



# Measuring and monitoring collective attention during shocking events

Xingsheng He<sup>1,2</sup> and Yu-Ru Lin<sup>2\*</sup>

\*Correspondence: [yurulin@pitt.edu](mailto:yurulin@pitt.edu)

<sup>2</sup>School of Computing and Information, University of Pittsburgh, Pittsburgh, PA 15260, USA

Full list of author information is available at the end of the article

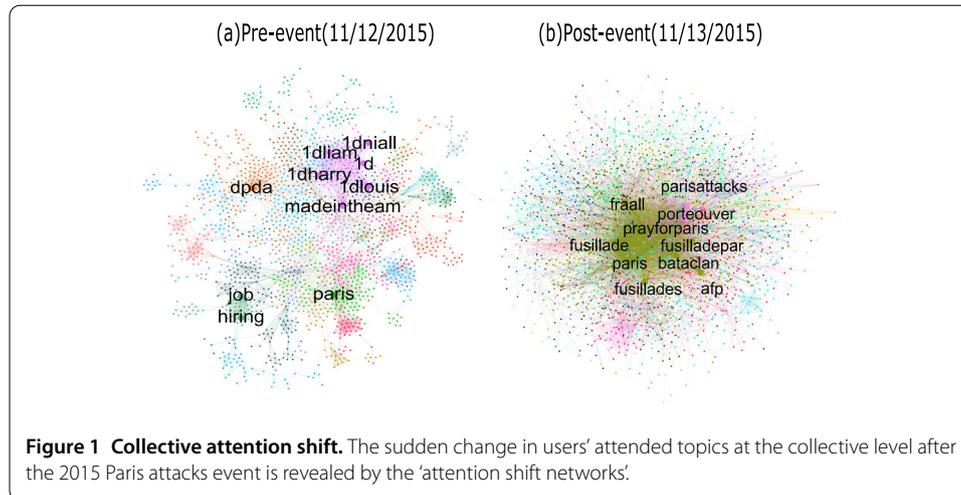
## Abstract

There has been growing interest in leveraging Web-based social and communication technologies for better crisis response. How might the Web platforms be used as an observatory to systematically understand the dynamics of the public's attention during disaster events? And how could we monitor such attention in a cost-effective way? In this work, we propose an 'attention shift network' framework to systematically observe, measure, and analyze the dynamics of collective attention in response to real-world exogenous shocks such as disasters. Through tracing hashtags that appeared in Twitter users' complete timeline around several violent terrorist attacks, we study the properties of network structures and reveal the temporal dynamics of the collective attention across multiple disasters. Further, to enable an efficient monitoring of the collective attention dynamics, we propose an effective stochastic sampling approach that accounts for the users' hashtag adoption frequency, connectedness and diversity, as well as data variability. We conduct extensive experiments to show that the proposed sampling approach significantly outperforms several alternative methods in both retaining the network structures and preserving the information with a small set of sampling targets, suggesting the utility of the proposed method in various realistic settings.

**Keywords:** collective attention; disaster response; collective intelligence; attention sampling

## 1 Introduction

The proliferation of Web-based social and communication technologies has provided an unprecedented opportunity for researchers to collect and study data about collective human behavior at large scales. In recent years, there has been growing interest in leveraging such technologies and data for better crisis response, with scope ranging across natural disasters [1], terrorist attacks [2], and political riots [3]. During these events, the flow and flood of information can easily lead to a poverty of attention and thus creates a need to allocate such attention efficiently for affected communities. Recent research also revealed the intrinsic heterogeneous levels of information load and showed that limited user attention can lead to a low discriminative capacity for users to identify better information from low-quality information [4]. *How might the Web platforms be used as an observatory to systematically understand the dynamics of the public's attention during disaster events? And how could we monitor such attention in a cost-effective way?* A systematic understanding of attention dynamics at the collective level within a disaster context serves as the basis for

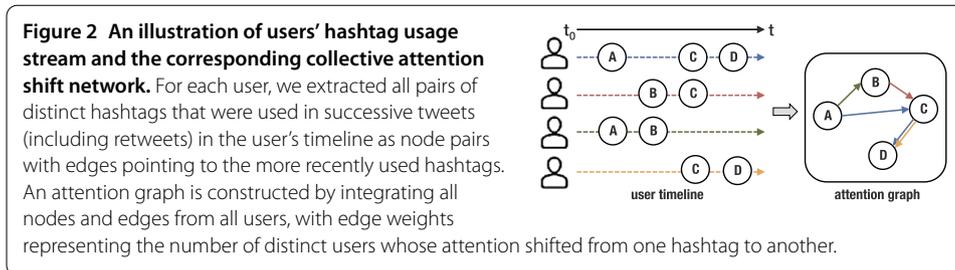


scheduling effective crisis communications, and facilitating timely crisis response such as just-in-time warning and evacuation.

In this work, we seek to quantitatively capture the collective attention shift under exogenous shocks, specifically terrorist events, by using Twitter users' communication streams. Figure 1 illustrates the collective attention *before* and *after* the 2015 Paris attacks event based on how Paris users shift their attention to various topics - captured by the use of different *hashtags*. Before the event, users' attended topics exhibited a salient community structure, reflecting their scattered attention among various topics. After the event happened, a few hashtags became the hubs that suddenly appeared in many users' tweets. Such sudden change in users' attended topics at the collective level is referred to as 'collective attention shift' in this work.

The vast digital traces available through social media platforms allow for exploring the patterns of collective attention; however, a systematic characterization of collective attention is non-trivial for the following reasons. First, the concept of 'collective attention' is not well defined, and a number of existing studies have related different quantities to collective attention [5–9], such as burstiness in tweet numbers [6] and popularity in news sharing [9]. Despite the abundant literature, there is a lack of formal definition that allows for operating the collective attention concept across different contexts. Second, collective attention dynamics correspond to the switching between the disordered and synchronized states of multiple individuals [10]. Characterizing such dynamics at the collective level and how it manifests under real-world exogenous shocks have not been explored. Furthermore, while social media data make the analysis of collective attention possible, the vast communication streams contributed from a large set of users have challenged the feasibility of monitoring collective attention in practice. While there have been works on retrieving event-specific information from the tweet streams [11] or distilling sub-topics from user posted content given a textual query [12], there has not been a systematic approach for tracking users' focal point and attention shift from the user timelines.

In this work, we introduce a new framework for capturing collective attention shift. We illustrate our framework by using a large corpus of twitter communications centered around multiple shocking terrorist attacks in 2015 and 2016. We employ hashtags as a proxy for users' attended topics, and utilize the hashtag adoption sequence from a user's tweet timeline as the trace of his/her attention shift process. A hashtag in a tweet usually



represents the central idea or the specific topic of the content. Hashtags are widely used under different settings, like event discussion, commercial promotion, disaster relief and so on. In our collection, we found that, on average, 25% of tweets contain hashtags (as shown in Additional file 1). The transition between hashtags for a user thus can be regarded as an instance for his/her attention shifts between topics. Also, users tend to adopt more hashtags when a large-scale event strikes (can be seen in Additional file 1). When we integrate the hashtag transitions from more users, we are more likely to cover sufficient details/on-going events of a community. As shown in Figure 2, we construct an *attention shift network* or *attention graph* to represent the attention shift process at the collective level.

The network representation allows us to capture the distinct patterns of collective attention shift during disaster events. Based on this formal representation, we can quantify the structural change of collective attention shift and further examine data sampling schemes that can capture the structural change in a cost-effective manner.

The main contributions of this work include:

- We propose a novel framework to represent and measure collective attention shift. Based on this, we systematically study the collective attention during multiple shocking terrorist attack events in 2015 and 2016 and reveal several properties of network structures and temporal dynamics that are consistent across events.
- We formulate a new problem for efficient monitoring of the collective attention dynamics, and we propose a cost-efficient sampling strategy that takes the users' hashtag adoption frequency, connectedness and diversity into account, with a stochastic sampling algorithm to cope with the variability of the sampling targets.
- We conduct extensive experiments and show that our proposed sampling approach significantly outperforms several alternative methods in both retaining the network structures and preserving the information with a small set of sampling targets, suggesting the utility of the proposed method in various realistic settings.

## 2 Related work

### 2.1 Collective attention

A growing number of existing studies have explored the idea of 'collective attention' and related it to different quantities extracted from user-generated data. For example, Lehmann et al. [5] studied collective attention by analyzing the temporal hashtag adoption patterns. Sasahara et al. [6] detected the burst-like increase in tweet numbers and semantic terms as a sign of emerging collective attention. Lin et al. [13] studied the conditions of shared attention as many users simultaneously tune in with the dual screens of broadcast and social media to view and participate. He et al. [14] discovered that the crowd was able to recognize and re-share relevant logistical messages from many irrelevant ones during a crisis

event. Wu and Huberman [9] modeled collective attention through the decay of popularity in news items shared in social media. Wang et al. [8] employed users' clicks stream on web forms as a proxy of collective attention. Other studies have focused on different aspects of collective attention, such as predicting future trending topics [7], classifying different types of collective attention [5], detecting the collective attention spam [15], and how information spreads in social networks with attention constraint [16].

The study of collective human behaviors has been a central topic in disaster management [17–19]. There have been both empirical studies [20] and modeling approaches [21] focusing on the collective human behavior under extraordinary events. Preis et al. [22] found that the number of photos uploaded to Flickr related to Hurricane Sandy bears a striking correlation to the atmospheric pressure in the US state New Jersey during that period. Borge-Holthoefer et al. [23] used an information theoretical approach to define and measure the temporal and structural signatures typical to collective social events as they arise and gain prominence.

