlation in a system, timing of actions should not matter, and therefore reshuffling the time stamps of the actions should not significantly change the amount of correlation. They also define a metric for measuring social correlation. If the social correlation measure does not change after shuffling, then social influence is ruled out as a cause of social correlation.

The authors of [8] present a systematic examination of a real network dataset using maximum likelihood estimation for exponential random graph models as well as new procedures to evaluate how well the models fit the observed networks. These procedures compare structural statistics of the observed network with the corresponding statistics on networks simulated from the learnt model[8]. All metrics in this paper measure some structural property of social networks. Attributes of nodes are not considered.

The microscopic evolution of social networks is investigated in [10]. The authors argue that using the likelihood is a more objective measure for comparing alternative generative models than using some of the many potential relevant network statistics. Only simple social networks without attributes and only the influence of transitivity are considered. The timing of the creation of nodes and edges is also modeled, which is beyond the scope of our work.

## 5. EXPERIMENTS

In this section we present our experiments and discuss the results. To evaluate the performance of our model, we build a generative model for each data set, learning the parameters using maximum likelihood and expectation maximization.

We used the version of the Epinions data set<sup>3</sup> published by the authors of [17]. The data set contains 22K users, 30K items, 108K ratings, and 117K social relations between users. The timestamps in the data set range from 2001/01/17 to 2002/02/01. Flickr.com is an online photo sharing website. We used the version of the Flickr data set<sup>4</sup> published by authors of [2]. The Flickr data set used in this paper contains 150K social relations and 30K rating actions expressed by 5.2K users on 5K items. The timestamps in the flickr data set are from 2206/11/02 to 2007/03/15.

In our experiments, we compare the model proposed in section 4, called the FullModel, against the following models which capture only some of the effects:

- *CrossModel*. This model considers only the effect of ratings on social relations (selection), and the effect of social relations on ratings (social influence). This model is similar to [3], and is derived by removing the parameters  $\Phi_{r,2}$  and  $\Phi_{t,1}$  from the FullModel (or setting these parameters to zero in the FullModel). It should be noted that this model considers social influence and social selection to be constant.
- *SocialOnly*. It ignores the effects of similarities, i.e. it just models the effects of social influence and transitivity. This model is derived by removing parameters  $\Phi_{r,2}$  and  $\Phi_{t,2}$ , from the FullModel. This setting simulates the model proposed by [18] in our framework. It should be noted that the model in [18] ignores the dynamic nature of effects.
- *SimilarityOnly*. It ignores the effects of social relations and just models selection and correlational influence. This model is derived by removing the parameters  $\Phi_{r,1}$ ,  $\Phi_{t,1}$  from the FullModel. Results of this model show whether considering only effects of similarity and correlation of rating patterns can capture the behavior of real data.
- *Baseline*. In this setting, we ignore all four bidirectional effects of social relations and rating patterns. This model is derived by removing the parameters  $\Phi_{r,1}$ ,  $\Phi_{t,1}$ ,  $\Phi_{r,2}$  and  $\Phi_{t,2}$  from the FullModel and is designed to investigate the power of randomness in SRN generation. This model is close to the preferential attachment model[15].

### 5.1 Experimental Results

In this section, we present our experimental results with respect to different evaluation metrics. It should be noted that we generate 10 SRNs using each model and take the average of each evaluation metric over these 10 generated SRNs. Each generated SRN consists of the same number of actions as the real data.

After learning the parameters, some interesting insights can be gained on the behavior of users in Epinions and Flickr. The estimated value of  $\phi$  is 0.52 for Epinions and 0.83 for Flickr. This indicates that in the Flickr data set, users tend to create more social relations compared to adding photos to their favorites list. The estimated value of  $\phi_{t,1}$  is 0.91 for Epinions and 0.9 for Flickr which means that the effect of transitivity is much more important than selection in both data sets. In other words, ignoring the effect of friends of friends on the creation of social relations (as done in [3]) is not realistic. For rating actions, we have  $\phi_{r,1} = 0.59$  for Epinions and  $\phi_{r,1} = 0.54$  for Flickr which shows that the strength of social influence is close to the strength of correlational influence with social influence being a little stronger in both data sets. In the following subsections we evaluate comparison partners against several evaluation metrics.

