Measuring Message Propagation and Social Influence on Twitter.com

Shaozhi Ye[1] and Felix Wu[2]

[1] Department of Computer Science, University of California, Davis
sye@ucdavis.edu
[2] Department of Computer Science, University of California, Davis
wu@cs.ucdavis.edu

Abstract. Although extensive studies have been conducted on online social networks (OSNs), it is not clear how to characterize information propagation and social influence, two types of important but not well defined social behavior. This paper presents a measurement study of 58M messages collected from 700K users on `Twitter.com`, a popular social medium. We analyze the propagation patterns of general messages and show how breaking news (Michael Jackson's death) spread through Twitter. Furthermore, we evaluate different social influences by examining their stabilities, assessments, and correlations. This paper addresses the complications as well as challenges we encounter when measuring message propagation and social influence on OSNs. We believe that our results here provide valuable insights for future OSN research.

1 Introduction

Online social networks (OSNs) have become a major Internet service for people to communicate with each other. To measure the information propagation process and the driving force behind it, social influence, this paper presents a detailed analysis of 58M messages over `Twitter.com`. More specifically, the following problems are investigated in this paper.

- How do messages propagate on Twitter?
- How can we measure social influence on Twitter?
- How do different influences correlate with each other?

To answer these questions, we first need to formally define message propagation. If a user u_0 sends a message M to a set of users $U = \{u_1, u_2, \dots\}$, we say M is *originated* from u_0 and *propagated* to U.

Furthermore, we are interested in message propagation over multiple hops, i.e., message cascading [2, 3]. For example, if a user $u_1 \in U$ also announces M after receiving it from u_0 and another user $v \notin U$ receives from u_1, then M is propagated by two hops, i.e., $u_0 \to u_1 \to v$. Each hop increases the chance for M to reach more users. Our analysis shows that OSNs are able to propagate messages quickly, which is an indicator of small-world effect (Section 3).

Different from the forwarding/broadcasting models in computer networks or peer-to-peer (P2P) networks, the content of M is often modified by users during the propagation over OSNs, which makes it hard to trace M. In this paper, we model the propagation of a message as a *message flow*, which is a tree of messages. Each node in the tree is a message, which is a reply to its parent node. When a message has multiple originators who receive this message outside of the social network of interest, we may model it as a forest by making each originating message the root of a tree. Section 3 presents our analysis of message flows on `Twitter.com`.

In some cases, especially in the event of breaking news, it is possible to identify a set of closely related messages by keywords. These messages are all about the same event although they may belong to different message flows. Section 4 presents our analysis of 500K messages related to Michael Jackson's death.

Furthermore, as previous studies have shown that during information propagation some users are more influential than others [1], we evaluate several social influences by developing a set of metrics to compare different influences. More specifically, we address the following problems.

- How to evaluate social influence metrics (Section 5.1)
- The change of social influence over time (Section 5.2)
- How to assess social influence (Section 5.3)
- Correlations between different influences (Section 5.4)

The rest of this paper is organized as follows. Section 2 describes how we collect the data set for analysis. Section 3 presents our measurement results for message propagation and Section 4 shows how the breaking news of Michael Jackson's death spread through the social network. Then Section 5 evaluates five social influence metrics. After reviewing prior work in Section 6, the paper concludes with Section 7.

2 Data collection

`Twitter.com` is a "real-time information network" for people to share news and communicate with others. A user (twitter) may send messages (tweets) via any of the following channels:

- Twitter website or its API
- Cellphone short message services (SMS)
- Instant messenger (IM)
- E-mail

Many Twitter clients have been developed on top of these channels. Messages are broadcast by Twitter to the sender's followers through the channels they choose. The way Twitter distributes messages provides a lot of flexibilities for users to send and receive messages. Twitter has almost doubled its users in 2009 and is estimated to have 70M users as of Jan. 2010 [12].

OSNs are built on top of social relationships or friendships while the definition of friendship varies from one OSN to another. The relationship between two users on `Twitter.com` is *following*. More specifically, if Alice is following Bob, Alice is able to receive all Bob's (public) messages. In this case, Alice is a *follower* of Bob and Bob is a *followee* of Alice.

In June 2009, the news of Michael Jackson's death spread all over the world. Many online social networks (OSNs) were flooded by messages/news related to this breaking event. We started collecting related messages from `Twitter.com` on June 27th, 2009, two days after the tragedy. The collection process was performed as follows.

