

Where are my tweeps?: Twitter Usage at Conferences

Edgardo Vega

Department of Computer Science
Virginia Tech, Blacksburg, VA, USA

evega@vt.edu

Ramanujam Parthasarathy

Department of Computer Science
Virginia Tech, Blacksburg, VA, USA

ramanuj@vt.edu

Josette Torres

Department of English
Virginia Tech, Blacksburg, VA, USA

josette@vt.edu

ABSTRACT

The microblogging service Twitter, founded in San Francisco, California in 2006, has become immensely popular in recent times. Apart from serving as a status message update service that provides the answer to the question "What's happening?" posed above the update box on the Web site, it has also served as a platform for developing new connections and a medium for building conversations. Twitter has found its place as an important tool in both professional and academic conference settings. In our study, we aim to analyze how Twitter usage patterns change significantly during a conference. We performed a series of data analysis tasks on a collection of user tweets obtained using the Twitter API from twenty different users for a five-week period. Our findings revealed that there was substantial increase in the Twitter usage during the conference week for the majority of the users.

1. INTRODUCTION

Twitter is a microblogging and social service which allows users to post 140 character updates, called tweets. Twitter was founded in 2006 and has grown to be one of the largest sites on the Internet. What has set Twitter apart in its short but fast development history is how its users have adapted Twitter for use in ways which were never intended by its developers. What was initially considered only a microblogging platform has quickly become a tool used by communities to gather around and discuss events and topics. For example, Twitter has been used as a political tool – as it was used in the last presidential campaign, a tool for gathering public opinions and survey results, a tool to physically and geographically organizing communities – as it is used for flash mobs, and many other uses. Each tweet is 140 characters or less in length; Twitter's simple architecture and interface makes it appear simplistic. Yet, the variety of uses is what makes Twitter so appealing yet challenging to study. Twitter is additionally used as a marketing channel to communicate with larger communities. For example, TechCrunch, a popular technology weblog, publishes links to all their posts on their official Twitter stream. This not only gives additional visibility for the content, but at the same time it helps in initiating conversations and comments from fellow Twitter users. Many other communities and media organizations use Twitter as a broadcasting platform, including CNN, the New York Times, the Washington Post, and Wired, just to name a few.

Initially, Twitter asked the question, "What are you doing?" with the intention of getting the current status update from the user. However, Mischaud [1] found in his study that more than half of the tweets he analyzed had nothing to do with the question. People have started using Twitter in many different ways and

scenarios. As a result, Twitter has changed the question to "What's happening?" as it makes more relevance.

One of the common uses of Twitter is in academic and professional conferences. Prior research has focused on how individuals have used Twitter as a communication platform. In the work of Java et al., boyd et al., and Krishnamurthy et al., they studied groups of Twitter users in an attempt to understand why individuals tweet [2-4]. From this work, the researchers were able to derive standard metrics for measuring a tweeter's behavior. These included measuring the number of tweets, retweets, followers, and others metrics that we plan to adapt for our own study. Some additional relevant findings from these studies are that the number of followers is not sequential, as previously assumed [3]. This indicates that community events, such as a conference, may stimulate different behavior. Similarly, in [2], they found that it was possible to aggregate a community's tweets in such a way to discover community intentions. This work indicates that communities can have an impact on individual behavior.

The discussion above has led us to study the use of Twitter when groups gather for an event, such as South by Southwest interactive (SXSWi) conference. These users share the experience of attending a conference as a community. The before, during, and after behaviors of users in these particular situations - in which we find the use of Twitter to be extremely interesting - has not been studied in depth.

2. BACKGROUND

Twitter.com is a popular microblogging service where the users post status updates, often referred to as "tweets." While Twitter does have a native web interface, a good percentage of the users use third party clients and applications to post updates.

