Studying the Spatio-Temporal Dynamics of Small-Scale Events in Twitter

Paul Mousset
IRIT, Université de Toulouse, CNRS
Toulouse, France
paul.mousset@irit.fr

Yoann Pitarch
IRIT, Université de Toulouse, CNRS
Toulouse, France
pitarch@irit.fr

Lynda Tamine
IRIT, Université de Toulouse, CNRS
Toulouse, France
tamine@irit.fr

ABSTRACT

Small-scale events are emerging as attractive objects of research. On Twitter, small-scale events represent weak sensors that report things happening in specific times and places. While previous work addressed the issue of detecting such events, very little is known so far about their inherent properties. In this paper, our main objective was to analyse the spatio-temporal peculiarities of small-scale events w.r.t different levels of location granularity, and to understand the general trend of their propagation along their lifetimes. Our findings suggest that (1) users involved in small-scale events mostly gravitate not significantly far from the geographical focus; (2) events do not exhibit major peaks; and (3) there exists distinct events that we can identify from users' posts that significantly differ from topic distribution, focus concentration and propagation distance perspectives across time.

KEYWORDS

Small-scale event; geo-tagged tweets; focus; entropy

1 INTRODUCTION

Microblogging platforms such as Twitter provide active communication channels and a gold-mine of timely real-world information which has been shown to be highly effective for gaining knowledge about people's profiles [22], and opinions [16] to cite just a few. In particular, for events such as festivals, political campaigns, pandemics and crisis situations, user-generated micro-posts play a crucial role as social sensors by allowing the monitoring of users' activities and the provision of timely responses and recommendations [12]. More specifically, event-related tweet streams provide valuable spatio-temporal data such as text messages with location mentions, the timestamp of the post and the geolocation of the user who posted the tweet. With the increasing connectivity of users through wireless networks and the wide use of mobile devices, geo-tagged tweets are currently created daily. This phenomena allowed the intensive use of Twitter data for detecting and monitoring both large-scale events (eg., earthquakes and epidemics) and small-scale events (eg., festivals, crimes and protests). An important body of early research focused on detecting, analysing and mining behavioural patterns from large-scale events (also called *global* events), which are bursty in the entire stream, impact a wide spatial area and trigger an important audience [6, 23]. Recently, there has been a growing research interest in detecting [1, 26, 30–32] and analysing [24, 25, 30] small-scale events. Unlike large-scale events, small-scale events (also called l*ocal* or *localized* events) are generally micro-phenomena that stimulate people to post a low number of messages for a certain period of time in a local region. Such events play the roles of *weak signals* which have potential in several applications such as public order protection and traffic road assistance. However, the literature review reveals that very little has been understood so far about the spatio-temporal dynamics of such events [24, 25, 30]. A prior study [30] focused on the analysis of user physical network structure during two micro-events, namely a parking garage collapse in Atlanta and a church shooting in Wichita. The results mainly revealed that the event-related structure of the networks is not particularly more dense than the Twitter network structure and that central Twitterers are geographically central particularly in more spatially narrowed events. In [24, 25], authors examined the users' posts during two incidents which refer to small-scale events that result in damage or injuries. The authors found that different types of users (eg., journalists, organisation and citizens) report on the incidents and that citizens are generally faster than official sources in propagating tweet posts.

In this paper we pursue this line of research and report the findings of analyses that are designed to investigate the spatio-temporal dynamics of small-scale events. By using the focus and entropy measures, we thoroughly study the spatial and timely tweet post distributions of such events based on a wide set of event types automatically identified in geo-tagged tweet streams. Moreover, we consider locations at varying levels of granularity, from the borough to the Point Of Interest (POI) level. The key differences between close previous work [24, 25, 30] and ours are the following: (1) previous work focused on the analysis and comparison of network structure in the Twitter network vs. event-related network [30] and

https://doi.org/10.1145/3209542.3209561

Table 3: Ratio of events w.r.t their duration.