1. Send the following two queries to Twitter's search service [3]: "Michael Jackson" and "MJ." Twitter allows us to send up to 20K queries per hour per whitelisted IP.
2. Parse the returned results (messages) to find the users who posted these messages. From June 27th, 2009 to August 16th, 2009, $716,588$ users were collected.
3. Crawl the profiles and follower/followee lists of the users we found in Step 2. We found $683,160$ users with valid and unprotected profiles [4].
4. Crawl all the messages these users have posted. Twitter allows us to collect up to $3,201$ most recent messages for each user. We crawled $58,500,320$ messages all together, i.e., each user posted 81 messages on average. Step 3 and 4 took us about 17 days.

Among all the messages we crawled, we select the tweets containing "Michael Jackson" or "MJ" as MJ related messages. To filter the false positives introduced by the query "MJ," which mostly are generated by URL shortening services such as `tinyurl.com` and `bit.ly`, we require that for each message there is no leading character before "MJ." After removing the noise, we found $549,667$ MJ related messages (about 1% of the entire data set we crawled) posted by $305,035$ users. $548,102$ messages were posted after Jun 25, 2009 12:21:04 p.m. (PST), when the 911 call was made to save Michael Jackson's life. We assume that these messages are related to the breaking event.

Thus we collected two datasets. One is the entire data set for analyzing the overall (or average) message propagation patterns and the other, referred as the "MJ" dataset, is used for breaking news propagation analysis.

In addition, to get the social graph for computing propagation distance, we crawled $61,907,902$ nodes with $1,520,959,208$ links, which covers 88% of the entire Twitter network according to the estimation given by Moore [12].

2.1 Message format

A message (status or tweet) provided by Twitter API contains the following fields.

[3] `http://search.twitter.com`

[4] A protected profile is not available to public thus our crawler can not crawl it.

- User_id: Unique identifier for the user who posted this message.
- Id: The message ID, which is unique for messages posted by the same user. Two messages posted by different users may share the same message ID.
- Text: The content of the message, up to 140 characters.
- Created_at: The creation time for this message.
- Source: The Twitter client software which was used to post the message.
- Truncated: Messages having more than 140 chars will be truncated by Twitter and have this field set. None of the 58.5M messages are truncated.
- In_reply_to_status_id: The message ID which this message replies to.
- In_reply_to_user_id: The user ID which this message replies to.
- Favorited: Indicating if the message is a favorited one, which actually is rarely used. 261 out of the 58.5M messages are marked as favorited.

2.2 Popular sources of messages

Among the 58.5M messages we crawled, the top three sources (Twitter clients) and the fractions of messages they contribute are as follows.

1. `Twitter.com` (38.0%);
2. TweetDeck (11.8%), a popular Twitter client;
3. TwitterFeed (5.7%), a service which enables users to send messages by posting to their blogs.

Besides, five major mobile Twitter clients, including Tweetie, UberTwitter, and TwitterFon, contribute to 21.1% together. In our data, there are 112 sources via which at least $10,000$ messages were sent.

For the MJ data set, 58.5% messages were posted through the Twitter website, and the five top mobile clients contributed only 11.0%. Further investigation in the data suggests that users are more likely to use mobile devices for sharing events related to themselves, for example, where they are or what they are doing.

2.3 Popular tweets

We find that popular tweets, i.e., the tweets being sent by many users, fall into two categories.

- Internet slang or short phrases, e.g., the top 10 most popular tweets in our data set shown as Table 1.
- Automatically generated messages, most of which are sent by virus/worms or online services for advertisement or spam purposes. For example, there is a website which allows users to customize the background of their Twitter pages. After a user enables this service, it automatically posts an advertisement on behalf of the user to his/her followers. There are also many messages advertising for websites which claim to boost the number of followers for users. Some of them are in fact performing phishing or “bait and switch” attacks. Extensive worms/epidemiology studies have been devoted to the propagation of such messages therefore we do not look further into it in this paper.

Table 1. Top 10 most popular tweets

Number of messages	Content
74,141	“LOL”
25,041	“:)”
19,455	“Thanks!”
18,005	“LMAO”
12,535	“” (Empty)
11,917	“LOL!”
10,295	“:D”
9,881	“:(”
9,861	“Thank you!”
9,207	“Thanks”

3 Measuring message propagation

On `Twitter.com`, a user may reply to a message sent by another user. Such reply message has the fields of “in_reply_to_status_id” and “in_reply_to_user_id” thus allows us to match it with the message it is replying to and further identify the message flow.