2.1 Conventions and Types

Twitter is being used in many different ways, not merely as just as a status update service. People use Twitter to share recommendations, make new friends, track events, and share pictures, among other uses. The variety of uses is what makes Twitter appealing and also challenging to study. The tweets, or updates, posted on Twitter consist of several different types. There are three different types of Twitter users, as described by Java et al., [2] namely the information sources, friends and information seekers. The "information sources" might be automated news feeds and tend to have a big follower base, while "friends" is a more general category, which includes all the common users. "Information seekers" are those users who generally do not post updates, but tend to keep track of updates posted by other users. While some of the types were present

natively, a good number of these features were added to the service, due to popularity among users. The following are the different types of tweets that one could encounter on a typical Twitter timeline.

1. Regular status updates
2. Replies – Messages in reply to another user (denoted by @ symbol)
3. Retweets – Messages originally posted by another user and then reposted (typically denoted by “RT”)
4. Followers – People that wish to receive updates from an individual user.
5. Following – People that the individual user “follows” and opts to see their updates.
6. Hashtags – A convention of prefixing a # symbol at the beginning of a keyword, making it easy to filter and search tweets.
7. Mentions – Referring to another Twitter user. Denoted by @ symbol and the Twitter handle.
8. Direct Messaging – A feature used for exchanging private messages between two users. Commonly referred to as DM.
9. Messages reposted from other services (location services, RSS feeds, etc.)

3. WHY STUDY TWEETS

Twitter is a huge information repository and provides an ample scope for us to look for varied information. It acts as an information sharing medium [5], a medium for social search [6], a phatic communication medium [7], a personal diary, and much more. Twitter search [8] is turning out to be a valuable tool for finding information, with one of the biggest advantages being real-time results. The search stream brings in content from millions of users and, unlike a normal Web search engine, there is no issue of indexing latency. This enables the user to get opinions on any given subject, not only from their personal network of friends, but also from public users. With location-based filtering, this will bring in more relevance.

3.1 Information Seeking

The users of Twitter generate over 50 million tweets in a single day, accounting for over 600 tweets per second [9]. One big task which might lie in the hands of the user is to find the information which he/she is looking for amidst thousands of tweets in the public timeline but, on the other hand, it is undoubtedly a collective source of intelligence that helps to obtain opinions, ideas and other relevant information with ease [10]. It serves as a medium that enables sharing of information with like-minded people. Haewoon, et al [5] from KAIST, Korea studied the influence of Twitter as an information sharing medium by crawling the entire Twitter site and obtaining a dataset containing 41.7 million user profiles and 106 million tweets, among other information. After ranking Twitter users and analyzing the trending topics and retweets, they have observed that the follower-

following distribution in Twitter has a deviation from that of a human social network.

Twitter facilitates the way in which people use the status message update feature to ask questions. Often referred to as “social search,” it involves finding answers with the help of friends in the social graph or even with unknown human resources. There are many motivations like trust, subjective nature of questions, and human expertise, amongst other reasons for using social search [6]. Over time, Twitter has shifted from a personal status sharing medium to an information sharing and conversational medium. The way we share and find information has changed with the introduction of microblogging networks, such as Twitter, to the Web. A study by Brooks and Churchill [6] involved recording how users search local news, find shopping deals and use Twitter for getting recommendations from members of their social graph. Twitter is like a personal social search service where the recommendations come from people who we know, unlike services like Aadvark (vark.com), which provides anonymous human responses. A more specific utilization in this context can be seeking time sensitive and location based information.

3.2 Tweet Archives

Each individual Twitter user sees updates from the people whom he or she follows on the personal timeline. The number of updates depends on the number of people the user follows; if it is on the higher side, the number of tweets that appear on the personal timeline could be overwhelming and could also cause information overload. For an individual who has specific tasks and information needs, the process of re-finding will be cumbersome. The Web interface of Twitter not only archives personal tweets, but also archives tweets from friends in the network. In April 2010, the U.S. Library of Congress acquired the entire public Twitter archive for preservation and research purposes [11].