Duration	Percentage
0 − 3 h	6.59%
3 − 6 h	66.83%
6 − 9 h	20.00%
9 + h	6.58%

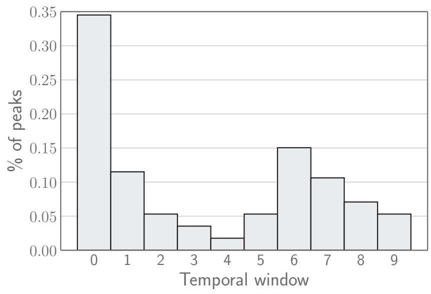

Figure 3: Ratio of peak occurrence in non-stationary events.

this objective, first we focus on the study of event temporalities to investigate the presence vs. absence of differences in event lifetimes and then cross the spatial and temporal dimensions to understand the event propagation trend.

4.2.1 Analysis of event temporalities. Our aim here is to analyse the temporal evolution of events. Accordingly, we first identify the relevant temporal window to be used in the study. Based on the average duration of events ($\sim 5\,h$) and standard deviation ($\sim 2\,h$), we split the events into intervals of 3h durations, as shown in Table 3. The results show that most events (66.83%) last between 3h and 6h and that very few are short (6.59%) or very long (6.58%). By cross-looking at the event size -in terms of number of posted tweets- per range of duration, we found a moderate positive correlation (Pearson coefficient correlation = 0.613). This suggests that tweet publications during an event have the same trend as the event does.

To gain a clear understanding of this observation, we split each event into 10 windows of equal-size and then, for each window, computed the number of tweets that were posted during the temporal interval. We studied the stationarity of the resulting time series using the Kwiatkowski Phillips Schmidt Shin (KPSS) test [17]. We found that only 28% of events are non-stationary ($p < 0.05$) which suggests the presence of peaks. For those non-stationary temporal series, we further determined the temporal windows within which the peaks occur (Figure 3). We can see that approximately 35% of the *candidate peaks* (which represent $\sim$ 9% of the overall events) appear at the *birth* of the event. This observation seems to be quite obvious because the latter is mechanically used to detect the event itself. To check this feeling, we computed the statistical differences in the propagation trends of events with peaks and those without peaks using focus, entropy and location-based feature values. We used the Welch's t-test [28] which does not assume equal population variance. Table 4 provides a summary of these feature values and the associated standard statistical indicators. The significance of the difference between feature means as determined by the obtained *p-value* and the level of significance are respectively reported in the last two rows of Table 4. We can observe that no significant difference has been reported for each of the studied features. Combining all these observations about temporal users' tweet publication, we hypothesize that, unlike for global events, the notion of peaks does not really make sense for small-scale events. Thus we consider all the events at the same level of interest in the following spatio-temporal analysis.

Table 4: Comparison of events with peak vs. without peak.

Level POI		# Events	Focus	Entropy	Distance User - User	Distance Focus - User
Events with peak		113	0.65	1.78	0.313	0.201
Events without peak		297	0.66	1.52	0.271	0,169
t-test	**p-value**	-	0.803	0.108	0.401	0.386
	Test significance	-	=	=	=	=

4.2.2 Analysis of spatio-temporal event trends. Our objective at this stage is to understand the spatio-temporal dynamics of events. In light of our objective, we split the events into 10 windows of equal-size and calculate the average entropy, $H^e(t)$, and focus, $F^e(t)$, for each temporal window. The results are shown in Figure 4. Looking specifically first at the entropy, we observe at the neighbourhood level, the entropy slightly increases (from 0.31 to 0.64) which indicates that small-scale events do not really propagate through different neighbourhoods. Therefore, events remain confined within less than 2 neighbourhoods, on average. At the POI level, when the events begin, tweets are posted from a limited number of POIs (less than 2 POIs on average since $H^e(0) = 0.59$). Then, the events tend to quickly propagate to approximately 2 locations in the first half of the event duration ($H^e(5) = 1.13$), before stabilizing thereafter at approximately 3 POIs ($H^e(9) = 1.59$). To measure the impact of the entropy increase on the concentration of tweets that are published within the same location, we turn our attention to the evolution of the average geographical focus. At the beginning of an event (i.e., during the first temporal window), the focus values are high: 88% and 78% at the neighbourhood and POI levels respectively. They slightly decrease as the event unfolds and stabilizes

