Credibility and the Dynamics of Collective Attention

TANUSHREE MITRA, Virgnia Tech
GRAHAM WRIGHT, Facebook
ERIC GILBERT, University of Michigan

Today, social media provide the means by which billions of people experience news and events happening around the world. However, the absence of traditional journalistic gatekeeping allows information to flow unencumbered through these platforms, often raising questions of veracity and credibility of the reported information. Here we ask: How do the dynamics of collective attention directed toward an event reported on social media vary with its perceived credibility? By examining the first large-scale, systematically tracked credibility database of public Twitter messages (47M messages corresponding to 1,138 real-world events over a period of three months), we established a relationship between the temporal dynamics of events reported on social media and their associated level of credibility judgments. Representing collective attention by the aggregate temporal signatures of an event's reportage, we found that the amount of continued attention focused on an event provides information about its associated levels of perceived credibility. Events exhibiting sustained, intermittent bursts of attention were found to be associated with lower levels of perceived credibility. In other words, as more people showed interest during moments of transient collective attention, the associated uncertainty surrounding these events also increased.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**;

Additional Key Words and Phrases: Credibility; social media; Twitter; collective attention; temporal; news events; misinformation

1 INTRODUCTION

Online social networks act as information conduits for real-world news and events [7], largely driven by collective attention from multiple social media actors [52]. Collective human attention drives various social, economic and technological phenomenon, such as herding behavior in financial markets [46], formation of trends [2], popularity of news [52], web pages [41], and music [43], propagation of memes [25], ideas, opinions and topics [42], person-to-person word-of-mouth advertising and viral marketing [25], and diffusion of product and innovation [4]. Moreover, it is the key phenomenon underlying social media reporting of emerging topics and breaking news [33]. However, unlike traditional news media—where information is curated by experienced journalists—social media news is unfiltered and therefore not subject to the same verification process

PACM on Human-Computer Interaction, Vol. 1, No. CSCW, Article 80. Publication date: November 2017.

80

as information presented by way of conventional sources. This naturally calls into question its credibility and the means with which to assess its credibility. Although scholars have increasingly expressed concern over the threats posed by digital misinformation, ranging from panic and violence incitement in society to libel and defamation of individuals or organizations [21], questions concerning the relationship between collective attention and information credibility have not been systematically quantified.

A fundamental attribute underlying any collective human behavior is how that behavior unfolds over time [3, 11]. Is there a relationship between allocation of collective attention and perceived credibility of events reported through social media? Do occasional bursts in collective attention—as more eyes and voices are drawn to the event's reportage—correspond to less certain information concerning the event? Uncovering the relationship between collective human behavior and information credibility is important for assessing the veracity of event reportage as it unfolds on social media. This relationship, if it exists, can provide insights into ways to disambiguate misinformation from accurate news stories in social networks—a medium central to the way we consume information [7] and one where digital misinformation is pervasive [12].

Empirical attempts at answering these questions in naturalistic settings have been constrained by difficulties in tracking social media posts in conjunction with judgments concerning the accuracy of the underlying information. Previous studies have instead focused on individual case studies involving specific news events [1, 29, 32], or have retrospectively studied a set of multiple prominent events [12, 14] which were known to contain misinformation. While useful, these approaches raise sampling concerns. In particular, they are based on the post-hoc investigation of events with known credibility levels, and thus select on the dependent variable [50]. Although these studies suggest the possibility of spikes in collective attention when false rumors propagate through social networks, the relation between collective attention and information credibility has not been systematically tested. We tested this relation by analyzing data from CREDBANK: "the first large-scale, longitudinal credibility corpus" [36]. The CREDBANK dataset was constructed by continuously tracking millions of public Twitter posts. Due to the microblogging nature of Twitter, tweeting activity comprises regular circadian rhythms [17] intertwined with irregular bursts of activities corresponding to real-world news events [33]. Thus, in recent years, Twitter has become an attractive source for disseminating information pertaining to news events [23].

CREDBANK's credibility corpus is based on tracking all large-scale, real-world events surfacing on Twitter between October 2014 and February 2015, followed by credibility assessment through a verified human annotation framework. This iterative framework resulted in an experimental setup that captured accurate human judgments of credibility of social media information soon after it gained collective attention. It is important to note that while this process cannot arrive at the *truth* of the event reporting (perhaps an impossible epistemological task), it does capture expert-level human judgment at the time the event unfolded. CREDBANK contains 1,377 social media events collected over a period of three months, along with 66 million tweets nested within the event topics. The uniqueness of the dataset is evident not only from the systematic collection process but also from the range of collected events. It contains, for example, controversial refereeing in soccer matches, announcement of celebrity deaths, as well as the emergence of Ebola in West Africa.

Although the nature of this data limits causal inference, we were able to test the correspondence between collective attention and the level of information credibility. After filtering out unique event instances, we were left with a pruned corpus of 1,138 real-world events spread over 47M tweets. Analyzing this massive dataset, we find that the amount of recurring collective attention bursts could be used to determine the level of perceived credibility of an event. Specifically, we demonstrate that multiple occasional bursts of collective attention toward an event is associated with lower

levels of perceived credibility. This finding opens a new perspective in the understanding of human collective attention and its relation to the certainty of information. In doing so, our results can have widespread implications in fields where predictive inference based on online collective interests dictates economic decisions, emergency responses, resource allocation or product recommendations [16, 53]; hence trusting the credibility of the collective reports is essential for an accurate anticipation by the predictive process.

2 RELATED WORK

Next, we outline research in three areas related to the work presented in this paper.

2.1 Social Media and Credibility

With social media's rise as an important news source [7], individuals are constantly relying on online social networks to consume and share information, without recourse to official sources. However, unlike traditional news sources, modern online social networks like Facebook and Twitter transmit both credible and less credible information, without attending to its veracity [14]. Moreover, users of these sites have been found to be poor judges of information credibility based on content alone [39]. Thus, scholars have increasingly become interested in assessing the credibility of social media content. Studies have examined specific events that were known to transmit misinformation, such as rumors disseminating during the 2011 Great East Japan earthquake [29], misinformation spread during the 2013 Boston marathon bombings [32], rumors propagating during the 2014 Sydney siege [1], and a case study of rumoring behavior in a Chinese microblogging community [27]. Studies have also engaged in extensive analysis of multiple historically reported cases of rumor, such as, automatically classifying rumor instances [55] or predicting the credibility level of tweets [6, 19, 40]. However, these studies are based on the retrospective investigation of popular historical events. Hence, they suffer from the *selection on the dependent variable* confound. On the contrary, our study overcomes this confound by grounding its results on CREDBANK's data [36] – a credibility corpus which asks human raters to assess the credibility of *all* social media events in near real-time.