28.1% messages in our dataset are replies. Huberman *et al.* [6] reported a similar portion of replies in their dataset (25.4%). The MJ dataset has a much smaller portion of replies, 9.4%, which suggests that it is more common for people to express their own feelings or opinions about MJ’s death instead of discussing it with their friends.

Message flows are identified as follows.

1. Sort all replies according to their timestamps with the earliest message on the top.
2. Walk through the sorted message list from the top to the bottom. For each message i, assuming j is the message which i replies to, i.e., j is i’s parent. If there exists a tree which has j, make i a child of j. Otherwise, create a new tree with j as its root, and attach i to j.
3. When we reach the end of the message list, output all message flows we have discovered.

We found $1,538,698$ message flows in the entire data set and each flow has 10.7 messages on average. Now we can analyze how these messages propagate by answering the following questions.

- How far away is a message propagated?
- How fast is a message replied?
- How long does a message flow last?

3.1 How far away is a message propagated?

On `Flickr.com`, a popular photo sharing website, users may choose to become a “fan” of a photo. Cha *et al.* [3] observed that even for popular photos, only 19%

of fans are more than 2 hops away from the uploaders. On Twitter, however, we find that 37.1% message flows spread more than 3 hops away from the originators, shown as Table 2. The large propagation distance indicates that Twitter is a better medium for propagating information. Meanwhile, text messages are probably easier to propagate than photos. The longest message flow in our data set consists of 1,555 replies made by 1,512 users.

Table 2. How far a message is propagated

Hops propagated	1	2	3	≥ 4
Fraction of messages	45.1%	9.4%	8.4%	37.1%

3.2 How fast is a message replied?

25% replies were made within 67 seconds and 75% were made within 16.5 minutes. This indicates that the communications on `Twitter.com` are mostly real time. The mean time to reply a message is 2.9 hours while the median is only 3.5 minutes, which indicates that there exists large delays for some replies. For example, we observed that a message got replied 20 months after it was sent.

3.3 How long does a message flow last?

25% message flows lasted less than two minutes and 75% lasted less than an hour. In other words, most conversations ended quickly. There is also a huge gap between its mean (6 hours) and median (8.9 minutes), indicating some outliers with long life time.

Therefore we conclude that on Twitter, messages get replied quickly and propagated relatively far away, although most conversations last for a short period of time.

4 Measuring the propagation of breaking news

As we have stated before, the MJ dataset has fewer replies (9.4%) than the entire dataset (28.1%). More importantly, all the messages in the MJ dataset are considered to be related to Michael Jackson's death. Therefore we can examine the aggregated propagation patterns of all these messages, which are more interesting than those of individual message flows.

Shown as Fig. 1, these messages cover about two months after the tragedy. Initially there were many speculations and shocks, which corresponds to the first spike. The initial spike would be larger if we started our crawling at Day 0 instead of Day 2. There is a spike in Day 3, probably because it is the first day we queried Twitter for entire 24 hours. The largest spike occurred during the day of Michael Jackson's memorial service, July 7th, 2009, i.e., 12 days after his death. Comparing to an average day within our data collection window, there were 10 times more MJ related messages posted at that day. After that spike, the number of messages kept steady between Day 16 and Day 60, although we stopped querying Twitter to find new users (not tweets) around Day 50. This "ripple" effect suggests that breaking news does not always disappear quickly, a counterexample of the "15 minutes of fame" theory. [5]