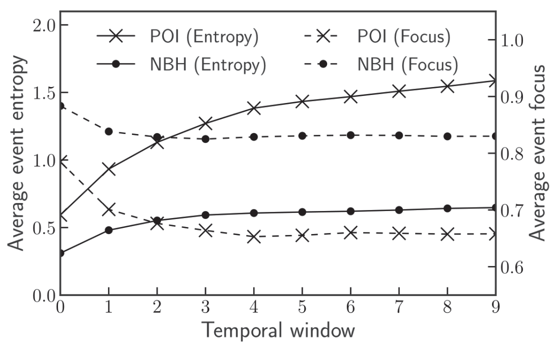

Figure 4: Average event entropy $H^e(t)$ and focus $F^e(t)$ evolution over time.

when reaching half of the event duration. Finally, 83% (resp. 66%) of tweets are posted from the focus at the neighbourhood (resp. POI) level. Moreover, we note that the coarser the level, the faster the focus values stabilize. Despite this drop in focus, the latter is still informative at any time of the event since it systematically attracts more than 50% of tweets ($F^e \geq 0.5$) regardless of the spatial level or the temporal window. Combining our observations about entropy and focus dynamics as highlighted from results, we conjecture that the more scattered an event is, the less a single location draws most of the users' attention.

4.3 Event types (RQ3)

Our practical objective here is twofold: (1) investigate wether specific types of events can emerge from the users' posts; (2) characterise the event types (if present) w.r.t topical and audience features.

4.3.1 Identification of event types. Building on previous results, we consider results at POI level only, and use the focus as a criterion for event categorization. More specifically, we split the events into 5 clusters according to their focus values and compute the average entropy and distance per event cluster and per temporal window as shown in Figure 5. Points labelled with 0 are associated with the first temporal window whereas points labelled with 9 are associated with the last window[5]. At a general glance, we can see that the clusters follow the same pattern: the distance values slightly increase during the event lifetime whereas the entropy quickly increases until half of the event duration and then remains stable. However a deeper analysis identifies three types of events from this result. The first type (solid line) consists of the 21% of events belonging to the 2 top clusters in Figure 5 that are associated with low focus values ($0 \leq F^e \leq 0.4$) and labelled as *Group A*. Events belonging to this type are spread between approximately 2 POIs from the beginning ($H^e(0) > 1.11$) and continue to propagate across 8 to 20 POIs during their lifetimes ($3.05 < H^e(9) < 4.34$). Moreover, they also spatially spread based on the increase in the average distance between users. These events are dynamic events that reach a wide audience since they propagate to both multiple locations and multiple geographic areas. The second type of events (dashed line) consists of the 37% of events belonging to the 2 median event clusters in Figure 5 that are associated with moderate focus values ($0.4 \leq F^e \leq 0.8$) and labelled as *Group B*. They globally remain concentrated within the same area, i.e., the distance slightly increases, but they spread over several locations. They arise in less than 2 POIs ($H^e(0) < 1$) and propagate quickly across 2 to 4 POIs ($1.14 < H^e(3) < 1.71$) for the first third of the event duration. For the remaining lifetimes, the events no longer propagate (i.e., their entropies remain stable). Finally, the third type of events (no line) consists of the 42% of events belonging to the cluster at the bottom in Figure 5 that have high focus values ($0.8 \leq F^e \leq 1$) and labelled as *Group C*. Events that fall within this group are very localized bringing people together in a single POI. Their entropies and distances do not change during the event.