2.2 Collective Attention

A phenomenon which is vital towards the spread of social media information is "collective attention" [41]. Hence, researchers have been attracted toward understanding how attention to new information propagates among large groups of people. While some studies have shown that dynamics of collective attention of online content is characterized by bursts signifying popularity changes [24, 41], others have demonstrated a natural time scale over which attention fades [52]. A study investigating the emergence of collective attention on Twitter, found that although people's attention is dispersed over a wide variety of concerns, it can concentrate on particular events and shift elsewhere either very rapidly or gradually [44]. Another parallel study focusing on spikes of collective attention in Twitter, analyzed the popularity peaks of Twitter hashtags [24]. They found that the evolution of hashtag popularity over time defined discrete classes of hashtags. Drawing on the progress of these studies, we ask: does the process of evolving collective attention reflect the underlying credibility of a social media story? Unraveling the relation between collective attention rhythms and corresponding credibility level is a complex empirical problem. It requires longitudinal tracking of collective mentions of newsworthy stories in social media along with their in-situ credibility judgments. To that end, CREDBANK [36] provides the most consistently tracked social media information and its associated credibility scores.

2.3 Time Matters

One useful way to understand the interplay between collective attention and information credibility is to examine user activity and information patterns through the lens of time. For decades social scientists have investigated the timing of individual activity to understand the complexity of collective human action. They have reported that timing can range from random [20] to well correlated bursty patterns [3]. The bursts in human collective action have not only led to social media reporting of emerging topics, but have also exhibited rich temporal dynamics of social media information spread [33]. For example, information diffusing through micro-blogging platforms like Twitter have demonstrated a short life span [52], with content rising and falling in popularity within hours; whereas, short quoted phrases (known as *memes*) have displayed several days to rise and fade away [26]. Moreover, general themes (like 'politics', 'economy', 'terrorism') have shown an even larger temporal life span [18, 51]. Social psychologists studying the spread of news and rumor have also noted the importance of temporal patterns in rumor transmission – different types of rumor mongering statements persist over varying temporal spans [5, 45]. However, despite the importance of temporal patterns in information diffusion and rumor transmission, there has been little work in understanding temporal trends in events and its associated credibility assessments. This paper is a step towards unraveling that relation.

3 METHOD

3.1 Data Description

The data investigated in this work was gathered from the CREDBANK corpus [36]. The corpus contains 1,377 events as they surfaced on Twitter between October 2014 and February 2015, their corresponding public tweets (a total of 66M messages) and their associated credibility ratings. The corpus was built by iteratively tracking millions of public posts streaming through Twitter; computationally detecting the underlying topics (i.e. clusters) of discussion in every block of million tweets; separating event-specific topics from non-event topics by asking independent human raters from Amazon Mechanical Turk (AMT); and then for each of these event topics gathering credibility ratings on a 5-point Likert scale ranging from 'Certainly Inaccurate' (-2) to 'Certainly Accurate' (+2). Each event is represented by a combination of top three terms from that event topic. For example, "chelsea", "game", "goal" refers to a football match event at a point in time when Chelsea scored a goal. Thirty independent human raters from AMT judged the accuracy level of an event by browsing through tweets in real-time, where each tweet contained all the top three topical terms. Such a task design ensured that the annotation task closely mimics the way a person would search Twitter to see information related to an event. Moreover, limiting tweets containing all top 3 topical terms ensured a balance between being too generic (by including fewer terms) and too specific (by including more terms), and also provided enough context to a human rater for performing the task.

To guarantee that the collected ratings were at par with expert level judgments, the creators of CREDBANK performed several controlled experiments before deciding on the strategy best suited for fetching quality annotations [37]. Finally, all tweets specific to the event topic were collected from the past seven days (a limit imposed by Twitter's search program interface). This did not seem to be a limitation because the experimental setup tracked recent events. Moreover, research has demonstrated that news-worthy topics on Twitter have an active period of at most a week [23]. Overall, this iterative framework allowed for a natural experimental setup with credibility of social media information being tracked soon after it gained collective attention. Additional detail of the data collection steps is outlined in the *Appendix*. A representative sample of events tracked during this three month period along with their credibility ratings is presented in Table 1. By listing the range of diverse events, our aim is to demonstrate the richness of the dataset and hence the

Event Terms	Representative Tweets	Start Time	End Time	Ratings	p_{ca}
Perfect Credibility: $0.9 \leq P_{ca} \leq 1$					
apple ipad air	Apple Introduces iPad Air 2 No surprises at Apple's unveiling of iPad Air 2 and Retina 5K display	2015-01-16 01:50	2015-01-23 01:50	[0 0 0 3 27]	0.90
hughes rip phil	RIP Phil Hughes you will remain in our hearts for ever!!! A very sweet tribute to Phil Hughes! RIP	2014-11-25 11:40	2014-11-28 09:00	[0 0 1 2 27]	0.90
High Credibility: $0.8 \leq P_{ca} < 0.9$					
court marriage supreme	Supreme Court legalizes gay marriage The Supreme Court has decided to rule regarding same-sex marriage!	2015-01-12 13:20	2015-01-20 06:00	[1 0 0 3 26]	0.87
romney mitt run	Must Read: Mitt Romney and His Decision Not to Run, by Peter Wehner Mitt Romney Will Not Run in 2016	2015-01-27 04:30	2015-02-03 22:50	[2 0 2 6 50]	0.83
Moderate Credibility: $0.6 \leq P_{ca} < 0.8$					
torture report cia	Senate Report Says Torture Program Was More Gruesome, Widespread Than CIA Claimed Senate report on CIA torture claims spy agency lied about 'ineffective' program	2014-12-11 06:10	2014-12-12 10:40	[1 1 0 4 24]	0.80
ebola york nyc	EBOLA IN NYC?! New York Officials Urge Calm Amid Ebola Diagnosis: New York Gov. Cuomo said	2014-10-17 18:50	2014-10-25 22:50	[1 1 5 3 20]	0.67
Low Credibility: $0 \leq P_{ca} < 0.6$					
school paris attack	There are reports of an attack at a school, in Paris – do you have anything? Third suspect was in school at time of Paris attack, classmates say	2015-01-05 10:50	2015-01-11 06:40	[0 1 3 9 17]	0.57
next coach michigan	The big question: Who will be Michigan's next football coach? My multiple sauces tell me Lane Kiffin will be the next Michigan coach.	2014-12-10 17:20	2014-12-18 11:40	[0 1 7 15 7]	0.23

Table 1. Example events from the four credibility classes along with representative tweets from CREDBANK. Events are represented with three event terms. Start and end times denote the time period during which tweets were collected using Twitter's search API. Rating shows the count of Turkers who selected an option from the 5-point Likert scale ranging from -2 ("Certainly Inaccurate") to +2 ("Certainly Accurate").

generalizability of our results ensuing from this dataset. For example, it contains events ranging from celebrity deaths to terroristic attacks.

3.2 Pruning Corpus for Sample Independence

During the iterative building of CREDBANK, if an event trended on Twitter for a sufficiently long time period, it is possible that the event is curated multiple times. For example, the event "arsenal", "win", "city" corresponds to the Arsenal's winning the football match against Stoke city. People on Twitter had active conversations about the event for several hours, resulting in the event being captured more than once in CREDBANK. However, our statistical analysis (discussed shortly) required sample independence. Occurrence of multiple instances of the same event will likely violate the independence assumption. Hence, we pruned our event sample to keep single distinct instances of each event. By matching the three terms in each event topic, we looked for duplicate event occurrences. Thereafter, if multiple instances of the same event existed, we picked the event which had the earliest curation time. Restricting events by earliest curation times ensured that we retained crowd worker annotations corresponding to the very first time that they performed the annotation task; hence preventing any potential prior knowledge bias. Our pruning step resulted in a dataset of 1,138 events spanning 47,000,127 tweets.