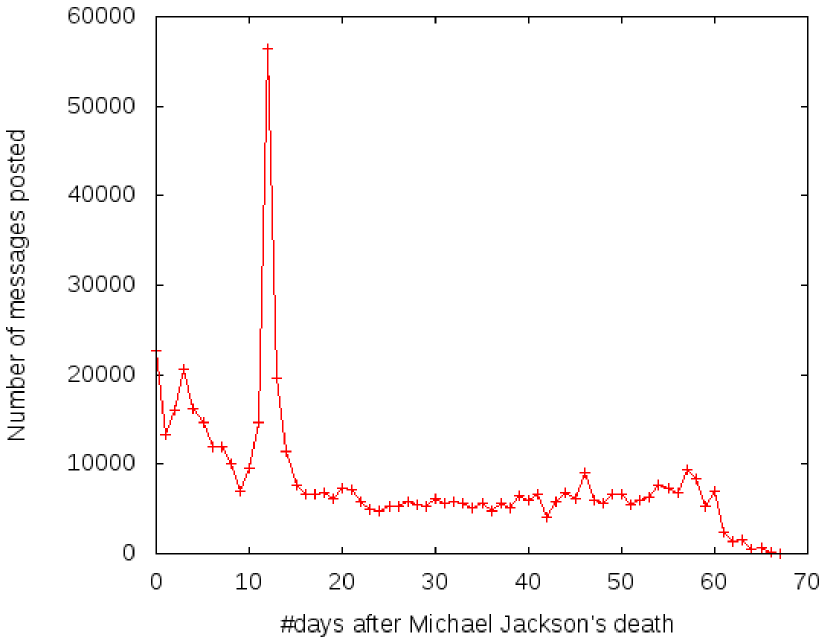

Fig. 1. Number of messages versus days after MJ's death

Propagation distance can not be used here because this message (MJ's death) was widely covered in the real world. There are many originators for this message and it is hard to tell if a user gets it from a followee or not. Thus we evaluate the coverage of the message within the Twitter network. More specifically, if a user posts a message related to MJ's death, all of his/her followers will receive this message, i.e., they are *covered* by this message. As a user may receive MJ related messages multiple times, we only consider the number of unique users who are covered.

Shown as Fig. 2, the number of covered users increases quickly in the first 15 days, with a large jump around the memorial service. Within 70 days about

[5] According to Wikipedia, "15 minutes of fame" is "short-lived, often ephemeral, media publicity or celebrity of an individual or phenomenon."

9M users were covered, roughly 12% of the entire Twitter graph. The number of covered users keeps increasing after the memorial service although with a slower rate. This suggests that these messages are sent by new originators. Examining the differences between initial shocks and afterward thoughts will be an interesting direction for future work.

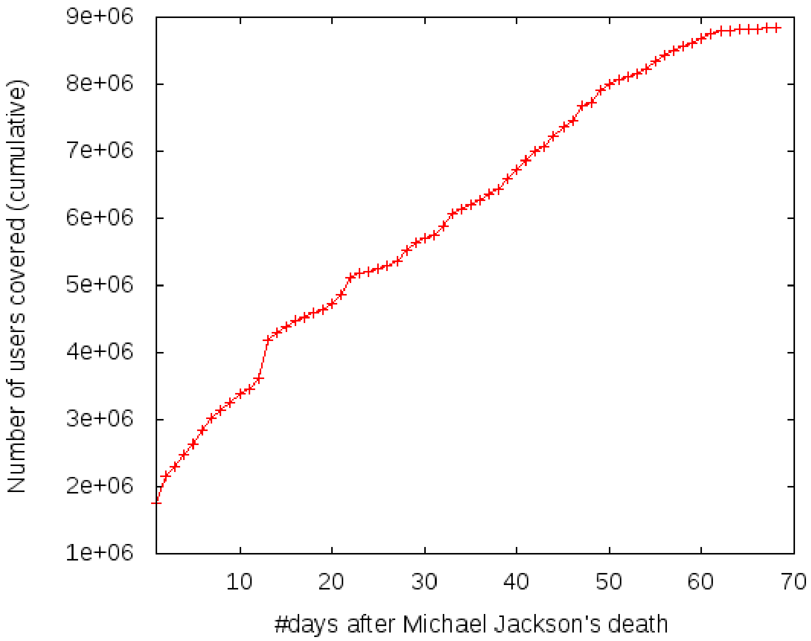

Fig. 2. Number of users covered (cumulative)

To further analyze the message propagation process, we define a *poster* as a Twitter user who posted at least one MJ related message. Fig. 3 shows the cumulative distributions for the number of of covered users and the number of posters. The first 5% posters covered about 20% of the users, which demonstrates the amplification power of OSNs as a medium, i.e., a small number of users may push a message to a large number of users. As more and more users post messages, few new users are covered since most users have already been covered by earlier posters. This is a strong signal of the small-world effect.

