

# Monitoring Academic Conferences: Real-time Visualization and Retrospective Analysis of Backchannel Conversations

Awalin Sopan

Department of Computer Science  
University of Maryland  
Email: awalin@cs.umd.edu

PJ Rey

Department of Sociology  
University of Maryland  
Email: pjrey@umd.edu

Brian Butler

College of Information Studies  
University of Maryland  
Email: bsbutler@umd.edu

Ben Shneiderman

Department of Computer Science  
University of Maryland  
Email: ben@cs.umd.edu

*Abstract*—Social-media-supported academic conferences are becoming increasingly global as people anywhere can participate actively through backchannel conversation. It can be challenging for the conference organizers to integrate the use of social media, to take advantage of the connections between backchannel and front stage, and to encourage the participants to be a part of the broader discussion occurring through social media. As academic conferences are different in nature, specialized tools and methods are needed to analyze the vast amount of digital data generated through the backchannel conversation, which can offer key insights on best practices. In this paper we present our two fold contribution to enable organizers to gain such insights. First, we introduce Conference Monitor (CM), a real time web-based tweet visualization dashboard to monitor the backchannel conversation during academic conferences. We demonstrate the features of CM, which are designed to help monitor academic conferences and its application during the conference Theorizing the Web 2012 (TtW12). Its real time visualizations helped identify the popular sessions, the active and important participants and trending topics during the conference. Second, we report on our retrospective analysis of the tweets about the TtW12 conference and the conference-related follower-networks of its participants. The 4828 tweets from 593 participants resulted in 8.14 tweets per participant. The 1591 new follower-relations created among the participants during the conference confirmed the overall high volume of new connections created during academic conferences. We also observed that on average a speaker got more new followers than a non-speaker. A few remote participants also gained comparatively large number of new followers due to the content of their tweets and their perceived importance to the conference followers. There was a positive correlation between the number of new followers of a participant and the number of people who mentioned him/her. The analysis of the tweets suggested that remote participants had a significant level of participation in the backchannel and live streaming helped them to be more engaged.

## I. INTRODUCTION

In many fields, social media (e.g., Twitter and blogs) communication has become an important aspect of the typical conference experience. Prior to the conference, organizers can provide important logistical details and up-to-minute reports on schedule changes through social media. Participants also use social media to document and actively discuss presentations (often interacting with the presenters themselves). Finally, after a conference, bloggers often review the activities

and offer important insights on how future conferences may be improved. The conversation occurring on digital media before, during, and after a conference is frequently referred to as ‘backchannel’. Backchannel communications help the conference participants make new connections and may stimulate conversation between peers. It can foster relationships between the people attending in person and the people who are following it remotely. Nonetheless, as conferences generally have short durations, the backchannel data must be immediately accessible in an organized fashion to be useful for the organizers. Real-time feedback can enable them to intervene when problems become apparent and to reinforce desirable behaviors. For example, conference organizers may wish to identify active and influential participants who may be reporting on the conference to a broader public. By reacting in a timely manner to concerns raised by such influential participants who often give voice to broader sentiments, organizers may be able to improve the overall impression of the conference. Spikes in activity may connote key events or issues. Organizers may observe the activity patterns to predict when interventions may be most effective and how discussion evolves throughout the conference. Our goal was to design a tool to help conference organizers better integrate social media into future conferences. With this goal in mind, we worked closely with the organizers of the Theorizing the Web 2012 (TtW12) conference, identified their needs during and after the conference, and developed the Conference Monitor (CM) web app. TtW12 was unique in that the organizers (one of whom is a second author to this paper) explicitly set forth to experiment with how social media could be used to improve the conference experience. Therefore, our collaboration with the TtW12 organizers helped us make the design decision for CM on what data to present for monitoring the conference and how to present them from an organizer’s point of view. The study led to the following contributions:

- 1) **Development of Conference Monitor (CM):** We developed a web app for real time visualization of the backchannel participation during a conference to address the organizational needs. CM is specifically de-

signed to capture the characteristics that are specific to conferences. The novelty of CM lies in its added utilities targeted for conference organizers, its support for annotation, comparisons of participants, comparisons of sessions and an easy-to-understand overview of the conversation. Its session-based and role-based filtering help both organizers and general users understand the activities regarding a particular session of interest or particular group of participants. It enables the organizers to annotate specific time points during the conference. Using CM's real time visualizations, TtW12 organizers identified the active people, the popular sessions, the recurring topics and the pace of tweets during the conference. They could also mark specific time points of interest.

- 2) **Retrospective analysis of the collected data:** We analyzed the collected data during TtW12, which provided detailed insights regarding the social network, the backchannel conversation content (retweets, URLs, etc.), and the participation of people during the conference. Conferences help create new connections among the participants but the growth of a social network due to a conference was not studied before. In our analysis, for the first time, we segregated the social network that is formed due to and during the conference and analyzed the evolution of the networks of the participants. We hypothesized about the network growth among the people who tweeted about TtW12 using the visualizations of CM and later tested those hypotheses with the whole dataset.