Cred Class	P_{ca} Range	Total Events	Distinct Events
Perfect	$0.9 \leq P_{ca} \leq 1.0$	421	342
High	$0.8 \leq P_{ca} < 0.9$	433	337
Moderate	$0.6 \leq P_{ca} < 0.8$	414	358
Low	$0.0 \leq P_{ca} < 0.6$	109	101

Table 2. **Credibility classes and corresponding event counts.** *Total Events* column shows event counts from CREDBANK. *Distinct Events* lists counts from the pruned corpus.

3.3 Credibility Classification

We measured an event's perceived credibility level based on how many human raters agreed that the event was "Certainly Accurate". Our approach is motivated from prior work which used CREDBANK to find language signals of credibility [38]. More formally, for each event we find the proportion P_{ca} of ratings marked as "Certainly Accurate".

$$P_{ca} = \frac{Count\ \text{"Certainly Accurate"}\ ratings\ for\ an\ event}{Total\ ratings\ for\ that\ event}$$

Next, we placed P_{ca} into four classes that cover a range of values (see Table 2). The class names are based on the perceived degree of accuracy of the event in that class. For example, events belonging to the "Perfect Credibility" class were rated as "Certainly Accurate" by almost all raters ($0.9 \leq P_{ca} < 1$). Representing P_{ca} over a four-class categorical variable instead of a continuous variable with a category corresponding to every value of P_{ca}, ensures that the comparisons across classes are reasonable and interpretive. Adapting the approach from [38], we validated our credibility classification by comparing classes generated by our P_{ca} method against those obtained via data-driven classification methods. We found high similarity between the two types of classification techniques, but we opted for our proportion-based (P_{ca}) technique, since it is much more interpretable, readily generalizable, and adaptable to domains other than Twitter on which CREDBANK was constructed.

4 STATISTICAL MEASURES

To understand the relation between collective attention and information credibility, we computed our measures using time-stamped tweets from the CREDBANK corpus, where groups of tweets corresponded to discussions of an event by multiple Twitter users over a certain time span. In designing our collective attention measures, we draw inspiration from prior work that have examined the relationship between collective attention and various facets of information, such as quality, novelty and popularity. In a seminal work, investigating the relation between collective attention and content novelty, Wu and Huberman modeled the shift in attention over a time scale [52]. One of the first studies to model collective human dynamics through the lens of temporal activity was Crane and Sornette's classic work on understanding the differences in attention bursts towards videos on YouTube [11]. Using time series of daily video views, they found three distinct classes of temporal trends associated with collective human dynamics. The trends also differed by video quality and type. For example, "junk" videos showed a sudden peak of activity followed by subsequent decay, whereas "viral" videos experienced bottom-up growth of viewing activity over time. Similar temporal treatment has been employed by Ratkiewicz and colleagues while modeling random collective shifts of attention among Wikipedia topics and Chilean Web pages [41]. Our rationale for investigating the link between temporal aspects of collective attention and credibility is largely based on the success of these prior studies.
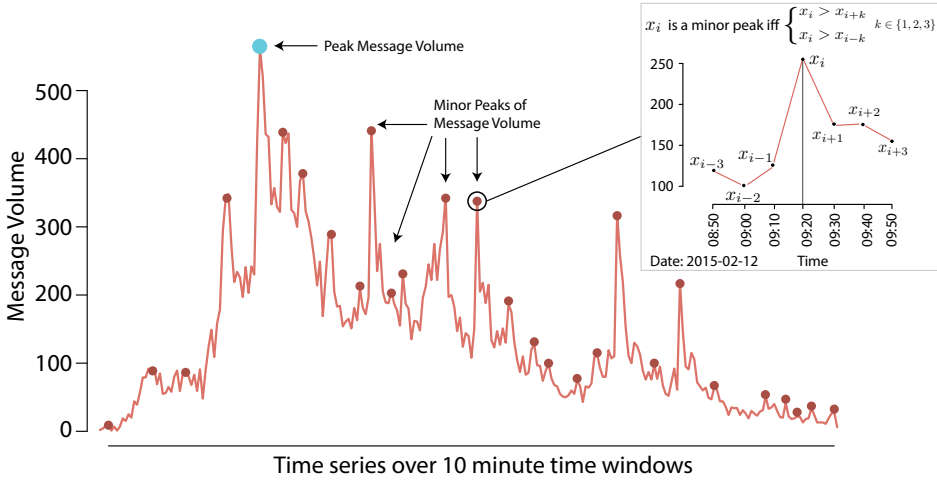
The inset diagram shows:

x_i is a minor peak iff $\begin{cases} x_i > x_{i+k} \\ x_i > x_{i-k} \end{cases} k \in \{1, 2, 3\}$

Fig. 1. **The time series of message volume for a sample event reported on Twitter. The event corresponds to Twitter discussions, where each tweet contained all three terms: "#chapellhillshooting", "muslim" and "white". The ● dot corresponds to the time window having maximum *message volume* while the ● dots correspond to the minor peaks observed in this volume. The inset diagram on the right side zooms in on one of the minor peaks, along with the rule triggering its designation.**

4.1 Collective Attention Metrics

Collective attention of each event reportage was measured using two metrics:

(1) *Message Volume*: message volume tracks the aggregate number of messages over time.
(2) *People Volume*: people volume records the aggregate count of unique users paying attention to the story over time.

Each measure is represented as a time series with message (or unique user) volumes aggregated over 10-minute time intervals. Our choice of a 10-minute window is supported by studies showing that Twitter acts as a medium for reporting breaking news and hence is characterized by fast diffusion of information [22, 23]. Moreover, we ran our statistical analysis by re-computing each event timeseries over a range of other time intervals (5, 20, and 30-minute time window sizes), and found consistent results. Thus, tracking collective attention on the order of 10-minutes is a reasonable representation of a rapidly evolving phenomenon. Each event may differ in the temporal dynamics of its collective attention; thus inferences drawn on a small set of events tracked for a few days may be confounded by temporal traits peculiar to certain news stories. However, by tracking news stories over several months and averaging over hundreds of such collective attention rhythms, our results represent the most consistent relations between the dynamics of collective attention and perceptions of information credibility. Our rationale for using *people volume*, in addition to *message volume*, as a collective attention metric is to ensure that the collective attention measured is not confounded by superfluous posting activity from potential Twitter bots, automated programs posing as human beings [9]. Since *people volume* corresponds to the unique number of individuals paying attention to the event over time, it aims to counteract any extreme posting activity by such bots.