4.3.2 Characterisation of event types. To gain better insights from the event types identified from the previous analysis, we performed a qualitative analysis at the topical level enhanced with a quantitative analysis of audience (in terms of number of users involved in events). Basically speaking, a topic is a common subject

[5]For the sake of readability, labels associated with intermediate windows are omitted.

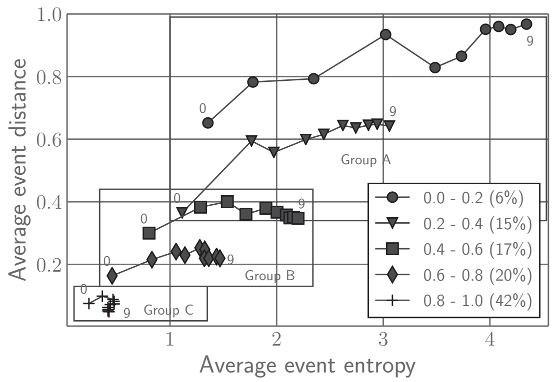

Figure 5: Evolution of the averaged distance and entropy values at POI level.

discussed in the Twitter stream. Given that the datasets used in our study are geo-tagged in New York City and that we are interested in event topics, we used the topic labels of the NY Times medium as already done in previous work on Twitter datasets [34]. The topic categories are *Arts, World, Business, Sports, Style, Technology and Science, Health, Education and Travel.* To perform the topic labelling task, we first built 3 event groups (*Group A, Group B, Group C*) by splitting the original event dataset per event type identified previously. Then we applied in each event group the Latent Dirichlet Allocation (LDA) model [8] to the meta-documents built from the tweets belonging to each event and then tuned the optimal number of topics using the perplexity measure [8]. We reached a minimal perplexity value of 27.6, 20.1 and 17.3 at 30 topics respectively for *Group A*, *Group B* and *Group C*. Each topic from the 90 automatically extracted LDA topics (30 topics extracted from each group) was labeled by 4 human assessors who were instructed to define topic labels w.r.t the NY Times topic categories if relevant and to assign to the '*Other*' topic category if no relevant NY topic category matched the LDA topic. Assessors' agreement was estimated using the Fleiss Kappa cœfficient and revealed a moderate agreement with value of 59.68%. A final topic category has been assigned to each event by applying the majority voting strategy. To have a picture of the group characteristics at the event level, we mapped the group topics to event topics by using the LDA inference algorithm [8] and then computed for each group the distribution of events and audience w.r.t each topic category, as shown in Figure 6. From a general view, we can see that apart from the '*Other*' category, the topics extracted from all the event groups are mostly related to *Arts* (resp. 48%, 28%, 28% for group A, B and C) and *Sports* (resp. 14%, 21%, 52% for group A, B and C). The observation about the relative high size of the '*Other*' category is consistent with previous work which have shown that Twitter streams give rise to specific topics that do not always fit with standard categorical topics [29, 34]. Thus, we further asked the annotators to assign Twitter labels as provided in [34] to the events belonging to the '*Other*' category. The annotation performed with a moderate Fleiss Kappa agreement of 56.87% showed that most of the topics belong to the *'Family and life'* category (resp. 55%, 56% and 97% events for Group A, B, and C) which is one of top hot topics addressed in Twitter including highly

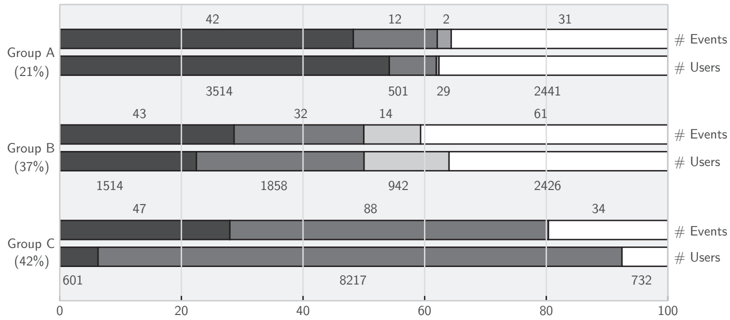

Figure 6: Distribution of number of users and number of tweets w.r.t event groups and topics.