## II. RELATED WORK

Analyzing online interactions and information sharing during popular events has opened a broad spectrum of research questions. McCarthy et al. [1] discussed the advantages and disadvantages of having backchannels during an ongoing event. This vast data of backchannel conversation can be understood if aggregated properly and made available in a way easier for domain experts to analyze. Tweetgeist [2] attempts to analyze the structure of broadcast events by visualizing the relevant tweet time line, though time line alone is not sufficient to understand an event as a whole. TwitInfo [3] leverages the understanding of participation in micro-blogging during an event by showing a map view and sentiment view along with the time line. Diakopoulos et al. [4] built a system to present tweets in the context of journalism. Another notable system is the visual backchannel [5] where the content flow is shown using a stream graph view. However these systems presenting overviews of the tweets do not attempt to identify the influential or important participants, or do not address the infrastructural features of academic conferences; they are targeted towards general users and general events. Conference Monitor, on the other hand, focuses on monitoring conferences from the organizers' point of view so they can take action during the conference and see the effects. Ebner et al. showed that tweets have limited ability to convey

meaningful information to the remote participants [6]. But again, they could not distinguish the in-person participants from the remote ones. Remote participants can help the event reach a broader audience. Starbird et al. showed in [7] that a substantial proportion of the people tweeting updates on the 2011 Egyptian Revolution were not on the ground in Cairo. Similarly, micro-blogging during natural disasters can bring together both people on the ground and people who are not, and create a social network of people providing them situational awareness [8]. The followers-networks and mentions-networks expand within a very short time during events. Since network growth (followers) for in-person and remote participants during an event might not follow the same pattern, we aim to analyze that fact in our study.

While in-person participants and remote participants are important groups to compare, even among the in-person participants, we can find people with different roles [9]. The organizers play an influential role by broadcasting information about their event, but identifying people with influential online personas can be useful for the organizers themselves. EventGraph [10] can help identify those people who become important hubs during events. The people and topics both face rapid change during events. Reinhardt et al. [11] surveyed how the use is different before, after and during conferences by organizers and participants. Their survey identified that the reason of participants' Twitter usage during conferences are sharing resources, keeping online presence, participating in parallel discussions, taking notes, communicating with others, and posing questions. Also analysis of citations via tweeting revealed how scientific works are cited during conferences [12]. Resource sharing and communicating are closely related with retweets and URL sharing. Suh et al. [13] showed retweetability of a tweet is correlated with presence of a URL and hashtags, and the number of followers and followees also influence retweetability. While their study was done with randomly sampled tweets, the result may differ for conference related tweets. Having a lot of followers may not indicate the importance of a Twitter user [14]; influence of a tweet or importance of a participant need to be measured in the context of the event itself rather than global metrics that we demonstrated in our analysis.

## III. BACKGROUND OF THE STUDY

The Theorizing the Web 2012 conference was held on April 14. The conference had a designated hashtag #TtW12 and Twitter account @TtW\_conf. After the success of the first Theorizing the Web conference in 2011 the organizers decided to run a follow-up event. They observed the participants' enthusiasm regarding the backchannel conversation, which was particularly active due to the fact that Internet researchers—almost by definition—are a highly-connected group. We categorized the participants (people who tweeted with the conference hashtag) in the conference backchannel into two roles:

- **In-Person Participants:** people who were present at the conference venue in person. They include both speakers

and non-speakers.

- Remote Participants: people not present at the venue in person but who tweeted about the conference.

The organizers created several strategies to better integrate the face-to-face and digital aspects of the conference. These strategies include ensuring that every in-person participant had free, high-speed WiFi access, providing particular hashtags for each session so the participants could mention which session they were tweeting about, live-streaming the entire event so that people could watch the talks remotely in real-time, including a “Twitter backchannel moderator” in each panel to identify questions from remote participants and ask them directly to the panelists, projecting the Twitter feeds in the hallways and the reception areas, distributing the speakers’ Twitter handles so that the participants could tweet with appropriate mentions, making the keynote a participatory interview rather than a top-down lecture so that incoming questions from backchannel participants (both in-person and remote) could be articulated to the speaker by the interviewer, and posting summaries of each panel to the conference blog prior to the event with images, so remote participants and in-person participants came well-prepared for the presentations.

We collected all the tweets having the word #TtW12 and from the account @TtW\_conf via a continuously running call to Twitter using streaming API. The tweets and metadata were parsed to retrieve the information about the participants, mentions, their followers-networks, their mentions-networks, the URLs shared via Twitter and the hashtag counts. The data set contained tweets with hashtag #TtW12 from March 14th (one month before the event) to April 27th (two weeks after the event), 4828 tweets in total and, on average 8.14 tweets per person.

*Session hashtags:* TtW12 had 4 sessions: session one (in morning), session two (before lunch), session three (afternoon), and finally session four (keynote session). The first three sessions each had three parallel tracks, located in different rooms. Each of the tracks were assigned a special hashtag so that the participants can tweet about the sessions using those hash tags. Session hash tags were in the pattern: room number followed by session number, so the track in room b during session 1 had the hashtag #b1, and the track in room c during session 1 had the tag #c1. Those hashtags were well broadcasted by the organizers through the conference website before the event itself. The session hashtags were also written in the conference rooms. This use of session hash tags in tweets related to the corresponding sessions helped compare the sessions.

*Identifying the in-person participants:* The conference organizers collected the Twitter handles of the in-person participants who tweeted about the conference and there were 92 of such people. This way, we succeeded to differentiate the in-person and remote participants and to compare their activities.

#### IV. CONFERENCE MONITOR AND ITS APPLICATION

In this section we describe the features of CM and the insights delivered by the corresponding features during TtW12.

CM’s web interface consists of four coordinated panels:

- Time line: A time line of tweets, the y-axis on the line shows number of tweets per 15 minute interval. It helps compare the sessions (e.g., which sessions generated more conversation, who were the active people during a session) and to observe the pace of tweets.
- Hashtags: The tag cloud of hashtags occurring in the tweets. Identifies trending and popular hashtags.
- User table: A table presenting relevant information about the participants. Identifies active participants (with more tweets) and popular participants (being mentioned and retweeted more).
- Tweet feed: Tweets occurring during the conference having the conference hashtag.