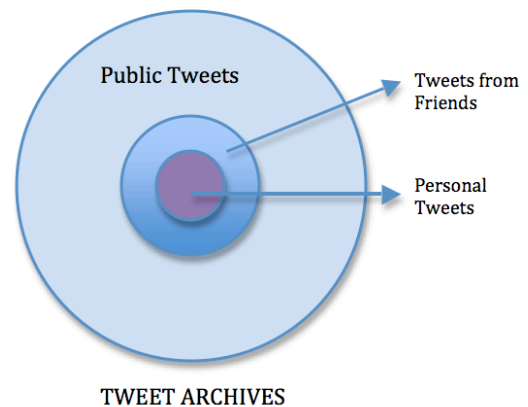


Fig 1. Representation of the different archives of tweets

The Twitter archive can be viewed from three different perspectives. The first one is the larger subset that contains billions of tweets from all the users of the service. The next level is the archive of updates from friends in the network. The final

one is the personal archive, or in other words, the tweets published by the individual user over a period of time. The level of personal information that can be obtained from this data set will increase in the outbound manner, with the individual archive containing more personal references.

There are a few drawbacks in the way of accessing tweet archives, or in other words, re-finding one's own tweets. The current Twitter search model only allows data to be searched from tweets that were published in the past seven days. Although the Twitter profile page has all the tweets, there is no organizational methodology; it is really tedious to look into past tweets. The Twitter API provides access to a greater number of tweets; there are also some external tools and plug-ins that will help to build a browsable and searchable archive of tweets. Also, Twitter's direct messaging capability stores private conversations between users; these are accessible only to the two parties involved in the conversation.

3.3 Conversations and Connections

Twitter also serves as a conversation tool and helps users to make new connections. Honeycutt and Herring [12] analyzed the collaboration and conversation aspects of Twitter by focusing on user replies and the coherence of exchanges. The @ sign used in the tweets facilitated conversations to a great extent. One other beauty of Twitter is that it develops an ambient intimacy among the users and serves as a phatic communication medium [7]. Kevin Makice studied this aspect by utilizing "TwitterSpaces," which are basically large displays streaming tweets from a closely-knit community. It gives a glimpse of the daily happenings of the members and brings in a community feeling. A similar setting is very common in many of the conferences where the tweets from the users attending the session are streamed live on a huge screen and this facilitates a persistent and mobilizing backchannel environment. McNely [13] notes that the low barriers to participation, SMS functionality, real time activity are some of the reasons why microblogging services like Twitter have enhanced the collaborative meaning-making capacity.

Oulasvirta et al [14] discuss the communication genre in microblogging and how it makes ordinary things visible to others. The study involved content analysis and qualitative analysis using a dataset that contained 400000 posts from Jaiku. They focused on categorizing responses for mundane and non-mundane responses and emphasized that microblogs are for individuality and serve as a medium for letting the public know about daily happenings in one's life. Contrastingly, microblogs are being used in many different scenarios in the past two years; particularly in ways in which it was not originally designed to be used, such as campaigning, customer interaction, and education, among others.

4. METHODOLOGY

This study was focused on finding how Twitter usage varied when the user was present in a conference setting. We identified users attending a specific conference. By using the Twitter API, we acquired five weeks' worth of tweets for those users and

performed qualitative and quantitative analysis on the anonymized tweet data.

4.1 Data collection

The first step was to choose the conference for which the study was to be done. We decided on analyzing the tweets from the SXSW Interactive festival, which took place from March 12-16, 2010 in Austin, Texas. The next step was to choose the users on whose tweets the analysis will be performed. We narrowed it down to twenty random users who attended the conference, with the condition that their tweets were publicly available. The users were selected based on the Klout influence score [13-<http://klout.com/kscore/>]. The Klout score takes into account the true reach, which is the numerical count of the engaged audience, the amplification probability, which is ability of the user to initiate conversations, and the network score. The scoring algorithm gives a score ranging between 1 and 100 for any given Twitter user, wherein a higher score on the scale signifies a stronger influence. For our study, we selected users whose scores ranged between 24-84 out of a possible 100 points. This distribution was made to ensure that the users whom we picked for our study had varied influence levels among the other twitter users. For instance, a celebrity user like Bill Gates (@billgates) has a Klout score of 87 of 100, which implies that he is a highly influential user on Twitter.

