Uncovering User Interaction Dynamics in Online Social Networks

Zhi Yang[§], Jilong Xue[§], Christo Wilson[*], Ben Y. Zhao[†] and Yafei Dai[§]
[§]Computer Science Department, Peking University, Beijing, China
[*]College of Computer and Information Science, Northeastern University, MA, USA
[†]Computer Science Department, U. C. Santa Barbara, CA, USA
{yangzhi, xjl, dyf}@net.pku.edu.cn, cbw@ccs.neu.edu, ravenben@cs.ucsb.edu

Abstract

Measurement studies of online social networks (OSNs) show that all social links are not equal, and the strength of each link is best characterized by the frequency of interactions between the linked users. To date, few studies have been able to examine detailed interaction data over time. In this paper, we first analyze the interaction dynamics in a large online social network. We find that users invite new friends to interact at a nearly constant rate, prefer to continue interacting with friends with whom they have a larger number of historical interactions, and most social links drop in interaction frequency over time. Then, we use our insights from the analysis to derive a generative model of social interactions that can capture fundamental processes underling user interactions.

Introduction

The last few years has seen the arrival of several measurement studies of user relationships and activities on popular online social networks, including Facebook (Wilson et al. 2009), Twitter (Kwak et al. 2010; Grier et al. 2010), LinkedIn (Leskovec et al. 2008), Renren (Jiang et al. 2010) and others (Schneider et al. 2009). A common observation made across many platforms is that the presence of a social link connecting two users is a poor estimate of the "relationship strength" between them. Fortunately, the gathering of interaction data has allowed to take into account the variation of link strength, enabling the research going beyond the purely topological point (Gilbert and Karahalios 2009; Kahanda and Neville 2009).

Although several studies examine the user interactions on OSNs, the mechanisms that drive how users create interaction in OSNs is still largely unknown. Prior works (Chun et al. 2008; Wilson et al. 2009) focus on a static view of interactions, and therefore only capture a small piece of the picture. The study (Viswanath et al. 2009) examines changing dynamics of user interactions in Facebook, but it fails to give a complete view of user interaction dynamics. A deeper understanding of user interactions requires the formulation of a generative model, which can intuitively capture the processes that drive user interaction events.

Prior works (Yook, Jeong, and Barabási 2001; Barrat, Barthélemy, and Vespignani 2004) provide models of traffic networks, and others (Zhao et al. 2011; Starnini, Baronchelli, and Pastor-Satorras 2013) present models for human face-to-face interactions. Although these models generate weighted network, they are not suitable in the context of today's OSNs due to different underlying dynamics and network properties. No generative graph model exists to explain properties observed in measured traces of user online interactions, or to construct realistic arbitrary-sized user online interaction traces.

In this paper, we seek to fill this void by building a model based on a large detailed trace of user interactions on *Renren*, the Chinese social network similar to Facebook in functionality. Our trace covers over a year in length, and contains data on the creation of 600+K users, 8+Million new links, and 29+ Million interaction events. We present detailed analysis of our interaction data, and extract three processes that drive dynamics of social interaction over the network:

Forgetting process: A particular pair of users slowly decrease their interaction rate over time. The potential reason is that users tend to forget each other as they cannot meet face to face on a regular basis, leading to the closeness between friends declined rapidly over time.

Reinforcement process: For each pair of users, the probability of continued interactions displays a memory reinforcement (inertia). In particular, the more two nodes interact with each other, the more it demonstrates a close relationship between them . Thus, the user would more likely to reinforce this relationship to counteract the forgetting process, which explains how users dynamically direct a finite set of resources (time and attention) towards the relationships.

Exploration process: In order to replace existing ones which are no longer attractive, users continuously explore new interaction relationships at a nearly constant rate, irrespective of their age or degree.

We propose a generative model for social interactions that is based on these three processes. Our model is important for understanding how the pairwise user interaction evolves. In addition, our model can be coupled with prior social models (Barabási and Albert 1999; Leskovec, Kleinberg, and Faloutsos 2007; Leskovec et al. 2008). This is an important application as we can capture the dynamical and full spectrum of relationship strengths in OSNs.

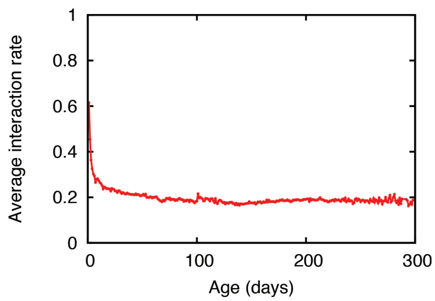

Figure 1: The average interaction rate over users of different age.