However, there has been limited work that systematically examines the changes in the patterns of collective attention induced by exogenous events. In this work, by using an attention shift network framework, we provide a novel definition of collective attention and systematically analyze the structural changes of collective attention during disasters.

## 2.2 Data sampling

In this work, we investigate an effective and efficient data sampling scheme in order to achieve timely monitoring of collective attention. Our work is relevant to but different from 'subgraph sampling.' In subgraph sampling, the goal is to construct a subgraph by selecting a small set of nodes while preserving the topological properties - such as degree distribution, path length [24], clustering coefficient and network diameter [25] - of the original graph. In this work, we propose a new sampling problem that samples the set of users while evaluating the sampling quality on their attention shift captured from the hashtag graphs. The sampling approaches discussed in this paper are all unsupervised.

Existing subgraph sampling methods can be roughly classified into two types: graph traversal methods and random walk based methods. In graph traversal methods, each node can only be visited once. Examples include breadth-first search, depth-first search, snowball sampling and forest-fire sampling [26], which are different in terms of the order of the visited nodes.

Random walk (RW) based methods [27, 28], in contrast, allow node re-visiting and are widely adopted due to their simplicity and efficiency. The node selection in an RW based method is inherently biased towards high degree nodes [27]. Such bias can be quantified by classic Markov Chain analysis and can be adjusted via re-weighting of the estimators [29]. To sample nodes with different desired properties, such as node attribute stratification, RW can be modified by using the Metropolis filter, known as Metropolis-Hasting Random Walk (MHRW) [28], to achieve a desired stationary distribution.

In the study of social diffusion, there have been works focusing on sampling the most influential users in a social network that could trigger future large cascades [30, 31]. Munmun et al. [32] studied different sampling methods in terms of their effectiveness in discovering information diffusion on Twitter.

Most of the existing sampling methods, including the aforementioned subgraph sampling, are based on a one-mode network setting, while in this work, we formulate a novel

sampling problem where the goal is to reconstruct the collective attention shift network (where nodes are hashtags) by selecting a small set of users. Such sampling relies on the underlying dynamic, bipartite relationship (how users tend to use hashtags), which has not been explored in existing research.

### 3 Characterizing collective attention under shocks

#### 3.1 Data collection

To study collective attention shift during disasters, we collect tweets around multiple shocking terrorist attacks in 2015 and 2016 - the Paris attacks<sup>a</sup> on 13 November 2015, the San Bernardino shooting<sup>b</sup> on 2 December 2015, the Brussels bombings<sup>c</sup> on 22 March 2016 and the Orlando nightclub shooting<sup>d</sup> on 12 June 2016. For each event, we consider users who have frequently posted tweets around the event location prior to the event occurrence. Such users are considered as members of *affected communities*. As a comparison with these directly affected communities, we choose two other cities New York City and London, as members of *observer communities* for the Paris attack event. This is done for two key reasons: First, the Paris attacks represent a significant terrorist event that shocked the world and has drawn global attention. It is of the interest of risk managers, researchers, and scientists to examine how the attention arouse and faded in other, indirectly affected communities. Second, New York City and London have been concerned about the potential terrorist attack risk due to the past events, and hence the Paris event is likely to draw significant attention in the two cities. Besides, the two cities are also mega cities that have compatible scale with Paris.

The data were collected using an iterative process: First, to identify a group of relevant users, we obtained a set of users who posted geo-tagged tweets in the event city within four weeks after the event occurrence. Then, for each user, we traced back his/her historical tweets, including original tweets and retweets, through the Twitter REST API. We included both the original tweets and retweets because the retweeting behavior itself is an explicit signal that the user considers the tweet to contain useful or interesting information [33]. We further retained the users who have posted geo-coded tweets and have used at least one hashtag before the event. Each event collection includes these users' tweets from approximately two weeks before the event and one week after the event.

To compare the attention shift patterns during man-made disasters with other non-emergency events, we also collected the followers' tweets for the two major presidential candidates during the 2016 US presidential election. To properly identify the supporters of each candidate, we first collected all the followers for all the major candidates, and then we removed the users that had followed more than one candidate and obtained the 'exclusive followers' - users who only followed one candidate exclusively. For example, the Trump followers only followed Donald Trump and no other presidential candidates. After identifying the exclusive follower set, we traced back their historical data during the presidential election for two weeks in November 2016.

In total, we collected eight datasets consisting of eight affected/relevant communities for the five events: Four man-made disasters and one non-emergency event (the 2016 US presidential election). Table 1 lists the basic information for the eight datasets.

#### 3.2 Representing collective attention shift

We propose an *attention shift network* (or *attention graph*) representation to represent the attention shift process at the collective level. We employ hashtags as a proxy for users'

**Table 1 Datasets used in this study**

Dataset	Duration	# of users	# of tweets
Paris attacks (Paris users)	10/27/2015-11/20/2015	13,439	2.58 million
Paris attacks (New York users)	10/27/2015-11/20/2015	20,022	2.63 million
Paris attacks (London users)	10/27/2015-11/20/2015	12,053	1.57 million
Brussels bombings	3/5/2016-3/29/2016	9,797	2.60 million
San Bernardino attack	11/17/2015-12/09/2015	10,109	1.12 million
Orlando shooting	5/26/2016-6/19/2016	6,157	1.26 million
Trump Followers	10/29/2016-11/11/2016	22,659	0.70 million
Clinton Followers	10/29/2016-11/11/2016	26,810	0.73 million

attended topics. Figure 2 illustrates the representation - in an attention graph, nodes represent hashtags contained in users' tweets, and the directed edges represent the shifting hashtag adoption - an edge from node  $A$  to node  $B$  indicates that the use of hashtag  $A$  followed by  $B$  has appeared in at least one user's timeline, and the edge weight reflects how many unique users have such shifting (from  $A$  to  $B$ ). The formal definitions are provided below.

Given a set of users  $U$  and their timeline tweets, we represent each *hashtag adoption event* as a triple  $\langle u, h_i, p_i \rangle$  indicating a user  $u \in U$  posted a tweet containing hashtag  $h_i$  at time  $p_i$ . We then define a user  $u$ 's *shift* of attention as a *transition* from one hashtag to another at different time points, as follows:

**Definition** (Attention shift) An attention shift event of a user  $u$  is a transition between two of  $u$ 's consecutive hashtag adoption events, denoted as  $s = (\langle u, h_j, p_j \rangle | \langle u, h_i, p_i \rangle)$ , where  $h_i \neq h_j$ ,  $p_i < p_j$ , and there is no adoption event  $(u, h_m, p_m)$  s.t.  $h_i \neq h_m$ ,  $p_i < p_m < p_j$ .

Each attention shift event  $s = (\langle u, h_j, p_j \rangle | \langle u, h_i, p_i \rangle)$  captures a transition from hashtag  $h_i$  to hashtag  $h_j$ . We consider attention shift events that occur within a limited period of time. Given a set of users  $U$ , we denote a set of attention shift events occurring within a time period  $t = (p_t - \delta, p_t]$  for  $\delta > 0$ , as  $S_U^{(t)} = \{(\langle u, h_j, p_j \rangle | \langle u, h_i, p_i \rangle) : \forall u \in U, \forall p_i \in t \wedge p_j \in t\}$ . The collective *attention shift network*, or *attention graph*, corresponding to  $t$  is defined as:

**Definition** (Attention graph) An attention graph is a weighted, directed graph  $A^{(t)} = \{H, E_S\}$  induced from attention shift events  $S_U^{(t)}$  with respect to a user set  $U$  and a time period  $t$ , where  $H$  is a set of hashtags existing in  $S_U^{(t)}$  and  $E_S \subset H \times H$  is a set of transitions between hashtags. Each transition edge  $e_{ij} \in E_S$  represents the existence of transition between hashtags  $h_i$  and  $h_j$  captured by an attention shift event  $s \in S_U^{(t)}$ , with a weight  $\kappa_{ij} \in \mathbb{R}$  indicating the relative frequency of the transition.

In our study, to construct attention graphs from data, we first chronologically sorted users' tweets and extracted hashtags that appeared in these tweets. For each user, we extracted all pairs of distinct hashtags that were used in successive tweets (including retweets) in the user's timeline as node pairs with edges pointing to the more recently used hashtags. Then we aggregated all nodes and edges from all users in each event dataset to construct the attention graph, with edge weights representing the number of distinct users whose attention shifted from one hashtag to another. To avoid boundary effects arising due to transitions across time periods and to smooth out short-term fluctuations, all graphs are built with a rolling time window. In this work, each graph is built on an hourly

basis, with an  $n$ -hour window length rolling average (averaging over the previous  $n$  hours). We use  $n = 4$  for the Paris, Brussels, New York and London datasets, and  $n = 8$  for the remaining four datasets due to data sparsity. The discussion for how to determine a proper time window is further provided in Additional file 1.

### 3.3 Measuring collective attention shift

The proposed attention graphs allow us to quantitatively capture the collective attention shift using well-developed network analysis. We draw several important network metrics from network science and social network analysis [13, 34, 35], and we include two additional metrics to capture the emergence aspect of collective attention. Below we only describe the key metrics for brevity.