#### 5.1.1 Comparing the likelihoods

As discussed in [10], using the likelihood of the real data is an objective measure for comparing alternative models. Table 1 lists the log-likelihood of Epinions and Flickr for all of the comparison partners. The FullModel achieves the clearly highest log-likelihood in both data sets, while the Baseline model, that ignores all four bidirectional effects, achieves the lowest likelihood.

#### 5.1.2 Evaluation of structural properties

As proposed in [8], we evaluate a generative model for social rating networks based on some structural properties of the generated graph. We compare the following measures of the generated SRN with those of the real SRN datasets in both Epinions and Flickr.

<sup>3</sup><http://alchemy.cs.washington.edu/data/epinions/>

<sup>4</sup><http://socialnetworks.mpi-sws.org/>

Model	Log-Likelihood	
	Epinions	Flickr
FullModel	-2.914E6	-2.732E6
CrossModel	-2.994E6	-2.791E6
SocialOnly	-2.977E6	-2.764E6
SimilarOnly	-3.031E6	-2.807E6
Baseline	-3.122E6	-2.829E6

**Table 1: Comparison of the log-likelihood of the Epinions data set and the Flickr data set using alternative models.**

- *Link Degree Distribution*: The distribution of the out-degree of nodes.
- *User Rating Degree Distribution*: The rating degree of a user is the number of items rated by him.
- *Item Rating Degree distribution*: The rating degree of an item is the number of users who have rated the item.

Our experiments show that all the above distributions follow a power law. For each model, we compute the *scaling exponent*<sup>5</sup> of each distribution and compare it to the corresponding scaling exponent in the real dataset. Table 2 compares the scaling exponent of different degree distributions for both Epinions and Flickr. Note that the scaling exponents for all models are very close. We believe this is mainly due to the fact that all models exploit the idea of "rich get richer". Thus all the comparison partners are generating SRNs close to the real data in terms of degree distributions.

Structural Measure	Epinions		Flickr	
	Model	Real	Model	Real
Link Degree	2.164	2.132	1.368	1.382
User Rating Degree	1.904	1.901	1.665	1.718
Item Rating Degree	2.68	2.76	1.998	1.969

**Table 2: Comparison of the scaling exponent of each degree distribution of the SRN created by FullModel and the real data for both Epinions and Flickr.**

Another metric we use is the effective diameter [15] of the social network. All comparison partners generate social networks with effective diameter of 6 in Epinions and 4.8 in Flickr, which are the same as the effective diameters of the real data set.

To conclude, evaluation on general structural properties showed that all models generate SRNs that are generally similar to the real data set.

### 5.1.3 Evaluation of social influence

We use an influence model proposed in [1] to compare the degree of influence in SRNs generated by the various models and the real data set.

At each timestamp  $t$ , each user  $u$  is exposed to some items. A user  $u$  is  $k$ -exposed to an item  $i$  at time  $t$  if that individual has not rated item  $i$  and has exactly  $k \geq 1$  direct neighbors who have rated  $i$  by time  $t$  and the latest rating of  $i$  among the direct neighbors of  $u$  is in the interval  $[t - \delta_0, t]$ . Each item  $i$  is exposed to  $u$  at timestamp  $t$  by  $a_{u,i,t}$  users. The probability  $P(u, i, t)$  of user  $u$  rating item  $i$  in a time window of size  $\delta$  after  $t$  is modeled as follows:

$$P(u, i, t) = \frac{e^{\alpha \ln(a_{u,i,t+1}) + \beta}}{1 + e^{\alpha \ln(a_{u,i,t+1}) + \beta}} \quad (16)$$

<sup>5</sup>[http://en.wikipedia.org/wiki/Power\\_law](http://en.wikipedia.org/wiki/Power_law)

Let  $A_{u,t}$  denote the set of items user  $u$  has not yet rated but is exposed to at time  $t$ . Then we can compute the likelihood of the given data  $D$  as follows:

$$P(D) = \prod_t \prod_u \prod_{i \in A_{u,t}} P(u, i, t)^{Y_{u,i,t}} (1 - P(u, i, t))^{1 - Y_{u,i,t}} \quad (17)$$

where  $Y_{u,i,t}$  is a boolean random variable with a value 1 if user  $u$  rates item  $i$  in the time window  $[t, t + \delta]$ <sup>6</sup>.