After deciding on the twenty users, we obtained their tweets using the Twitter API for a five-week time frame. All tweets published by the users, starting two weeks before the conference, during the conference week, and two weeks after the conference were scraped using the Twitter API. For the twenty users put together, we obtained a total of 10,157 tweets. In addition to the actual tweets, we also recorded additional information such as time the tweet was published, device from which the tweet was made, presence of hashtags, inclusion of links, etc.

4.2 Quantitative Analysis

The quantitative tweet analysis was performed to obtain statistical results from the data that we collected. The analysis helped to understand the change in the usage pattern among the users by using the numerical information. The following observations were made by noticing changes in the tweet counts during the five-week period.

- Distribution of tweets
- Inclusion of hashtags
- Distribution of retweets
- Number of replies
- Number of mentions
- Presence of external links

4.3 Qualitative Analysis

Following a grounded theory approach we performed qualitative analysis, and the 10,157 tweets were manually coded. We assigned the individual tweets into fifteen specific categories. We first identified the sentiment of the tweet and marked it as positive, negative or neutral. The second attribute was to see if the tweet was related to the conference or if it was a personal tweet.

Each tweet was also tagged with one or more of the following thirteen additional categories:

1. **Emotion:** Represents a thought, feeling or behavior. E.g., @jennIRL @GabiKachman Glad your books showed up -- enjoy!
2. **General Quotes:** Reference to a common quotation relevant to the context or an own quotation. E.g., "It doesn't make economic sense to have full-time reviewers"
3. **Observational:** A personal observation that mostly results due to some recent activity. E.g., Twitter Launches A New Dynamic Homepage (and it's pretty cool)
4. **Informational:** Provides information that is helpful for other users. Might be a broadcast to all followers or an individual response. E.g., @kitson technical issue. working on that. thanks for the feedback.
5. **Promotion:** Used to promote a session or a personal affiliation. E.g., RT @umairh: i'm going to be doing a keynote interview with @ev from twitter at sxsw. hope to see many of you there
6. **Location:** Makes reference to a physical location. E.g., at LAX - 3 hours til the next flight leaves.... I'm getting there.... slowly
7. **Location Service:** Updated from a location-based service, such as Foursquare or Gowalla. E.g., Maybe you've heard there is a rock show here tonight? (@ Stubb's BBQ w/ 115 others) <http://4sq.com/4xtWYc>
8. **Organization:** Mentions an establishment, like a company or a hotel. E.g., Just arrived at the convention center. Thank goodness four seasons had umbrellas for guests
9. **Recommendations:** Giving recommendations about sessions to attend, places to visit, etc., to other followers. E.g., @AmandaMooney You should goto Nick's Crispy Tacos in on Broadway and Polk. There are a ton of great bars there too.
10. **Retweets:** Re-posting tweets posted by another user. E.g., RT @joestump: Truth is SimpleGeo is developing AR for NES games. <http://bit.ly/bUDtP2>
11. **Reporting Quotes:** Positing quotations from a session. E.g., On web standards: "The minute everyone understands something is important all progress stops" - @cshirky #sxswi
12. **Satisfaction:** Displaying a positive tone about a session, restaurant, etc. E.g., Community funded reporting panel from @digidave alone made this trip worthwhile for me. #sxswi
13. **Reference:** Making a mention of another Twitter user. E.g., Excited to meet @probblogger again at SXSWi

5. FINDINGS

Our findings from the study were two fold. The first one was from the quantitative analysis and second one was from the qualitative analysis.