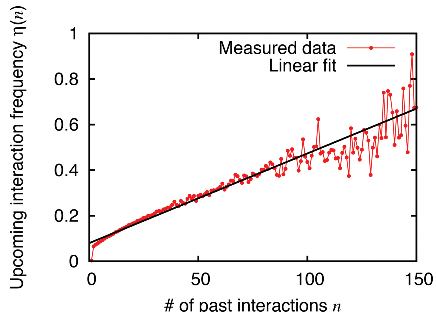

Figure 2: Upcoming interaction frequency $\eta(n)$ for an interaction edge that already has n interactions.

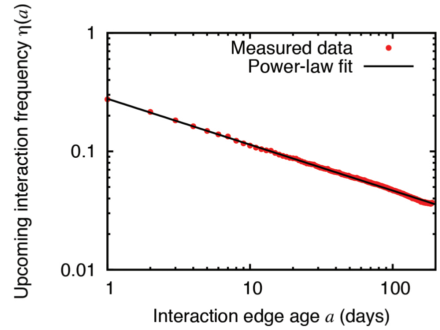

Figure 3: Upcoming interaction frequency $\eta(a)$ for an interaction edge of age a.

Experimental Dataset

With 120 million users, Renren is the largest and oldest online social network in China, and provides functionality and features similar to Facebook. The most popular interaction between users in the early evolution of Renren is writing wall posts to each other.

Our interaction dataset includes all the 29,506,068 wall posts occurring in our measurement period. To guarantee user privacy, we only get the anonymized IDs of sender and receiver for each wall post, without knowing the content. Since our goal is to characterize edge strength based on user interactions, we ignore the wall posts not along edges (e.g., greeting messages between strangers). As a result, we focus on the remaining 23,000,141 wall posts created along edges, representing the friendship maintenance effort of users (accounting for nearly 80% of the total wall posts).

To better measure mutual relationship (tie strength) between users, we use the term *interaction* to mean a pair of reciprocal wall posts. For example, if node u sends m messages to v but receives n messages from v, the number of interactions between them is $\min(m, n)$. The wall posts that have not been replied to are pruned. This definition means that u and v cannot be supposed to have strong mutual relationship if one sends many messages to the other but rarely receives replies (e.g., u trusts user v, but not necessarily vice versa). So we use the the number of interactions as a conservative estimate on the edge strength, instead of the total number of wall posts that occurred along the edge. Based on this definition, we transform wall posts into 7,639,488 interactions.

The interaction definition also allows us to represent the Renren *interaction network evolution* as a series of undirected edge-weighted graphs $G_1, \ldots, G_T$, so that a snapshot $G_t = (V_t, E_t, W_t)$ consists of the nodes, edges and corresponding interactions that have arrived by time t. The term *interaction edge* represents the friendship edge along which at least one interaction is generated. We define the creation time of an interaction edge as the time when its first interaction is generated. For example, we say a node u creates an interaction edge (u, v) at time t if u and v begin to interact with each other at that time. Table 1 summarizes our interaction dataset. We see that interaction edges only accounts for a small fraction (32%) of the total edges, meaning users only interact with a small subsect of their friends.

Period	2005.11 ∼ 2006.12
# of wall posts	29,506,068
# of wall posts along edges	23,000,141
# of interactions	7,697,270
# of users having interactions	420,978
Mean # of interactions per user	18
# of edge having interactions	2,623,040
Mean # of interactions per edge	3

Table 1: Summary of Renren interaction data

Analysis of User Interactions

In this section, we analyze the Renren interaction data to uncover the temporal pattern of user interaction evolution. In the following, we study how users create interaction edges and generate interactions along these edges.

Interaction Partners Invitation

Intuitively, making friends in an OSN is very easy, since the click of "add as friend" button does not need any energy cost. In contrast, interaction relationship requires more effort to create and maintain, e.g., a certain amount of time and energy used for reading and writing wall posts. Such energy cost will limit the rate at which users add new interaction partners, since they only have a finite amount of resources (e.g., time and energy).

Given a node u of age a, we define its *interaction rate* $r_u(a)$ as the ratio of the number of interaction edges $n_u(a)$ it has created to its current age a, i.e., $r_u(a) = n_u(a)/a$. Interaction rate measures the speed at which users request new friends to interact with. To examine the temporal pattern of interaction edge initiations, we examine $R(a)$, the average interaction rate of nodes achieved age a during our measurement period:

$$R(a) = \frac{\sum_{t=1}^{T} \sum_{u \in S_t(a)} r_u(a)/|S_t(a)|}{T} \quad (1)$$

where $S_t(a) = \{u|t - t_0(u) = a\}$ is the set of nodes achieve age a at time step t. Here, $t_0(u)$ is the arrival time of a node u. As shown in Figure 1, people tend to be more interactive immediately after they join, but the effect quickly wears off as the interaction rate $R(a)$ converges to a constant just after a week.