Figure 1 shows the unfiltered view of the system. The organizers can select tweets of interest using faceted filtering provided with it. All the views can be filtered based on time, session, user and hashtag. This cascaded filtering can be used in combination with one another to generate complex queries regarding the tweets. The filters are:

- Time frame (selected from the time line view)
- Session (selected from the session list: 1, 2, 3 or 4)
- Conference role (in-person, remote, speakers)
- Hashtag (chosen from the hashtag view)
- Participant (chosen from the user table)

After filtering, the tweet pace of that filtered set can be added in the main time line view as a separate line. Each of these lines is drawn with a different color and is superimposed over the existing ones. This enables the simultaneous comparison of different subsets of tweets.

##### A. Time line of tweets

The time line view shows the pace of the tweets. This view can be zoomed and then only the tweets from that selected time frame are shown. The hash tag view also changes to show the tag clouds from tweets of that time frame and the user table shows the participants who tweeted at that time frame. There are also predefined temporal selection options as a drop down list, from where a particular session can be selected to filter the time frame of only that session, or the last 24 hours, or last week or last month.

*Comparing TtW12 sessions using the time line:* The tweet pace pattern in the time line view showed (in Figure 1) that lunch and other break times had low tweet volume as expected. Interestingly, the tweet volume was higher in sessions before the lunch break and it declined after the break. In both cases the rate of tweets were similar but the number of participants dropped. It might be an indicator that some people left after lunch, just got tired of tweeting, or, maybe, they were paying close attention to the talk at the expense of tweeting.

As conferences usually have several sessions covering specific topics, it is often desirable to follow tweets from a specific session rather than going through all of the tweets. CM’s session-based filtering shows the tweets, participants, and hashtags during a selected session. Fig 2 (Top) shows

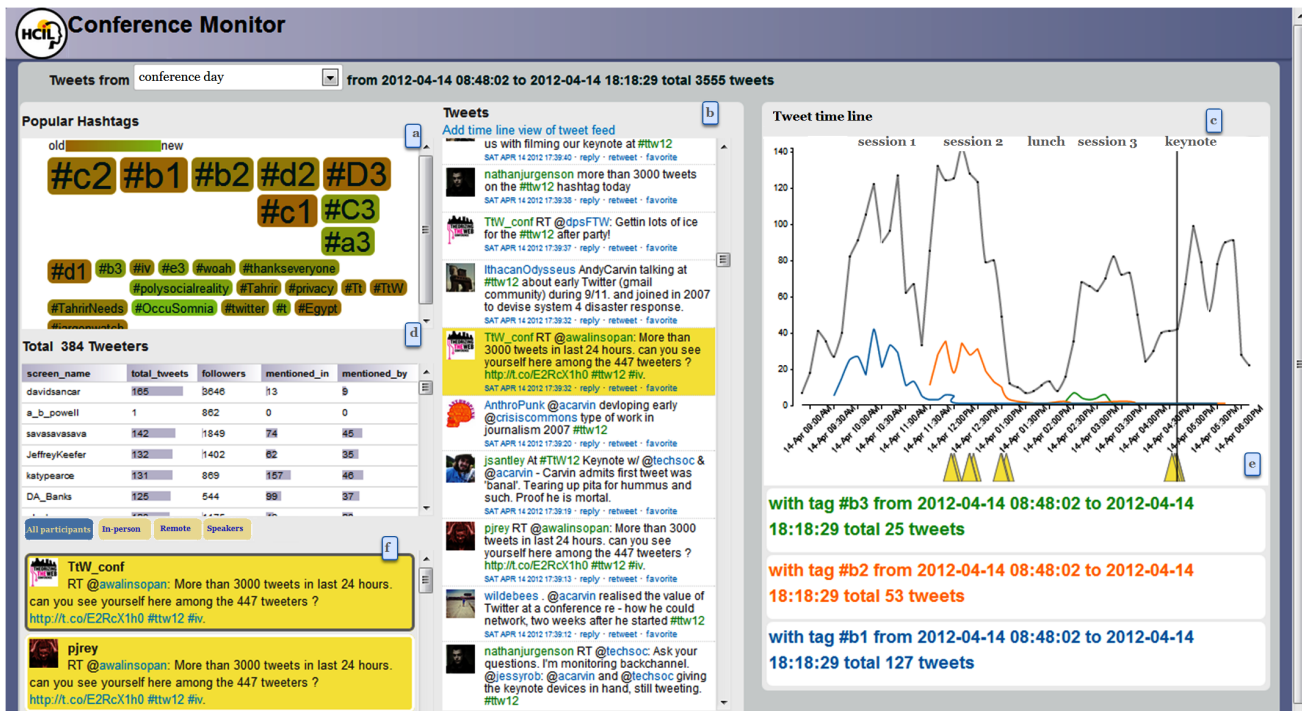


Fig. 1. Conference Monitor web app user interface. a) Hash tag view b) Tweet feed view c) Time line view, d) User table, e) Tweet feed, e) Annotation marked by organizers shown with triangles, and f) Highlighted tweets from organizers. The role-based filter buttons are below the user table. The tweets from the organizers are highlighted in the Tweet feed view as well.