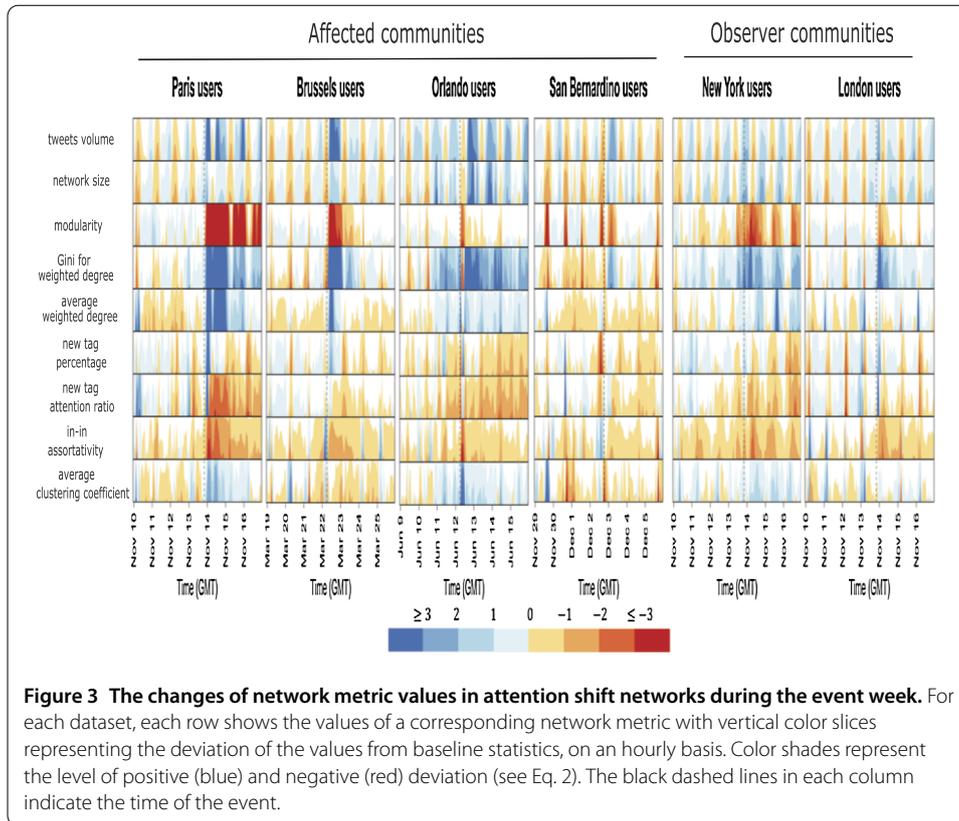
- *Network size*: the number of hashtags in the network.
- *Modularity*: the community structure exhibited in hashtag connections. A high modularity value indicates a clearly separated community structure. We leverage the Infomap algorithm [36] to compute the modularity of a directed, weighted network.
- *Average weighted degree*: the edge weights capture the number of users whose attention shifted from one hashtag to another; hence the average weighted degree of a network for a particular period of time reflects the attention shift frequency or rate.
- *Gini coefficient for weighted degree*: measures the level of degree concentration, denoting whether a few hashtags have become dominant in connecting with other hashtags. Gini coefficient ranges from 0 to 1, with 1 representing the highest concentrated attention. Like the power-law exponent, Gini coefficient can be used to measure the preferential patterns but in a more general way. We use the weighted distribution instead of the unweighted distribution as the weighted one allows for capturing the number of unique users that have shifted their attention.
- *Assortativity*: the tendency for a node to attach to others that are similar in terms of node degree. For a directed network, there are four types: in-in, in-out, out-in and out-out assortativity. In this work, we use the weighted in-in assortativity as defined in [37].
- *Average clustering coefficient*: the tendency of nodes to form triangles. In an attention shift network, this reflects the degree to which the collective attention is likely to shift at a local scale.
- *New tag percentage*: the number of newly emerged hashtags relative to the total number of hashtags in the network [38]. We consider a hashtag to be *new* if it has not been used within one week prior to the time of the network.
- *New tag attention ratio*: the percentage of weighted degrees given by the newly emerged hashtags. Specifically, the ratio  $r$  is defined as:

$$r = \frac{\sum_{j \in H_{\text{new}}} k_{\text{in}}^j}{\sum_{i \in H} k_{\text{in}}^i}, \quad (1)$$

where  $k_{\text{in}}^i$  is the weighted in-degree of node  $i$ , and  $H_{\text{new}}$  and  $H$  are the set of newly emerged hashtags and all hashtags in a network, respectively.

### 3.4 Observation: collective attention shift around terrorist attacks

When an unexpected event like disaster strikes, enormous event-related information suddenly draws public attention. How does social media users' collective attention change in



response to external shocks such as terrorist attacks? In this section, using the aforementioned metrics, we examined the collective attention shift patterns around several terrorist attacks.

For each of the six datasets, we construct time-dependent attention graphs for a one-week period centered on the event occurring time. We compute all the network metrics for each graph, and normalize the metric values against their *baseline* values to see how collective attention after shocks deviated from its pre-event state.

Let  $x_i$  represent a time-dependent value of a metric for time  $t_i$ . The normalized value is given by:

$$z_i = \frac{x_i - \mu_b}{\sigma_b}, \tag{2}$$

where  $\mu_b$  and  $\sigma_b$  are the mean and standard deviation of the metric values, respectively, measured in the baseline period prior to the event time. In this work, we consider the week prior to the event week as the baseline period.

As shown in Figure 3, we summarize the collective attention shift patterns from various time-dependent metric values by using *horizon graphs* [39]. A horizon graph allows for comparing and contrasting a large number of time series simultaneously. It divides a time-series chart into colored bands with hues differentiating the positive and negative values, and layers the bands with positive and negative values to the same region to create a nested form. In the plots, we discretize the normalized values into four positive bands (in blue) and four negative bands (in red). Bands with darker colors indicate that the corresponding regions are *significantly* deviated from the baseline values.

We highlight below the most salient patterns:

*Expansion:* increases in *tweet volume*, *network size* and *new tag percentage*. As shown in Figure 3, exogenous shocks triggered immediate spikes in the volume of tweets. Such escalations were more pronounced and lasted longer in the affected communities than in the observer communities. The attention graphs expanded as users tweeted with more hashtags (resulting in rises in network size) and more newly emerged hashtags (rises in new tag percentage) immediately after the onset of the events. This is consistent with prior work (e.g., [38]) that new hashtags are more likely to appear in users' timelines when a major event happens. During the Orlando shooting event, as the attack occurred late at night, the fraction of new hashtags remained relatively high for multiple hours until the next day.

*Agglomeration:* decrease in *modularity*. A higher modularity value indicates a more discernible community structure. In all datasets, we observe relative high modularity values from the attention graphs constructed before the events - e.g.,  $0.85 \pm 0.02$  for Paris users and  $0.83 \pm 0.05$  for Brussels users. This suggests users' attention tends to scatter over different topics in the normal state. As illustrated in Figure 1(a), the labeled hashtags represent different topical interests that users attended to before the attacks. After the attacks (Figure 1(b)), the community boundaries of these different topics blurred as many users began to adopt the set of event-related hashtags, resulting in a significant drop of modularity -  $0.64 \pm 0.15$  for Paris users and  $0.73 \pm 0.14$  for Brussels users.

*Concentration:* increase in *Gini coefficient*. After the attacks, there is an increase of Gini coefficient in all communities - e.g., from  $0.51 \pm 0.03$  to  $0.65 \pm 0.09$  for Paris users,  $0.51 \pm 0.06$  to  $0.57 \pm 0.05$  for Orlando users and  $0.55 \pm 0.02$  to  $0.58 \pm 0.03$  for New York users. This suggests that a small fraction of hashtags drew a disproportionately large amount of attention in both the affected and observer communities and became the focal points in the users' conversation on social media.

*Mutableness:* increase in *average weighted degree*. In attention graphs, edge weights represent the number of unique users who switch to use a hashtag from another. A higher value of average weighted degree reflects a relatively high frequency of hashtag switching in the network. We observe a significant increase in average weighted degree in almost all communities (e.g., Paris users:  $6.91 \pm 0.96$  to  $10.11 \pm 3.02$ ; Brussels users:  $12.33 \pm 1.59$  to  $14.56 \pm 4.76$ ), except for the San Bernardino user set ( $5.81 \pm 1.38$  before the attack and  $5.34 \pm 0.62$  after the attack) due to the relatively low activity in this community, as well as the London user set ( $7.64 \pm 1.32$ ,  $7.86 \pm 1.17$  before and after the attack, respectively) that represents an observer community. The increase in degrees suggests that users became more liable to switch hashtags in their tweet sequence immediately following the event.

*Re-mixing:* decrease in *assortativity*. A higher assortativity value indicates users are more likely to switch from a highly connected hashtag to another highly connected hashtag, or from a less connected hashtag to another less connected hashtag. We discover an evident decline in assortativity in almost all communities - especially for the affected communities (e.g., from  $0.02 \pm 0.05$  to  $-0.12 \pm 0.05$  for Paris users and  $-0.05 \pm 0.11$  to  $-0.13 \pm 0.08$  for Orlando users), suggesting that after the events, users became more likely to switch from a less connected hashtag to a highly connected hashtag, compared with the pre-event use. This pattern is consistent with the observation of post-event concentration.

