Forwarding Links without Browsing Links in Online Social Networks

Jing Jiang‡, Xiao Wang†, Li Zhang§, Yafei Dai†
‡State Key Laboratory of Software Development Environment Beihang University, Beijing, China
†Department of Computer Science and Technology, Peking University, Beijing, China

{jiangjing, lily}@buaa.pku.edu.cn, {wangxiao, dyf}@pku.edu.cn

Abstract—Online social networks (OSNs) become important platforms of information diffusion. People forward links pointing to interesting content, and share them with friends. Browsing links is generally considered as the previous step for forwarding links. However, few studies have confirmed this hypothesis before. The deep understanding of forwarding links can help OSNs to improve current forwarding mechanism, and allow people to read more information and forward more links.

In this paper, we study the relationship between forwarding links and browsing links. We crawled a connected graph component of 42.1 million users, and 2.1 billion links forwarded by them. Moreover, we collected browsing records of 50,000 active users. Based on these datasets, we observe that 42% of users forward more than 50% of links without browsing links. Browsing links is often unnecessary for forwarding links. This finding changes the traditional concept of link dissemination in OSNs. We also find that when users directly forward links, they have more time to leave more comments and forward more links, which promotes the spreading of links. OSNs should encourage people to read snippets and directly forward links.

Keywords—Forwarding links, online social networks, measurement

I. INTRODUCTION

Online social networks have brought to the public a new style of social lives parallel to offline activities, and they have the potential to alter the way people interact with the internet. The popularity of online social networks (OSNs) makes them major platforms of information diffusion. 1 million links are shared in Facebook every 20 minutes.1

The explosive growth in OSNs has given researchers opportunity to obtain massive quantities of data for empirical analysis. Initial studies have measured the characteristics of information diffusion in Flickr [1], Facebook [2], [3], Twitter [4], [5] and other online social networks. Some works focus on people and analyze user’s uploading or retweeting activities [6], [7], [8], [9].

As shown in Figure 1, normal steps of link diffusion are as follow [2]: Firstly, users primarily interact with information through an aggregated history of their friends’ recent activities, which is called as the news feed. Majority of these activities contain links to content, which are identified by URLs. According to the content, snippets are automatically generated and displayed below links. People may also find links by external events, such as e-mails or instant messages. Secondly, users read snippets and decide to browse some links. Finally, if the content is interesting, people forward links and recommend them to friends.

Search engines display snippets of search results. People sometimes find enough information from snippets, and do not take a further step to visit web pages. Since OSNs provide similar snippets, we ask the question: Do users often forward links after browsing links, or directly forward links without visiting the links? To our knowledge, no one has studied the relationship between browsing links and forwarding links before. Measurement results can help OSNs to improve current forwarding mechanism, and allow people to read more information and forward more links.

Our study. We collect and analyze large-scale traces in Renren social network2, one of the largest and most popular OSNs in China. We crawl an exhaustive snapshot of 42.1 million users and capture 2.1 billion links forwarded by these people. We also obtain browsing records of 50,000 active users. These datasets give us an opportunity to gain valuable insights in OSNs.

We provide a number of insights into the relationship between forward links and browsing links, including:

- 42% of users forward more than 50% of links without browsing links.
- From 6 AM to 8 AM, people forward 68% of links without browsing links. Users prefer to directly forward links in the morning, when they are busy at work.

1http://www.statisticbrain.com/facebook-statistics/

2http://www.renren.com
• When people forward links without browsing links, they leave more comments and forward more links.

Our findings change traditional concept of link dissemination: Normal steps in Figure 1 do not apply for all links. In fact, browsing links is sometime unnecessary before forwarding links. Moreover, forwarding links without browsing content is beneficial for link dissemination. It saves user’s time of browsing web pages, and provide users more time to read snippets, forward links and leave comments. OSNs should generate good snippets, provide enough information, save user’s time of browsing links, and allow them to forward more links.

II. COLLECTION METHODOLOGY

Before diving into detailed measurement, we provide background information about the Renren social network. Then we introduce our collection methodology.

Renren is one of the biggest and oldest OSNs in China [10]. Renren can be best characterized as Facebook’s Chinese twin, with a similar user interface and most or all of Facebook’s features [11]. Users maintain personal profiles and establish bidirectional friend relationships. They publish different types of content, such as blogs, photos and albums. When people find interesting content, they forward links of the content. This function is similar to the ‘share’ function in Facebook. In user’s news feed, Renren automatically displays snippets below links. Snippets are generated based on the content type, which is similar to the search engine. More specifically, a small part of texts are extracted for blogs, and thumbnails are generated for photos and albums.

