A Tempest in a Teacup? Analyzing Firestorms on Twitter

Hemank Lamba*, Momin M. Malik*, and Jürgen Pfeffer
{hlamba,momin.malik,jpfeffer}@cs.cmu.edu
School of Computer Science
Carnegie Mellon University

***Abstract*—**

'Firestorms,' sudden bursts of negative attention in cases of controversy and outrage, are seemingly widespread on Twitter and are an increasing source of fascination and anxiety in the corporate, governmental, and public spheres. Using media mentions, we collect 80 candidate events from January 2011 to September 2014 that we would term 'firestorms.' Using data from the Twitter decahose (or gardenhose), a 10% random sample of all tweets, we describe the size and longevity of these firestorms. We take two firestorm exemplars, #myNYPD and #CancelColbert, as case studies to describe more fully. Then, taking the 20 firestorms with the most tweets, we look at the change in mention networks of participants over the course of the firestorm as one method of testing for possible impacts of firestorms. We find that the mention networks before and after the firestorms are more similar to each other than to those of the firestorms, suggesting that firestorms neither emerge from existing networks, nor do they result in lasting changes to social structure. To verify this, we randomly sample users and generate mention networks for baseline comparison, and find that the firestorms are not associated with a greater than random amount of change in mention networks.

I. Introduction

On Twitter, firestorms (or shitstorms, or Twitterstorms) have become an object of fascination and anxiety. They are one of the major topics in discussions of Twitter in the realm of public relations and brand management [1]. Individual events frequently receive media coverage, and the phenomenon as a whole receives coverage as well; political comedian John Oliver did a segment critiquing corporations' use of Twitter on the September 15, 2014 episode of his HBO television show, *Last Week Tonight*, featuring many examples of firestorms. Online magazine *Slate* dubbed 2014 the 'Year of Outrage' in an eponymous special feature, listing one example for every day of the year (each with an accompanying tweet to illustrate the outrage) alongside reflection articles such as 'The Life Cycle of Outrage.'

The term 'firestorm' refers to an event where a person, group, or institution suddenly receives a large amount of negative attention [2]. Any sudden controversy or expression of outrage may be termed a firestorm, although we are interested in a firestorm as something more specific: a case where the sudden negative attention is in response to a recent action or statement of the target entity (rather than without a specific trigger, such as in a premeditated protest or prank) and arises spontaneously (rather than through prior coordination, such as from a group prepared for mobilization). Furthermore, we are interested in when this attention exhibits network effects: the initial negative attention causes more people to learn of the action or statement, and these people then contribute their own negative attention. Such cases are examples of negative word-of-mouth dynamics. We focus on firestorms targeting public figures, businesses, and institutions, where consequences are public; we do not consider firestorms targeting private individuals [3], as the consequences there are in terms of the individuals' experiences which we consider a different topic.

The ultimate question is if participation in or consumption of firestorms has an effect outside of Twitter, such as through purchasing decisions, voting behavior, attendance at protests, or even participation in violence, either directly (by firestorm participants) or indirectly (by people influenced by firestorm participants). However, such information is impossible to collect directly at scale, and difficult even to indirectly infer from Twitter data. Instead, we draw on literature about the *biographical consequences of activism* [4] to ask, can we detect a change in firestorm participants as a result of the event? We look specifically at social ties of firestorm participants and form the research question: what is the relationship between social ties and firestorm participation? I.e., do the people who participate in a firestorm know each other beforehand? Do they communicate during? And do they continue to communicate after? If there is a discernible change in social ties over the course of a firestorm, it suggests a social impact that could lead to long-term consequences. On the other hand, if firestorms arise from existing social ties, it would point to firestorms being a consequence rather than a cause of other action, and if there is no relation to social ties, it would be inconclusive but, as social actions are embedded in networks of social ties, it would suggest firestorms are of little importance.

II. Background and Related Work

A. Firestorms

There have been several papers directly on Firestorms. We summarize these in tables I and II.

Much of this literature is about the problem specification, with conclusions being very preliminary. What has been found

*Equally contributing authors.

NAME	ARTICLE(S)
Crises	Bruns & Stieglitz 2012 [5], Park et al. 2012 [6], Rajasekera 2010 [7]
Scandals	Bruns & Stieglitz 2012 [5]
Bad news	Park et al. 2012 [6]
Firestorms	Mochalova & Nanopoulos 2014 [8], Pfeffer et al. 2014 [2]
Shitstorms	Stieglitz & Krüger 2014 [9]

TABLE I. SIMILAR TERMS TO FIRESTORMS IN LITERATURE.

YEAR	FIRESTORM	ARTICLE(S)
2009	Domino's employees prank video	[6]
2010	Toyota recall	[10], [7], [11]
2011	Playstation Network hack and shutdown	[1]
2012	#QantasLuxury campaign after labor dispute	[5], [2], [1], [10]
2012	Papa Johns "lady chinky eyes" receipt	[6]

TABLE II. FIRESTORMS CONSIDERED IN THE LITERATURE.

so far is that external events such as statements do affect the firestorm [1], [9], but that there is a time lag in the diffusion of an apology [6], and that a small number of users are responsible for the vast majority of the tweets [5] just as in Twitter activity in general [12].

Twitter's culture makes brands particularly vulnerable to firestorms. Van Dijck [13] discusses the "paradox of Twitter," one aspect of which is that the thing that gives Twitter value for marketers—the authenticity and openness of social interaction—is destroyed when marketers try to intervene to capitalize on that value. Nitins and Burgess [1] write that early on, brands saw social media as instant and free access to consumers around the world. But they failed to consider the culture of social media, importing their standard one-to-many communication models. They quickly found that consumers also now have the ability "to 'talk back' to companies—even very large global corporations—[and] to do so in public; they can share their pleasure, or displeasure, with potentially millions of other consumers without significant effort," and that they often resented the intrusion of companies. Because of this, they continue, "Twitter users frequently delight in 'gotcha' moments". Note that this scenario may be very different for certain celebrities and brands who develop a strategy of appealing to iconoclasm, for whom frequent negative attention may be beneficial and Twitter may be an easy way to garner this (e.g., potentially Kenneth Cole and Urban Outfitters).