Other metrics, such as the clustering coefficient, exhibit less prominent patterns consistent across all datasets. In sum, we observe a set of consistent properties of collective

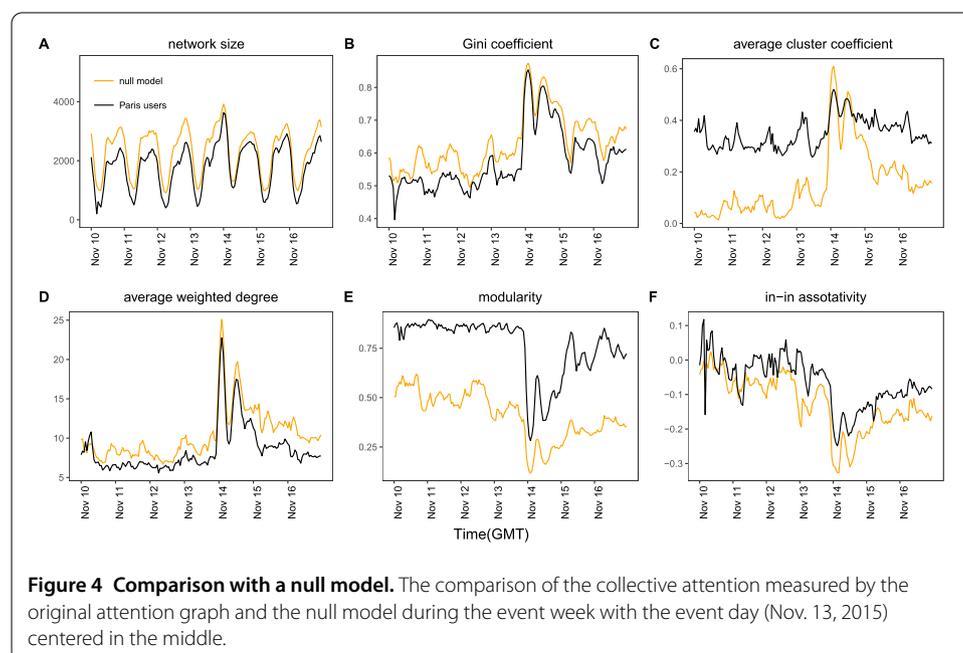
attention shift during these events, referred to as *CREAM* - Concentration, Re-mixing, Expansion, Agglomeration, and Mutableness. The acronym highlights the consistent characteristics observed during exogenous events in five aspects and it could help emphasize the multi-aspect characteristics of collective attention dynamics. The attention graphs allow for qualitatively characterizing the collective attention shift process, with evident statistical and structural changes under exogenous shocks.

### 3.5 Comparison with a null model

To evaluate whether the aforementioned patterns emerge due to users 'collectively shift' attention, or simply due to the change of individual activities, we compare the observed values with a 'baseline' created based on a null model - in which the frequencies of both user activities and the hashtags remain but the selection of hashtags are randomized. The experiment here is to examine whether the observed patterns differ from the baseline.

The null model is created as follows: For a fixed set of users in a period of time, we record the timestamp and the number of hashtags in each of their tweets, and then we extract all the hashtags they have used and randomly shuffle the hashtags and assign them to each user at each time of using a hashtag. After the shuffling procedure, the users' tweeting frequency and the frequency of a hashtag being used by all users remain the same, but which hashtags were used by which users were randomized.

Figure 4 shows the collective attention measured by the original attention graph and from the null model (through the shuffled hashtag selection) for the Paris users in the event week. The result indicates that the attention graphs constructed from the original tweet stream differ from those constructed from the null model in almost every aspect: The shuffling tends to result in more nodes and edges in the attention graphs, because the process tends to reduce the occurrences of isolated nodes - e.g.,  $2,327 \pm 780$  for the null model and  $1,793 \pm 756$  for the original attention graph in terms of *network size*,  $10.41 \pm 3.27$  for the null model and  $8.49 \pm 3$  for the original attention graph in terms of *average weighted*



*degree*. The shuffling also leads to a larger attention concentration:  $0.63 \pm 0.08$  compared with  $0.57 \pm 0.09$  in terms of *Gini coefficient*. The *Gini coefficient* is largely influenced by a few popular hashtags that are likely to appear in users' consecutive tweets, representing the hovered attention on a specific topic. However, the null model uniformly distributes these popular hashtags across different periods of time, making them more connected to other hashtags and resulting in a larger attention concentration. The attention graphs from the null model are also less clustered, they have a much smaller *modularity* ( $0.41 \pm 0.12$  for the null model and  $0.76 \pm 0.15$  for the original attention graph), smaller *average clustering coefficient* ( $0.16 \pm 0.14$  for the null model and  $0.35 \pm 0.06$  for the original attention graph) and smaller *assortativity* ( $-0.12 \pm 0.08$  for the null model and  $-0.07 \pm 0.07$  for the original attention graph).

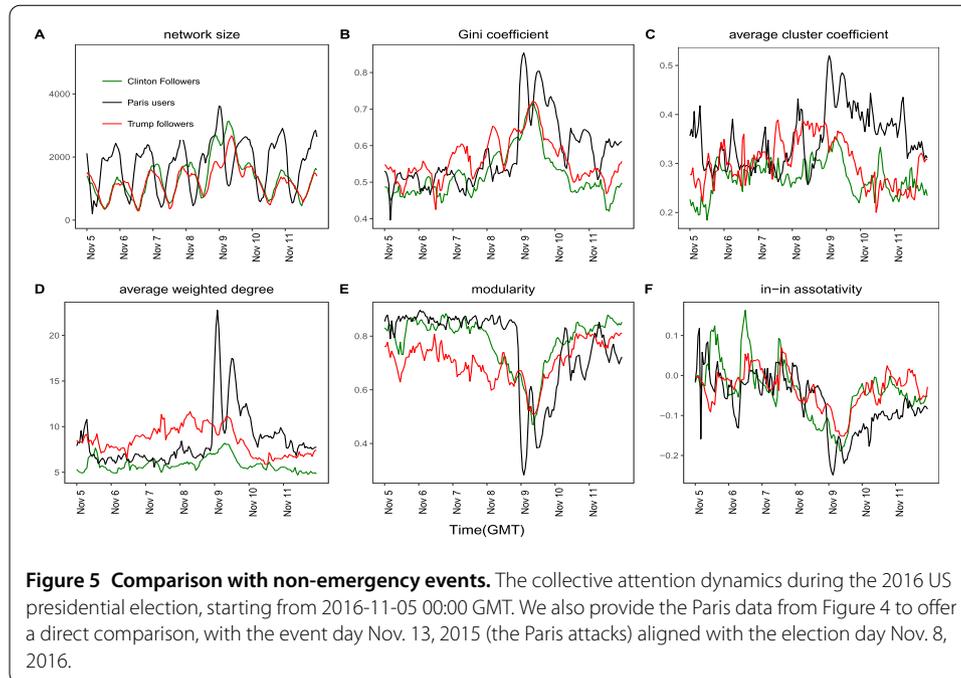
We observe similar patterns across all datasets (see Additional file 1 for comparisons for all datasets). On the surface, the temporal trends of original attention graphs and the randomized ones seem to be quite similar, which may suggest that the increased frequency of changing hashtags (during an event) would be sufficient to result in the 'attention graph agglomeration' (as reflected by the decreasing modularity). However, this can be readily explained by considering the construction of the null model – the null model has kept the frequencies of hashtags the same as those in the original graphs, and more user activities during the events lead to more connections around the popular hashtags, which is reflected by the higher values of Gini coefficient (and a sharp increase in average clustering coefficient) in the randomized graphs. Yet, the lowest modularity values obtained from the original attention graphs still tend to be much higher than those in the randomized graphs, suggesting that even during a shocking event where users' attention was greatly distracted from their normal focus, the original attention graphs remain more structured than completely randomly rewiring graphs.

The comparison between the originally observed attention graphs and the baseline reveals the following key phenomenon of the attention dynamics. Normally, collective attention exhibits a heterogeneous structure where people's focal attention is localized into various clusters organized by different topics. When encountering an exogenous event, it acts as an external force that direct the shift of people's attention – that is, the overall attention appears to converge into a more homogenized structure at a particular period. The choice of hashtags – where users pay attention, is not random, suggesting that the process is collective.

### 3.6 Comparison with non-emergency events

To better understand the collective attention dynamics under different contexts, we also study the attention shift patterns during a planned event: The 2016 US presidential election. Several representative network metrics for the two follower user sets are shown in Figure 5. The displayed time window spans one week during the 2016 US presidential election and is centered on the election day, November 8, 2016.

In Figure 5, we also provide the Paris data from Figure 4 to offer a direct comparison with the non-emergency event, with the event day Nov. 13, 2015 (the Paris attacks) aligned with the election day Nov. 8, 2016. From Figure 5, we can observe some consistent properties that we find during a terrorist attack: the increase in Gini coefficient, the agglomeration of network structure and the decrease in assortativity. Notably, during the election, the values of different metrics jointly exhibit a gradual build-up of attention before reaching

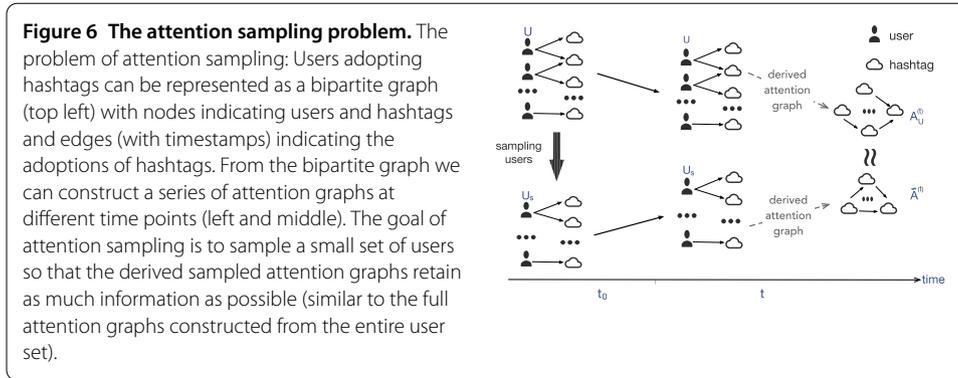


their peaks. In Figure 5, corresponding to an unexpected event, a sudden increase (or decrease) of activity followed by a gradual relaxation of concentrated attention while the metric values of the planned event (election) show both a progressive build-up of attention and a gradual relaxation after the peak. Taking *Gini coefficient* for example, for Trump and Clinton followers, it takes 15 and 16 hours to rise from  $\mu_b + \frac{1}{2}(p_{\max} - \mu_b)$  to  $p_{\max}$  while it only takes 2 hours for the Paris users, where  $\mu_b$  is the mean value in the week prior to the event week and  $p_{\max}$  is the peak value in the event week.