5.1 Quantitative Results

After performing the quantitative analysis we found a significant variation in the many of the parameters that we analyzed. Almost 40% replies were made during the third week (conference) and it dropped to less than half of that during the next week. This implied that many more conversations happened during the conference time period. The same trend was applicable to retweets and mentions also. The average number of retweets made by a user was around 100 during the conference week whereas it was between 40 and 60 during the other four weeks. Over 2000 mentions were made for all the twenty users put together during the conference week and it was significantly lower for the other weeks. The important metric, which we were interesting in looking at, is the number of tweets made by the user. As indicated by the figure 2, it skyrocketed during the third week, which implied that the user was using Twitter to post updates more often.

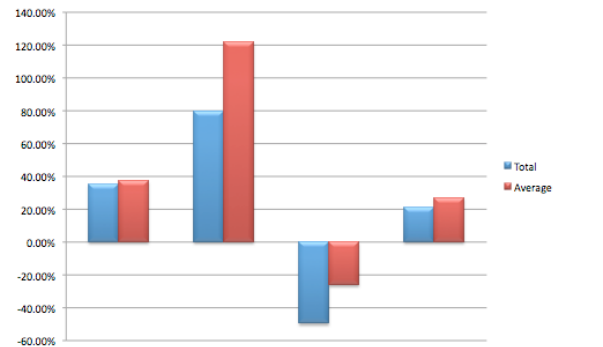


Fig 2. Distribution of the number of tweets

Overall, there was significant increase in all the parameters that we measured and the results are indicated in figure 3

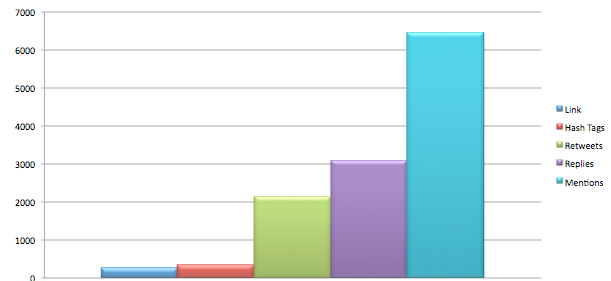


Fig 3. Overall quantitative analysis distribution

5.2 Qualitative Results

In the qualitative analysis part, we manually tagged over a 10000 tweets into fifteen different categories. This was performed to see the context in which the tweet was made. While measuring the sentiment we found that majority of the tweets were neutral and only a few were made either in a positive or negative sentiment (fig.4)

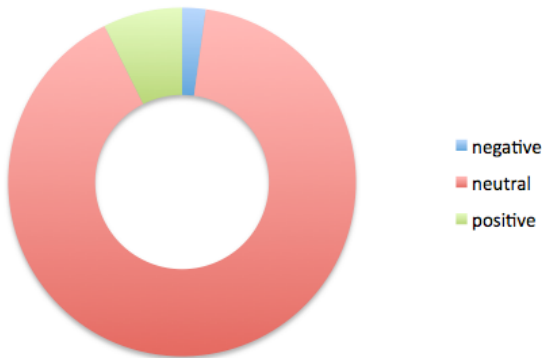
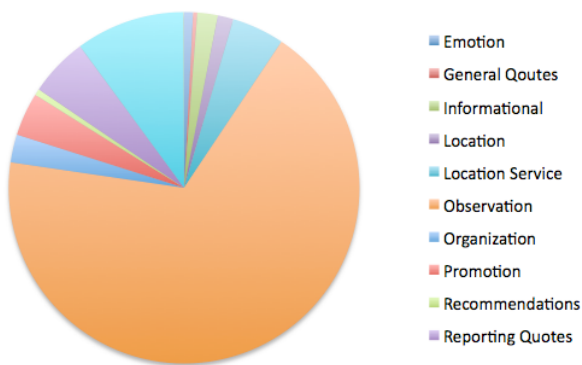


Fig 4. Sentiment of the tweets

The major part in the qualitative analysis was to figure out how the tweets fit in the thirteen pre-defined categories, which we allotted. We found that 68% of the tweets were observations of the happenings. Retweets were next in line and they constituted 10% of the total tweets made. Location based services and reporting quotes constituted 5% each of the total volume. Figure 5 shows the complete distribution of the qualitative analysis.