In terms of modeling firestorms, there is relevant literature on 'media hypes' or 'media storms,' and on news cycles. Vasterman [14] characterizes media hypes as being *self-reinforcing*, potentially being driven less by external events after the triggering event and more by discussion about itself [14], [15] until the issue is crowded out by another topic. Vasterman suggests a smooth left-skewed distribution as a model, while Wein and Elmelund-Præstekær [15] find evidence of a decreasing oscillatory pattern. For news cycles, Leskovec et al. [16] found a 'saw-toothed' shaped increase followed by an exponential decay, which they were able to reproduce in a simulation model that combined an imitation effect and a recency effect.

It is often assumed that firestorms have an effect, and are therefore important. However, whether or not this is so is an open question. Kimmel and Kitchen [17] argue that while word-of-mouth, including negative word-of-mouth, is significant in shaping consumer attitudes and behavior, its power is also frequently oversold. They urge caution towards claims of the impact of word-of-mouth on social media. More generally, the question of whether low-commitment online protest and activism has an impact is hotly debated. Many argue that 'slacktivism' [18] or 'clicktivism' [19] are not effective at achieving their aims, while others argue that 'hashtag activism' [20] is better than nothing.

A more nuanced way of looking for an impact from firestorms is to consider the effects on participants themselves. McAdam's famous 1989 work [4] introduced the idea that participating in activism has an impact on individuals, and even if the given activism itself is not successful, it has 'biographical consequences.' Individuals influenced by earlier participation go on to do further actions that are significant. Indeed, looking at activists rather than campaigns shows that online activism plays a role in larger movements even when specific campaigns have no impact [21], and this is important to consider when judging the effectiveness of online action [22].

In order to address our research question of the relationship between social ties and firestorm participation, we look at mention networks. Merritt et al. [23] have shown in another context that discussion is an effective proxy for friendship ties. We would go further to say that while we cannot measure exposure from friendship data as in Myers and Leskovec [24] from the available data, using mentions as the measure of social ties is a stronger mark of connection and thus a more meaningful measure. This is different from Granovetter's [25] concept of a strong tie, but a mention is nonetheless a strong*er* tie than a following relationship and we would expect its impact to be greater than just a follower relationship.

B. Twitter

Amidst the enormous recent academic literature involving Twitter [26], researchers are increasingly beginning to appreciate that studying the microblogging platform is not necessarily the same as studying human behavior in general [27], [28]. There are multiple barriers to generalization, including that Twitter demographics are non-representative [29], [30]; that the possibility of making money from link farming [31] or from selling bots to inflate metrics [32] means there is widespread spam [33] that Twitter is not able to entirely or immediately filter out [34], and this spam may distort research findings [31]; that there are idiosyncratic conventions of Twitter [35], [36], [37] and a specific culture and ideology that is anti-establishment [1], [13] and, as shown by what drives adoption, focused on celebrity culture [38]; that even beyond spam, the Twitter social graph [39] has non-random patterns of adoption that potentially give it a topology vastly different from that of the underlying social network [40]; that the most accessible channel of data, the Streaming API, is unreliable within certain parameters [41], which causes difficulty in studying meso- and macro-scale phenomena [42]; and that Twitter is itself neither globally uniform [43] nor a static, stable environment across years [44], [13].

However, Twitter is host to many firestorms in itself. That is, there are frequent cases where a tweet sets off a firestorm, where an apology is given via tweet, or where a protest is organized under a hashtag. This, combined with how Twitter has become a critical channel of communicating and cultivating brand reputation and identity [1], [9], [45], means that firestorm behavior on Twitter is of interest in and of itself without needing to be representative of larger social behavior.

C. Specific firestorms

While we ultimately find 80 candidate examples of firestorms, we select two of them to examine more closely. The first is #CancelColbert, a hashtag started by activist Suey Park in reaction to a tweet quoting a skit on the satirical news program *The Colbert Report*, from American political comedian Stephen Colbert. The hashtag took off, and was soon followed by a reaction against the hashtag. We use it as an example of a firestorm that potentially comes from an already well-connected community, as initially it would only have been users following Park who would have seen her call to trend #CancelColbert.

Second is #myNYPD, a campaign started by the New York Police Department (@NYPDnews) to collect positive stories about the NYPD. However, it was 'hijacked' and used it as an opportunity to highlight grievances around police brutality: alongside sarcastic comments about the kindness of police, users posted pictures of NYPD officers grabbing, kicking, beating, and otherwise abusing people. The campaign was widely considered a failure and embarrassment for the NYPD, and is an excellent example of hashtag hijacking and public relations gone wrong.

III. Method

A. Firestorm identification

As we note above, the link between activity on Twitter and larger societal phenomena is complex and difficult to disentangle from all the confounding factors. Thus, while people take to Twitter over practically every controversy or outrage, we decided that only firestorms that have some substantive connection to Twitter would be meaningful to study with Twitter data. We developed inclusion criteria, that a controversy first must have had some media mention, and second it must meet at least one of the following conditions:

- The controversy began around a tweet or series of tweets;
- The entity at the center of the controversy posts a apology, retraction, non-apology, or otherwise major statement on Twitter; or
- A specific hashtag, that we were able to find through searching through media, is associated with the controversy.

We also choose to exclude cases where it is obvious that something sent from a professional account was meant to be posted from a personal account, as we find that in such cases Twitter users are generally more amused than angry. We include cases of social media account managers failing to use proper discretion (i.e., intending to post what they did, but being mistaken that it was appropriate).

We limit the period under consideration to the middle of September 2014, as per our data. We only found several firestorms in 2009 and 2010, and no earlier examples, so we choose to consider only firestorms in 2011 onwards.