Time	Total tweets	Participants	Average tweets per participant
Before lunch: 8 am to 1 pm	2151	281	7.6
After lunch: 1 pm to 6 pm	1435	197	7.8

TABLE I  
PARTICIPATION BEFORE AND AFTER LUNCH

the view after filtering to session 1. The morning sessions in rooms b, c and d were held in parallel, but the room d session (hashtag d1) had the least Twitter activity. Similarly, b2, c2 and d2 were hashtags for parallel sessions in different rooms during session 2, and all had comparable levels of activity. The afternoon sessions b3, c3 and d3 had comparatively low backchannel activity (Fig 2). Sorting the user table by tweet count showed which participants were more active during the selected session. During the keynote conversation there were 698 tweets, 283 of those mentioned the keynote speaker and 175 mentioned the interviewer, either quoting the speakers or asking them direct questions. The interviewer (Zeynep Tufekci, @techsoc read, synthesized, and responded to the Twitter stream while asking questions to the interviewee (Andy Carvin, @andycarvin). The speaker himself made 15 more tweets after the conference in reply to the tweets from the conference participants.

### B. Hashtag view

The hashtag view shows the tag cloud of the top 20 most frequent hashtags appearing in the conference tweets. The size of the hashtag indicates the frequency and the color indicates its age (time spent since its first occurrence). Older hashtags are displayed in brown and newer ones are in green. With this visualization, conference organizers can see if a hashtag is very popular in spite of its recency. After selecting a hashtag, the tweet feed view shows only the tweets with that hashtag, the user table shows only the participants who tweeted with that hashtag. A time line can be added in the main time line view to show the propagation of tweets with the selected hashtag and compare it with other hashtags' time times in the same view.

The tag cloud view for TtW12 showed that the session tags were the most used hashtags (for example the tags c1, b3 etc.), indicating that the participants welcomed this idea of session tags and cared enough to use them. For the parallel sessions, we could see which room had more activity by the size of its hashtag in the hashtag view (see Fig 2). TtW12 organizers estimated that, on average, 25 people were physically present in each of the tracks.

### C. User table

The user table provides an overview of the participants during the conference. Some participants may not be active with tweeting but can be mentioned many times during the conferences, especially if s/he is a speaker. Therefore only tweet counts do not give the full picture of the importance



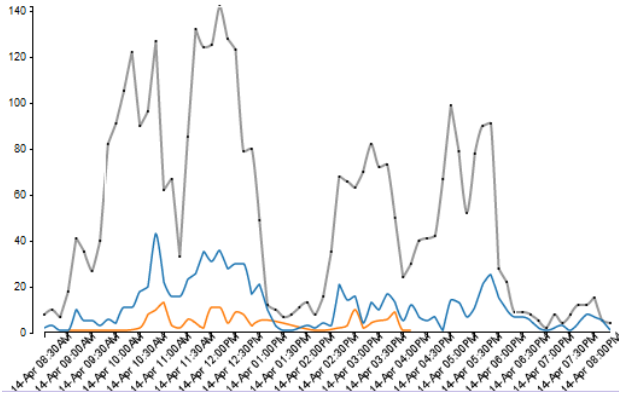


Fig. 4. Time line view after filtering down to remote participants only. Blue line is for the tweets from all the remote participants, and orange line is for the tweets only from Jeffery Keefer.

of that participant. With CM, active and popular participants can be identified easily from the total tweets, 'mentioned in' and 'mentioned by' columns. Unlike other existing tools, ours presents the metrics relevant in the context of the conference. Participants in conferences assume various roles, and some might want to follow tweets only from the speakers or the organizers. One novel feature of CM is its role-based filtering associated with the table: it can be used to separately examine the activity of in-person participants, speakers, and remote participants. If the remote participants filter is chosen, then only the remote participants are shown in the table and all the other views also change accordingly. Also after choosing a participant from this table the tweet feed view shows the tweets from and mentioning that account providing an easy understanding on what s/he is talking and what is being talked about him or her. As Twitter itself does not show conversation as a thread, the tweet feed view in CM, showing both tweets from and about a participant, helps understand the whole context of a conversation.

*Identifying active participants:* After keeping only the speakers, the table showed that among the speakers, the most active participant was @KatyPearce, who tweeted and got mentioned many times. Selecting this participant (Figure 3) showed the hashtags appeared on this participant's tweets in the hashtag view and the tweets from and about her in the tweet feed view. This participant's time line (added in the main time line view) demonstrated her active participation throughout the whole event.

Among the remote participants Jeffery Keefer was the most active one (Fig 4) with the most number of tweets. With total 139 tweets he got 18 new followers while watching the live streaming of TtW12. His tweet stream in the time line showed that he was active most of the time and became inactive in the backchannel during the keynote speech.

#### D. Annotation

The visualizations of CM gives an overview of the conference to anyone but an organizer might also want to annotate specific moments of conference for future reference and go

	In-person participants	Remote participants	Total
Total tweets	3728 (77%)	1100 (23%)	4828
Retweets	1142 (66%)	585 (34%)	1727
Total participants	92 (15.5%)	501 (85.5%)	593
Average tweets per participant	40.2	2.2	
Average retweets per participant	12.4	1.2	

TABLE II  
PARTICIPATION OF IN-PERSON AND REMOTE PARTICIPANTS.

back to that time point for further analysis. Therefore we added annotation capability for the organizers. They can annotate significant happenings during the conference on the time line view. For example, they can save a note such as: "Emailed the person with most number of followers at 12 am." and annotate the time when they performed that action. Also, CM automatically annotates the tweets from the organizers and their timestamps. These time points of annotation and tweets from the organizers are marked with yellow triangles below the main time line view and hovering over them shows the time and annotation at that point. Selecting a triangle indicating tweets from the organizers highlights the corresponding tweet in the feed. Organizers can select time frame before and after the actions to assess the effect of those actions.

#### V. RETROSPECTIVE ANALYSIS

In this section we discuss the second part of our contribution, our retrospective analysis of TtW12 backchannel, focusing on the topics, contents, and network growth of participants during the conference.