Interaction Distribution

Next, we examine how users distribute new interactions over their existing interaction partners. We analyze the interaction distribution from two perspectives: first, what is the effect of *intensity*, *i.e.* is their correlation between the number of times friends have interacted in the past, and the number of times they will interact in the future?

Fig. 2 plots $\eta(n)$, the average number of new interactions between friends that already have n interactions:

$$\eta(n) = \frac{\sum_{t=0}^{T} \sum_{e \in S_n(t)} I_e(t)/|S_n(t)|}{T} \tag{2}$$

where $S_n(t) = \{e| \sum_{k=0}^{t-1} I_e(k) = n\}$ is the set of interaction edges that already have n interaction before time t. We observe that $\eta(n)$ is proportional to the number of past interactions across the edge in the network. Intuitively, this means that the interactions between friends *reinforce* their relationship, leading to more future interactions.

Second, what is the effect of *time*, *i.e.* do friends tend to interact more or less over time?

Fig. 3 plots $\eta(a)$, the average number of new interactions created along edges of age a:

$$\eta(a) = \frac{\sum_{t=0}^{T} \sum_{e \in S_a(t)} I_e(t)/|S_a(t)|}{T} \tag{3}$$

where $S_a(t) = \{e|t - t_0(e) = a\}$ is the set of interaction edges with age a at time t, and $I_e(t)$ is the number of new interactions generated along edge e at time t. We see that $\eta(a)$ is inversely proportional to edge age a in the network. Intuitively, this means that a given pair of users tends to interact less over time. One possible explanation for this is users tend to forget each other as they cannot meet face to face on a regular basis, leading to the closeness between friends declined rapidly over time. Interestingly, our observation on online relationships is consistent with the ecology model on real-life relationships. Prior work (Burt 2000) investigated four annual surveys of colleague relationships for 345 bankers in a large financial organization, and found that the liveness of relationships decay over time and decay is also a power function of time.

User Interaction Model

Besides befriending with others, nodes also request a certain number of friends to interact with, and distribute interactions over their interaction friends. Based on the insights on user interaction behavior in the OSNs, we now introduce our interaction model, as shown in Figure 4.

Intuitively, not all the users are simultaneously present in system. Thus we assume that users can be in an active or an inactive state. If an user is active, she interacts with her

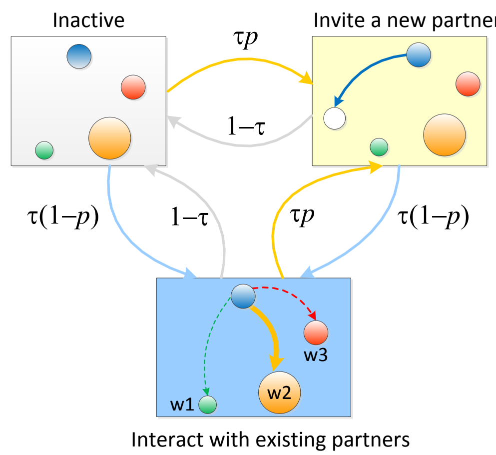

Figure 4: User interaction model in OSNs.

friends; otherwise she simply rests without interacting. According to empirical observations (e.g., constant interaction frequency), we assume that, at each time step, one inactive user can become active with a probability r, while one active user can become inactive with probability $1 - r$. In practice this means that the user activity pattern, while stochastic, will display some regularity in time, interaction events following each other on average at $1/r$ steps, very long interevent times are exponentially rare.

Once the user is active, she would communicate one of her friends. The empirical observations show that the user invites new friends to interact at a constant rate, irrespective of node age. Therefore we assume that, once the node is active, with probability p, she chooses to communicate with a new interaction partner from her friends that she have not interacted with yet. With the complementary probability $1-p$, the user chooses to communicate with one of her existing partners that she has previously interacted with. However, interactions are biased by the interpersonal attraction built up over time. The more interest she raises in a partner, the more likely she will interact with this partner (inertia). Based on the empirical observations (e.g., effects of intensity and time), we measure the appeal η_{uv} of a partner v to an user u by $\eta_{uv} = n_{uv}/a_{uv}^{\tau}$, where n_{uv} is the current number of interactions between users u and v, and a_{uv} is the current age of this interaction relationship. Thus, if an user u chooses to interact with an existing partner, she will choose the partner v with a probability proportional to η_{uv}.

The interaction model captures the fundamental fact that the interaction relationships require that we invest time to keep them alive, especially once it becomes physically difficult for friends to meet face to face on a regular basis. In particular, each user has a *forgetting* behavior: the attraction between a pair of users declined rapidly when they lose contact (captured by the decay factor τ). To counteract the effects of forgetting, each user exhibits a *reinforcing* behavior: she wants to keep the important relationships alive. Thus, with limited time to use, she biases towards relationships of more interactions. Also, each user has a *exploring* behavior: she continuously explores new interaction relationships (captured by the probability p), in order to replace existing ones which are no longer attractive.