By maximizing the above likelihood we estimate the values of  $\alpha$  and  $\beta$ . As discussed in [1], the coefficient  $\alpha$  is considered as the measure of influence by the neighbors for selecting an item to rate:

$$\ln\left(\frac{P(u, i, t)}{1 - P(u, i, t)}\right) = \alpha \ln(a_{u,i,t}) + \beta$$

We call this influence coefficient the *item adoption influence* since only the event of rating an item (but not the value of the rating) is taken into account. We can also model the probability  $P(u, i, t)$  of user  $u$  rating item  $i$  within distance 0.5 of the average of ratings of his neighbors on  $i$ . We call this influence coefficient the *rating adoption influence*. The social influence measures for different models and the real data set are presented in table 3. Note that since ratings in Flickr have binary values, rating adoption influence is not applicable to Flickr. As shown in this table, FullModel has the social influence measures that are closest among all comparison partners to those of the real data set in both Epinions and Flickr. Surprisingly, the SimilarityOnly model that uses only the effect of similar users comes fairly close to the real data set in terms of social influence coefficients in both data sets. All other models exhibit social influence coefficients that are much smaller than in the real data set. These results show that correlational influence plays a very important role in rating behavior of users and neglecting it is not realistic.

Model	Epinions		Flickr
	item adoption influence	rating adoption influence	item adoption influence
Real Data	2.1436	2.195	1.41
FullModel	2.186	2.252	1.32
CrossModel	1.170	1.314	1.14
SocialOnly	0.966	1.079	1.21
SimilarityOnly	1.951	2.008	1.56
Baseline	0.687	0.820	0.74

**Table 3: Influence coefficients for item and rating adoption in the social rating network generated by different models and the real data set in Epinions and Flickr.**

### 5.1.4 Measuring the effect of social relations on the growth of similarity of rating patterns

When user  $u$  creates a social relation to user  $v$ , then according to social influence,  $u$  will consider ratings of  $v$  in his future actions. Thus, the similarity of ratings of users  $u$  and  $v$  should grow

<sup>6</sup>For the sake of efficiency, we consider only 12 timestamps in our experiment to compute the exposures in Epinions (2001/02/17, 2001/03/17, ... 2002/01/17).  $\delta$  and  $\delta_0$  are set to 30 days in our experiments with Epinions. In Flickr, we consider 8 timestamps (2006/02/15, 2006/03/01, ..., 2007/03/01).  $\delta$  and  $\delta_0$  are set to 15 days



after creation of the social relation. The more actions they do, the more similar they are supposed to get. The difference between the similarity of  $u$  and  $v$  at the end of the evolution of the SRN and at the moment of creating the social relation is called the *similarity growth* for users  $u$  and  $v$ . In our SRN, users may perform different numbers of actions after creating a social relation. We group pairs of users  $(u, v)$  based on the number of rating actions  $u$  performs after creating the social relation to  $v$ . Figure 6 shows the similarity

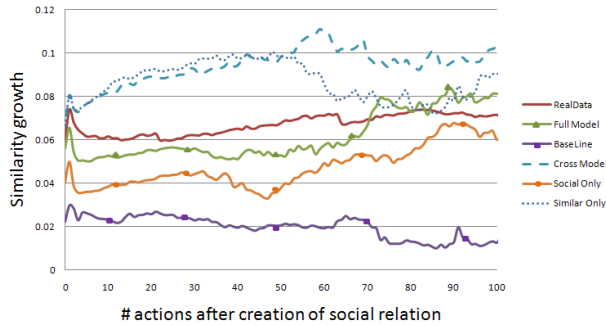


Figure 6: Comparison of Similarity Growth after creation of social relation for different models in Epinions.