B. Firestorms collection

Our search method consisted first of web searches for "Firestorms" and Twitter, "shitstorms" and Twitter, and other variations; we went through any lists of "social media fails" we came across; and lastly, searched through tags like "Twitter" and "PR" on technology- and culture-oriented blogs and aggregation sites such as Buzzfeed, Mashable, and The Verge. We used other aggregated lists, such as on KnowYourMeme.com, John Oliver's segment, and the Slate feature on the 'Year of Outrage'. For each firestorm we collected the start date, the date of any apology or retraction if it exists, and all hashtags and handles associated with the firestorm (some firestorms are centered around a particular user, others around a hashtag).

C. Data source

Our data source is an archive of the Twitter decahose, a random 10% sample of all tweets. This is a scaled up version of Twitter's Sample API, which gives a stream of a random 1% sample of all tweets. As found by Morstatter et al. [41], the Sample API (unlike the Streaming API) indeed gives an accurate representation of the relative frequencies of hashtags over time. We assume that the decahose has this property as well, with the significant benefit that it gives us more statistical power to estimate the true size of smaller events.

The decahose, like the Sample API, does not allow queries regarding the social graph, thus preventing us from modeling individual exposure to information [46]. And because information about when links were formed is not stored by Twitter, it is difficult to reconstruct the state of the social graph at a previous point in time [39]. However, following the demonstration in [23], we use mentions as a proxy for ties. We recognize that any given mention has only a 1/10 chance of being in our data, but this means that we are, on average, capturing ties that consist of at least 10 mentions. Since many of the mention networks we find are fairly dense, keeping only ties consisting of at least ten mentions would be justifiable even as a filtering strategy.

D. Data extraction

We do pre-processing on the decahose data to simplify the computational task, extracting (1) daily summaries of copresent entities and the user who posted that tweet (e.g., if @user tweets "@alter #tag1 #tag2", we would record the copresence of @user, @alter, #tag1, and #tag2), and (2) daily tabulations of hashtag and mention frequencies for all such entities in the data by day. The aggregation by day is by the UTC timestamp in the tweets, which potentially splits firestorms across days as experienced by firestorm participants. Fine-grained extraction may be appropriate for future work. For each candidate hashtag or mention, we extracted the daily frequency from -5 days (for a baseline) to +60 days (for a tail) from the start date. We found that the tails died off well within 10 days, such that a smaller period would be sufficient for future extractions.

In addition to extracting frequency plots for all firestorms, and often for multiple entities (hashtags and mentions) for each firestorm to see if the firestorm was better captured by one entity or another, we extracted the full text and metadata of tweets for the 20 firestorms with the highest volume of tweets on their respective peak days. For these 20 we also constructed *mention networks* of all firestorm participants. This consisted of taking all usernames found in the firestorms (i.e., all uses who included in at least one tweet with the entity by which we identified the given firestorm) and extracting all mentions between them during the firestorm (including tweets not containing the firestorm entity). We did not consider mentions by or of users not participating in the firestorm. In order to do a pre- and post-firestorm comparison, we similarly collected all mentions between firestorm participants going back to two weeks before the firestorm, and forward to two weeks after the firestorm. We aggregated these into networks by one-week intervals. In order for mentions of the target's Twitter handle (when there is a clear target) during the firestorm to not drown out other structure, we remove the node of the target handle during the firestorm week. For consistency, we remove the target handle from other weeks as well.

For spam filtering, we first did qualitative investigation of the data. Most of what we identified were tweets that contained a URL and a string of unrelated hashtags, e.g.,

> NEW F O L L O W E R S=>[a URL here] #DescribeYourCrushIn3Words,#Brentto600k, #CancelColbert,#ULTRALIVE,#HowOldAreYou,Napier an

This led us to investigate a rule-based filtering system similar to that employed by Kwak et al. [36]. However, as it turned out, spam tweets of this form accounted for less than half a percent of the total volume of tweets. Investigating the top hashtags for that day revealed no overlap, suggesting that the spam captured in our data was from spambots employing a minority strategy of tweeting out all currently trending topics. Because the volume of spam was negligible, and investigation showed the top tweeters in both of our case studies were indeed humans, we decided to not employ any filtering.

IV. Results

A. Firestorms

For each of the 80 firestorms, we identified the one entity that best represents the firestorm, the number of tweets posted related to it in the first 7 days of the event, and the number of unique users who participated in this event. The day on which maximum activity in number of tweets was observed is referred to as the day of peak activity. We found that most of our firestorms have an estimated peak volume of below 50,000[1] (fig. 1). The outlier with over 200,000 tweets is #WhyImVotingUKIP, although we didn't investigate why this might have been so large compared to other firestorms.

By comparing the date of the initial event to the date of the peak, we can see how quickly the firestorm reached peak activity; this was generally on the start date, for 64 out of the 80 cases. We can see from figure 3 that the larger firestorms in our collection do not take longer to decay, suggesting a phenomenon not tied to scale.

As mentioned above, we focus on the 20 firestorms with the largest number of tweets on their peak day. Brief descriptions

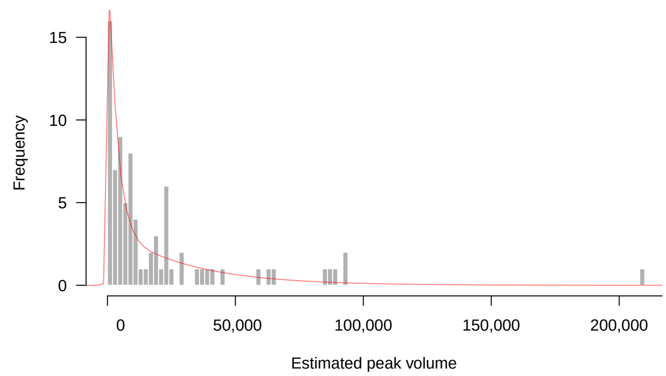

Fig. 1. Histogram of peak sizes of collected firestorms, with a scaled fitted logspline. The x-axis is the estimated volume of tweets reached on the peak day, and the y-axis is the number of firestorms reaching that volume.