personal and opinionated tweets. We can also observe from Figure 6 the following: (1) *Group A* is characterised with the lowest number of events (21%) and the highest average audience per event[6] (74). *Group A* seems to represent a set of few important events since they are moreover spatially spread as shown in the previous analysis. (2) In *Group B*, the average audience is less important (45) than in *Group A* but users are involved in a higher number of events (37%) and address more diverse topics including *Art*, *Sport*, *Politic* and *Other*. Combining this observation with the spatial analysis, we expect that events in this group are more likely to be less important events in wide-open spaces. (3) *Group C* includes the highest proportion of events (42%) with an average audience per event higher than in *Group B* (56, but still lower than in *Group A*), with however comparable topic diversity. Combining these observations with the spatial narrowness of users involved in this group of events suggests that *Group C* includes numerous and topically diverse micro-events with a low spatial impact. A qualitative annotation of a sample of events allowed us to confirm our expectations. For instance the *Global citizen festival* and the *Race of the cure* events which are well known periodic events in the US fall into the *Group A*, the *Tennis US Open* and the *NY Comic Con* fall into the Group B, while we found numerous private concerts and soccer matches in the *Group C*.

5 CONCLUSION AND IMPLICATIONS

In this paper, we analysed the spatio-temporal dynamics of small-scale events. Our primary objective was to determine the perimeters of their geographical social impacts at different levels of location granularity, and to gain understanding of their audience and the general trends of their propagation along their lifetimes. Our results suggest the following trends:

- In response to **RQ1**, the results show that the focus is a significant origin location from which users post their tweets, particularly at coarser levels of location granularity. Moreover, even if events propagate over several locations, they mostly reach narrow regions. Building on these findings, one relevant practical implication that we envision is the design of information seeking algorithms that are able to timely and automatically enlarge the event propagation diameter by rooting event mentions to users who are located in narrowed regions. User's location, if not explicitly provided, could either be inferred using improved state-of-the art algorithms for tweet geo-location [15]. This would increase the situational awareness particularly during security incidents.
- In response to **RQ2**, we found that the temporal series of events are mostly stable which suggests the absence of significant peaks. We also found that evently timely evolve from diverse locations and quickly stabilize not significantly far from the focus. A relevant research opportunity that arises from this study is to examine these findings alongside previous research findings about large-scale event detection [6] to design novel algorithms that can jointly detect both weak and strong signals in Twitter streams considering appropriate spatio-temporal distribution and density of posts. Such general detectors can provide means for monitoring people' activities (eg., for public order maintenance purpose).
- In response to **RQ3**, we found that we can detect distinct types of events with evolutions that are significantly different according to audience, focus concentration and propagation distance trends over time. Based on these findings, the implications for further theoretical investigation is to develop models for predicting event type based on the event-related features. Event type prediction would be a prior step to the development of an automatic visual summarization method that would give a high-level picture of what is happening in a region.

Our study has some limitations. First, we only used the focus, entropy and distance metrics to report the analysis results. Although these measures are the primary metrics that are used for the spatio-temporal analysis of events, they are still insufficient for revealing other relevant facets such as propagation rate. Second, enlarging the spatio-temporal scope of our study to other cities and during different periods might give better insights about the generalisability of our findings. This investigation is planned for future work.

[6] Ratio between total number of users and total number of events in the group.

ACKNOWLEDGMENTS

This research was supported by IRIT and ATOS research program under ANRT CIFRE grant agreement.