As an example illustrating our data and methods, Figure 1 shows aggregate message volume for an event reported on Twitter where every message contained the terms "#chapellhillshooting", "muslim", and "white". On February 10, 2015, three Muslim students in Chapel Hill, North Carolina were shot to death by a white neighbor and speculations concerning the motives of the shooter surrounded the event [15]. While authorities suggested the motive to be an ongoing dispute between neighbors over a parking space, many social media users suggested a hate crime as the motive. Twitter messages concerning this specific topic on February 12 blamed the media for ignoring the coverage of an event involving Muslim killings and suggested the shooting was an act of terrorism and so a hate crime. It was not until February 13 that authorities opened an investigation to determine if the shooting was in fact a hate crime. Credibility rating distributions showed that less than 50% of raters agreed that the social media reportage of the event was "Certainly Accurate", thereby questioning the alleged terror claims underlying the act.

4.2 Temporal Measures of Collective Attention

To quantify the importance of the time when collective attention maximized, we first computed the strict global maximum in the time series [48]. We call this the *peak attention*. This is the ratio of messages (or unique people) within the peak time window to the total cumulative volume of messages (or unique people) over the entire event time series. Similar peak fraction measures have been used in prior work to model the collective human viewing behavior of online videos [11].

$$PeakAttention = \frac{max(x_1, \cdots, x_n)}{\sum\limits_{i=1}^{n} x_i} \tag{1}$$

where x_i is the count of messages (or unique people) in time window i in an event time series $x_1, x_2, \cdots, x_n$.

Our choice of the above measure is based on the success of prior studies using peak-fraction based metrics to successfully characterize herding behavior over time [11, 54]. To illustrate how peak attention measure can characterize variations in time series, consider the example of an event marked by a sudden spike in collective attention followed by a subsequent drop. The lack of precursory growth suggests that most of the attention was concentrated on the peak, thereby resulting in high *peak attention* (Figure 2c and 2d). Whereas, an event with steady growth in collective attention, followed by a gradual decay would imply a relatively smaller fraction of attention in the peak, thus leading to lower value of *peak attenion* (Figure 2b).

While *peak attention* captures the importance of the maximal burst in collective attention, it does not take into account the presence or absence of spikes in the precursory growth and in the subsequent decay following the burst. Hence, we define a measure to quantify the spikiness in collective attention. We detect all strict local maxima [48] in each of the event time series. A strict local maxima corresponds to an instance in the time series when the volume of messages (or unique people) is larger than the volume in the neighboring time windows. We define this neighborhood as three time windows on either side of the local maxima and call these local maxima *minor peaks*. Thus, the attention in a minor peak is higher than the attention 30 minutes (i.e., three time windows times 10-minute window size) before and after the occurrence of a peak.

$$x_i \text{ is a minor peak iff } \begin{cases} x_i > x_{i+k} \\ x_i > x_{i-k} \end{cases}, k \in \{1, 2, 3\} \tag{2}$$

We then define minor peak attention as the ratio of messages (or unique people) in the local maxima relative to the total cumulative volume of messages (or unique people) over the entire event time series. Formally, if $\mathcal{M}$ is the set of all minor peak indices in an event's message (or unique people)
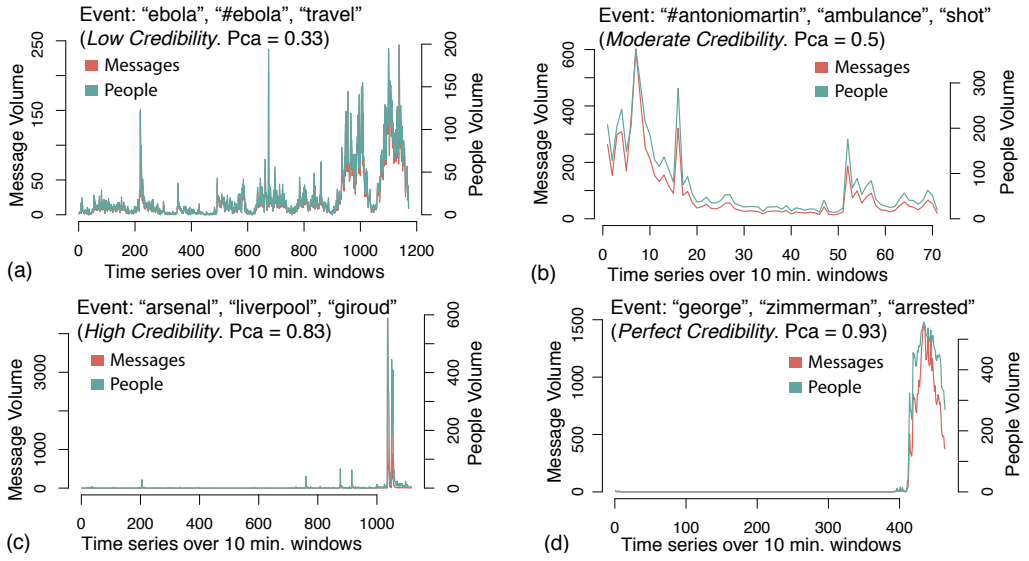
Fig. 2. Time series of collective attention metrics (message volume and unique people volume) for example events in each credibility class. The examples show representative behavior of collective attention metrics in each credibility class. While events in all four classes are marked by peak attention with respect to both message and people volume, events in the low and moderate credibility classes exhibit multiple minor peaks, signifying that persistent attention is characteristic of lower credible social media events.

time series, then minor peak attention is defined as follows:

$$Minor\ Peak\ Attention = \frac{\sum\limits_{j \in \mathcal{M}} x_j}{\sum\limits_{i=1}^{n} x_i} \qquad (3)$$

The inset diagram in Figure 1 shows a local maximum. While the peak attention captures the maximum momentary interest that an event acquires during its lifetime on Twitter, the points representing minor peak attention reflect renewed and ongoing recurrences of momentary interest. Additionally, both these measures have two important properties: both are invariant with respect to scaling and shifting [54]. First, since both measures are proportions based on cumulative collective attention, they are invariant to the overall volume of attention. Hence, two event time series having similar peaky shapes but different total attention volumes would be treated similarly. Secondly, both measures are computed independent of the maxima position on the time axis. Thus, if two event time series peaks occur at different times but possess a similar peaky structure, the measures will be invariant to the translations on the time axis. Hence, both these measures—despite being simple representations of temporal dynamics—are useful in interpreting the relationship between collective attention rhythms and event credibility across a range of different events exhibiting high variability in overall popularity and time of popularity.