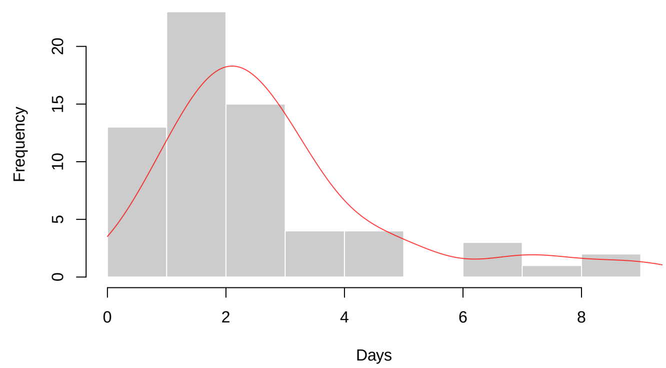

Fig. 2. Histogram of distribution of the number of days it took the collected firestorms to decay to 90% of peak volume, with scaled fitted density.

of each of the firestorms, along with their respective numbers of tweets and dates, is given in Table III. We excluded from the table and from fig. 2 any cases where the time to decay was more than 10 days. In such cases, the firestorm tweets did not exceed one-tenth of the average total volume of the given entity; in such cases, the firestorm likely would matter little to the target and may not even be noticeable, an indication that all things called firestorms are not necessarily meaningful to call as such.

B. Case studies

The volumes of our two case studies, #CancelColbert and #myNYPD, are shown in figures 4 and 5 respectively. For these, we chose to show 16-day period because, after that, there was negligible activity (there was also very little activity in #myNYPD after the first week, but we included the same length of time for comparison). In both plots, the bimodal initial peak corresponds to the hours from late night to early morning.

Since the number of new users is almost identical to that of the total tweet frequency, we also provide a log-log plot of the distribution of tweets per user in figure 6 which shows a typical heavy-tailed distribution (here we provide raw decahose numbers as scaling by 10 would lose the head of the distribution, where most of the mass is).

[1]The estimate is number of tweets observed from the decahose multiplied by 10 (sampling rate)

Firestorm Hashtag/Mention	Start Date	Tweets	Users	Source with description of the firestorm
#whyimvotingukip	2014-05-20	39,969	32,376	knowyourmeme.com/memes/events/whyimvotingukip
#muslimrage	2012-09-17	15,722	11,947	buzzfeed.com/ryanhatesthis/newsweeks-muslim-rage-cover-sparks-immediate-ba#.ljGJy1ykw
#CancelColbert	2014-03-27	13,277	10,349	newyorker.com/news/news-desk/the-campaign-to-cancel-colbert
#myNYPD	2014-04-22	12,762	10,362	knowyourmeme.com/memes/events/mynypd
#AskThicke	2014-06-30	11,763	9,699	knowyourmeme.com/memes/events/askthicke
@TheOnion	2013-02-24	9,959	8,802	hollywoodreporter.com/news/onion-calls-quvenzhane-wallis-c-424113
@KLM	2014-06-29	8,716	8,050	utsandiego.com/news/2014/jun/29/klm-adios-amigos-twitter-mexico/
#qantas	2011-10-26	8,649	5,402	reuters.com/article/2011/11/22/us-qantas-idUSTRE7AL0HB20111122
@David_Cameron	2014-03-05	7,096	6,447	theguardian.com/politics/2014/mar/06/celebrities-parody-camerons-on-the-phone-to-obama-selfie-tweet
@celebboutique	2012-07-20	6,679	6,189	mashable.com/2012/07/20/celebboutique-misguided-aurora-tweet-sparks-twitter-outrage/
@GaelGarciaB	2014-06-29	6,646	6,234	utsandiego.com/news/2014/jun/29/klm-adios-amigos-twitter-mexico/
#NotIntendedToBeAFactualStatement	2011-04-13	6,261	4,386	knowyourmeme.com/memes/events/not-intended-to-be-a-factual-statement
#AskJPM	2013-11-06	4,321	3,418	dealbook.nytimes.com/2013/11/13/after-twitter-fail-jpmorgan-calls-off-q-and-a/?_r=0
@SpaghettiOs	2013-12-06	2,890	2,704	huffingtonpost.com/2013/12/07/spaghettios-pearl-harbor-tweet_n_4404397.html
#McDStories	2012-01-18	2,374	1,993	businessinsider.com/mcdonalds-twitter-campaign-goes-horribly-wrong-mcdstories-2012-1
#AskBG	10-17-2013	2,221	1,933	bbc.com/news/business-24563421
#QantasLuxury	2011-11-22	2,098	1,657	reuters.com/article/2011/11/22/us-qantas-idUSTRE7AL0HB20111122
#VogueArticles	2014-09-10	1,894	1,819	washingtonpost.com/blogs/style-blog/wp/2014/09/10/vogues-celebration-of-big-butts-falls-flat-and-inspires-voguepitches/
@fafsa	2014-06-25	1,828	1,692	nbcnews.com/business/personal-finance/fafsa-im-poor-tweet-sparks-online-backlash-n140356
@UKinUSA	2014-08-24	142	140	washingtonpost.com/blogs/post-politics/wp/2014/08/25/british-embassy-apologizes-for-tweet-commemorating-the-burning-the-white-house-but-not-for-the-actual-burning-of-the-white-house/

TABLE III. Top 20 firestorm events from Feb, 2011 to September, 2014, sorted by the number of tweets.

