Tracing the German Centennial Flood in the Stream of Tweets: First Lessons Learned

Georg Fuchs Natalia Andrienko Gennady Andrienko

Sebastian Bothe Hendrik Stange

Fraunhofer IAIS
Schloss Birlinghoven
Sankt Augustin, Germany
{firstname}.{lastname}@iais.fraunhofer.de

ABSTRACT

Social microblogging services such as Twitter result in massive streams of georeferenced messages and geolocated status updates. This real-time source of information is invaluable for many application areas, in particular for disaster detection and response scenarios. Consequently, a considerable number of works has dealt with issues of their acquisition, analysis and visualization. Most of these works not only assume an appropriate percentage of georeferenced messages that allows for detecting relevant events for a specific region and time frame, but also that these geolocations are reasonably correct in representing places and times of the underlying spatio-temporal situation. In this paper, we review these two key assumption based on the results of applying a visual analytics approach to a dataset of georeferenced Tweets from Germany over eight months witnessing several large-scale flooding situations throughout the country. Our results confirm the potential of Twitter as a distributed 'social sensor' but at the same time highlight some caveats in interpreting immediate results. To overcome these limits we explore incorporating evidence from other data sources including further social media and mobile phone network metrics to detect, confirm and refine events with respect to location and time. We summarize the lessons learned from our initial analysis by proposing recommendations and outline possible future work directions.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Experimentation, Verification

GEOCROWD '13:, November 05 - 08 2013, Orlando, FL, USA

http://dx.doi.org/10.1145/2534732.2534741

Keywords

Visual Analytics, Spatio-temporal Reasoning and Analysis, Social Media, Reality Monitoring

1. INTRODUCTION

The popularity of microblogging services such as Twitter in conjunction with the widespread proliferation of personal mobile devices that are able to provide location information has led to the availability of massive, realtime streams of geolocated information (e.g. status updates). For the Twitter service alone, users worldwide generate in excess of 400 million tweets each day [1]. Realtime analysis of microblogs is interesting for a number of applications, from the validation of socio-economic theories, location-based micro-marketing strategies, to their use as a form of highly distributed 'social sensors' that represent an invaluable data source to emergency and disaster management. The latter gained much attention in the recent flood control management in Germany and is generally considered a new information source in early response and crisis management [15].

Consequently, a considerable number of works has dealt with issues of their acquisition, analysis and visualization. Common objectives are the rapid detection of single critical events and reconstructing complex situations for example to establish a continuously updated situation awareness. Most of these works not only assume an appropriate percentage of georeferenced messages from the message stream in order to detect abnormal or unusual events, but also that these geolocations are reasonably correct in representing places and times of the spatio-temporal (complex) events or even situations.

In this paper, we review these two key assumptions based on the results of applying a visual analytics approach to a set of geolocated tweets from Germany over an eight-month period (22 November 2012 – 24 July 2013), with 'ground truth' being available for the data – in particular, on the severe floodings throughout Germany in the summer of 2013 as well as on several conventions in major towns. The motivation for our examination is two-fold: first, to assess to what extend freely available Twitter data from Germany – where personal privacy and data protection issues are perceived significantly more acutely than in many other countries – can be utilized for the purpose of detecting and localizing