5 STATISTICAL ANALYSIS AND RESULTS

We tested the differences in collective attention measures across the credibility classes using the Wilcoxon Rank Sum or Mann-Whitney U test. Mann-Whitney test is a non-parametric test which
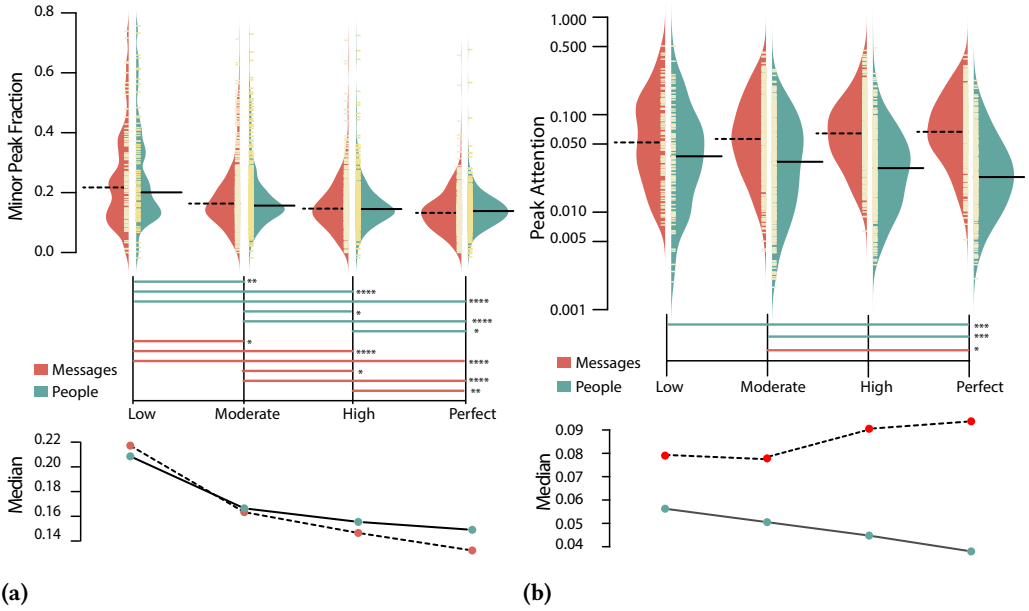
(a)   (b)

Fig. 3. **Collective attention shown as a beanplot distribution. The shape of each half of the asymmetric bean represents the Gaussian density estimation of the distribution. The lines (in yellow) are the actual data points; the dotted long bean line is the median corresponding to the message volume, the solid line shows the median for people volume. The * denotes pairwise significant differences between cluster medians after correcting for familywise error-rate. (a). Proportion of minor peak fractions are statistically different across all credibility class pairs for both message and unique people volume. (b). Peak attention is significantly different across "Low" and "Perfect", and "Moderate" and "Perfect" credibility classes for unique people volume, and "Perfect" and "Moderate" classes for message volume. The line charts at the bottom panel show the median trends across the credibility classes.**

does not make any assumption about the distribution of the response variable and is ideal for ordinal responses such as ours. For each temporal measure (peak and minor peak attention) and for each collective attention metric (message and people volume), we performed pairwise Wilcoxon Rank Sum tests, followed by Bonferroni corrections [13] to control for potential inflation of the family-wise error rate by multiple test comparisons. We found that, for both *message volume* and *people volume*, differences in the minor peak fraction are statistically significant ($p < 0.00833$ after Bonferroni corrections and using Wilcoxon Rank Sum tests). As shown in Figure 3, median *minor peak attention* decreases as credibility level increases from "Low" to "Perfect". We also found a significant moderate degree of negative correlation between P_{ca} and minor peak fraction for both *message volume* ($r = -0.33$) and *people volume* ($r = -0.33$). These results suggest that an event attracting renewed interest is associated with lower perceived credibility. On the contrast, *peak attention* of messages was only statistically different between "Perfect" and "Moderate" credibility classes. Peak attention for *people volume* could only provide coarse-grained information separating "Low" and "Perfect", and "Moderate" and "Perfect" credibility classes. These results indicate that *peak attention* is not a useful signal for event credibility. To ensure that these ratio-based, collective attention measures described above are not sensitive to event duration, which can affect the

	P	H	M	L
P		0.00012**	5.4e-12****	4.6e-11****
H	0.002639*		0.001703*	3.6e-06****
M	1.0e-08****	0.004221**		0.00547*
L	8.4e-11****	1.5e-06****	0.001981**	

(a) Minor peak attention pairwise statistical differences

	P	H	M	L
P		ns	0.00322*	ns
H	ns		ns	ns
M	0.00093***	ns		ns
L	0.00045***	ns	ns	

(b) Peak attention pairwise statistical differences

Table 3. Pairwise statistical significance after Wilcoxon Rank Sum tests. P, H, M, L correspond to Perfect, High, Moderate and Low credibility classes. The top half of the diagonal corresponds to *message volume*, while bottom half shows pairwise differences in *people volume*. ns stands for non-significance.

denominator (cumulative volume), we performed pairwise Wilcoxon Rank Sum test comparisons of event duration across the credibility classes. We found no significant difference, indicating that event duration does not skew the collective attention metrics for a particular credibility class. Moreover, to ensure that our collective attention metrics are independent observations over time–a criteria necessary for the validity of our statistical analysis–we performed Ljung-Box Q (LBQ) tests [30]. We were able to reject null hypothesis for each of our LBQ tests; thus confirming that our collective attention measures for both message and people volume are independent over time.

In order to establish robustness of our results, in addition to these comparative tests, we also performed predictive analytics. By running a series of regression models, we checked whether credibility can be predicted by collective attention measures after including additional controls as features. We found that the minor peak for message volume and the interaction effects between minor peak fraction and peak attention corresponding to both message and people volume were significant in our predictive models. The results are summarized in the appendices.

6 DISCUSSION

By investigating a large-scale longitudinal credibility corpus (CREDBANK), we were able to test the relationship between an event's perceived credibility level and the temporal dynamics of its collective attention. According to our findings, moments of renewed collective attention are associated with event reportage marked by decreased levels of perceived credibility. Perhaps the underlying uncertainty of an event is what keeps people yearning for additional information and be engaged in collective sense-making. Research has shown that situations of uncertainty are often marked by collective problem solving and sense making by people to understand their environment [45, 47]. This bears resemblance to *milling* behavior exhibited during crisis situations [28]. Milling occurs following an initial warning message—a period during which people move about aimlessly trying to make sense of the situation [35]. Moreover, incomplete information following a warning message, leads to additional milling activity as people attempt to make sense of the crisis situation [34]. Recent studies have also demonstrated online milling activity in Twitter during disaster events [47, 49]. At the wake of a crisis, after the initial warning message, serial message transmission

via retweeting resulted in message amplification and inidividuals being involved in sense-making activities before taking protective action [47, 49].

Do frequent peaks in collective attention lead to lower perceived credibility? Or do reduced levels of credibility spark the continued interest in the event? Our current study cannot establish the causal direction of this relation. However, we are able to establish that the persistence of collective attention peaks is a reliable temporal signature for an event's perceived credibility level. Despite being central to the spread of news events, collective attention and its relation to veracity of news has not been studied. We envision that these results will lay a new foundation for the field to further investigate the tripartite relation between time, collective attention, and credibility of online information. Here we propose a research agenda outlining broad categories of questions that we hope will spark a more robust study of this relation, and in turn offer new insights about the misinformation ecosystem.

- Theoretical underpinnings: What are the theoretical underpinnings of credibility and collection attention? Under what circumstances is temporal collective attention enough to accurately predict credibility? What happens when we consider information content along with temporality?
- Information novelty: Research on collection attention have noted that attention to novel items eventually fades as it propagates further [52]. How does temporal decay of information-novelty relate to information accuracy? Can we build simple models that characterize this relation?
- Ways of tracking attention: Our work is based on tracking collective attention from the past seven days after an event occurs. What happens when you have a more constrained view of the attention? For example, imagine a scenario where the event has just trended and we only have a snapshot of the attention. Can we determine how the attention and the associated certainty or uncertainty of the event will change in the next hour?