##### A. Topics

Our word-frequency based analysis showed that the most frequent words in the conference tweets were about social networking services Twitter, Facebook and Pinterest as well as journalism, politics and privacy. The contents of tweets also varied during the conference. Before the conference, the tweets were more about the conference location, housing, getting there, food, etc. Tweets from the organizers relevant to those topics got more retweets before the conference. This highlights the importance of social interaction even before the conference. Tweets that are not directly related to the conference topics also played important role for the participants. For example, the live streaming was not working for one of the sessions and the organizers readily fixed the problem after reading a tweet about the situation. When one conference participant was looking for housing near the venue, that tweet was retweeted by other conference followers. The backchannel moderators in TtW12 helped identify those tweets and made an immediate organizational decision that was appreciated by the participants.

##### B. Participation of people

Among all the 593 participants in the backchannel, 523 people mentioned or replied to or retweeted other people's tweets, which indicates the overall trend to communicate with

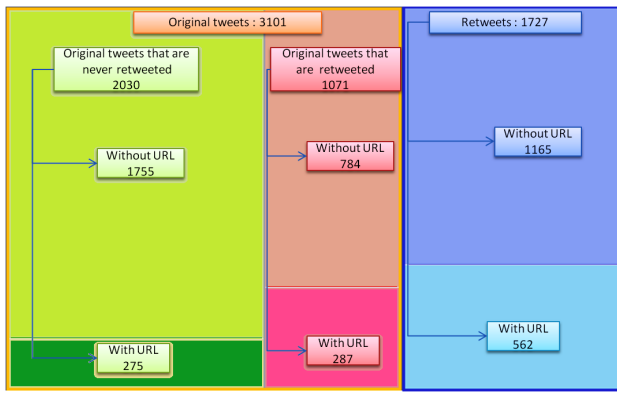


Fig. 5. Tree map showing the division of tweets by retweets and URL. 1071 tweets were retweeted at least once, among them 287 had at least one URL. 287 tweets with URL were shared 562 times, averaging 1.9 retweets per tweet. On the other hand, 784 tweets without URLs were retweeted 1165 times, averaging 1.5 retweets per tweet.

others. Among all the 92 in-person participants, 62 were non-speakers. All of the 30 speakers tweeted about the conference. The remote participants followed the event via reading the Twitter feed or watching live streaming. In total 24 people watched the live streaming remotely and contributed to the backchannel conversation with 236 tweets in total (9.8 tweets on average) and 43 retweets (1.8 retweets on average). In general, the activity level of the live-streamers was higher than other remote participants and lower than the in-person participants (table II).

### C. Retweets and URLs

Fig 5 shows the breakdown of tweets into retweets and URL-containment using a tree map, depicting that most of the non-retweeted tweets did not have any URLs and 26.7% of the tweets with a URL were retweeted. Retweets with URLs mostly had links to blog posts on relevant topics, live stream video, online flier of the conference, list of participants and panel spotlight. On the other hand retweets without URLs were mostly relevant to the topics discussed in the conference (specially during the keynote conversation), positive feedback about the conference, controversial speech (e.g., one participant used the term “digital native” in a tweet, which incited criticism from several other participants).

In total 338 URLs were posted via 1124 tweets averaging 3.3 per URL. Among them 48% URLs were not reshared or retweeted. Those were mostly part of conversations, for example sending a picture or video URL to another person or containing a personal photo at the venue.

boyd et al. showed in general, 52% of retweets contained a URL and 22% tweets contained a URL [15], whereas in our case 31.9% of retweets contained a URL (different from the previous studies of overall tweets) and 23.28% of tweets contained a URL (similar to previous work). Also, in Twitter only 3% of tweets were retweets, whereas in the case of this conference, 35.77% of tweets were retweets. In our data set, a tweet with URL had a 51% probability of being retweeted.

### D. Network growth in context of the conference

By attending a conference, individuals expect to expand their network of contacts. While Twitter followers are not synonymous with an individual’s professional network, gaining followers is a form of expanding one’s network. Remote participants might also be able to engage in such network development by joining the backchannel conversation. This leads to questions of whether remote participation through the conference backchannel, at least to some degree, can lead to significant network development for the individuals. For this analysis we collected the followers-network data of the participants from April 9 to April 15. We filtered out the followers who did not tweet about the conference; therefore, the remaining network consisted only of the followers who were also explicitly interested in the conference.

Follower growth of a participant

$$= \text{The number of conference related new followers} \\ = \text{Number of total new followers} - \\ \text{number of new followers who did not tweet} \\ \text{about the conference}$$

There were 858 follower connections present on the 9th April and, as the conference day approached, the number of new followers grew (see Fig 6). Finally, until April 15, 1591 new connections were made among 364 people. The followers-network growth had a sharp drop on the day before the conference (people were traveling to attend the conference and were not active online) and most of the new connections were established on the day of the conference. From the pie chart we can see that 33% of the new connections were made to the speakers indicating people’s interest towards the speakers.

Although the remote participants gained fewer average number of new followers than the in-person participants, a few of them managed to get as many as 18 new followers in spite of their sparse activity. Their tweets were either retweeted by the conference organizers or they were mentioned in the organizers’ tweets. In this case, being mentioned by important people who have many followers related to the conference gave them more visibility and helped expand their network. One remote participant (Marc Smith) gained 23 new followers. The only 2 tweets he made were sharing the event graph of the conference, which generated immediate retweets. The general conception to gain new followers is to tweet more but on the context of conference quality of tweets is as important as their quantity. This emphasizes the importance of analyzing tweets and network growth differently in the context of a conference. Observing the follower growth pattern, we came up with three hypotheses and performed an ANOVA test controlling for the number of followers and number of conference community followers. The hypotheses and their results are presented in table III and in Fig 7.