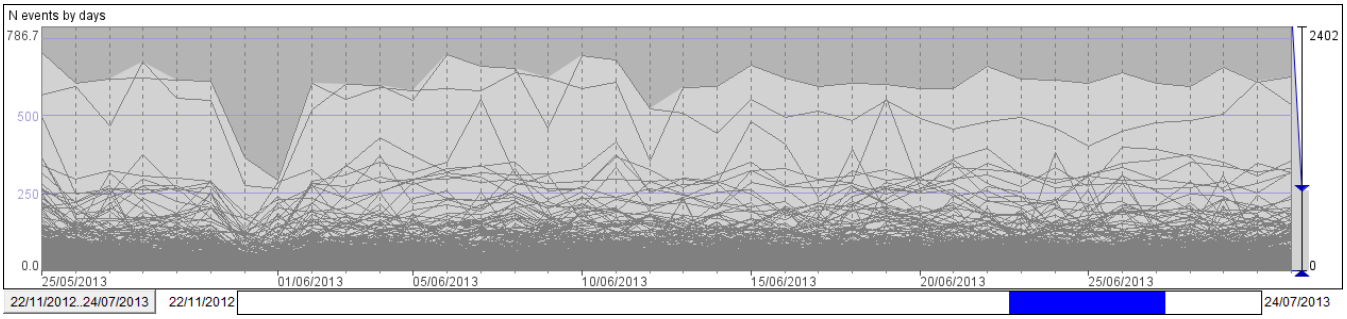

Figure 1: The time series graph of tweet frequencies, zoomed in to the time interval 25 May – 01 July 2013. Each line represents the number of tweets per daily interval for one Voronoi cell of the spatial tessellation.

such events; second, how to interpret and evaluate immediate findings from Twitter as the primary source, potentially including subsidiary data sources to confirm and possibly refine locations and time extent of these events.

2. RELATED WORK

Micro- and social blogging have been investigated by researchers in computer science, social science, and other disciplines dealing with data analysis. Twitter in particular has been used as a source for as diverse activities as event detection and tracking [6] to sentiment analysis [13].

However, the analysis of this unstructured source is quite challenging: tweets that may be of interest to an analyst are buried in a very large amount of non-related messages, and contents of individual tweets typically contain many abbreviations, slang and misspellings. This high ratio of noise combined with the brevity of individual tweets makes many traditional natural language processing tasks, such as part-of-speech tagging [8], named entity recognition [10], topic detection [20] and sentiment analysis [13] much more challenging. Yet, this kind of processing is typically required to detect relevant tweets and to extract higher-level, meaningful information from them.

Most approaches, therefore, either use specific task-tailored content models [12, 18] limiting their immediate transferability to other application areas, or relatively simple means usually based on message or term frequency to classify tweets into 'related' vs. 'unrelated' [14]. This has the advantage that no or only few assumptions on the underlying model are needed. Zhang et al. [19] describe a geospatial extension of the generic frequency–inverse document frequency (*tf–idf*) model [9] for tag-based queries of geolocated web resources. Thom et al. [16] propose the inverse document density as a continuous variant of this geospatial *tf–idf* model.

However, all message or term frequency/density based methods work on the assumption that a suitable number of geolocated messages can be obtained, and that at least the majority of these Tweets are in close proximity, both spatially and temporally, to the events and situations indicated by them. While this has been shown to reasonably hold for very abrupt events like earthquakes in Japan [14] and the US [7, 17] where a large percentage of active Twitter users allow transmission of their locations, this is not necessarily the case in countries like Germany with a much more acute perception of potential privacy issues arising from publishing one's tweeting location. In fact, measurements we conducted for Germany indicate that as low as 1% of tweets on average come with geographical coordinates, compared to the 11.03%–13.04% ("firehose" vs. streaming API, respectively) for Europe as a whole [11].

3. TWITTER DATA USED

The data for our experiments were collected by querying the Twitter streaming API (via the 1% "garden hose" access) for tweets with geographical coordinates within the bounding rectangle of Germany. In this study we only used coordinate-level geolocated tweets but did not (yet) attempt to include tweets with toponym-derived geolocations. Since the streaming API sample rate matches the average ratio (for Germany) of coordinate-level geolocated tweets we actually account for almost 100% of these within the selected region. Obtained tweets have been further filtered by their country code and whether the coordinates are indeed within said bounding rectangle, since the collected data also contained messages from the Netherlands and other countries, as well as messages with the country code of Germany but located outside Germany (which means that the Twitter API may not provide exactly what you request). The final dataset that has thus been explored comprises 5,806,223 records (tweets) from 196,357 unique users and covers the time period from 22 November 2012 to 24 July 2013.

The aim of the subsequent investigation is to check whether relevant events can be detected from the Twitter data. In particular, we wanted to evaluate how a known disaster with significant spatial and temporal extent – namely, the floods of June 2013 – is reflected in the available data. This is in contrast to sudden disasters like earthquakes with immediate but relatively local effects [7, 14]. Instead, we here concentrate on an evolutionary disaster situation comprised of multiple individual spatio-temporal events, like road blockages or a crevasse.