In order to study the relationship between forwarding behavior and browsing behavior, we find opportunities to collect datasets from Renren: Firstly, people forward links in Renren, and these links are organized into completely open pages. Therefore, we collected 2.1 billion links forwarded by 42.1 million users. We use these datasets to understand browsing behavior. Secondly, we build long-term collaboration with Renren. We directly obtain browsing records of 50,000 active users from Renren, and use them to understand browsing behavior. Different from forwarding records, browsing records are not displayed in profiles and they cannot be collected by traditional measurement techniques. Therefore, few works have studied whether browsing links is necessary before forwarding links. The collaboration with Renren provides us this special opportunity.

Crawling forwarding records. Friend relationships were public in the past. This allowed us to perform an exhaustive crawl of the largest connected component with 42.1 million users [10]. We seed crawlers with these 42.1 million users and proceed to crawl their links. We record the username, the link identifier, the forwarding time and the number of comments on this link. Renren labels the link’s type, such as blog, photo and album. The type of link is gathered too. In total, 42,115,509 users forward 2,052,205,513 links. Some links are same, because they are forwarded for by different people. After filtering out same links, we gather 118,228,021 unique links.

Collecting browsing records. Our group has built long-term collaboration with technical teams in Renren. We randomly select 50,000 active users who forwarded at least 20 links from January to June of 2012. We manually check 100 users, and none of them are spammers. Then Renren provides browsing records of 50,000 users between January and June of 2012. The browsing record includes the username, the link identifier and the browsing time. These datasets provide us an opportunity to gain valuable insights of user behavior.

We focus on user behaviors and do not need any actual content. Therefore, we wait for crawls to complete, then go through data to anonymize user IDs and strip any private data from our datasets to protect user privacy. Moreover, all user IDs are hashed to random IDs, link addresses are mapped to hash values. All timestamps are replaces with relative sequence numbers.

III. FORWARDING LINKS WITHOUT BROWSING LINKS

In user’s news feed, OSN automatically displays snippets of links recommended by friends. People read snippets and choose to visit some web pages pointed by links. Then people find some interesting web pages and forward their links. Figure 1 describe these normal steps of link dissemination. However, snippets already provide summary of web pages, and we doubt whether users must browse links before forwarding them. We ask the question: Do users often forward links after browsing links, or directly forward links without visiting links? In order to answer this question, we study the relationship between forward links and browsing links.

A. Definition and Basic analysis

As shown in Figure 2, we classify forwarding behaviors into three modes:

• After-browsing: The user browses the content of the link, and then forwards the link. This mode is considered as normal.

• Before-browsing: The user forwards the link, and visits the link later. When the user reads the interesting snippet, he or she may not have enough time to browse the link
immediately. Therefore, the user forwards the link, and later visits the link.

- Without-browsing: The user directly forwards the link, without browsing the link. If snippet provides enough information, the user immediately shares the link with friends, but never visits the original web page.

If the user browses the same link for several times, we consider the first visit here. For example, a user browses the link at first, then forwards the link, and later browses the link again. This forwarding behavior is still classified as the after-browsing mode.

As described in section II, our datasets include forwarding records and browsing records of 50,000 active users. We use these datasets and classify forwarding behaviors into appropriate forwarding modes. According to username and the link identifier, we match forwarding record and browsing record. If the forwarding record does not match any browsing record, it means that the link is directly forwarded without browsing content. Therefore, the forwarding activity belongs to without-browsing mode. If the forwarding record matches a browsing record, then we compare the forwarding time and the browsing time. If the the forwarding time is before the browsing time, the forwarding activity belongs to before-browsing mode, otherwise the forwarding activity belongs to after-browsing mode.

For every user, we compute the percentage of links forwarded in each mode. Then we aggregate all users the percentage and plot results in Figure 3. 49% of users forward less than 50% of links after browsing links, while other 51% of users forward more than 50% of links after browsing links. The after-browsing mode is common in online social networks. In contrast, the forwarding before browsing mode is rare. 92% of users forward less than 10% of links before browsing links. Surprisingly, 42% of users forward more than 50% of links without browsing links. This important finding shows that the forwarding without browsing mode is common in online social networks. Some forwarding activities do not follow normal steps. Users read snippets of links, and then directly share links without browsing links. This finding has significant implication: Snippet generation becomes important in the forwarding mechanism. Good snippets can provide enough information and replace content pointed by links.