5 Measuring social influence

According to Wikipedia, *social influence* "occurs when an individual's thoughts or actions are affected by other people." When it comes to specific fields or

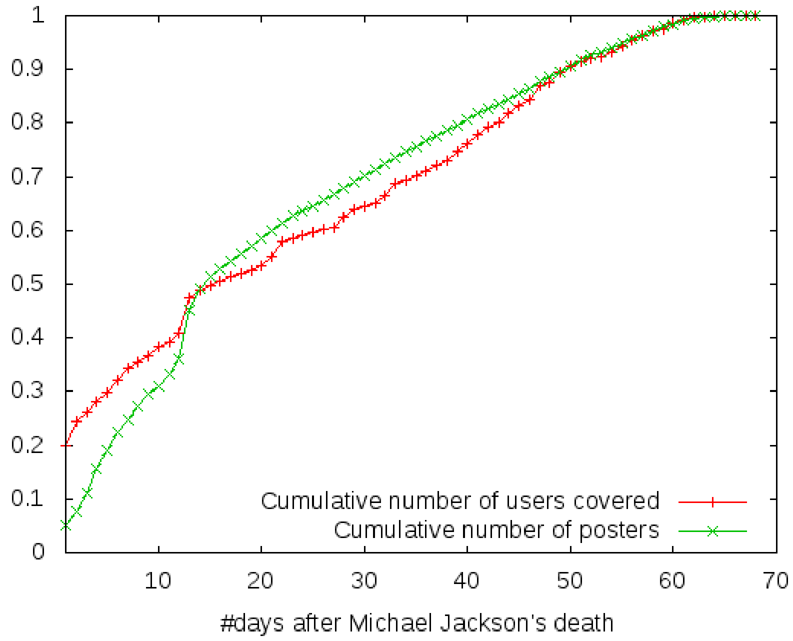

Fig. 3. Covered users versus posters (CDF)

application scenarios, this vague definition needs to be clarified, i.e., what is the *action* and how to determine whether an action is *affected by other people*.

For `Twitter.com`, we examine the following metrics for social influence.

- **Follower influence** (F): The action here is receiving messages (following). The more followers a user has, the more influential this user is. It is also known as *degree influence*, which corresponds to the size of the audience a user has.
- **Reply influence** (R): The action here is replying. The more replies a user receives, the more influential he/she is. This influence can be quantified by the number of replies (R_M) the user receives or the number of users who make these replies (R_U). R_U is less biased towards the users who make lots of replies. We evaluate both of them in this paper.
- **ReTweet influence** (RT): The action here is retweeting. Similarly, the more frequently the user's messages are retweeted by others, the more influential this user is. This can also be quantified by the number of retweets (RT_M) or the number of users who retweet (RT_U).

5.1 How to evaluate social influence metrics

A social influence metric gives a score for each user being measured. For example, F gives the number of followers of a user. The score itself does not tell us how influential a user is, while when we compare the scores of two users, the one with larger score is likely to be more influential. Thus what really matters is the relative order of the users according to the metric, which can be represented by a list of users ranked by the metric. With ranking lists, it is possible to compare social influence given by two metrics which capture different actions, for example, the number of followers (F) and the number of replies (R_M).

To evaluate how stable these metrics are, we split the dataset into two sets according to when the messages were posted and compare their ranking lists. If with a certain social influence metric, the ranking lists of these two sets are close to each other, this metric is relatively stable.

Secondly, we also compare the ranking lists given by different metrics to examine their correlations. If two ranking lists are close to each other, these two metrics are similar. We want to identify the metrics which are similar and those which are unique.

To quantify the difference between two ranking lists, the following two measures are widely used. Given n different items and their two permutations (ranking lists) x and y,

- Spearman's rank correlation coefficient (ρ) [13]:

$$\rho = 1 - \frac{6\sum(x_i - y_i)^2}{n(n^2 - 1)} \tag{1}$$

where x_i and y_i are the ranks of item i in list x and y respectively.
- Kendall Tau rank correlation coefficient (τ) [7]:

$$\tau = \frac{n_c - n_d}{0.5n(n-1)} \tag{2}$$

where n_c is the number of concordant pairs and n_d is the number of discordant pairs. Given two items i and j, if $x_i > x_j$ and $y_i > y_j$ (or $x_i < x_j$ and $y_i < y_j$), i and j are a concordant pair, otherwise i and j are a discordant pair.

Both ρ and τ are inside the interval $[-1, 1]$ and assume the value:

- -1 if a ranking list is the reverse of the other;
- 0 if the ranking lists are completely independent;
- 1 if the ranking lists are the same.

Furthermore, we are interested in the top k most influential users. It is problematic to compare users with small influentials. On OSNs many users are likely to have small influence, i.e., the long tail effect. Small variance may change their rankings a lot. For example, lots of users have only a couple of replies, getting one more reply may improve a user's rank by several thousands or more. In this paper we evaluate the ranking distance with top $1,000$, $5,000$, and $10,000$ most influential users.