4. EXPERIMENT 1: DETECTION BASED ON LOCAL TWEET FREQUENCY

Hypothesis: A disaster situation, such as flooding, causes an increase in the overall number of georeferenced tweets posted at the immediate place, or at least in close vicinity to where the disaster is located. This can be used for detecting the approximate locations of disasters.

4.1 Data preparation

The territory of Germany has been divided into cells (Voronoi polygons). For this purpose, a 2% random sample of the

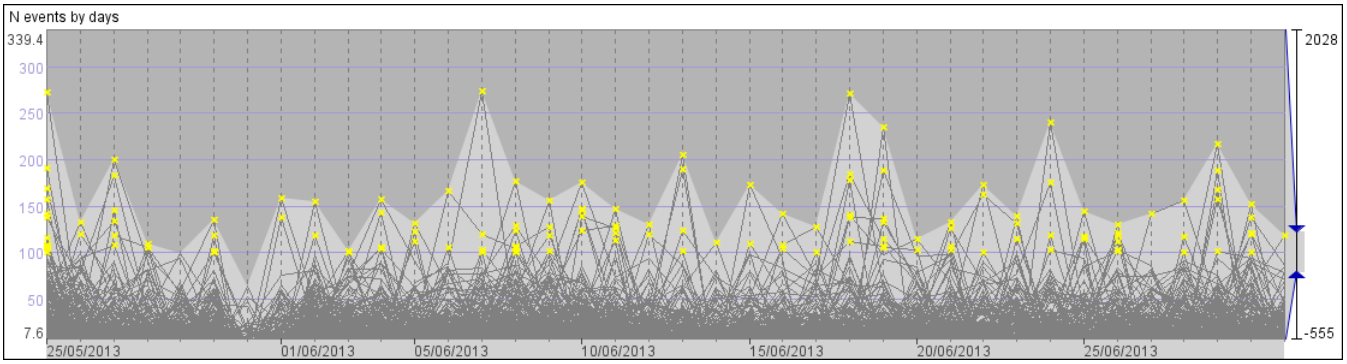

Figure 2: The time graph shows the differences (residuals) of tweet frequencies to their mean values for the previous 14 days. The yellow crosses mark differences of 100 or more.

tweets has been taken from the database. The points have been grouped into convex spatial clusters, and the centers of the clusters have been taken as generating seeds for the Voronoi tessellation; for details refer to [3].

The entire dataset has then been spatio-temporally aggregated by counting the number of tweets for each cell and day. The result of this analysis step is a set of time series of counts of tweets, each comprising 244 daily intervals and being associated with a specific cell of the territory division.

4.2 Exploration of the spatial time series

In this experiment we analyzed the distribution of total counts of tweets over a time period of 8 months. Based on the stream of tweets we have conducted our analysis with focus on Germany where there is only little understanding of the impact and usage of Twitter in crisis situations compared to other countries.

Generally, the most prominent peaks (i.e., sharp increases in the number of tweets) can easily be detected visually as well as algorithmically [2] from the time graph. To interpret their meaning we have extracted the most frequent words and word combinations from tweets posted in the respective cells and time intervals corresponding to those peaks, setting the minimal frequency (tweet count) threshold to 5. Initial results showed us that these peaks do not directly correspond to disaster events.

The highest peak, up to 2,402 messages per day, occurred in Berlin from 6 to 8 May 2013. The frequency champion is `re:publica`, which occurred 511 times on May 6, 475 times on May 7, and 394 times on May 8. `re:publica` refers to a European conference on social media, blogs, and digital society that took place in Berlin in this time period. Similarly, the second and third highest peaks – in Bochum on 24–25 November 2012 and in Hamburg on 27–29 December 2012, respectively – were found to coincide with a political party convention and the Chaos Communication Congress, one of the largest European hacker conventions. These findings indicate a large noise-to-relevant-event ratio for the explored time interval.