B. Transitions between different modes

In this subsection, we take a further step and explore transitions between different modes. We ask the question: If the user shares the link in a forwarding mode, will this user forward the next link in the same mode? For every user, we compute the transition probability from a forwarding mode to a forwarding mode. For example, the user forwards 5 links, and their forwarding modes are: after-browsing, after-browsing, without-browsing, after-browsing, without-browsing. The user has 3 after-browsing behaviors. 1 after-browsing behavior is followed by the after-browsing behavior, and 2 after-browsing behaviors are followed by without-browsing behaviors. Therefore, the transition probability from the after-browsing mode to the after-browsing mode is 33%, and the transition probability from the after-browsing mode to the without-browsing mode is 67%. Since the user does not forward links before browsing links, we do not compute the transition probability from the before-browsing mode to any modes.

For all users, we calculate these transition probabilities. Then we compute average value of all users and plot results in Figure 4. The transition probability from the after-browsing mode to the after-browsing mode is 65%. If people forward links in the after-browsing mode, they are likely to forward next links in the after-browsing mode. The transition probability from the without-browsing mode to the without-browsing mode is 68%. Results show that both after-browsing and without-browsing activities are stable. When people are free, they may continually forward links after they browse links; when they are busy, they may also successively forward links without browsing links. The before-browsing mode has 39% probability of moving to the after-browsing mode, and 51% probability of moving to the without-browsing mode. The before-browsing mode is unstable, and quickly moves to other modes.
C. Time Distribution

Next, we study the time distribution of forwarding modes. First of all, we study daily patterns of links forwarded in different modes. We group links by the hour of day, and plot the percentage of links forwarded in each mode in Figure 5. For example, 0 in the horizontal axis stands for links forwarded between 0 AM and 1 AM. From 0 to 4, the percentage of links in the without-browsing mode is slightly larger than that of links in the after-browsing mode. From 5 to 11, the percentage of links in the without-browsing mode is much larger than that of links in the after-browsing mode. More specifically, people forward 68% of links in the without-browsing mode from 6 AM to 8 AM, while people only forward 28% of links in the after-browsing mode from 6 AM to 8 AM. From 12 to 17, the percentage of links in the without-browsing mode is slightly smaller than that of links in the after-browsing mode. From 17 to 23, the difference value between without-forwarding curve and after-forwarding curve is bigger than 5%. The percentage of links in the before-browsing mode is the smallest all the day. We guess this phenomenon is probably because that people are busy at work in the morning, and they prefer to quickly read snippets and forward links directly. People are free in the evening and have enough time. They prefer to browse links and forward links. The time of the day has the significant effect on forwarding modes.

Some links are are forwarded after browsed. For these links, the interval time is calculated as the forwarding time minus the browsing time. For example, the interval time is $t_2 - t_1$ in Figure 2 (a). For every user, we compute the average interval time of all links which are forwarded after browsed. Then we aggregate all users the interval time and plot results in Figure 6. 63% of users have the interval time less than 20 minutes, but 23% of users have the interval time more than 300 minutes (5 hours). A majority of people forwards links immediately after they browse links, while a small part of people share links a long time after they visit links. This minority of users do not forward links immediately after they browse links. Their activities are also strange and need further study in future work.

Some links that are forwarded before browsed. For these links, the interval time is calculated as the forwarding time minus the browsing time. In Figure 2 (b), the interval time is $t_4 - t_3$. Figure 7 shows the CDF of interval time for links that are shared before visited. 25% of users have the interval time less than 1 day, but 26% of users have the interval time more than 10 days. People may forward links to make bookmarks. They firstly forward links and add links as favorites. Then they visit links when they are free.

D. Activity

In this subsection, we analyze activities related to different forwarding modes. When people forward links, they sometimes write comments to express opinions and attract friends’ attention. For every user, we compute the number of comments for links forwarded in each mode, divided by the number of links forwarded in each mode. The result is the average number of comments per link. Then we compute the average value of all users and plot results in Table I. For links in the after-browsing mode, people only leave 0.15 comments per link. For links in the before-browsing mode, people leave 0.18 comments per link. For links in the without-browsing mode, people leave 0.27 comments per link, which is
nearly twice as the value of links in the after-browsing mode. In the without-browsing mode, people do not spend time in browsing links, but have more time to express opinions and leave comments. In news feed, comments are displayed in links’ snippets and attract friends’ interest. Results show that the without-browsing mode promotes the spreading of links in OSNs.

Next, we explore how forwarding modes impact the frequency of forwarding behaviors. For every user, we compute the percentage of links forwarded in each mode. We find the maximum percentage and the corresponding forwarding mode. Then we classify the user into the type of maximum forwarding mode. For example, the user forwards most of links in the without-browsing mode, and the user is classified as the without-browsing type. For users in each type, we plot the number of forwarded links in Figure 8. People seldom forward links before browsing, and only 3 people are classified into the before-forwarding type. For the after-forwarding type, 63% of users forward less than 100 links, and other 37% of users forward more than 100 links. For the without-forwarding type, only 39% of users forward less than 100 links, and other 61% of users forward more than 100 links. Users in the without-forwarding type share more links than users in the after-forwarding type. When people prefer to forward links directly, they spend little time in visiting links, and have additional time to read more snippets and share more links with friends. Results further show that the without-browsing mode is beneficial for the link dissemination in OSNs.