We remark that by using a simple proportion based classification technique, we identified robust and scalable credibility classes; hence it is also potentially applicable to other online settings where user's collective attention drives popularity of content. Moreover, by using simple and interpretable parameters computed on times series of minute-wise user and message attention, we revealed vital temporal indicators associated with information credibility. Contrary to other sophisticated methods which require the estimation of power-law exponents for unraveling collective attention dynamics, or the calculation of costly correlations between activity time series, the parameters employed here can easily be computed in a scalable way. Although devoid of any predictive power, these measures can support the discovery of collective attention patterns in large-scale records of human activity. Using our results one can envision an application which does temporal tagging of social media event streams, showing streams which are marked by multiple minor peaks, thereby implicitly hinting at the underlying credibility of the information stream or making regions characterized by *milling* activity.

On the basis of these results, we envision that organizations struggling to handle the propagation of online misinformation [31] can harness the temporality of collective attention to predict the level of credibility. We may be able to subsequently design interventions aimed at controlling the spread of false information or cautioning social media users to be skeptical about an evolving topic's veracity, ultimately raising an individual's capacity to assess credibility of information. Imagine a news reporting tool which shows social media discussions highlighting areas which witnessed multiple minor peaks of human activity, or think of a fact-checking system that compares temporal regions of high minor peak attentions to those with fewer attention peaks, or consider temporal tagging of scientific discourse or medical records emphasizing areas that garnered intermittent

temporal popularity. We foresee that our findings can lead to a new class of such temporally aware systems which underscore degrees of information uncertainty based on temporal signals of collective attention. Finally, our study has practical implications in the field of computational social science where inferences about human social behavior are based on reports of online interactions [10] and trusting the credibility of those reports is crucial for any downstream analysis. For example, imagine a health researcher investigating the spread of Ebola via social media reports or a financial trader gauging market volatility based on breaking news and citizen reports on social media; veracity of those reports will affect the subsequent inferences.

7 CONCLUSION

To study the dynamics of collective attention and its relation to information credibility in a natural setting, we analyzed the temporal patterns of 47M Twitter messages spread across 1,138 social media events along with their in-situ credibility ratings. We do so by multiple statistical comparison tests over parameters computed on the time series of collective attention of messages and distinct users. Although simple, this approach provides fundamental insights about collective attention and information credibility that would otherwise be missed by more complicated predictive analysis methods.

ACKNOWLEDGMENTS

8 APPENDIX: ROBUSTNESS CHECK

8.1 Predictive Analytics to check robustness

While comparative tests allow us to inspect whether collective attention differs across credibility groups, it does not provide any predictive analytics. We turn to two types of regression models to check whether credibility can be predicted by collective attention measures. Table 4 presents a series of models.

8.1.1 Cumulative Link Models. In our first set of models, we treat our response variable as an ordered list of credibility levels ("Low" < "Moderate" < "High" < "Perfect") and use *cumulative link models* [8] to fit our data. Cumulative link models (CLM) are ordinal regression models which provide the perfect framework for rightfully modeling an ordinal dependent variable. While there are multiple ordinal regression models, CLMs are the most popular and the most widely used [8]. We implemented the cumulative link model using R's clm function from the ordinal[1] package. The CLM model takes in a vector of explanatory variables as input (e.g., minor peak attention for message volume) and outputs the cumulative probabilities denoting whether an observation falls in a particular ordered response category (e.g., "Moderate" credibility). Model 1 in Table 1 shows the results after fitting clm using all the collective attention measures as explanatory variables. The β coefficients and odds ratios determine the relative power of the variables in predicting credibility levels. To measure interaction effects we include interaction terms between minor peak fraction and peak attention corresponding to both message and people volume.

Next, we examine the effect of collective attention measures after including additional features as controls. We add the following seven features. These features have also been used in previous research for automatic credibility assessment [6, 38].

---

[1]https://cran.r-project.org/web/packages/ordinal

	CLM (Ordinal Regression)						LM (Linear Regression)			
	DV: Credibility Levels (Low <Moderate <High <Perfect)						Dep. Var.: Average Credibility			
	Model 1 - CLM			Model 2 - CLM			Model 3 - LM		Model 4 - LM	
	β	O.R.	S.E.	β	O.R.	S.E.	β	S.E.	β	S.E.
(Intercept)							0.00159	0.00 ***	0.00160	0.00 ***
Minor Peak Attention$_{Message}$	-8.20	(0.04)	2.18 ***	-8.06	(0.03)	2.23 ***	-0.00081	0.00 **	-0.00069	0.00 **
Peak Attention$_{Message}$	-0.74	(0.34)	1.33	-1.06	(0.30)	1.35	-0.00015	0.00	-0.00013	0.00
Peak Attention$_{People}$	-2.50	(1.33)	1.90	-2.84	(1.19)	1.95	-0.00031	0.00	-0.00038	0.00 .
Minor Peak Attention$_{People}$	4.23	(0.00)	2.28 .	3.82	(0.00)	2.33	0.00023	0.00	0.00011	0.00
MinorPeakAtten$_{Msg}$ X PeakAtten$_{Msg}$	53.46	(0.48)	10.16 ***	52.78	(0.95)	10.26 ***	0.00620	0.00 ***	0.00580	0.00 ***
MinorPeakAtten$_{People}$ X PeakAtten$_{People}$	-60.29	(0.08)	12.09 ***	-60.23	(0.88)	12.31 ***	-0.00650	0.00 ***	-0.00614	0.00 ***
Total Tweets				-0.54	(0.35)	0.42			-0.00003	0.00
Word count				0.59	(0.06)	0.37			0.00006	0.00 *
Uppercase				0.13	(45.80)	0.31			0.00000	0.00 **
Hashtags				-0.04	(0.58)	0.11			-0.00003	0.00
Quotations				-0.05	(1.81)	0.08			-0.00001	0.00
Questions				-0.13	(1.14)	0.10			-0.00002	0.00
Total users				-0.17	(0.96)	0.09 .			-0.00001	0.00
	LogLikelihood		-1391.12	LogLikelihood		-1384.01	R2	0.1574	R2	0.1733
	AIC		2800.24	AIC		2800.03	Adj. R2	0.1529	Adj. R2	0.1637

$* * * p < 0.001$ $* * p < .01$ $* * p < .01$ $* p < .05$ $\cdot p < .10$ $N = 1138$

Table 4. Models predicting credibility. Models 1 and 2 are cumulative link regression models which predict credibility as an ordinal outcome variable with five ordered levels (Low < Moderate < High < Perfect). Models 3 and 4 are linear regression models which predict average credible. Both models 2 and 4 control for additional features to test the relative importance of the temporal measures of collective attention. Coefficients and standard errors are shown. Odds ratios are also shown for the CLM models.

(1) Total tweets - count of tweets in each event topic.
(2) Word count - total number of words in each event topic.
(3) Uppercase - count of capital letters in all tweets of an event.
(4) Hashtags - total hashtag occurrences in the event topic.
(5) Quotations - count of quoted content in tweets of an event.
(6) Questions - count of question marks in each event topic.
(7) Total users - number of people tweeting and retweeting about the event.