The comparison between unexpected events and planned events suggests that collective attention patterns are robust indicators of the social semantics: Specifically, the gradual build-up of attention before reaching a peak usually can be seen in scheduled social events, representing an expected behavior. Collective attention with asymmetric activity patterns before and after the peak is more likely to be associated with unexpected events.

#### 4 Attention sampling

Attention graphs allow for systematically observing and quantifying the collective attention dynamics. A straightforward way for a thorough investigation of attention dynamics of an interested community is to gather data from as many individuals as possible. However, this is often prohibited in practice. In particular, collecting data from social media platforms is limited by the API rate or sampling limits. For example, using the Twitter Streaming API with certain parameters, applications are allowed to get up to 5,000 users (as well as about 1% of all tweets being tweeted at that moment). With such a data collection constraint, including users who are less active in posting tweets, or users who share overlapping interests with many others, would be a waste of streaming resource. The data sampling considerations require understanding how different users attend to new information, leading to an interesting and realistic problem: how might we sample a subset of users from the originally available user set to accurately and efficiently retain the attention dynamics? Below we formalize the problem.



### 4.1 Problem formulation

Let  $A_U^{(t)}$  denotes the attention graph induced from attention shift events  $S_U^{(t)}$  with respect to a user set  $U$ . We seek to construct an attention graph  $\tilde{A}^{(t)}$  similar to  $A_U^{(t)}$  but from a smaller set of users. The problem of *attention sampling* is defined as follows:

**Problem** (Attention sampling) Given a set of attention shift events  $S_U^{(t_0)}$  with respect to a user set  $U$  occurring within a time period  $t_0$  prior to  $t$ , the goal of attention sampling is to find a subset of users  $U_s \subset U$  such that the *partial* attention graph  $A_{U_s}^{(t)}$  derived from the later attention shift events of users  $U_s$  is *similar* to the *full* attention graph  $A_U^{(t)}$  derived from the later attention shift events of users  $U$ .

Generally, the *similarity* between the partial and full graphs ( $A_{U_s}^{(t)}$  and  $A_U^{(t)}$ ) can be quantified in terms of a given metric, such as the aforementioned network metrics.

Unlike the traditional subgraph sampling problem primarily focusing on an unipartite network, the attention sampling problem concerns the set of bipartite relationships ( $U \times H$ ) captured in the set of attention shift events  $S_U^{(t_0)}$  occurring before the time of attention graph construction  $t$ . In other words, the attention sampling seeks to build attention graphs in a cost-effective manner based on users' hashtag adoption patterns. Figure 6 illustrates the attention sampling problem.

### 4.2 Sampling approach

We decompose the attention sampling problem into two sub-problems. First, what criteria should we look for when sampling users? We propose heuristics to select users more likely with certain hashtag adoption tendencies so as to construct a partial attention graph that is closer to the full attention graph. Second, as a user's adoption behavior may vary across time, how should we cope with the data variability and generate samples for reliable measurement? We propose a stochastic sampling method that allows for selecting users with a small perturbation.

#### 4.2.1 Sampling criteria: who should we include in a sample?

We conjecture that an ideal sample set should include users with the following tendencies:

- *Activeness*: the extent to which a user will actively mention various topics of interest in their tweets. We consider users who tend to tweet with hashtags at a relatively high frequency as primary sampling candidates as they are more likely to tweet with hashtags during the time of interest.

- *Connectedness*: the extent to which a user will diversely cover the topics of interest of many other users. We consider users who tend to tweet with hashtags commonly used by others as desirable sampling candidates as their hashtag use is more likely to cover the use of a broader set of users.
- *Adaptiveness*: the extent to which a user will adaptively attend to rare or new topics of interest. We consider users who tend to tweet with novel hashtags as desirable sampling candidates as they are more likely to attend to new topics or information about newly emerging events.

While each of these criteria may be captured by a particular measure, considering each criterion to rank users separately is sub-optimal as the rankings by different criteria may not agree with one another. To simultaneously capture all these criteria, we introduce a new weighted scheme that scores users based on their weighted degree in a user-user graph. Specifically, given a graph  $G = \{U, E_U\}$ , where  $U$  is a set of users and  $E_U \subset U \times U$  represents co-adoption relationships among users. Each edge  $e_{xy} \in E_U$  represents how the hashtags used by the two users  $u_x$  and  $u_y$  is *important* or *informative* in the population. Let  $p_i$  denote the probability of a hashtag  $h_i$  used by any user, and  $H_x$  the set of hashtags used by a particular user  $u_x$ . An edge weight  $w_{xy} \in \mathbb{R}$  is given by:

$$w_{xy} \propto - \left[ \sum_{i \in H_x} \log p_i + \sum_{j \in H_y} \log p_j + \sum_{k \in H_C} \log p_k \right], \quad (3)$$

iff  $H_C \neq \emptyset$ , otherwise  $w_{x,y} = 0$ , where  $H_C = H_x \cap H_y$ .  $p_i$  denotes the empirical probability (relative frequency) of a hashtag  $i$  used by any user, calculated based on the observed data. For example, for the Paris users, in the two weeks before the Paris attacks, in total 123,475 unique hashtags had been used 301,059 times, hashtag like #paris had been used 2,090 times. Thus the relative frequency  $p_i$  of #paris is 2,090/301,059. The equation aims to capture both the similarity of the two users (through  $H_C$ ) and the coverage for both users (through  $H_x$  or  $H_y$ ). Hence, a common hashtag accounts for both similarity and user's coverage.

Based on Eq. 3, two users who use less frequent hashtags (with smaller  $p_i$ ), use more hashtags individually (with larger  $H_x$  or  $H_y$ ), and more hashtags in common (with larger  $H_C$ ), thus their edges are given more weights. We call the proposed weighted scheme *CoP-plexity* as it simultaneously captures the frequency, connectedness and diversity of users' hashtag usage.

#### 4.2.2 Sampling algorithms: how should we make a stochastic sample?

Given a user-user co-adoption graph  $G$ , our next step is to construct stochastic samples from the graph such that the sampled attention graph is more robust to the variability and uncertainty of user tendency. We first consider two sampling algorithms widely adopted in the sampling literature: Random Walk (RW) and Metropolis-Hastings Random Walk (MHRW). An RW based approach [27] allows node re-visiting and is well-known for its simple, resource-efficient properties; however, RW is inherently biased towards high degree nodes. The MHRW [28], on the other hand, is useful in achieving a desired stationary distribution by design. However, an MHRW based approach tends to include too many low-connectivity nodes in the sampled network - such nodes are less likely to cover broadly the hashtags of many other users.

**Algorithm 1:** Power-MHRW sampling

---

```

Input : a user-user weighted graph  $G$ , a rejection power parameter  $\alpha \in [0, 1]$ 
1   $sample\_size \leftarrow n$ ;
2   $C = \emptyset$ ;
3   $v \leftarrow \text{initial node}$ ;
4  while  $|C| < sample\_size$  do
5    Select a node  $u$  from the neighbors of  $v$  with probability proportional to edge weight  $w_{vu}$ ;
6    Generate uniformly at random a number  $0 \leq p \leq 1$ ;
7    if  $p \leq k_v k_u^{-\alpha}$  then
8       $v \leftarrow u$ ;
9       $C = C \cup \{u\}$ ;
10   else
11     Stay at  $v$ ;
12   end

```

---

To compromise the two regimes (RW's high-degree dominance vs. MHRW's stratified sampling), we propose an extension to the MHRW algorithm. In our algorithm, called *Power-MHRW*, an exponent parameter  $\alpha$  (i.e., the *power* parameter) is introduced to adjust the probability of selecting a user between the two regimes. By tuning  $\alpha$ , we can appropriately modify the transition probabilities when walking in the user-user network, trying to balance the trade-off between obtaining high-degree users and a diversified sampled user set. The proposed algorithm is listed in Algorithm 1.

In Algorithm 1,  $n$  denotes the desirable sampled user size, given as an input for the sampling algorithm. This number determines the cost of monitoring collective attention - the more users, the more accurate attention trends can be captured but more computational cost.  $k_v$  and  $k_u$  are weighted degrees of node  $v$  and node  $u$ . Nodes are sampled iteratively (steps 4-11): at each iteration, the algorithm randomly picks a node  $u$  from the neighbors of current sampled node  $v$ , with probability proportional to the edge weight  $w_{vu}$ . The 'neighbors' are defined as users who co-attended to common hashtags. Different weighted schemes such as the *CoPerplexity* defined in Eq. 3 can be used to define the edge weights. After picking a node  $u$ , the algorithm decides to include the node into the sample with probability  $\min(1, k_v k_u^{-\alpha})$ . When  $\alpha = 0$ , the algorithm always accepts the picked node, which is identical to the original RW algorithm. When  $\alpha = 1$ , the algorithm always accepts the picked node if it has smaller weighted degree than the current sampled node, which results in more uniform weighted degree distribution in the sampling as in the original MHRW algorithm. When  $0 < \alpha < 1$ , we trade off between the bias towards high weighted degree nodes, as in the RW algorithm, and the desired uniform distribution of weighted degrees, as in the MHRW algorithm.