growth versus the number of actions for the different models on the Epinions data set. The *FullModel* follows the pattern of similarity growth in the real data more closely than other models. There are some interesting and surprising results in Figure 6. The *SimilarOnly* and *CrossModel* tend to have higher growth of similarity compared to the real data, while *SocialOnly* tends to have a lower growth of similarity compared to both real data and *FullModel*. *SimilarOnly* and *CrossModel* use *selection* for link creation, but *SocialOnly* uses the *transitivity* for link creation. The similarity growth versus the number of actions for the different models in the Flickr data set is depicted in figure 7. Again *FullModel* has the closest pattern of similarity growth to that of the real data. In a way, the results show that users who are already similar before creating a social relation, tend to get more similar compared to users who are not that similar when creating the social relation. *FullModel* is a tradeoff between these two cases and our experiments show that it behaves much more similar to the real data.

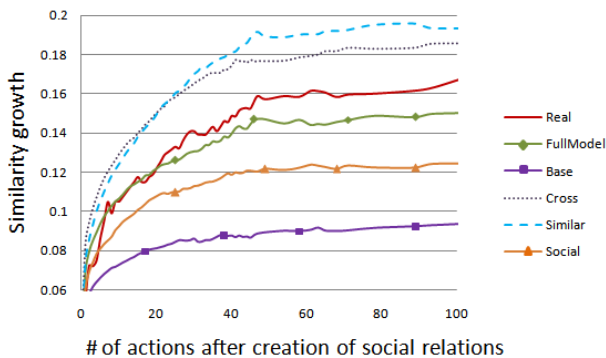


Figure 7: Comparison of Similarity Growth after creation of social relation for different models in Flickr.

### 5.1.5 Measuring the transitivity influence

Transitivity of social relations is an important property which affects the creation of social relations between users. The distribution of the shortest distance between users  $u$  and  $v$ , at the moment when  $u$  creates a social relation to  $v$ , can measure the transitivity effect. If the probability of short distances (i.e. 2) is high, then the transitivity has been an influential factor in trust creation. We compute the average shortest distance of pairs of users at the moment of creating a social relation between these two users. Table

Model	$\overline{d_{min}}$	
	Epinions	Flickr
Real Data	3.93	2.73
FullModel	4.05	2.86
CrossModel	5.59	3.636
SocialOnly	3.24	2.68
SimilarOnly	5.59	3.64
Baseline	5.21	4.76

Table 4: Average shortest distance of user pair at the moment of creating social relations ( $\overline{d_{min}}$ ) in different models and the real data set for Epinions and Flickr.

4 compares the average shortest distance of user pairs at the moment of creating social relations ( $\overline{d_{min}}$ ) in different models and the real data set for both Epinions and Flickr. Note that pairs of users who are not reachable at the time of the creation of a social relation are considered to have distance 7. As shown in this table, the average shortest distance for FullModel is the closest one to that of the real data set. SocialOnly achieves lower average distance since it only relies on neighbors and does not consider the selection effect. CrossModel, SimilarityOnly, and Baseline have a higher average shortest distance, likely because they are ignoring transitivity which according to our experiments is a very important factor in the creation of social relations.

### 5.1.6 Measuring the effect of selection

To evaluate the effect of selection, we compute the average similarity of users when creating a social relation. Table 5 presents the average similarity of user pairs at the moment of creating social relations for different models and for the real data in both Epinions and Flickr. In this table, we only consider the pairs of users for which we can compute the similarity. Table 5 also shows the percentage of user pairs for which we can compute the similarity in each comparison partner. Again, FullModel is clearly the closest to the real data set in both Flickr and Epinions. CrossModel and SimilarOnly have higher average similarity since they only rely on similar users for creating social relations and ignore the transitivity effect. The SocialOnly model, however, also has an average similarity close to that of the real data set. This seems to be due to the fact that strength of selection in Epinions and Flickr is very low compared to transitivity. The FullModel is also closest to the real data set in terms of the percentage of user pairs creating social relation for which we can compute the similarity.