1) *Mentions*: We also observed a positive correlation (0.66) between the number of times participants were mentioned and

Hypothesis	F, p	Result
H1: Speakers, general in-person participants, and remote participants will differ with respect to how many conference related new followers they gain. Specifically: follower growth for speakers > follower growth for non-speaker in-person participants > follower growth for remote participants	$F(2, 599) = 22.495$ , $p < 0.001$	Statistically significant. We expected in-person participants to have more detailed, topically-relevant tweets than remote participants, thus increasing the attractiveness of their tweets and attracting more followers from the conference community than individuals who are just remotely tweeting about the conference. Participants' role in the conference significantly influenced their conference related network growth.
H2: Participants who create tweets using the conference hashtag will gain more conference-related followers than those who create fewer tweets using the conference hashtag. Tweeting about the conference will affect the number of followers an individual gains from the conference community.	$F(1, 599) = 12.284$ , $p < 0.001$	Statistically significant effect. Participants with more conference related tweets gained more new conference related followers than participants with fewer tweets.
H3: Individual's conference related tweet volume will have a greater impact on the number of followers they gain from the conference community if they are present at the venue.	$F(2, 618) = 4.290$ , $p < 0.05$	Speakers who tweeted more got more new followers than speakers who tweeted just once. In-person participants also gained more followers by tweeting more, but the same effect was not observed for remote participants. Ninety-eight of the remote participants tweeted more than once but they did not gain significantly more new followers than the other remote participants who tweeted just once.

TABLE III  
FOLLOWER GROWTH HYPOTHESES AND ANOVA RESULT (ADJUSTED  $R^2 = 0.589$ ).

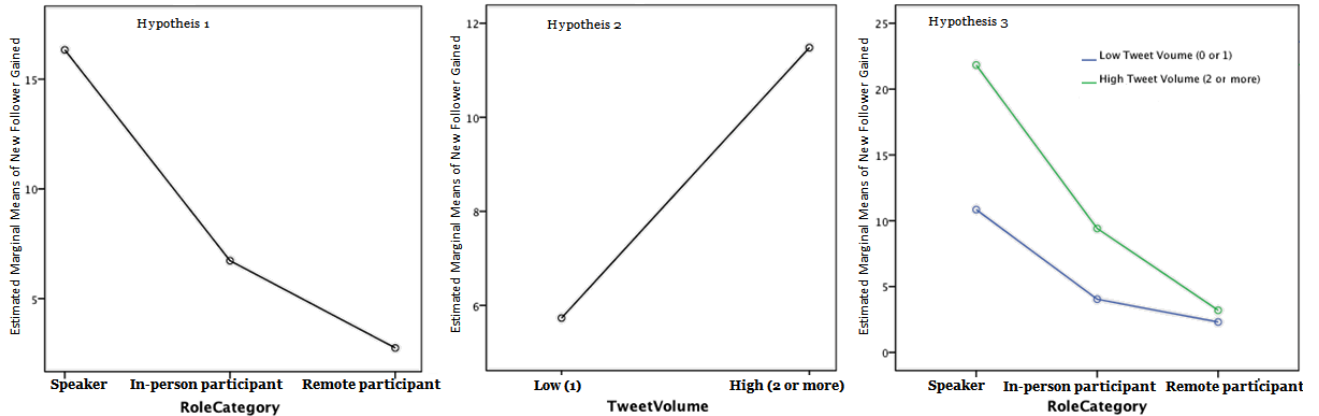


Fig. 7. ANOVA test results of hypotheses 1, 2 and 3. Covariates appearing in the model: followers count = 2104.88

their total tweets. The mention-network diagram generated with NodeXL [16] in Fig 8 shows that the most mentioned people were often also the central people in their clusters. They were @acarvin (keynote speaker), @pjrey (organizer), @nathanjurgensen (organizer), @TtW\_conf (the conference account), @dfreelon, @racialicious, and @katyperce (most active speaker). @techsoc (interviewer of the keynote speaker), @TahrirSupply and @SultanaAlQassemi were mentioned mostly during the keynote speech and they belonged to the same cluster. Twitter usage during the Arab Spring was a key discussion topic during that session as keynote Andy Carvin covered that event as a journalist.

Speakers were repeatedly mentioned and they also gained more followers. We hypothesized that being mentioned by more people would bring more new followers for a participant. Our analysis showed that the number of new followers of a participant and the number of people mentioning him/her had a positive correlation of 0.6258 supporting our hypothesis.

## VI. DISCUSSION ON THE ANALYSIS

We summarize our findings from the analysis on Twitter usage during TtW12 in this section.

### A. Remote participants can contribute and get benefit

The tweets revealed that the remote viewers embraced the idea of live streaming, and some remote participants followed the whole event, tweeting about it the whole day and making new connections. Live-streamers tweeted questions to the keynote speaker and his interviewer. Live-streamers felt engagement with the conference, as demonstrated by their active tweeting and creating new connections. One tweet from a live-streamer was:

*“Thanks to everybody #TtW12 who made those of us not present in corporality feel so welcome and included. A model for conferences @TtW\_conf”.*

Since visibility is one of the primary benefits of a conference, marginalizing remote participants in conferences usually