## 5 Experiments

This section presents experiments for evaluating the effectiveness of the proposed sampling approach.

### 5.1 Experiment setup

We use four datasets in our experiments: the Paris users and the London users for Paris attacks, the Brussels users and the San Bernardino users. We select the user sets that exhibit more prominent patterns before and after the attacks (Paris users and Brussels users)

as well as a less obvious one (San Bernardino users). We also select a user set (London users) from the observer communities for comparison. We compare our proposed sampling approach with several baseline methods, and evaluate their effectiveness based on two different aspects.

*Baseline.* We compared our approach with the following baseline methods: (1) *random*: randomly samples a portion of users from the entire user set. (2) *connectedness*: samples users with the probability proportional to their number of followers and followees which we refer to as *follower* and *followee*.<sup>e</sup> (3) *activeness*: samples users with the probability proportional to the number of tweets they posted within the two weeks before the attack. (4) *CoPerplexity*: the proposed method based on Algorithm 1, with edge weight given in Eq. 3. In addition, we compared the proposed sampling algorithm with alternative algorithms, *CoPerplexity\_RW* (when  $\alpha = 0$ ), *CoPerplexity\_MHRW* (when  $\alpha = 1$ ) and refer our algorithm as *CoPerplexity\_PRW* (when  $0 < \alpha < 1$ ). (5) *CommonTag*: a variant of the proposed method, but instead of using the *CoPerplexity* edge weight, we simply consider the number of common tags between two users as edge weight. In this method,  $\alpha$  is fixed to be 0 and we refer to it as *CommonTag\_RW*.

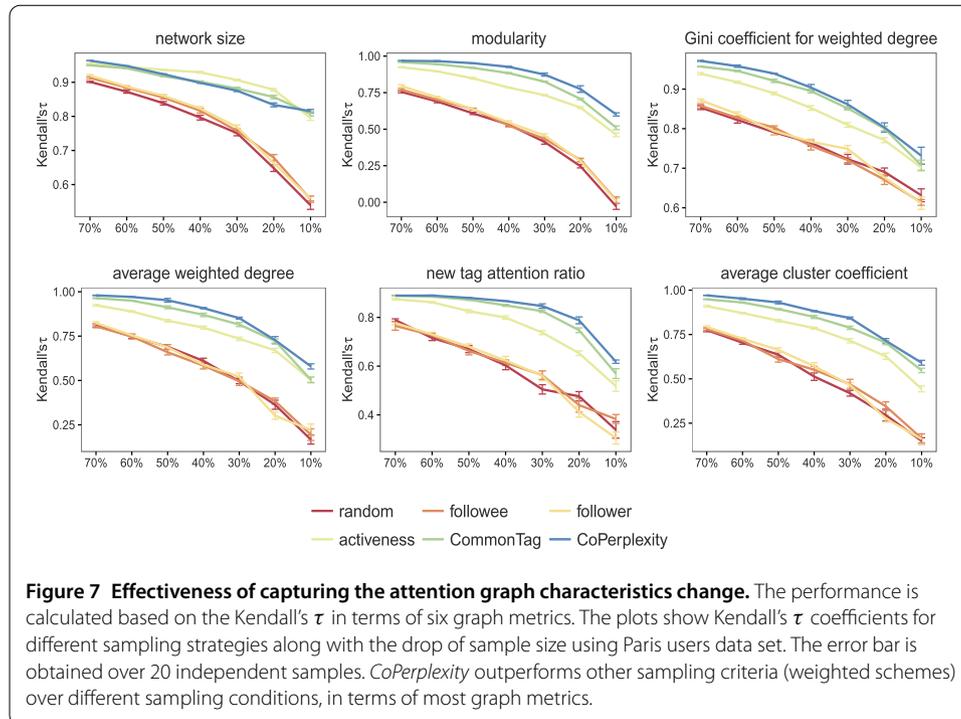
*Evaluation metrics.* We evaluated the effectiveness of a sampling approach in capturing the collective attention shift, based on two aspects: (1) *To what extent the sampling method captures the attention graph characteristics change?* We use Kendall's  $\tau$  to measure the non-parametric association between the graph metrics derived from the original attention graphs and the sampled graphs. Unlike other correlation metrics like Pearson's  $\gamma$ , Kendall's  $\tau$  does not make the parametric assumption (e.g., linearity) in the variables that is unrealistic for our empirical observations. Here, we focus on the 'trend' of the collective attention, not the absolute value of a graph metric, because we are interested in the fundamental characteristics of social dynamics of attention during the course of a disaster event – whether, or to what extent, the structure of the collective attention expands, shrinks, changes density, condenses, and so on. (2) *How much attended information was retained by the sampling method?* We evaluated this based on leveraging the *precision@L* to compare the top  $L$  hashtags in the attention graphs generated by different user sets. Precision measures how many of the most attended hashtags from the sampled user set were also in the top  $L$  hashtags list of the original attention graphs. By 'most attended,' we sorted the hashtags in each attention graph in a descending order in terms of *weighted in-degree* and the top  $L$  ones are the most attended hashtags.

For an event week, we constructed an attention graph for every hour  $t$  using tweets between  $[t - n, t]$ , where  $n$  is the time window, thus we have  $24 \times 7 = 168$  original attention graphs in a week. After we have sampled users, we constructed the sampled attention graphs in a similar way. So for each metric, we have 168 values at each time point for the original attention graph as well as the sampled attention graph. We compute the correlation between the two time series of graphs. A higher Kendall's  $\tau$  suggests similar collective attention dynamics from the sampled users and a higher *precision@L* suggests the sampled user sets better retain the attended information.

## 5.2 Results

The experiments aim to answer two questions:

- Sampling criteria - What kind of users should be included in a sample for capturing the collective attention shift in a larger population?



- Sampling algorithms - What stochastic sampling algorithm is effective for monitoring the dynamics of collective attention?

*Evaluation on sampling criteria.* Figure 7 shows the performance of different weight schemes for sampling users, over different sampling conditions where the sampling fraction ranges between 10-70%. Each condition is tested with 20 independent samples and the mean and standard deviation are reported. The RW based sampling approach is used in *CoPerplexity* and *CommonTag*, with  $\alpha$  fixed to be 0. The performance is calculated based on the Kendall's  $\tau$  in terms of six graph metrics. When sample size = 100%, we have Kendall's  $\tau = 1$  for each metric (not shown in the plots). As expected, Kendall's  $\tau$  values monotonically decrease when the sample sizes decrease. Figure 7 shows that, *CoPerplexity* outperforms other sampling criteria (weighted schemes) over different sampling conditions, in terms of almost all graph metrics, except for *network size*. The performance of different sampling criteria is distinguishable, where *CoPerplexity*, *CommonTag* and *activeness* perform much better than the other three, suggesting that users' hashtag adoption tendency in the past is decisive in attention sampling. Figure 8 summarizes the performance over different schemes for sampling fractions 30% and 50%. We observe that the Kendall's  $\tau$  values for *CoPerplexity* in terms of all metrics remains relatively high - between 0.8 and 0.9 under the 50% sampling fraction, and above 0.75 under the 30%. This suggests that *CoPerplexity* is a robust method for sampling attention graphs while retaining their topological characteristics.

*Evaluation on sampling algorithms.* We further examine the effectiveness of different sampling algorithms and their capability in retaining the most attended information. Based on the results of sampling criteria, we focus on the most effective weighted schemes *CoPerplexity*, *CommonTag* and *activeness* and further study the performance of *CoPerplexity* with different sampling algorithms, namely, *CoPerplexity\_RW*, *CoPerplexity\_MHRW*, and *CoPerplexity\_PRW*. The performance is measured based on *precision@L*.