## 6. CONCLUSION

Social rating networks (SRN) are becoming more and more important in reviewing websites such as Epinions.com or online sharing websites like Flickr.com. In this paper we analyzed and modeled the temporal behavior of users in an SRN using bidirectional effects of rating patterns and social relationships. Our model is based on our observations of the behavior of user while express-

Model	Epinions		Flickr	
	Average Similarity	% of non-zero similarity pairs	Average Similarity	% of non-zero similarity pairs
Real Data	0.381	3.07	0.618	6.2
FullModel	0.375	3.21	0.627	5.4
CrossModel	0.524	11.78	0.703	16.8
SocialOnly	0.365	2.01	0.716	4.3
SimilarOnly	0.558	12.03	0.681	15.7
Baseline	0.253	2.40	0.728	5.09

**Table 5: Average similarity of user pairs at the moment of creating social relations in different models and the real data set for both Epinions and Flickr.**

ing ratings or creating social relations. While existing models for other types of social networks have captured some of the factors, our model is the first one to represent all four factors, i.e. social relations-on-ratings (social influence), social relations-on-social relations (transitivity), ratings-on-social relations (selection), and ratings-on-ratings (correlational influence).

Base on our observations, these effects are dynamic. In other words, the strength of each effect is not constant throughout the evolution of an SRN. We analyzed and modeled the temporal dynamism of these effects by defining them to be functions of the dynamic features of an evolving SRN.

Given the sensitive nature of social network data, there are only very few public SRN datasets. This motivates the development of a model to capture the temporal dynamics of users' behaviors and to create such synthetic datasets for research purposes. To evaluate the accuracy of our model, we compared the synthetic SRN generated by our model with the real data and with SRN generated by other models that have been proposed in the literature.

Our experimental study on the Epinions dataset and the Flickr dataset demonstrated that the proposed model produces social rating networks that agree with real world data on a comprehensive set of evaluation criteria much better than existing models. These criteria included the relevant degree distributions and measures of the four major influence factors that we identified for social rating networks. In addition, our experimental study provided interesting insights into the factors driving the evolution of SRNs. In particular, we found that transitivity plays a much bigger role in the creation of social relations than selection, although the most closely related generative model [3] does not even consider transitivity.

We would like to mention a few directions for future research. As most existing generative graph models, we adopt a discrete time concept. However, a continuous time concept and a corresponding model of the inter-action times, similar to that proposed in [10] and [18], may further improve the faithfulness of the generated data. Furthermore, the use for inference purposes, e.g. for link prediction, is one of the motivating applications of generative models and should be explored.

## 7. REFERENCES

- [1] A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlatino in social networks. In *KDD 2008*.
- [2] M. Cha, A. Mislove, and K. P. Gummadi. A measurement-driven analysis of information propagation in the flickr social network. In *Proceedings of the 18th Annual World Wide Web Conference (WWW'09)*, Madrid, Spain, April 2009.
- [3] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In *KDD 2008*.
- [4] N. E. Friedkin. *A Structural Theory of Social Influence*. Cambridge University Press, 1998.
- [5] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 1992.
- [6] P. Holland and S. Leinhardt. Some evidence on the transitivity of positive interpersonal sentiment. *American Journal of Sociology*, pages 1205–1209, 1972.
- [7] P. Holme and M. Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74:056108, 2006.
- [8] D. R. Hunter, S. M. Goodreau, and M. S. Handcock. Goodness of fit of social network models. *Journal of the American Statistical Association*, 103(481), 2008.
- [9] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *KDD 2006*.
- [10] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins. Microscopic evolution of social networks. In *KDD 2008*.
- [11] Y. Matsuo and H. Yamamoto. Community gravity: measuring bidirectional effects by trust and rating on online social networks. In *WWW 2009*.
- [12] M. Mcpherson, L. S. Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444, 2001.
- [13] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 2007.
- [14] M. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.
- [15] M. Newman, A.-L. Barabási, and D. J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, 2006.
- [16] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: an open architecture for collaborative filtering of netnews. In *CSCW '94: ACM conference on Computer supported cooperative work*, 1994.
- [17] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *KDD 2002*.
- [18] E. Zheleva, H. Sharara, and L. Getoor. Co-evolution of social and affiliation networks. In *KDD 2009*.