When comparing the top k items of two ranking lists, there may exist $x_i < k$ while $y_i > k$ hence the two top k lists are not two permutations of the same set of items anymore. To resolve this problem, we perform a matching process to compute $\rho_k(x, y)$ and $\tau_k(x, y)$ as follows.

1. Select the top k items from list x, denoted as $\{u_i\}$, $i = 1, 2, \dots, k$;
2. For each u_i, get its ranking in list y, denoted as y_i.
3. Let y'_i be the rank of u_i in the list of $\{y_i\}$.
4. Compute the rank distance between the top k lists of x and y'.

This matching process generates two lists with the same set of k items while it is asymmetrical to list x and y. In other words, by switching the processing order of x and y, we may get different set of k items. To get symmetrical results, $(\rho_k(x, y) + \rho_k(y, x))/2$ and $(\tau_k(x, y) + \tau_k(y, x))/2$ are reported as ρ and τ respectively in the rest of this paper unless explicitly stated.

We also report the overlap between the top k lists, i.e., the number of items which rank top k in both lists. The overlap is normalized by k for comparison between different ks.

5.2 The change of social influence over time

To show the change of social influence over time, we sort the messages according to their timestamps and split them into two by selecting the first 50% messages as Dataset I and the second 50% messages as Dataset II. For each dataset we compute ρ, τ, and overlap for their top k lists, shown as Table 3. The results on F is not available because we only have one snapshot of `Twitter.com` thus do not know the number of friends of each user before and after a certain time.

Table 3. The change of social influence over time

Metric	Top 1,000			Top 5,000			Top 10,000		
	ρ	τ	Overlap	ρ	τ	Overlap	ρ	τ	Overlap
R_U	0.722	0.526	**73.9%**	0.688	0.496	**72.8%**	0.669	0.470	**71.9%**
R_M	0.614	0.433	57.8%	0.541	0.374	54.8%	0.521	0.358	55.7%
RT_U	**0.802**	**0.997**	54.1%	**0.778**	**0.999**	53.0%	**0.777**	**1.000**	53.2%
RT_M	0.634	0.451	65.6%	0.608	0.419	62.5%	0.574	0.393	61.1%

The maximum of each column is highlighted. For ρ, τ, and overlap, we always have the following order:

$$R_U > RT_M > R_M$$

The overlap of two lists ranked by R_U is the largest among all 4 metrics being evaluated here, i.e., the set of users with most repliers (not replies) is most stable. Its corresponding ρ and τ are the second largest (only smaller than RT_U), implying that R_U is a reliable metric. Huberman *et al.* [6] found that R_U is also a good estimator for the number of messages a user posts. Hence R_U seems to be a promising candidate for measuring social influence.

The large ρ and τ for RT_U indicates that the action of *retweeting* is more consistent over the time compared to other actions. As reported by Cha *et al.* [1], the most retweeted users were content aggregation services, businessmen and news sites. Our results suggest that such users are likely to keep their influence stable. On the other hand, users with large number of replies are mostly celebrities, whose influence fluctuates as R_M being the most unstable metric.

Except RT_U, both ρ and τ decrease as the list gets longer (i.e., k gets larger). This validates the motivation to use the top k lists instead of the entire list, i.e., the top k users are more stable than the users with lower ranks.

5.3 Assessing influence: messages versus users

Both the number of messages and the number of users can be used to assess the influence. To see the difference, we compare the top k lists generated by the same influence metric with different assessments.

Table 4. Rank distance between RT_M and RT_U

Metric	Top $1,000$	Top $5,000$	Top $10,000$
ρ	0.833	0.817	0.795
τ	0.655	0.628	0.604
Overlap	82.5%	81.6%	82.1%

Shown as Table 4, the difference between RT_M and RT_U is small, and more importantly, it does not get much larger as k increases.

Table 5. Rank distance between R_M and R_U

Metric	Top $1,000$	Top $5,000$	Top $10,000$
ρ	0.749	0.574	0.533
τ	0.571	0.409	0.369
Overlap	61.1%	55.5%	57.2%

R_M and R_U, on the other hand, are relatively far away from each other. We find that the gap between them gets larger as k increases, shown as Table 5. This suggests that the distribution of replies is highly skewed, i.e., some users make many more replies than others. Checking the pair of users involved in a reply, we find that 47% replies are between the top 10% most frequently communicated user pairs.

Hence we conclude that for retweet influence, either the number of messages (RT_M) or users (RT_U) can be used whereas for reply influence, it would be prudent to evaluate the choice of R_M or R_U with the specific application scenario.