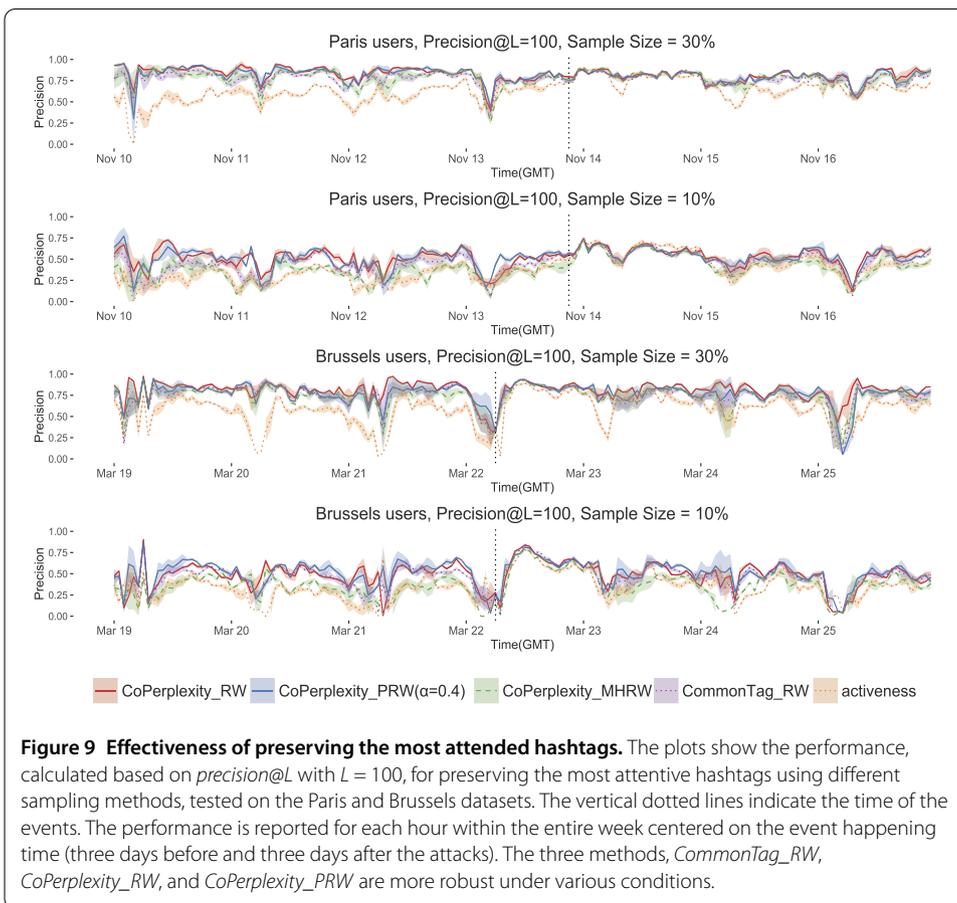
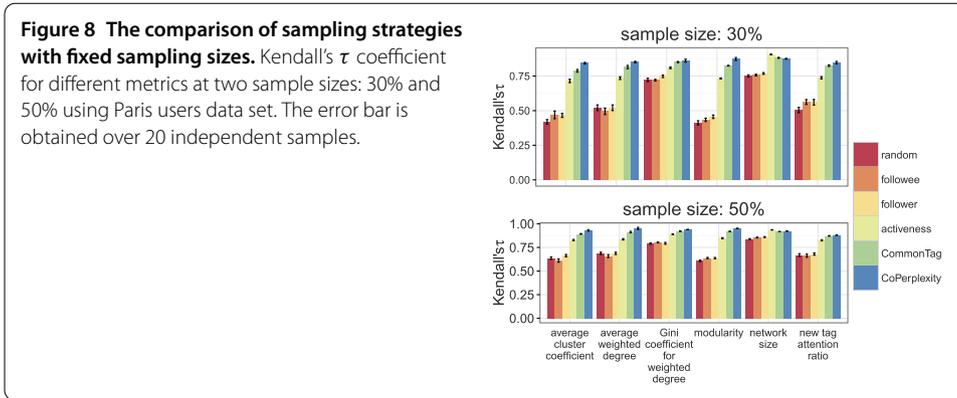


Figure 9 shows performance results for different sampling methods under different sampling conditions, on the four selected datasets, with  $L = 100$ . The performance is reported for each hour within the entire week centered on the event happening time (three days before and three days after the attacks). Under the relatively stringent sampling conditions (10% and 30%), we observe that attention sampling in normal time is more difficult compared with that during the exogenous events. Specifically, several hours after the attacks, we observe a sharp increase of precision for all sampling methods and the precision values remain relatively high for next 24 hours, reflecting the consistent concentrated attention towards the event-related topics after the attacks. In Figure 9, there are several time points

**Table 2 Performance of the top three sampling methods in capturing the most attentive hashtags during the event week (168 hours), evaluated based on *Precision@100* each hour over different sampling conditions (10% and 30% of the complete user set)**

Sample	<i>CoPerplexity_PRW</i> ( $\alpha = 0.4$ )			<i>CoPerplexity_RW</i>			<i>CommonTag_RW</i>		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Paris attacks(Paris users)									
10%	<b>51.19%</b>	33.93%	14.88%	43.45%	43.45%	13.10%	8.33%	23.21%	68.45%
30%	46.43%	43.45%	10.12%	<b>52.38%</b>	39.88%	7.74%	7.14%	17.26%	75.60%
Paris attacks(London users)									
10%	<b>53.57%</b>	36.90%	9.52%	39.88%	47.02%	13.10%	10.12%	22.66%	72.62%
30%	32.73%	49.40%	17.86%	<b>59.52%</b>	28.57%	11.90%	17.86%	26.79%	55.36%
Brussels bombings									
10%	<b>57.14%</b>	25.00%	17.86%	30.95%	36.90%	32.14%	14.88%	40.48%	44.64%
30%	22.02%	44.05%	33.93%	<b>68.45%</b>	22.62%	8.93%	12.50%	35.12%	52.38%
San Bernardino shooting									
10%	<b>44.04%</b>	36.31%	19.64%	21.43%	47.02%	31.55%	40.47%	19.05%	40.47%
30%	22.02%	39.29%	32.14%	33.93%	46.43%	31.55%	<b>50.60%</b>	23.21%	26.19%

For each method, the fraction of the method ranked in the top three places are reported.

where the precision values reach zero, which indicates at those time points, no attended hashtags have been obtained by the sampled users. This usually happened during the late nights and it is reasonable when the tweet volume is extremely small and the sampled user size is very small. In general, attention sampling based on *activeness* (biased towards active users) and MHRW based sampling (stratified sampling) appears to be less effective, especially in normal time where collective attention scatters over diverse topics.

The three methods, *CommonTag\_RW*, *CoPerplexity\_RW*, and *CoPerplexity\_PRW* are more robust under various conditions. For *CoPerplexity\_PRW*, we have evaluated different values of  $\alpha$  and obtained optimal results when  $\alpha$  ranges from 0.3-0.5. Due to the space limit, we only report the results of *CoPerplexity\_PRW* for  $\alpha = 0.4$ . Table 2 summarizes the results of the best three sampling methods presented in Figure 9 by reporting the fraction of the methods' rankings (labeled as 1st, 2nd, and 3rd) during the entire 168 hours. When two (or more) methods tie for the same position in the ranking, they obtain the same label; thus the sum of fractions at certain rankings may exceed 100%. Taking the first line in Table 2 for example, starting from the left, 51.19% indicates, during the event week, there were 86 hours ( $86/168 = 51.19\%$ ) during which the method *CoPerplexity\_PRW* ( $\alpha = 0.4$ ) performs the best; 33.93% indicates that 57 hours of the week ( $57/168 = 33.93\%$ ), the method *CoPerplexity\_PRW* ( $\alpha = 0.4$ ) performs the second best and so on. The results show that *CoPerplexity\_RW* performs better than *CommonTag\_RW* for most of the time, further supporting the effectiveness of the proposed *CoPerplexity* weighted scheme. *CoPerplexity\_PRW* and *CoPerplexity\_RW* have comparable performance, but *CoPerplexity\_PRW* ( $\alpha = 0.4$ ) outperforms *CoPerplexity\_RW* under a lower sampling fraction (10%). This suggests that the combination of *CoPerplexity* and an RW based algorithm can effectively identify a set of users for collective attention monitoring. When smaller sampling size is desirable, *CoPerplexity\_PRW* has the advantage to include more diversified users. Together with Table 2, the results suggest that different sampling schemes are more distinguishable when the users' collective attention exhibit sharp contrast before and after the event: for the Paris attacks and the Brussels shooting, *CoPerplexity\_RW* based method is 3-6 times better than *CommonTag\_RW* when considering the fraction of the 1st rank.

Under a smaller collective attention shift scenario (the San Bernardino shooting), different weighting schemes become comparable: when sample size is 30%, *CommonTag\_RW* even outperforms *CoPerplexity\_RW* and *CoPerplexity\_PRW*. When the sample size further drops to 10%, *CoPerplexity\_PRW* performs the best but is comparable to *CommonTag\_RW* (44.04% to 40.07% in terms of the fraction of the 1st rank).

The results we obtained can help guide the data collection strategy in practice. For example, considering collecting data using the Twitter Streaming API and assuming each application is allowed to retrieve up to 5,000 users' timelines in real time. This constraint corresponds to about 37.2% of the Paris users, or about 51% of the Brussels users. Based on Figure 9, we can see if the API limit was imposed, *CoPerplexity\_PRW* and *CoPerplexity\_RW* can retain over 85% attended information for both data sets. In terms of retaining similar graph structure changes, both achieve at least 0.84 in Kendall's  $\tau$ . Moreover, if the practitioners choose to trade off between the accuracy and the sampling user size, *CoPerplexity\_PRW* and *CoPerplexity\_RW* can reasonably capture the attention dynamics with even smaller user sets. For example, we can retain about 75% attended information with only 25% sampled users.

## 6 Discussion and future work

In this work, we presented a large-scale study on collective attention as well as an investigation of data sampling for monitoring collective attention. We provide a new definition of collective attention and systematically studied the changing patterns of collective attention during disasters, and salient features - including five types of changes in attention network structure - were observed from multiple events. Our study design based on network statistics can be used to systematically observe collective attention dynamics across various conditions. Our examination of the attention sampling problem suggested that, while it is not difficult to capture collective attention under exogenous shocks due to the public's concentration on attentive topics, monitoring collective attention dynamics across normal and event periods could be challenging. We proposed a novel sampling scheme that considers several aspects of how individuals normally share interest with their social network (activeness, connectedness, and adaptiveness) and demonstrated the utility of our approach through an extensive experimentation. We further provide a practical data sampling strategy such that the desirable monitoring effectiveness can be achieved given various sampling constraints imposed by the data sources. Moreover, the study framework as well as the sampling scheme we proposed can be applied to other social media platforms that have a hashtag feature, like Facebook and Instagram. Our model can also be applicable to non-social-media data. For example, we can trace the shift between keywords of papers to study the collective attention shift in research interests over time. For e-commerce systems where products have tags, we can study the users' collective purchasing patterns by tracing the shifts between different tags from their purchasing history.