REFERENCES

[1] Hamed Abdelhaq, Christian Sengstock, and Michael Gertz. 2013. Eventweet: Online localized event detection from twitter. *PVLDB Endowment* 6, 12 (2013), 1326–1329.

[2] Charuc C. Aggarwal and Karthik Subbian. 2012. Event Detection in Social Streams. In *Proceedings of the 2012 SIAM International Conference on Data Mining (SDM'12)*. SIAM, 624–635.

[3] Paolo Arcaini, Gloria Bordogna, Dino Ienco, and Simone Sterlacchini. 2016. User-driven geo-temporal density-based exploration of periodic and not periodic events reported in social networks. *Information Sciences* 340-341 (2016), 122 – 143.

[4] Sebastien Ardon, Amitabha Bagchi, Anirban Mahanti, Amit Ruhela, Aaditeshwar Seth, Rudra Mohan Tripathy, and Sipat Triukose. 2013. Spatio-temporal and Events Based Analysis of Topic Popularity in Twitter. In *The 22nd ACM International Conference on Information and Knowledge Management (CIKM '13)*. ACM, 219–228.

[5] Sitaram Asur and Bernardo A. Huberman. 2010. Predicting the Future with Social Media. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '10)*. IEEE, 492–499.

[6] Farzindar Atefeh and Wael Khreich. 2015. A survey of techniques for event detection in twitter. *Computational Intelligence* 31, 1 (2015), 132–164.

[7] Hila Becker, Mor Naaman, and Luis Gravano. 2011. Beyond Trending Topics: Real-World Event Identification on Twitter.. In *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM'11)*. AAAI, 438–441.

[8] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. JMLR'03. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3 (JMLR'03), 993–1022.

[9] Anders Brodersen, Salvatore Scellato, and Mirjam Wattenhofer. 2012. YouTube Around the World: Geographic Popularity of Videos. In *Proceedings of the 21st World Wide Web Conference 2012 (WWW'12)*. ACM, 241–250.

[10] Zhiyuan Cheng, James Caverlee, and Kyumin Lee. 2010. You Are Where You Tweet: A Content-based Approach to Geo-locating Twitter Users. In *Proceedings of the 19th ACM Conference on Information and Knowledge Management (CIKM '10)*. ACM, 759–768.

[11] Gabriel Pui Cheong Fung, Jeffrey Xu Yu, Philip S. Yu, and Hongjun Lu. 2005. Parameter Free Bursty Events Detection in Text Streams. In *Proceedings of the 31st International Conference on Very Large Data Bases (VLDB '05)*. ACM, 181–192.

[12] Bo Hu and Martin Ester. 2013. Spatial Topic Modeling in Online Social Media for Location Recommendation. In *Seventh ACM Conference on Recommender Systems (RecSys '13)*. ACM, 25–32.

[13] Allan James, Carbonell Jaime, Doddington George, Yamron Jonathan, and Yang Yiming. 1998. Topic Detection and Tracking Pilot Study Final Report. In *DARPA Broadcast News Transcription and Under standing Workshop (BNTUW '98)*. Morgan Kaufmann Publishers, 194–218.

[14] Krishna Y. Kamath, James Caverlee, Kyumin Lee, and Zhiyuan Cheng. 2013. Spatio-temporal Dynamics of Online Memes: A Study of Geo-tagged Tweets. In *The 22nd International World Wide Web Conference (WWW '13)*. ACM, 667–678.

[15] Sheila Kinsella, Vanessa Murdock, and Neil O'Hare. 2011. "I'M Eating a Sandwich in Glasgow": Modeling Locations with Tweets. In *Proceedings of the 3rd International CIKM Workshop on Search and Mining User-Generated Contents (SMUC '11)*. ACM, 61–68.

[16] Efthymios Kouloumpis, Theresa Wilson, and Johanna D Moore. 2011. Twitter sentiment analysis: The good the bad and the omg! *Icwsm* 11 (2011), 538–541.