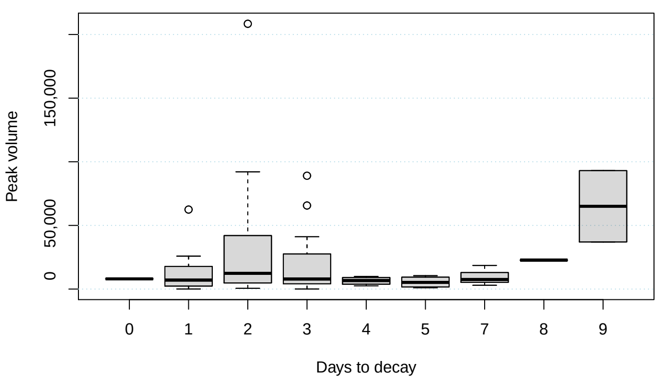

Fig. 3. Boxplot of peak volume versus the days it took to decay to 90%.

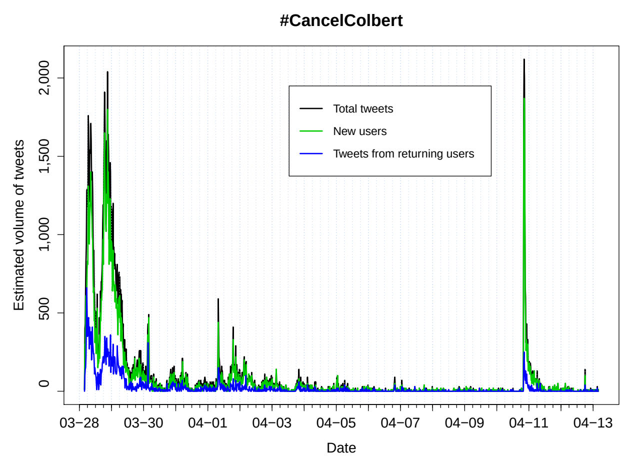

Fig. 4. Estimated number of tweets with hashtag #myNYPD (case-insensitive) over 16 days in 2014, plotted at the midpoints of bins with a width of 16.67 minutes. Ticks are at 6 hour intervals, UTC. The number of tweets from users using the hashtag for the first time is used to measure the number of new users. Since the total number of tweets is very close to this number, we include the difference between the two, which are the tweets from returning users.

For #CancelColbert, the first peak visible in the plot is the initial Twitter discussion after the tweet from @suey_park. Colbert discussed the campaign on his show on the night of March 31st in a segment entitled, "Who's Attacking Me Now? - #CancelColbert", but this had little impact. The second peak is, interestingly, from April 10th, the day of a press release from CBS (later also covered on *The Colbert Report* that night) announcing that Colbert would be leaving his show to become the host of the famous American late-night show *The Late Show*. Much of the content of that spike were jokes about #CancelColbert having worked. The vast majority of these tweets were from users who had not participated in the initial firestorm (see fig. 4) and, further analyzing the tweets, there were almost no additional mentions of the users who were heavily mentioned before: @StephenAtHome has an estimated 14,190 mentions in the first 13 days and 1,610 in the next 13 days, and @suey_park has an estimated 11,970 mentions in the first period and only 980 in the second. This suggests far lower levels of interaction for the event that was not a firestorm than in the event that was, a topic for future exploration.

In #myNYPD, the users with the highest tweet volume appear to be members of the public, with the exception of @Copwatch, an activist network. The profiles with the highest indegree (most mentions) is revealing: while @NYPDnews is most mentioned, with an estimated 15,180 mentions (almost all in the initial 13-day period), second-most is @Occupy-WallStreetNYC, with 10,860, followed by @YourAnonNews with 5,620, @Copwatch with 4,390 and @VICE with 3,580. The frequent mentions of Occupy show linking back to recent police action at Zuccotti Park against protestors from the Occupy Wall Street movement. The frequent mentions of @VICE are tweets linking to an article [2] about the firestorm

[2]"Disastrous #myNYPD Twitter Campaign Backfires Hilariously," https://news.vice.com/article/disastrous-mynypd-twitter-campaign-backfires-hilariously

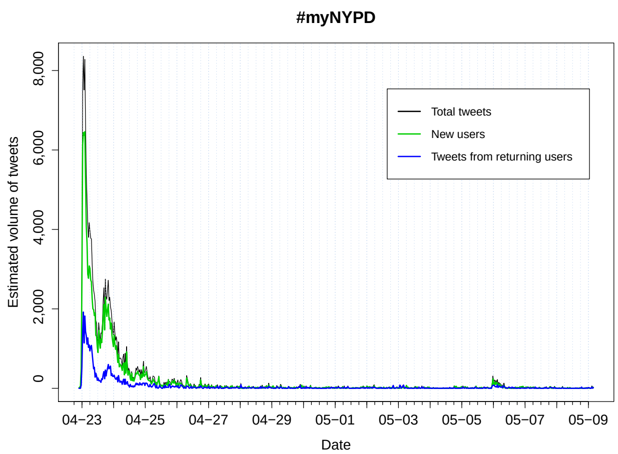

Fig. 5. Estimated number of tweets with hashtag #CancelColbert (case-insensitive) over 16 days in 2014, with other details identical to that in fig. 4.

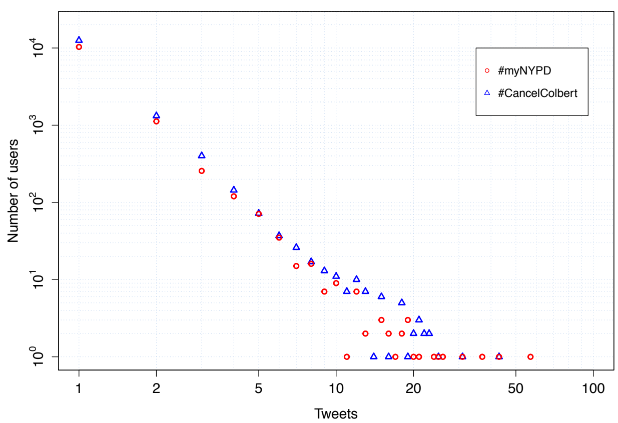

Fig. 6. Distribution of tweets per user in the decahose (log-log scale).

quickly published on the website of *Vice* magazine, giving one possible example of a feedback loop between media and firestorms.