Heterogeneous Model

The activation probability r represents user activeness in the social interaction. To this end, we have assumed that all users have the same tendency to be active, that is, the activation probability r does not depend on the user who is interacting. Real social systems display however additional complexity since the social behavior of individuals may vary significantly across the population. For example, individuals vary widely in the total time spent accessing OSNs (Benevenuto et al. 2009), and may devote different amount of energy to interaction.

A natural extension of the model presented above consists therefore of making the probability r dependent on the user who is interacting. To this aim, we assign to each user u a parameter r_i that characterizes his/her propensity to form social interactions. In real networks this propensity will depend on the features of the users. In the model we assume that this propensity, that we call "sociability", is a quenched random variable randomly chosen from a prefixed distribution $\zeta(r)$ characterizing the system's heterogeneity, which is assigned to each agent at the start of the dynamical evolution and remains constant.

Conclusion

In this paper, we develop an interaction model that generates interactions across social links. The insights behind our model are derived from a large scale dataset from Renren. This data reveals that users invite new friends to interact at a nearly constant rate, prefer to interact with friends with whom they have many past interactions, and gradually lose interest in interacting with old friends. We believe that these observations not only affect the design of network interaction models but also have broader implications in other areas, such as friend recommendation, information diffusion, and news feed ranking.

There are several directions to extend the current study. First, we can combine our interaction model with social graph model to generate interaction graphs. Another direction is to accommodate more attributes of nodes to improve the accuracy of the model. Recent works (Allamanis, Scellato, and Mascolo 2012) begin to examine influence of spatial attribute on the temporal evolution of friendship links, but how these factors affect interaction evolution remains unknown.

Acknowledgments

We thank the anonymous reviewers for their comments. This work was supported by the National Science Foundation for Young Scholars of China (Grant No.61202423).

References

Allamanis, M.; Scellato, S.; and Mascolo, C. 2012. Evolution of a location-based online social network: Analysis and models. In *Proc. of IMC*.

Barabási, A., and Albert, R. 1999. Emergence of scaling in random networks. *Science* 286(5439):509–512.

Barrat, A.; Barthélemy, M.; and Vespignani, A. 2004. Modeling the evolution of weighted networks. *Physical review E*.

Benevenuto, F.; Rodrigues, T.; Cha, M.; and Almeida, V. 2009. Characterizing user behavior in online social networks. In *Proc. of IMC*.

Burt, R. S. 2000. Emergence of scaling in random networks. *Social Networks* 1(22):1–28.

Chun, H.; Kwak, H.; Eom, Y.; Ahn, Y.; Moon, S.; and Jeong, H. 2008. Comparison of online social relations in volume vs interaction: a case study of Cyworld. In *Proc. of IMC*.

Gilbert, E., and Karahalios, K. 2009. Predicting tie strength with social media. In *Proc. of CHI*.

Grier, C.; Thomas, K.; Paxson, V.; and Zhang, M. 2010. @spam: the underground on 140 characters or less. In *Proc. of CCS*.

Jiang, J.; Wilson, C.; Wang, X.; Huang, P.; Sha, W.; Dai, Y.; and Zhao, B. Y. 2010. Understanding latent interactions in online social networks. In *Proc. of IMC*.

Kahanda, I., and Neville, J. 2009. Using transactional information to predict link strength in online social networks. In *Proc. of ICWSM*.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. B. 2010. What is twitter, a social network or a news media? In *Proc. of WWW*.

Leskovec, J.; Backstrom, L.; Kumar, R.; and Tomkins, A. 2008. Microscopic evolution of social networks. In *Proc. of KDD*.

Leskovec, J.; Kleinberg, J. M.; and Faloutsos, C. 2007. Graph evolution: Densification and shrinking diameters. *TKDD*.

Schneider, F.; Feldmann, A.; Krishnamurthy, B.; and Willinger, W. 2009. Understanding online social network usage from a network perspective. In *Proc. of IMC*.

Starnini, M.; Baronchelli, A.; and Pastor-Satorras, R. 2013. Modeling human dynamics of face-to-face interaction networks.

Viswanath, B.; Mislove, A.; Cha, M.; and Gummadi, K. P. 2009. On the evolution of user interaction in facebook. In *Proc. of WOSN*.

Wilson, C.; Boe, B.; Sala, A.; Puttaswamy, K. P. N.; and Zhao, B. Y. 2009. User interactions in social networks and their implications. In *Proc. of EuroSys*.

Yook, S. H.; Jeong, H.; and Barabási, A.-L. 2001. Weighted evolving networks. *Physical review letters*.

Zhao, K.; Stehlé, J.; Bianconi, G.; and Barrat, A. 2011. Social network dynamics of face-to-face interactions. *Physical review E*.