Since the values of each of these variables depend on how much an event trended, they may differ widely across events with varying degrees of popularity. Thus to place values on an approximately common scale, we standardize variables by centering at the mean and dividing by the standard deviation.

We find that minor peak fraction for *message volume* and the interaction terms are significant in both models. Surprisingly, none of the control variables are significant in Model 2. The coefficient estimates for CLMs are given in ordered log odds. Looking at the significant β estimates, we find that for a one unit increase in the minor peak attention for *message volume*, there will be a 8.20 (8.06 with controls) decrease in the credibility level on the log odds scale. We can also interpret the log odds coefficients in terms of odds ratios. An odds ratio less than one indicates that as the independent variable increases, the dependent variable decreases, while the converse is true for an odds ratio greater than one. In both models 1 and 2, we find that the odds ratio for minor peak fraction corresponding to *message volume* is less than 1 and significant, indicating that the higher the minor peak attention for *message volume*, the lower the likelihood that the event is credible.

The coefficient of the interaction effect between minor peak and peak attention for *message volume* is significant and positive, denoting that the effect of minor peak for *message volume* is greater when peak attention for messages increases. This indicates that even when there is a maximal burst in collective attention of messages, the presence of spikiness in collective attention of messages has an increasing effect on credibility. It is curious to note that while minor peak attention for *people volume* is not significant, there is a significant interaction effect between the corresponding minor peak and peak attention measures. The coefficient of the interaction effect is negative. This indicates that the effect of minor peak attention for *people volume* is greater when peak attention for *people volume* is low. In other words, when the maximal burst in collective attention by distinct users is low, the spikiness in collective user attention has a larger effect on the credibility level.

8.1.2 Linear Models. In order to check the robustness of our results and to ensure that the observed effects is not an artifact of our model or how we measure credibility, we repeat our predictive analytics using average credibility as the outcome variable.

$$Average\ Credibility = \frac{\sum_{i=-2}^{2} i * number\ of\ raters\ who\ rated\ i}{total\ number\ of\ raters} \tag{4}$$

With a continuous variable as our dependent variable, we use linear regression to test how well the collective attention measures can predict average credibility. We fit two sets of models—model 3 with only the collective attention measures, while model 4 includes additional controls. As before, we find that the minor peak for *message volume* and the interaction effects are significant in both models. The only controls which are significant are the number of words and the amount of uppercase letters present in the tweets for an event. However, the coefficient estimates are fairly small.

REFERENCES

[1] Ahmer Arif, Kelley Shanahan, Fang-Ju Chou, Yoanna Dosouto, Kate Starbird, and Emma S. Spiro. 2016. How Information Snowballs: Exploring the Role of Exposure in Online Rumor Propagation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, 466–477. https://doi.org/10.1145/2818048.2819964

[2] Sitaram Asur, Bernardo A. Huberman, Gabor Szabo, and Chunyan Wang. 2011. Trends in Social Media: Persistence and Decay. In *International AAAI Conference on Web and Social Media (ICWSM '11)*.

[3] Albert-Laszlo Barabasi. 2005. The origin of bursts and heavy tails in human dynamics. *Nature* 435, 7039 (2005), 207–211.

[4] M Bass. 1969. Frank. *Management Science* 15 (1969), 215 – 227.

[5] Prashant Bordia and Ralph L Rosnow. 1998. Rumor Rest Stops on the Information Highway Transmission Patterns in a Computer-Mediated Rumor Chain. *Human Communication Research* 25, 2 (1998), 163–179.

[6] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information Credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web (WWW '11)*. ACM, New York, NY, USA, 675–684. https://doi.org/10.1145/1963405.1963500

[7] Andrea Caumont. 2013. *12 trends shaping digital news* (Pew Research Center). http://www.pewresearch.org/fact-tank/2013/10/16/12-trends-shaping-digital-news. (2013).

[8] Rune Haubo Bojesen Christensen. 2010. ordinal-regression models for ordinal data. *R package version* 22 (2010).

[9] Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. 2010. Who is Tweeting on Twitter: Human, Bot, or Cyborg?. In *Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC '10)*. ACM, New York, NY, USA, 21–30. https://doi.org/10.1145/1920261.1920265

[10] Scott Counts, Munmun De Choudhury, Jana Diesner, Eric Gilbert, Marta Gonzalez, Brian Keegan, Mor Naaman, and Hanna Wallach. 2014. Computational Social Science: CSCW in the Social Media Era. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW Companion '14)*. ACM, New York, NY, USA, 105–108. https://doi.org/10.1145/2556420.2556849

[11] Riley Crane and Didier Sornette. 2008. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences* 105, 41 (2008), 15649–15653.

[12] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. 113, 3 (2016), 554–559. https://doi.org/10.1073/pnas.1517441113

[13] Olive Jean Dunn. 1959. Estimation of the Medians for Dependent Variables. *The Annals of Mathematical Statistics* 30, 1 (1959), 192–197.

[14] Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor Cascades. In *International AAAI Conference on Web and Social Media (ICWSM '14)*.

[15] Sam Frizell. 2015. *3 Muslim Students Murdered in North Carolina*. time.com/3704759/ muslim-students-murdered-chapel-hill. (2015).

[16] Sharad Goel, Jake M. Hofman, SÃĺbastien Lahaie, David M. Pennock, and Duncan J. Watts. 2010. Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences* 107, 41 (2010), 17486–17490. https://doi.org/10.1073/pnas.1005962107

[17] Scott A. Golder and Michael W. Macy. 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science* 333, 6051 (2011), 1878–1881. https://doi.org/10.1126/science.1202775

[18] Daniel Gruhl, R. Guha, David Liben-Nowell, and Andrew Tomkins. 2004. Information Diffusion Through Blogspace. In *Proceedings of the 13th International Conference on World Wide Web (WWW '04)*. ACM, New York, NY, USA, 491–501. https://doi.org/10.1145/988672.988739

[19] Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo, and Patrick Meier. 2014. TweetCred: Real-Time Credibility Assessment of Content on Twitter. In *Social Informatics: 6th International Conference, SocInfo 2014, Barcelona, Spain, November 11-13, 2014. Proceedings*, Luca Maria Aiello and Daniel McFarland (Eds.). Springer International Publishing, Cham, 228–243. https://doi.org/10.1007/978-3-319-13734-6_16

[20] Frank A Haight and Frank A Haight. 1967. Handbook of the Poisson distribution. (1967).

[21] Lee Howell. 2013. Digital wildfires in a hyperconnected world. *World Economic Forum* (2013).