C. Mention networks

As discussed earlier, we create mention networks to study change in social interactions pre- and post-firestorm. When investigating the mention networks, we saw a characteristic pattern: the network of the week of the firestorm looked dramatically different from the others (fig. 7), with far more concentration and far less of a distributed network structure. Some of the normally present conversational structure seemed to disappear. We investigated a number of global network metrics (density, centralization, clustering coefficient/transitivity, reciprocity), but even when a given metric changed across networks from week to week, the change was not so great that 95% confidence intervals from week to week did not overlap. However, we found that there were far fewer edges in common between the firestorm week and the other weeks. We measured this formally with a Jaccard index (size of intersection divided by size of union) on the directed edges of the mention networks.

Figure 8 shows the distributions over the 20 top firestorms between each pair of weeks; we add vertical lines at the mode as it makes the difference more noticeable than lines at the means, but t-tests for comparisons of means still show that the difference in means between a non-firestorm week and a firestorm week is significant in all cases, and between any two non-firestorm weeks is non-significant. Surprisingly, even the pre-firestorm weeks and post-firestorm weeks were more similar to *each other* than to the firestorm, indicating that there is a minimum underlying social structure of discussion, relatively constant in time, but from which a firestorm departs.

The similarity between pre- and post-firestorm weeks' mention networks, and the dissimilarity between all of these networks and the firestorm mention networks, still does not show whether or not a firestorm had an effect on the network structure. To investigate this, we constructed comparison 'panels' by randomly sampling from the decahose users who tweeted during the week of the event but not about the firestorm itself. We again generated mention networks across five weeks for these users. This time, we looked at the Jaccard indexes between weeks -1 and +1, and between weeks -2 and +2, and compared the distribution of these Jaccard indexes from networks of randomly sampled users to networks of firestorm participants. We found that the difference in means was not significant. That is, the way in which firestorms may change the mention networks of participants is not significantly different from the churn in networks that we would expect by random chance.

V. Discussion

Our research question was about the relationship between social ties and firestorm participation. We find first that the mention networks pre- and post-firestorm are more similar to each other than to the mention network of the week of the firestorm. If firestorms emerged from existing networks, we would expect to find more similarity between firestorm mention graphs and pre-firestorm mention graphs. Conversely, if firestorms had created lasting links among participants, we would expect to find more similarity between firestorm mention graphs and post-firestorm mention graphs. Instead, we found low similarity between firestorms and other weeks. We further find by comparison to a randomly sampled group that we cannot find the firestorm had *any* discernible impact on patterns of discussion. Going back to our theoretical motivations, it seems that at least among the firestorms we sample, we see no evidence of the type of social change associated with action that has biographical consequences on participants. This suggests that, at least along this dimension, firestorms should not be a source of anxiety for targets nor a source of satisfaction for opponents; firestorms in general do not create the conditions to lead to larger and more long-term actions, at least among the mass of participants.

VI. Conclusion

We have identified that across events identified as 'firestorms,' there is a departure from otherwise regular patterns of social interactions. Since both pre- and post-firestorm mention networks are different from firestorm mention networks, but the pre- and post-firestorm networks are similar to each other, it seems that the firestorms do not have a significant

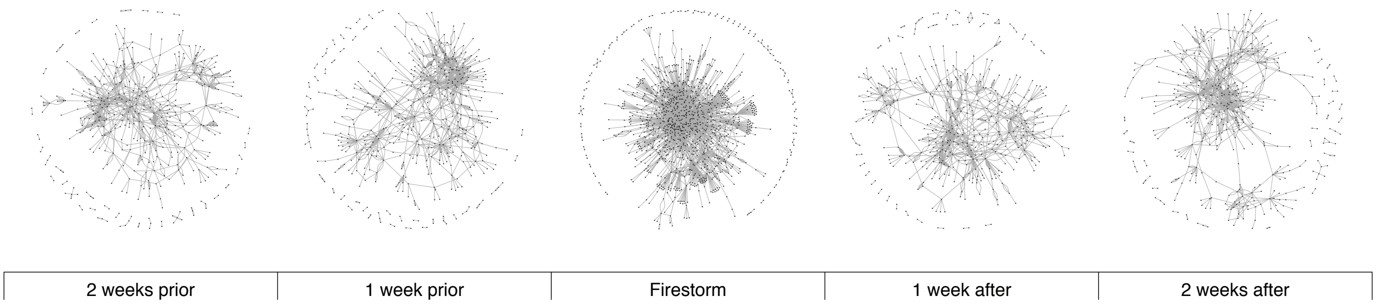

Fig. 7. Network of mentions between firestorm participants, in this case for #askJPM, aggregated by week, before, during, and after the event.

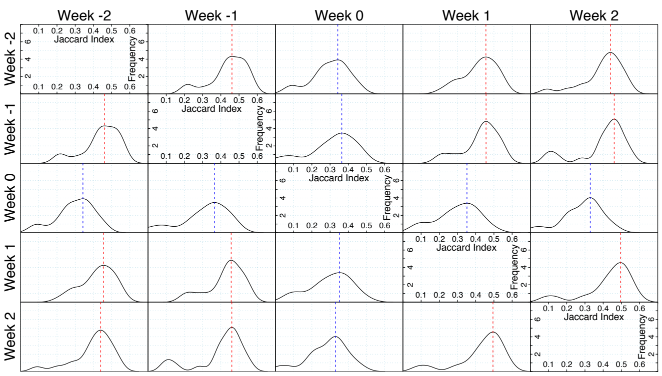

Fig. 8. Distributions of the Jaccard index of edges between the mention networks two weeks before, one week before, during, one week after, and two weeks after the firestorms. Vertical lines are put at the mode of each distribution. The matrix is symmetric; this redundancy is provided for ease in vertical comparisons. The figure shows that the networks after the firestorms resemble much more the networks before the firestorm than during the firestorm.

impact on communities. From our theoretical background, this finding suggests that firestorms will generally have little long-term impact. We believe that there are still interesting future research directions, including for basic research, including:

- Distinguishing sarcasm using ties and temporal clues;
- Firestorms as subsets of the Twitter ecosystem with different spam dynamics;
- Event detection and decay modeling of a specific, emotional and social type of event;
- Feedback effects on firestorms of simultaneous media coverage;
- Identifying the target of negative statements; for example, negative #CancelColbert tweets may be angry with Colbert or with the campaign.