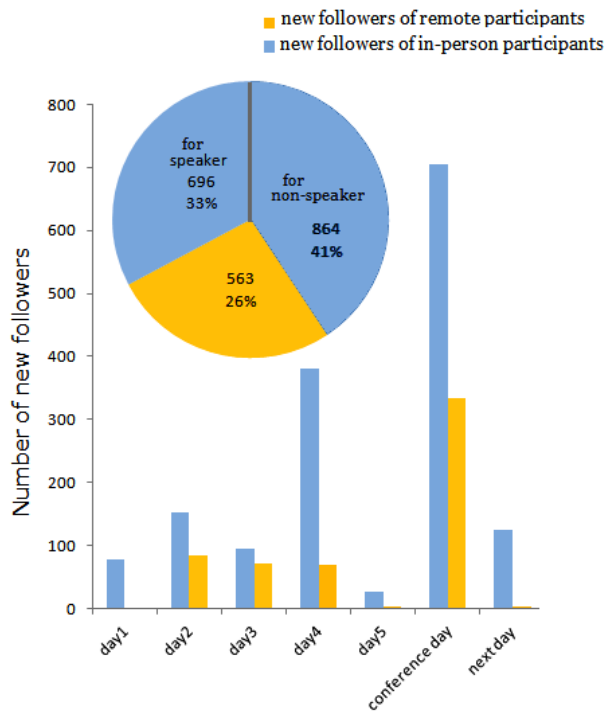


Fig. 6. Addition of new followers during the conference week. Speakers got 696 new followers, non-speaker in person participants got 864 and remote participants, 563.

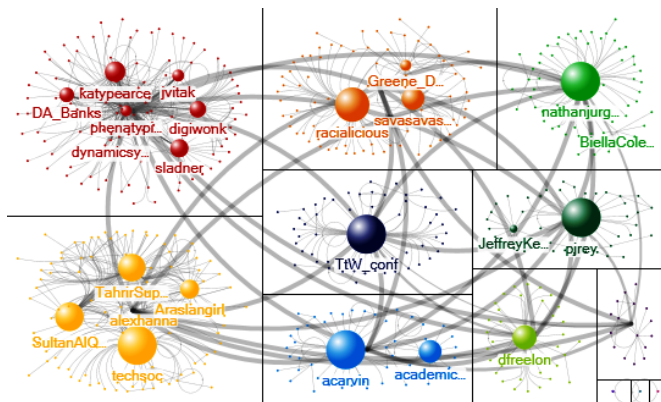


Fig. 8. Several clusters are formed within the mention network. Node size is proportional to the number of times they were mentioned. Edges connecting two clusters are drawn with thick lines.

discourages their participation. In contrast, TtW12 organizers got repeated feedback on how integrated remote participants felt. One tweeted that he set an alarm to wake up at 6 am in Australia to be part of it. In total, 34% of the retweets were from the remote participants. A high volume of tweets made the conference “Trending” in DC-area tweets. Part of this possibly happened due to the actions taken by the organizers. Such visibility can help conferences reach a broader audience who might have a tangential interest in the conference topic.

### B. Network growth depends on participant’s role

Followers-network started to grow even before the conference and the highest number of new connections occurred

on the conference day. We observed that a participant’s total number of followers and total number of conference related followers were not correlated. The organizers already had many followers who were related to the conference, whereas some participants started with smaller followers-network before the conference and their network expanded during the conference. Significant growth of the followers-network for the speakers (hypotheses 1 and 3) indicates the strong interest in speakers. This justifies our filtering by role. One tweet had request for the speakers’ Twitter handles in each of their slides. On average each speaker gained 23 new conference related followers during the conference week whereas the number for each non-speaker is 13.9 and for each remote participant is 4.6. Some remote participants with just a few tweets managed to get higher numbers of new followers because the organizers mentioned them in tweets.

During the conference, we interviewed speaker Katy Pearce and got her perspective on the use of Twitter. She mentioned that she keeps her followers up to date about her publications and recent works, and that helps her get more citations for her work. Her suggestion about how to make new connections with other researchers via Twitter was to write thoughtful comments about their work to get their attention, so that they would also follow back.

### C. Tweets about conferences should be analyzed differently

A person with more followers has a broader network of influence in general, but in the context of the conference, his effective network of influence is basically determined by his followers, who are also concerned about the conference. Their tweets regarding a conference might not be interesting to all of their followers and might not get retweeted as many times as expected. The total followers count and conference related followers count did not show any positive correlation in our analysis. Therefore, total followers, tweets and mentions cannot be relevant for conferences. Keeping this in mind, we used the conference-related metrics for measuring their follower growth. We propose that the conference related influence metrics can be their conference related followers count and mention counts.

### D. Session hashtags are useful

Use of session tags delivered meaningful comparisons of the sessions and analysis of parallel sessions. The real-time view showed tweets about a session before, during and after the sessions. Participants could identify who was attending which session by looking at the session tags in the tweets and got updates about other sessions from corresponding tweets.

### E. More participants in the morning sessions

The total number of participants decreased after the lunch break, which reduced the total activity in the backchannel. But the average tweets per participant did not decrease that much. This indicates some participants were very active throughout the whole day and some stopped being active after the morning sessions.

## F. Data Clean-up and Limitation

One complicating factor was the same hashtag was used both for Theorizing the Web and Teen Tech Week, which happened in March. We rendered the network diagram of the people tweeting with TtW12 and saw two separate clusters of people: one with people from Theorizing the Web and another with Teen Tech Week. We removed tweets and user information of the people in the other cluster to clean up the noisy data before our analysis.