Thus, to find out whether there are any peaks that can be attributed to the flood events of June 2013, we have specifically looked at the parts of the time series for the time period 25 May – 01 July 2013 (Figure 1). It can be observed that there are no prominent peaks in this period. Knowing that the area of Dresden was affected by the flood on river Elbe we specially looked at the time series of the cells covering Dresden center and suburbs. The time series do not show any significant peaks although a small local maximum in tweet counts occurs in the center of Dresden on June 3. The flood-relevant hashtag `#hochwasser` occurs among most frequent words and combinations for this place and day, but its frequency is only 9 while the highest frequencies are 61 for `#linkebpt` (again, refering to a political party convention), 41 for "Dresden" and 39 single or in-phrase occurrences of the word "Neustadt", a district of Dresden.

We have also tried to find potentially interesting events by comparing the frequency values in the time series with the mean values for the previous 14 days. The time graph of the differences (Figure 2) shows very many peaks. From 25 May to 1 July 2013, frequencies of 138 (cell, day) combinations differ from their respective 14-days mean by 100 or more. For these peaks in *relative* frequencies we extracted frequent words and combinations, which surprisingly do not contain the term "Hochwasser" at all. In other words, none of the short-term surges in tweeting activity shown by Figure 2 is associated with this key term.

4.3 Experiment 1: Conclusion

Our hypothesis was that a disaster event may cause a noticeable increase in the number of tweets – and by extension, georeferenced tweets – posted in the disaster-affected area. Based on our empirical experiment, we have to reject this hypothesis. Note that the hypothesis was based on the georeferenced tweets.

Significant increases in tweet frequency are mostly caused by public gatherings, such as conferences and party conventions, but not by disaster events. Moreover, at least for Germany, peak-generating gatherings address quite specific public such as people interested in social media, political parties, computer hackers, etc.

We have also seen that the sets of tweets posted in places that are known to be disaster-affected may contain too few occurrences of relevant terms; hence, relevant messages can be easily lost in the bulk of posted tweets. This means that, in order to detect possible disasters, it is indeed necessary to look specifically for related tweets. One approach to this is creating a specific vocabulary of relevant terms and related words which should appear these messages.

5. EXPERIMENT 2: EXPLORATION OF FLOOD-RELATED TWEETS

Hypothesis: Spatio-temporal clusters of tweets containing relevant keywords may indicate disaster-affected places. The clusters do not need to be large in terms of message counts, but several spatio-temporally co-located messages may deserve inspection while a single tweet may be not indicative.

5.1 Data preparation

To further refine the set of georeferenced tweets to include only messages potentially relevant to flood events, we filter messages to compulsory include substrings "hoch" and "wasser". Although this yields a certain amount of "false positives", such as "Gerade ein Foto hochgeladen @ Wasserschloss Haus Kemnade", it allows capturing flood-relevant messages where "hoch" and "wasser" are separated, for example "Wooow!! Das Wasser vom Rhein ist ziemlich hoch!!". There are also several re-occuring misspellings of "hochwasser" both as word and hashtag, typically due to duplicate or skipped letters. To be more lenient towards skipped double letters in particular we further added "waser" (missing second 's') as matching substring after reviewing frequent terms.

This final query retrieved 2,443 messages for the whole territory of Germany. After removing "false positive" messages (e.g., containing "hochgeladen", "hochzeit") 2,429 messages remain. This message set spans the time period from November 25, 2012 05:40:52 to July 25, 2013 11:45:16.

5.2 Spatio-temporal distribution of tweets

A map in Figure 3 shows the spatial distribution of flood-related messages represented by purple dots. Concentrations of tweets can indeed be observed in the areas of Dresden, Magdeburg and along Elbe river, all of which were affected by the June 2013 floods (see Figure 4).

However, there is also a concentration in Berlin, which was not affected. Closer inspection of the latter revealed these messages mention remote flood events, actions of politicians and flood relief efforts, or flood-caused traffic problems to and from Berlin. This highlights that concentrations of disaster-related Twitter messages need to be interpreted with caution. People can and do tweet about remote events as well as indirect or collateral event effects.