VII. Acknowledgements

This material is supported by the Army Research Office under Award No. W911NF1410481, LexisNexis and HPCC Systems. Momin is supported by an award from the ARCS Foundation.

References

[1] T. Nitins and J. Burgess, "Twitter, brands, and user engagement," in *Twitter and society*, K. Weller, A. Bruns, J. Burgess, M. Mahrt, and C. Puschmann, Eds. New York: Peter Lang, 2014, pp. 293–304.

[2] J. Pfeffer, T. Zorbach, and K. M. Carley, "Understanding online firestorms: Negative word-of-mouth dynamics in social media networks," *Journal of Marketing Communications*, vol. 20, no. 1–2, pp. 117–128, 2014. [Online]. Available: http://dx.doi.org/10.1080/13527266.2013.797778

[3] J. Ronson, "How one stupid tweet blew up Justine Sacco's life," *New York Times Magazine*, February 2015. [Online]. Available: http://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html

[4] D. McAdam, "The biographical consequences of activism," *American Sociological Review*, vol. 54, no. 5, pp. 744–760, 1989. [Online]. Available: http://www.jstor.org/stable/2117751

[5] A. Bruns and S. Stieglitz, "Quantitative approaches to comparing communication patterns on Twitter," *Journal of Technology in Human Services*, vol. 30, no. 3–4, pp. 160–185, 2012. [Online]. Available: http://dx.doi.org/10.1080/15228835.2012.744249

[6] J. Park, M. Cha, H. Kim, and J. Jeong, "Managing bad news in social media: A case study on Domino's pizza crisis," in *ICWSM '12: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2012, pp. 282–289. [Online]. Available: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4672

[7] J. Rajasekera, "Crisis management in social media and digital age: Recall problem and challenges to toyota," Research Institute, International University of Japan, Working Papers EMS_2010_06, 2010. [Online]. Available: http://www.iuj.ac.jp/im-info/Publications/pdf/IM-2010-02.pdf

[8] A. Mochalova and A. Nanopoulos, "Restricting the spread of firestorms in social networks," in *ECIS 2014: Twenty Second European Conference on Information Systems*, 2014.

[9] S. Stieglitz and N. Krüger, "Public enterprise-related communication and its impact on social media issue management," in *Twitter and society*. New York: Peter Lang, 2014, pp. 281–292.

[10] S. Stieglitz and N. Krüger, "Analysis of sentiments in corporate Twitter communication: A case study on an issue of Toyota," in *ACIS 2011: Australiasian Conference on Information Systems*, 2011. [Online]. Available: http://aisel.aisnet.org/acis2011/29/

[11] D. Fan, "Tweets are not public opinion but can be used to predict public opinion," University of Minnesota, Working Papers, 2012. [Online]. Available: http://www.mapor.org/confdocs/absandpaps/2012/2012_papers/1B1_Fan_paper.pdf

[12] B. Heil and M. Piskorski, "New Twitter research: Men follow men and nobody tweets," *Harvard Business Review*, 6 2006. [Online]. Available: https://hbr.org/2009/06/new-twitter-research-men-follo/

[13] J. van Dijck, "Chapter 4: Twitter and the paradox of following and trending," in *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, 2013, pp. 68–88.

[14] P. L. Vasterman, "Media-hype: Self-reinforcing news waves, journalistic standards and the construction of social problems," *European Journal of Communication*, vol. 20, no. 4, pp. 508–530, 2005. [Online]. Available: http://dx.doi.org/10.1177/0267323105058254

[15] C. Wien and C. Elmelund-Præstekær, "An anatomy of media hypes: Developing a model for the dynamics and structure of intense media coverage of single issues," *European Journal of Communication*, vol. 24, no. 2, pp. 183–201, 2009. [Online]. Available: http://dx.doi.org/10.1177/0267323108101831

[16] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *KDD '09: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2009, pp. 497–506. [Online]. Available: http://dx.doi.org/10.1145/1557019.1557077

[17] A. J. Kimmel and P. J. Kitchen, "WOM and social media: Presaging future directions for research and practice," *Journal of Marketing Communications*, vol. 20, no. 1-2, pp. 5–20, 2014. [Online]. Available: http://dx.doi.org/10.1080/13527266.2013.797730

[18] E. Morozov, "From slacktivism to activism," *Foreign Policy*, September 2009. [Online]. Available: http://foreignpolicy.com/2009/09/05/from-slacktivism-to-activism/

[19] M. White, "Clicktivism is ruining leftist activism," *The Guardian*, August 2010. [Online]. Available: http://www.theguardian.com/commentisfree/2010/aug/12/clicktivism-ruining-leftist-activism

[20] D. Carr, "Hashtag activism, and its limits," *New York Times*, March 2012. [Online]. Available: http://www.nytimes.com/2012/03/26/business/media/hashtag-activism-and-its-limits.html

[21] H. Christensen, "Political activities on the internet: Slacktivism or political participation by other means?" *First Monday*, vol. 16, no. 2, 2011. [Online]. Available: http://firstmonday.org/ojs/index.php/fm/article/view/3336

[22] M. Sauter, "'LOIC will tear us apart': The impact of tool design and media portrayals in the success of activist DDOS attacks," *American Behavioral Scientist*, vol. 57, no. 7, pp. 983–1007, 2013. [Online]. Available: http://dx.doi.org/10.1177/0002764213479370

[23] S. Merritt, A. Jacobs, W. Mason, and A. Clauset, "Detecting friendship within dynamic online interaction networks," in *ICWSM '13: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 2013, pp. 380–389. [Online]. Available: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6049

[24] S. A. Myers and J. Leskovec, "The bursty dynamics of the Twitter information network," in *WWW '14: Proceedings of the 23rd International Conference on World Wide Web*, 2014, pp. 913–924. [Online]. Available: http://dx.doi.org/10.1145/2566486.2568043

[25] M. Granovetter, "The strength of weak ties: A network theory revisited," *Sociological Theory*, vol. 1, pp. 201–233, 1983.