## VII. FUTURE WORK AND CONCLUSION

Technology and society have always been enmeshed, and digital information is now part of our social lives. The fluid nature of social media communication lets the conversation flow beyond the conference venue and out to the broader public that may have interest in the topics being discussed. Henceforth, the digital aspects of conferences should not be ignored. Our case study with TtW12 showed CM could help organizers monitor this digital dimension during the conference in real-time. CM enables the identification of interesting participants and tweeting patterns via simple interactions. Filtering views by time identifies which participants are more active in which sessions and who is mentioned more during a given time frame. The user table can identify the influential and active people in real time. Also the time line view shows how long influential people are active and when they are mentioned. These insights are not easily identifiable otherwise. The real-time visualizations help formulate hypotheses about the conference. CM made possible to follow the topics, sessions or speakers of interest by selecting a hashtag, a participant or a session instead of linearly reading all the tweets. One improvement suggestion from sociologist Marc Smith was to add simple network statistics in the real-time visualization, which we aim to include in future version. The case study also revealed that it is both possible and beneficial to create a bridge between the in-person and remote participants of a conference through the use of technology and social media. We observed a high volume of remote participation and analyzed the network growth in the context of the conference. While prior works studied types and volume of tweets, they did not address the network evolution during such events. The novelty of this analysis lies in the fact that it accounts for only the relevant followers who are also related to the conference. To do this, we also needed to identify which part of the followers-networks was formed due to a person's involvement in the conference. Our method of this identification might have missed some new followers who did not actually tweet about the conference. Further study can shed more light on identifying these silent followers. We wish to use CM for other conferences and evaluate its usefulness in those cases. Our broader goal is to create a platform to monitor online social media conversation and make it useful for situation awareness and intervention in the cases of natural disasters, and national and international events.

## ACKNOWLEDGMENT

We cordially thank Catherine Plaisant, Anne Rose, Jae-wook Ahn, Rajan Zachariah, Robert Gove, Krist Wong-

sphasawat and Sarah Webster for their valuable contribution for this study. This work is supported by the NSF grant IIS – 0968521.

## REFERENCES

- [1] J. McCarthy *et al.*, "Digital backchannels in shared physical spaces: experiences at an academic conference," in *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 2005, pp. 1641–1644.
- [2] D. Shamma, L. Kennedy, and E. Churchill, "Tweetgeist: Can the twitter timeline reveal the structure of broadcast events?" in *CSCW 2010*, 2010. [Online]. Available: <http://research.yahoo.com/pub/3041>
- [3] A. Marcus, M. Bernstein, O. Badar, D. Karger, S. Madden, and R. Miller, "Twitinfo: Aggregating and visualizing microblogs for event exploration," in *Proceedings of the 2011 annual conference on Human factors in computing systems*. ACM, 2011, pp. 227–236.
- [4] N. Diakopoulos, M. De Choudhury, and M. Naaman, "Finding and assessing social media information sources in the context of journalism," in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: ACM, 2012, pp. 2451–2460. [Online]. Available: <http://doi.acm.org/10.1145/2208276.2208409>
- [5] M. Dork, D. Gruen, C. Williamson, and S. Carpendale, "A for large-scale events," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 16, no. 6, pp. 1129–1138, 2010.
- [6] M. Ebner, H. Mhlburger, S. Schaffert, M. Schiefner, W. Reinhardt, and S. Wheeler, "Getting granular on twitter: Tweets from a conference and their limited usefulness for non-participants," in *Key Competencies in the Knowledge Society*, ser. IFIP Advances in Information and Communication Technology, N. Reynolds and M. Turcsnyi-Szab, Eds. Springer Boston, 2010, vol. 324, pp. 102–113. [Online]. Available: [http://dx.doi.org/10.1007/978-3-642-15378-5\\_10](http://dx.doi.org/10.1007/978-3-642-15378-5_10)
- [7] K. Starbird and L. Palen, "(how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ser. CSCW '12. New York, NY, USA: ACM, 2012, pp. 7–16. [Online]. Available: <http://doi.acm.org/10.1145/2145204.2145212>
- [8] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: what twitter may contribute to situational awareness," in *Proceedings of the 28th international conference on Human factors in computing systems*, ser. CHI '10. New York, NY, USA: ACM, 2010, pp. 1079–1088. [Online]. Available: <http://doi.acm.org/10.1145/1753326.1753486>
- [9] C. Ross, M. Terras, C. Warwick, and A. Welsh, "Enabled backchannel: conference twitter use by digital humanists," *Journal of Documentation*, vol. 67, no. 2, pp. 214–237, 2011.
- [10] D. Hansen, M. Smith, and B. Shneiderman, "Eventgraphs: charting collections of conference connections," in *Forty-Forth Annual Hawaii International Conference on System Sciences*, 2011.
- [11] W. Reinhardt, M. E. G. G. Beham, and C. Costa, "How people are using twitter during conferences."
- [12] K. Weller, E. Dröge, and C. Puschmann, "Citation analysis in twitter. approaches for defining and measuring information flows within tweets during scientific conferences," in *Proceedings of Making Sense of Microposts Workshop (#MSM2011). Co-located with Extended Semantic Web Conference, Crete, Greece*, 2011.
- [13] B. Suh, L. Hong, P. Pirulli, and E. Chi, "Want to be retweeted? large scale analytics on factors impacting retweet in twitter network," in *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, aug. 2010, pp. 177–184.
- [14] M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi, "Measuring user influence in twitter: The million follower fallacy," in *4th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2010.
- [15] D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," in *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*. IEEE, 2010, pp. 1–10.
- [16] M. Smith, B. Shneiderman, N. Milic-Frayling, E. M. Rodrigues, V. Barash, C. Dunne, T. Capone, A. Perer, and E. Gleave, "Analyzing (social media) networks with NodeXL," in *C&T '09: Proc. fourth international conference on Communities and Technologies*. New York, NY, USA: ACM, 2009, pp. 255–264. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=1556460.1556497>