[17] Denis Kwiatkowski, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54, 1 (1992), 159 – 178.

[18] Kathy Lee, Ankit Agrawal, and Alok Choudhary. 2013. Real-time Disease Surveillance Using Twitter Data: Demonstration on Flu and Cancer. In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13)*. ACM, 1474–1477.

[19] Deepa Mallela, Dirk Ahlers, and Maria Soledad Pera. 2017. Mining Twitter Features for Event Summarization and Rating. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, 615–622.

[20] Andrew J. McMinn, Yashar Moshfeghi, and Joemon M. Jose. 2013. Building a Large-scale Corpus for Evaluating Event Detection on Twitter. In *The 22nd ACM International Conference on Information and Knowledge Management (CIKM '13)*. ACM, 409–418.

[21] Stuart E Middleton, Lee Middleton, and Stefano Modafferi. 2014. Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems* 29, 2 (2014), 9–17.

[22] Alan Mislove, Bimal Viswanath, Krishna P. Gummadi, and Peter Druschel. 2010. You Are Who You Know: Inferring User Profiles in Online Social Networks. In *Proceedings of the Third International Conference on Web Search and Web Data Mining (WSDM '10)*. ACM, 251–260.

[23] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, 851–860.

[24] Axel Schulz, Eneldo Loza Mencía, and Benedikt Schmidt. 2016. A Rapid-prototyping Framework for Extracting Small-scale Incident-related Information in Microblogs. *Information Systems* 57, C (2016), 88–110.

[25] Axel Schulz and Petar Ristoski. 2013. The Car that Hit The Burning House: Understanding Small Scale Incident Related Information in Microblogs. In *Proceedings of the Seventh International Conference on Weblogs and Social Media (ICWSM'13)*. AAAI, 11–14.

[26] Axel Schulz, Benedikt Schmidt, and Thorsten Strufe. 2015. Small-Scale Incident Detection Based on Microposts. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media (HT'15)*. ACM, 3–12.

[27] Sarah Vieweg, Amanda L. Hughes, Kate Starbird, and Leysia Palen. 2010. Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI '10)*. ACM, 1079–1088.

[28] B. L. Welch. 1947. The Generalization of Student's Problem When Several Different Population Variance Are Involved. *Biometrika* 34, 1-2 (1947), 28–35.

[29] David Wilkinson and Mike Thelwall. JASIST'12. Trending Twitter topics in English: An international comparison. *Journal of the American Society for Information Science and Technology* (JASIST'12), 1631–1646.

[30] Sarita Yardi and danah boyd. 2010. Tweeting from the Town Square: Measuring Geographic Local Networks. In *Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM'10)*. AAAI, 194–201.

[31] Chao Zhang, Liyuan Liu, Dongming Lei, Quan Yuan, Honglei Zhuang, Tim Hanratty, and Jiawei Han. 2017. TrioVecEvent: Embedding-Based Online Local Event Detection in Geo-Tagged Tweet Streams. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17)*. 595–604.

[32] Chao Zhang, Guangyu Zhou, Quan Yuan, Honglei Zhuang, Yu Zheng, Lance Kaplan, Shaowen Wang, and Jiawei Han. 2016. GeoBurst: Real-Time Local Event Detection in Geo-Tagged Tweet Streams. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR '16)*. ACM, 513–522.

[33] Kaiqi Zhao, Gao Cong, and Aixin Sun. 2016. Annotating Points of Interest with Geo-tagged Tweets. In *Proceedings of the 2016 ACM on Conference on Information and Knowledge Management (CIKM'16)*. ACM, 417–426.

[34] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan, and Xiaoming Li. 2011. Comparing Twitter and Traditional Media Using Topic Models. In *Proceedings of the 33rd European Conference on Advances in Information Retrieval (ECIR'11)*. ACM, 338–349.