It was also worthy to note that only 21.8% of the tweets came from the web interface of Twitter. The rest of the tweets were made from third party clients like Tweetie, Tweetdeck etc. We were not able to see if it was published from a mobile device as the API doesn't provide complete information on this front. After analysis of all the data, we concluded that Twitter usage at conferences is significantly more elevated than typical usage.

6. FUTURE WORK

Future research on this topic should explore whether elevated Twitter use is consistent across conferences in different disciplines, such as the arts and humanities, government, business, and so on. More studies including different conference types should be conducted. Also, interviews with Twitter users who attend conferences should be conducted so as to gain a better understanding of individual Twitter usage.

7. ACKNOWLEDGMENTS

Our thanks to Professor Manuel Pérez-Quiñones of the Virginia Tech Department of Computer Science and Professor Carlos Evia of the Virginia Tech Department of English for their assistance in the completion of this research. Thanks to all our classmates from the Personal Information Management class at Virginia Tech for providing feedback on our work. We also thank Easy Chair Coffee shop, Blacksburg for being the perfect environment for all our team meetings.

8. REFERENCES

- [1] Mischaud, E. (2007). Twitter: Expressions of the whole self. An investigation into user appropriation of a web-based communications platform. London: Media@lse. Retrieved May 20, 2008 from http://www.lse.ac.uk/collections/media@lse/mediaWorkingPapers/MScDissertationSeries/Mischaud_final.pdf
- [2] Boyd, Danah, Scott Golder, and Gilad Lotan. 2010. "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter." *HICSS-43. IEEE: Kauai, HI, January 6.*
- [3] Java, A., Song, X., Finin, T., and Tseng, B. 2007. Why we twitter: understanding microblogging usage and communities. In *SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis. ACM, 56-65.*
- [4] Krishnamurthy, B., Gill, P., and Arlitt, M. 2008. A few chirps about twitter. In *Workshop on online Social Networks. ACM, 19-24.*
- [5] Kwak, Haewoon, Lee, Changhyun, Park, Hosung, and Moon, Sue. (2010). What is Twitter, a Social Network or a News Media?. *The 19th World-Wide Web (WWW) Conference.* Raleigh, North Carolina. (Conference paper)
- [6] Andrew L Brooks, Elizabeth Churchill, Tune In, Tweet on, and Twit out: Information snacking on Twitter, *CHI 2010 Workshop on Microblogging*
- [7] Makice, Kevin. (2009). Phatics and the design of community. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems.* (pp. 3133-3136). (conference paper)
- [8] Gene Golovchinsky, Miles Efron, Making sense of Twitter Search, *CHI 2010 Workshop on Microblogging*
- [9] Measuring Tweets, <http://blog.twitter.com/2010/02/measuring-tweets.html>

- [10] Cheong, Mark, Lee Vincent, Integrating web-based intelligence retrieval and decision-making from the twitter trends knowledge base, 2009, *Proceeding of the 2nd ACM workshop on Social web search and mining*
- [11] Tweet Preservation , <http://blog.twitter.com/2010/04/tweet-preservation.html>
- [12] Honeycutt, C. and S. Herring. Beyond Microblogging: Conversation and Collaboration in Twitter. *Proc 42nd HICSS*, IEEE Press (2009).
- [13] McNely Brian, Backchannel persistence and collaborative meaning-making, Proceedings of the 27th ACM international conference on Design of communication , Bloomington, Indiana, USA
- [14] Oulasvirta, Antti, Lehtonen, Esko, Kurvinen, Esko, and Raento, Mika. (2009). Making the ordinary visible in microblogs. *Personal and Ubiquitous Computing*. Special issue on Social Interaction and Mundane Technologies. (journal article)
- [15] Spector, A. Z. 1989. Achieving application requirements. In *Distributed Systems*, S. Mullender, Ed. ACM Press Frontier Series. ACM, New York, NY, 19-33. DOI=<http://doi.acm.org/10.1145/90417.90738>.