5.4 Correlations between different influences

In this section, we examine the correlations between these social influence metrics. More specifically, let L_i and L_j be the ranking lists generated by social influence metric i and j respectively, and $D(i, j)$ be the distance between L_i and L_j where $i, j \in \{F, R_U, R_M, RT_U, RT_M\}$ and $D \in \{\rho, \tau, \text{overlap}\}$. For each metric i, we compute $\eta = \sum_{j \neq i} D(i, j)$, i.e., summation of the distance between the ranked list generated by metric i and the ranking lists generated by the other 4 metrics. A large η means that the corresponding metric i is close to the rest of 4 metrics.

To our surprise, for all D, η of 5 metrics follows the same order, i.e.,

$$R_M > R_U > RT_U > RT_M > F$$

This order also holds for all ks we have tested. The first four metrics are close to each other with F being far away, which shows that F is a poor estimator for other social influences. Meanwhile, it suggests future social influence studies that analysis on any of the first four metrics (R_M, R_U, RT_U, and RT_M) might apply to the other three but F is likely to be an outlier thus needs to be examined carefully.

One promising application for such correlation analysis is to reduce the computation cost for expensive social influence metrics, such as the ones introduced in Section 6.3. Although some work has been proposed to compute social influences efficiently such as [15], some metrics are still expensive for large scale OSNs. Fortunately in many cases, only top influential users (i.e., users with large x) are concerned. If we can identify two close metrics x and y, where x is the expensive one we are interested in and y is a simple metric such as R_U, then y can be used as an estimator for x. More specifically, as users with large x probably have large y too, we may select top k users with metric y and compute metric x for these k users. As long as x and y are close to each other, we will not miss many users with large x.

6 Related work

There is a large body of literature on information propagation and social influence, most of which looks at traditional social networks. Here we limit our discussion to OSN related models and measurements.

6.1 Content generation

Guo *et al.* [5] measured how user content is generated on three popular OSNs (the name of these OSNs are not revealed in their paper). They found that the number of active posters (i.e., users who post a lot of messages) is much larger than that of a power-law distribution, which suggests that the network is not dominated by a small number of users.

The Web Ecology Project [8] examined 1.8M tweets about Michale Jackson's death and show how users express emotion on Twitter. This work focuses on content and semantic analysis and provides us with insights for how users encode emotional content with a small number of words. Our analysis in Section 4 complements their work.

6.2 Information propagation

Huberman *et al.* [6] defines *friends* of a user A as the set of users with whom A has exchanged directed messages. By examining 300K Twitter users, they observed that the correlation between the number of messages and the number of friends is larger than the correlation between the number of messages and the number of followers/followees. The friend network is shown to be much sparser than the follower/followee network, which may alleviate the scalability problem for OSN studies.

Kwak *et al.* [9] examined the propagation of 106M retweets and found that up to 1,000 followers, the average number of users being covered by a retweet is not affected by the number of followers the originator has. Another interesting observation is that although the median time to retweet for the first hop is large (about an hour), that of the second hop is quite small (less than 10 minutes).

Lerman and Ghosh [10] measured how popular news spread on Digg (a social news aggregator) and Twitter. Voting on Digg and retweeting on Twitter are used to identify related messages. The number of votes/retweets is shown to increase quickly within a short period of time and saturates after about a day. The distribution of story sizes (number of votes/retweets a story has) is approximately a normal distribution instead of the widely observed power-law distribution. Their observations are based on the aggregated patterns of many stories thus does not focus on breaking news. A key difference is that breaking news such as Michael Jackson's death lasts much longer time and covers many more users. Lerman and Hogg [11] further developed a model to estimate the popularity of news. It is not clear if their model is able to predict the spread pattern for breaking news as such stories tend to be outliers compared to average popular stories.

6.3 Social influence

Kwak *et al.* [9] ranked Twitter users by the number of follows and by PageRank and found that these two rankings are similar. The ranking by retweets differs from the previous two rankings, which is similar to our findings in Section 5.1. A variation of τ is used to compare these rankings. Their comparison is preliminary as only these two influences are considered. They also did not look at the change of influences over time.

Weng *et al* [14] examined top 1,000 Singapore-based twitters plus their followers and followees. They found the following relationship is highly symmetrical and developed a topic-sensitive PageRank-like influence measure, TwitterRank. The basic idea is that given a topic, the social influence of a twitter is the sum of the social influences of his/her followers. It would be interesting to extend TwitterRank by substituting following with replying/retweeting.