E. Content type

Links point to different types of content. We wonder whether people prefer to forward links belonging to the same type. 97.8% of links belong to blogs, photos or albums, while only 2.2% of links belong to other types. Therefore, we focus on these 3 types and analyze their impacts on forwarding modes.

First of all, we examine whether people mainly forward a certain type of links. For every user, we rank the type based on the number of links forwarded in each type, and calculate the percentage of links in the biggest type. For example, a user forwards 3 blogs, 10 photos and 7 album. The biggest type is photo, and the percentage of links in the biggest type is 50%. We compute the percentage of forwarded links in the biggest type, and plot results in Figure 9. 30% of users forward less than 50% of links in the biggest type, while other 70% of users forward more than 50% of links in the biggest type. People concentrate links into the biggest type.

We take a further step and find the biggest type for every user. Then we calculate the percentage of users who like each type most and plot results in Table II. The blog is the most popular type. 80% of people forward most of links pointing to blogs, 12% of users prefer to spread photos, and other 8% of users like albums most. In order to satisfy user’s interests, OSNs can identify user’s favorite type and recommend links belonging to the corresponding type.

Finally, we study the distribution of forwarding modes in various content types. For every user, we compute the number of blogs’ links forwarded in the after-forwarding mode, divided by the total number of blogs’s links of this user. We also compute the percentage of blogs’ links in other forwarding modes. Then we calculate the average value of all users. For links pointing to photos and albums, we also compute the percentage of links forwarded in each forwarding mode, and plot results in Figure 10. For links pointing to photos, 57.8% of links are in the after-forwarding mode, while 36.8% of links are in the without-forwarding mode. Users are more likely to visit links and then forward links. For links pointing to albums, the percentage of links in the after-forwarding mode is as high as 66.5%. We compare different types and observe that people mostly prefer without-browsing mode for blogs, and like after-browsing mode for albums. The type of content has obvious

![Fig. 8. Number of forwarded links for different types of users](image1)

![Fig. 9. Concentration of content type](image2)

<table>
<thead>
<tr>
<th>Type</th>
<th>% of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blog</td>
<td>80%</td>
</tr>
<tr>
<td>Photo</td>
<td>12%</td>
</tr>
<tr>
<td>Album</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Table II**

The Distribution of the Biggest Type
impacts on forwarding activities.

IV. RELATED WORK

Much effort has been put into understanding of information propagation in online social networks. Photo’s spread in Flickr [1], [12], dissemination of pages in Facebook [13], [3], product adoption in Instant Messaging [14], individual invitation in games [15], ‘gesture’ diffusion in Second Life [16], linking pattern through the blog [17], propagation of tweets in Twitter [4], [5] and forwarding emails in a social sensor system [18] all report on information dissemination. These works mainly explore characteristics of information cascades, such as dissemination scope and speed.

Some studies analyze user’s uploading or forwarding behaviors. Ding et al. examine the uploading behavior of YouTube users [6]. Yang et al. understand retweeting behaviors in Twitter [7]. Previous works [8], [9] compare different kinds of content recommendation, and study their effects on user’s forwarding behaviors.

We measure the relationship between forwarding behaviors and browsing behaviors. To our knowledge, this is the first study of forwarding links without browsing links in OSNs.

V. CONCLUSIONS

In this paper, we study the relationship between forwarding links and browsing links. We observe that 42% of users forward more than 50% of links without browsing links. People sometimes read snippets and directly forward links, without browsing links. This finding changes the traditional concept of link dissemination in OSNs. We take a further step and observe characteristics of the without-browsing mode: 1) Users prefer to directly forward links in the morning, when they are busy at work. 2) In the without-browsing mode, people leave more comments and forward more links. These findings suggest that without-browsing mode is beneficial for link dissemination. Users do not spend time to browse links, and have more time to read snippets and forward links, especially when they are busy at work. It implies that snippet generation becomes important in the forwarding mechanism. OSNs should generate good snippets to provide enough information of links, save user’s time of browsing links, and encourage the without-browsing mode.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation of China under the National Basic Research Program of China grant No. 2011CB302305, National Natural Science Foundation of China under Grant No.61300006, and the State Key Laboratory of Software Development Environment under Grant No.SKLSEDE-2013ZX-26.

REFERENCES