*Limitations.* There are several limits in the current work. First, the empirical finding of collective dynamics during disasters is limited to the context of hashtag use. Second, the major event instances we included in this study are man-made disasters. Specifically, they were all induced by unexpected mass violent attacks, which often lead to acute reaction in the communities. The results of our study may not be generalized to other types of events. Third, the study relied on users' geotagged tweets to derive event related datasets. The users considered in this study were mostly tech-savvy mobile and social media users,

but they were not representative of the general population. Forth, the sampling approaches discussed in this paper were all unsupervised. It could be advantageous to investigate supervised algorithms with respect to the given sampling criteria.

To properly sample users for studying behavioral trends is important for scientists (especially in social science, cognitive and behavioral science, marketing, etc.) to learn human behaviors from big data. This is different from data infrastructure or engineering problems but will become increasingly important in data science field.

As part of the future work, we plan to address some of the aforementioned limitations. In particular, we plan to conduct a more comprehensive study on attention network patterns following different types of events including natural disasters, with different scopes and populations. The results from such study will not only be valuable for developing disaster management systems, but also offer insights in human sense making processes under conditions of uncertainty and rapid change.

### Additional material

**Additional file 1: Supplementary Information.** (pdf)

### Acknowledgements

This work is part of the research supported by NSF Grants #1423697, #1634944, #1637067, #1739413, and the CRDF & CIS at the University of Pittsburgh. Xingsheng He also thanks the support of China Scholarship Council. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of the funding sources.

### Abbreviations

RW, random walk; MHRW, Metropolis-Hasting random walk; PRW, power random walk.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

Conceived & designed: YRL, XSH; Analyzed the data: XSH, YRL; Contributed analysis tools: XSH, YRL; Wrote the manuscript: XSH, YRL. All authors read and approved the final manuscript.

### Author details

<sup>1</sup>Department of Electronic Science and Technology, University of Science and Technology of China, Hefei, 230000, China.

<sup>2</sup>School of Computing and Information, University of Pittsburgh, Pittsburgh, PA 15260, USA.

### Endnotes

- <sup>a</sup> [https://en.wikipedia.org/wiki/November\\_2015\\_Paris\\_attacks](https://en.wikipedia.org/wiki/November_2015_Paris_attacks)
- <sup>b</sup> [https://en.wikipedia.org/wiki/2015\\_San\\_Bernardino\\_attack](https://en.wikipedia.org/wiki/2015_San_Bernardino_attack)
- <sup>c</sup> [https://en.wikipedia.org/wiki/2016\\_Brussels\\_bombings](https://en.wikipedia.org/wiki/2016_Brussels_bombings)
- <sup>d</sup> [https://en.wikipedia.org/wiki/2016\\_Orlando\\_nightclub\\_shooting](https://en.wikipedia.org/wiki/2016_Orlando_nightclub_shooting)
- <sup>e</sup> Due to the heterogeneity in users' followers and followees (some users may have millions of followers while some only have several), the probability is based on a log-transformed follower (followee) count.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 30 May 2017 Accepted: 12 November 2017 Published online: 22 November 2017

### References

1. Acar A, Muraki Y (2011) Twitter for crisis communication: lessons learned from Japan's tsunami disaster. *Int J Web Based Communities* 7(3):392-402
2. Burnap P, Williams ML, Sloan L, Rana O, Housley W, Edwards A, Knight V, Procter R, Voss A (2014) Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack. *Soc Netw Anal Min* 4(1):1-14
3. Vis F (2013) Twitter as a reporting tool for breaking news: journalists tweeting the 2011 UK riots. *Digit Journalism* 1(1):27-47

4. Qiu X, Oliveira DF, Shirazi AS, Flammini A, Menczer F (2017) Limited individual attention and online virality of low-quality information. *Nat Hum Behav* 1:0132
5. Lehmann J, Gonçalves B, Ramasco JJ, Cattuto C (2012) Dynamical classes of collective attention in Twitter. In: *WWW 2012*, pp 251-260. ACM, New York
6. Sasahara K, Hirata Y, Toyoda M, Kitsuregawa M, Aihara K (2013) Quantifying collective attention from tweet stream. *PLoS ONE* 8(4):e61823
7. Roy SD, Lotan G, Zeng WK (2015) The attention automaton: sensing collective user interests in social network communities. *IEEE Trans Netw Sci Eng* 2(1):40-52
8. Wang C-J, Wu L (2016) The scaling of attention networks. *Phys A, Stat Mech Appl* 448:196-204
9. Wu F, Huberman BA (2007) Novelty and collective attention. *Proc Natl Acad Sci USA* 104(45):17599-17601
10. Yang J, Leskovec J (2011) Patterns of temporal variation in online media. In: *WSDM 2011*, pp 177-186. ACM, New York
11. Mahata D, Talburt JR, Singh VK (2015) From chirps to whistles: discovering event-specific informative content from Twitter. In: *Proceedings of the ACM web science conference*, p 17. ACM, New York
12. Zhang H, Setty V (2016) Finding diverse needles in a haystack of comments: social media exploration for news. In: *Proceedings of the 8th ACM conference on web science*, pp 286-290. ACM, New York
13. Lin Y-R, Keegan B, Margolin D, Lazer D (2014) Rising tides or rising stars?: dynamics of shared attention on Twitter during media events. *PLoS ONE* 9(5):e94093
14. He X, Lu D, Margolin D, Wang M, Idrissi S, Lin Y-R (2017) The signals and noise: actionable information in improvised social media channels during a disaster. In: *Proceedings of web science 2017*
15. Lee K, Caverlee J, Kamath KY, Cheng Z (2012) Detecting collective attention spam. In: *Proceedings of the 2nd joint WICOW/AIRWeb workshop on web quality*, pp 48-55. ACM, New York
16. Weng L, Flammini A, Vespignani A, Menczer F (2012) Competition among memes in a world with limited attention. *Sci Rep* 2:335
17. Yin J, Lampert A, Cameron M, Robinson B, Power R (2012) Using social media to enhance emergency situation awareness. *IEEE Intell Syst* 27(6):52-59
18. Lin Y-R, Margolin D (2014) The ripple of fear, sympathy and solidarity during the Boston bombings. *EPJ Data Sci* 3(1):31
19. Lin Y-R, Margolin D, Wen X (2017) Tracking and analyzing individual distress following terrorist attacks using social media streams. *Risk Anal* 37(8):1580-1605
20. Sano Y, Yamada K, Watanabe H, Takayasu H, Takayasu M (2013) Empirical analysis of collective human behavior for extraordinary events in the blogosphere. *Phys Rev E* 87(1):012805
21. Bassett DS, Alderson DL, Carlson JM (2012) Collective decision dynamics in the presence of external drivers. *Phys Rev E* 86(3):036105
22. Preis T, Moat HS, Bishop SR, Treleaven P, Stanley HE (2013) Quantifying the digital traces of Hurricane Sandy on Flickr. *Sci Rep* 3:3141
23. Borge-Holthoefer J, Perra N, Gonçalves B, González-Bailón S, Arenas A, Moreno Y, Vespignani A (2016) The dynamics of information-driven coordination phenomena: a transfer entropy analysis. *Sci Adv* 2(4):e1501158
24. Rusmevichientong P, Pennock DM, Lawrence S, Giles CL (2001) Methods for sampling pages uniformly from the world wide web. In: *AAAI fall symposium on using uncertainty within computation*, pp 121-128
25. Hübler C, Kriegel H-P, Borgwardt K, Ghahramani Z (2008) Metropolis algorithms for representative subgraph sampling. In: *ICDM 2008*, pp 283-292. IEEE
26. Leskovec J, Faloutsos C (2006) Sampling from large graphs. In: *ACM SIGKDD 2006*, pp 631-636. ACM, New York
27. Lovász L (1993) Random walks on graphs. *Combinatorics* 2:1-46
28. Gjoka M, Kurant M, Butts CT, Markopoulou A (2010) Walking in Facebook: a case study of unbiased sampling of OSNs. In: *Infocom, 2010 proceedings IEEE*, pp 1-9. IEEE
29. Volz E, Heckathorn DD (2008) Probability based estimation theory for respondent driven sampling. *J Off Stat* 24(1):79-97
30. Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: *ACM SIGKDD 2003*, pp 137-146. ACM, New York
31. Chen W, Wang Y, Yang S (2009) Efficient influence maximization in social networks. In: *ACM SIGKDD 2009*, pp 199-208. ACM, New York
32. De Choudhury M, Lin Y-R, Sundaram H, Candan KS, Xie L, Kelliher A (2010) How does the data sampling strategy impact the discovery of information diffusion in social media? In: *ICWSM*, vol. 10, pp 34-41
33. Uysal I, Croft WB (2011) User oriented tweet ranking: a filtering approach to microblogs. In: *Proceedings of the 20th ACM international conference on information and knowledge management*, pp 2261-2264. ACM, New York
34. Newman ME (2003) The structure and function of complex networks. *SIAM Rev* 45(2):167-256
35. Rivera MT, Soderstrom SB, Uzzi B (2010) Dynamics of dyads in social networks: assortative, relational, and proximity mechanisms. *Annu Rev Sociol* 36:91-115
36. Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci USA* 105(4):1118-1123
37. Foster JG, Foster DV, Grassberger P, Paczuski M (2010) Edge direction and the structure of networks. *Proc Natl Acad Sci USA* 107(24):10815-10820
38. Lin Y-R, Margolin B, Keegan A, Baronchelli A, Lazer D (2013) #bigbirds never die: understanding social dynamics of emergent hashtag. In: *ICWSM 2013*
39. Heer J, Kong N, Agrawala M (2009) Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In: *SIGCHI 2009*, pp 1303-1312. ACM, New York