Ghosh and Lerman [4] proposed a model to predict the number of votes a user's post generates on Digg by considering graph properties of the network, which decide how news can be delivered. Voting influence on Digg is similar to replying and retweeting influence on Twitter. Applying their model to Twitter would be another direction for future work.

Cha *et al.* [1] investigated three social influence metrics (in-degree, retweets, and mentions) across topics and time with the messages of 6M Twitter users, which is probably the closest work to this paper. They found: 1) Popular users with large number of followers may not necessarily get more retweets or mentions. 2) Users can hold influences over multiple topics. 3) Limiting tweets to a single topic may help users gain influences.

The differences between their work and our analysis in Section 5 are as follows.

- Cha *et al.* [1] did not have similar analysis as Section 5.3 and 5.4.
- We discuss in detail how to evaluate social influence metrics, such as the matching process to compare the top k lists.
- Cha *et al.* [1] only considered ρ and overlap.

7 Conclusions

This paper presents a measurement study of 58M messages sent by 700K users on `Twitter.com`.

First, by examining message flows, we find that replies arrive quickly and a significant portion of messages propagate far away from the originator, i.e., the discussions are not restricted to his/her followers. Further, we show how messages related to Michael Jackson's death spread through the network, which demonstrates the power of Twitter as a social medium.

Secondly, we evaluate different social influences by examining how they change over time, how to assess them, and how they correlate with each other.

The discussions in this paper reveal the complications we have to deal with to characterize message propagation and evaluate social influence. We believe that the analysis and measurements presented here pave the way for systematic measurements and investigations of OSNs.

Acknowledgements

We thank `Twitter.com` for whitelisting our crawlers during the past year. Min Liu provided insights to some of our observations. Dr. Norm Matloff made valuable comments on an earlier version of this paper. We are also grateful to anonymous reviewers for the suggestion of using tree/forest to model message flows and pointers to related work.

References

1. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.P.: Measuring user influence in Twitter: The million follower fallacy. In: ICWSM'10: Proceedings of International AAAI Conference on Weblogs and Social Media (2010)
2. Cha, M., Mislove, A., Adams, B., Gummadi, K.P.: Characterizing social cascades in Flickr. In: WOSP '08: Proceedings of the First Workshop on Online Social Networks. pp. 13–18 (2008)
3. Cha, M., Mislove, A., Gummadi, K.P.: A measurement-driven analysis of information propagation in the Flickr social network. In: WWW '09: Proceedings of the 18th International Conference on World Wide Web. pp. 721–730 (2009)
4. Ghosh, R., Lerman, K.: Predicting influential users in online social networks. In: SNA-KDD: Proceedings of KDD Workshop on Social Network Analysis (2010)
5. Guo, L., Tan, E., Chen, S., Zhang, X., Zhao, Y.E.: Analyzing patterns of user content generation in online social networks. In: KDD '09: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 369–378 (2009)
6. Huberman, B., Romero, D., Wu, F.: Social networks that matter: Twitter under the microscope. First Monday 14(1–5) (Jan 2009)
7. Kendall, M.G.: A new measure of rank correlation. Biometrika 30(1/2), 81–93 (June 1938)
8. Kim, E., Gilbert, S., Edwards, M.J., Graeff, E.: Detecting sadness in 140 characters: Sentiment analysis and mourning Michael Jackson on Twitter. Tech. Rep. 3, Web Ecology Project (2009)
9. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? In: WWW '10: Proceedings of the 19th International Conference on World Wide Web. pp. 591–600 (2010)
10. Lerman, K., Ghosh, R.: Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. In: International AAAI Conference on Weblogs and Social Media (2010)
11. Lerman, K., Hogg, T.: Using a model of social dynamics to predict popularity of news. In: WWW '10: Proceedings of 19th International World Wide Web Conference. pp. 621–630 (2010)
12. Moore, R.J.: New data on Twitter's users and engagement (2010), `http://themetricsystem.rjmetrics.com/2010/01/26/new-data-on-twitters-users-and-engagement/`
13. Spearman, C.: The proof and measurement of association between two things. The American journal of psychology 15, 72–101 (1904)
14. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twitterrank: finding topic-sensitive influential twitterers. In: WSDM '10: Proceedings of the Third ACM International Conference on Web Search and Data Mining. pp. 261–270 (2010)
15. Xie, B., Kumar, A., Ramaswamy, P., Yang, L.T., Agrawal, S.: Social behavior association and influence in social networks. In: Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing. pp. 434–439 (2009)