[22] Mengdie Hu, Shixia Liu, Furu Wei, Yingcai Wu, John Stasko, and Kwan-Liu Ma. 2012. Breaking News on Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2751–2754. https://doi.org/10.1145/2207676.2208672

[23] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a Social Network or a News Media?. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, New York, NY, USA, 591–600. https://doi.org/10.1145/1772690.1772751

[24] Janette Lehmann, Bruno Gonçalves, José J. Ramasco, and Ciro Cattuto. 2012. Dynamical Classes of Collective Attention in Twitter. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 251–260. https://doi.org/10.1145/2187836.2187871

[25] Jure Leskovec, Lada A. Adamic, and Bernardo A. Huberman. 2007. The Dynamics of Viral Marketing. *ACM Trans. Web* 1, 1, Article 5 (May 2007). https://doi.org/10.1145/1232722.1232727

[26] Jure Leskovec, Lars Backstrom, and Jon Kleinberg. 2009. Meme-tracking and the Dynamics of the News Cycle. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09)*. ACM, New York, NY, USA, 497–506. https://doi.org/10.1145/1557019.1557077

[27] Qinying Liao and Lei Shi. 2013. She Gets a Sports Car from Our Donation: Rumor Transmission in a Chinese Microblogging Community. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 587–598. https://doi.org/10.1145/2441776.2441842

[28] Michael K Lindell and Ronald W Perry. 2003. *Communicating environmental risk in multiethnic communities*. Vol. 7. Sage Publications.

[29] Fang Liu, Andrew Burton-Jones, and Dongming Xu. 2014. RUMORS ON SOCIAL MEDIA IN DISASTERS: EXTENDING TRANSMISSION TO RETRANSMISSION. In *PACIS 2014 Proceedings*.

[30] Greta M Ljung and George EP Box. 1978. On a measure of lack of fit in time series models. *Biometrika* 65, 2 (1978), 297–303. https://doi.org/10.1093/biomet/65.2.297

[31] Victor Luckerson. 2014. *Fear, Misinformation, and Social Media Complicate Ebola Fight*. Time Inc. time.com/3479254/ ebola-social-media/. (2014).

[32] Jim Maddock, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. 2015. Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 228–241. https://doi.org/10.1145/2675133.2675280

[33] Michael Mathioudakis and Nick Koudas. 2010. TwitterMonitor: Trend Detection over the Twitter Stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data (SIGMOD '10)*. ACM, New York, NY, USA,

1155–1158. https://doi.org/10.1145/1807167.1807306

[34] Dennis S Mileti and Colleen Fitzpatrick. 1991. *Communication of public risk: Its theory and its application.*

[35] Dennis S Mileti and John H Sorensen. 1990. *Communication of emergency public warnings: A social science perspective and state-of-the-art assessment.* Technical Report. Oak Ridge National Lab., TN (USA).

[36] Tanushree Mitra and Eric Gilbert. 2015. CREDBANK: A Large-Scale Social Media Corpus with Associated Credibility Annotations. In *International AAAI Conference on Web and Social Media (ICWSM '15).*

[37] Tanushree Mitra, C.J. Hutto, and Eric Gilbert. 2015. Comparing Person- and Process-centric Strategies for Obtaining Quality Data on Amazon Mechanical Turk. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15).* ACM, New York, NY, USA, 1345–1354. https://doi.org/10.1145/2702123.2702553

[38] Tanushree Mitra, Graham P. Wright, and Eric Gilbert. 2017. A Parsimonious Language Model of Social Media Credibility Across Disparate Events. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17).* ACM, New York, NY, USA, 126–145. https://doi.org/10.1145/2998181.2998351

[39] Meredith Ringel Morris, Scott Counts, Asta Roseway, Aaron Hoff, and Julia Schwarz. 2012. Tweeting is Believing?: Understanding Microblog Credibility Perceptions. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12).* ACM, New York, NY, USA, 441–450. https://doi.org/10.1145/2145204.2145274

[40] Vahed Qazvinian, Emily Rosengren, Dragomir R. Radev, and Qiaozhu Mei. 2011. Rumor Has It: Identifying Misinformation in Microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '11).* Association for Computational Linguistics, Stroudsburg, PA, USA, 1589–1599.

[41] Jacob Ratkiewicz, Santo Fortunato, Alessandro Flammini, Filippo Menczer, and Alessandro Vespignani. 2010. Characterizing and Modeling the Dynamics of Online Popularity. *Phys. Rev. Lett.* 105 (Oct 2010), 158701. Issue 15. https://doi.org/10.1103/PhysRevLett.105.158701

[42] Daniel M. Romero, Brendan Meeder, and Jon Kleinberg. 2011. Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. In *Proceedings of the 20th International Conference on World Wide Web (WWW '11).* ACM, New York, NY, USA, 695–704. https://doi.org/10.1145/1963405.1963503

[43] Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts. 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* 311, 5762 (2006), 854–856. https://doi.org/10.1126/science.1121066

[44] Kazutoshi Sasahara, Yoshito Hirata, Masashi Toyoda, Masaru Kitsuregawa, and Kazuyuki Aihara. 2013. Quantifying collective attention from tweet stream. *PloS one* 8, 4 (2013), e61823. https://doi.org/10.1371/journal.pone.0061823

[45] Tamotsu Shibutani. 1966. *Improvised news: A sociological study of rumor.* Ardent Media.

[46] Sitabhra Sinha, Arnab Chatterjee, Anirban Chakraborti, and Bikas K Chakrabarti. 2010. *Econophysics: an introduction.* John Wiley & Sons.

[47] Kate Starbird, Emma Spiro, Isabelle Edwards, Kaitlyn Zhou, Jim Maddock, and Sindhuja Narasimhan. 2016. Could This Be True?: I Think So! Expressed Uncertainty in Online Rumoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* ACM, New York, NY, USA, 360–371. https://doi.org/10.1145/2858036.2858551

[48] James Stewart. 2015. *Calculus: Early transcendentals.* Cengage Learning.

[49] Jeannette Sutton, Emma S. Spiro, Britta Johnson, Sean Fitzhugh, Ben Gibson, and Carter T. Butts. 2014. Warning tweets: serial transmission of messages during the warning phase of a disaster event. *Information, Communication & Society* 17, 6 (2014), 765–787. https://doi.org/10.1080/1369118X.2013.862561

[50] Zeynep Tufekci. 2014. Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. In *International AAAI Conference on Web and Social Media (2014).*

[51] Xuanhui Wang, ChengXiang Zhai, Xiao Hu, and Richard Sproat. 2007. Mining Correlated Bursty Topic Patterns from Coordinated Text Streams. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '07).* ACM, New York, NY, USA, 784–793. https://doi.org/10.1145/1281192.1281276

[52] Fang Wu and Bernardo A. Huberman. 2007. Novelty and collective attention. *Proceedings of the National Academy of Sciences* 104, 45 (2007), 17599–17601. https://doi.org/10.1073/pnas.0704916104

[53] Ye Wu, Changsong Zhou, Jinghua Xiao, JÃijrgen Kurths, and Hans Joachim Schellnhuber. 2010. Evidence for a bimodal distribution in human communication. *Proceedings of the National Academy of Sciences* 107, 44 (2010), 18803–18808. https://doi.org/10.1073/pnas.1013140107

[54] Jaewon Yang and Jure Leskovec. 2011. Patterns of Temporal Variation in Online Media. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining (WSDM '11).* ACM, New York, NY, USA, 177–186. https://doi.org/10.1145/1935826.1935863

[55] Li Zeng, Kate Starbird, and Emma S Spiro. 2016. # Unconfirmed: Classifying Rumor Stance in Crisis-Related Social Media Messages. In *International AAAI Conference on Web and Social Media (ICWSM '16).*