[26] S. A. Williams, M. M. Terras, and C. Warwick, "What do people study when they study Twitter? Classifying Twitter related academic papers," *Journal of Documentation*, vol. 69, no. 3, pp. 384–410, 2013. [Online]. Available: http://dx.doi.org/10.1108/JD-03-2012-0027

[27] D. Ruths and J. Pfeffer, "Social media for large studies of behavior," *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014. [Online]. Available: http://www.sciencemag.org/content/346/6213/1063.short

[28] Z. Tufekci, "Big questions for social media big data: Representativeness, validity and other methodological pitfalls," in *ICWSM '14: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 2014. [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8062

[29] A. Mislove, S. Lehmann, Y.-Y. Ahn, J.-P. Onnela, and J. Rosenquist, "Understanding the demographics of Twitter users," in *ICWSM '11: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011. [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2816

[30] M. Malik, H. Lamba, C. Nakos, and J. Pfeffer, "Population bias in geotagged tweets," in *ICWSM '15: Papers from the 2015 ICWSM Workshop on Standards and Practices in Large-Scale Social Media Research*, 2015, pp. 18–27, AAAI Technical Report WS-15-18. [Online]. Available: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10662

[31] S. Ghosh, B. Viswanath, F. Kooti, N. K. Sharma, G. Korlam, F. Benevenuto, N. Ganguly, and K. P. Gummadi, "Understanding and combating link farming in the Twitter social network," in *WWW '12: Proceedings of the 21st International Conference on World Wide Web*, 2012, pp. 61–70. [Online]. Available: http://dx.doi.org/10.1145/2187836.2187846

[32] J. Donath, "Signals in social supernets," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 231–251, 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1083-6101.2007.00394.x

[33] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson, "Trafficking fraudulent accounts: The role of the underground market in Twitter spam and abuse," in *USENIX Security '13: Proceedings of the 22nd USENIX Security Symposium*. Washington, D.C.: USENIX, 2013, pp. 195–210. [Online]. Available: https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/thomas

[34] K. Thomas, C. Grier, D. Song, and V. Paxson, "Suspended accounts in retrospect: An analysis of Twitter spam," in *IMC '11: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, 2011, pp. 243–258. [Online]. Available: http://dx.doi.org/10.1145/2068816.2068840

[35] d. boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter," in *HICSS '10: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, 2010, pp. 1–10. [Online]. Available: http://dx.doi.org/10.1109/HICSS.2010.412

[36] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *WWW '10: Proceedings of the 19th International Conference on World Wide Web*, 2010, pp. 591–600. [Online]. Available: http://dx.doi.org/10.1145/1772690.1772751

[37] A. Java, X. Song, T. Finin, and B. Tseng, "Why we Twitter: Understanding microblogging usage and communities," in *WebKDD/SNA-KDD '07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, 2007, pp. 56–65. [Online]. Available: http://dx.doi.org/10.1145/1348549.1348556

[38] E. Hargittai and E. Litt, "The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults," *New Media & Society*, vol. 13, no. 5, pp. 824–842, 2011. [Online]. Available: http://dx.doi.org/10.1177/1461444811405805

[39] M. Gabielkov and A. Legout, "The complete picture of the Twitter social graph," in *CoNEXT Student '12: Proceedings of the 2012 ACM Conference on CoNEXT Student Workshop*, 2012, pp. 19–20. [Online]. Available: http://dx.doi.org/10.1145/2413247.2413260

[40] G. Schoenebeck, "Potential networks, contagious communities, and understanding social network structure," in *WWW '13: Proceedings of the 22nd International Conference on World Wide Web*, 2013, pp. 1123–1132. [Online]. Available: http://dl.acm.org/citation.cfm?id=2488388.2488486

[41] F. Morstatter, J. Pfeffer, and H. Liu, "When is it biased?: Assessing the representativeness of Twitter's Streaming API," in *WWW Companion '14: Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion*, 2014, pp. 555–556. [Online]. Available: http://dx.doi.org/10.1145/2567948.2576952

[42] A. Bruns and H. Moe, "Structural layers of communication on Twitter," in *Twitter and society*, K. Weller, A. Bruns, J. Burgess, M. Mahrt, and C. Puschmann, Eds. New York: Peter Lang, 2014, pp. 15–28.

[43] B. Poblete, R. Garcia, M. Mendoza, and A. Jaimes, "Do all birds tweet the same? Characterizing Twitter around the world," in *CIKM '11: Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, 2011, pp. 1025–1030. [Online]. Available: http://dx.doi.org/10.1145/2063576.2063724

[44] Y. Liu, C. Kliman-Silver, and A. Mislove, "The tweets they are a-changin: Evolution of Twitter users and behavior," in *ICWSM '14: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 2014. [Online]. Available: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8043

[45] N. Krüger, S. Stieglitz, and T. Potthoff, "Brand communication in Twitter: A case study on Adidas," in *PACIS 2012: Pacific Asia Conference on Information Systems*, 2012. [Online]. Available: http://aisel.aisnet.org/acis2011/29/

[46] D. M. Romero, B. Meeder, and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on Twitter," in *WWW '11: Proceedings of the 20th International Conference on World Wide Web*, 2011, pp. 695–704. [Online]. Available: http://dx.doi.org/10.1145/1963405.1963503