Figure 5 shows the spatio-temporal distribution of the flood-related tweets from Figure 3 in a space-time cube (STC). The vertical dimension represents the time, with the oldest tweets located at the bottom and the most recent at the top of the cube. Viewing direction in Figure 5 is south-west to north-east, i.e., the STC's bottom-right edge corresponds to geographic south. The balls representing tweets have further been color-coded according to four distinct time intervals.

A very noticeable visual feature in Figure 5 is the vertical column, denoting a constant stream of messages from the exact same geographic location. This turned out to be Konstanz (at Lake Constance); the messages, most probably automatically generated, report the lake's water level. The highest reported level is 467.5cm, reached on 03 June 2013. Besides the column one observes three major layers (i.e., time periods), color-coded in Figure 5 for better visual separation (see legend). These four periods include 69, 69, 39, and 2,252 tweets, respectively. The highest number of tweets (306) was again reached on 03 June 2013.

The three maps in Figure 6 show the spatial distribution

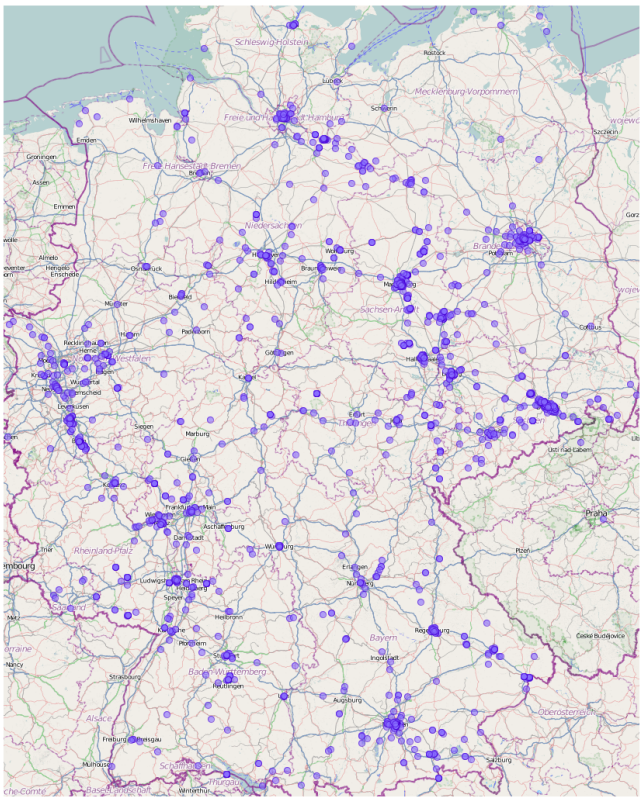

Figure 3: The map shows the spatial distribution of the tweets containing flood-relevant substrings.

Figure 4: Overview map of flood-affected rivers in Germany including dates (source: Wikipedia.org)

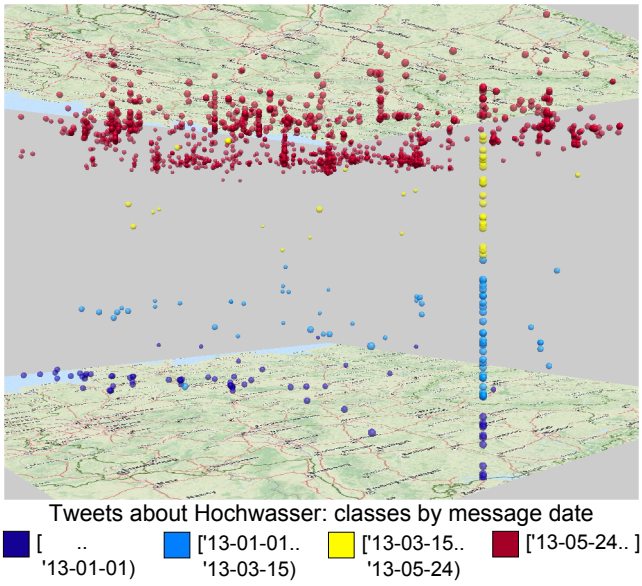

Figure 5: The space-time cube shows the spatio-temporal distribution of the flood-related tweets.

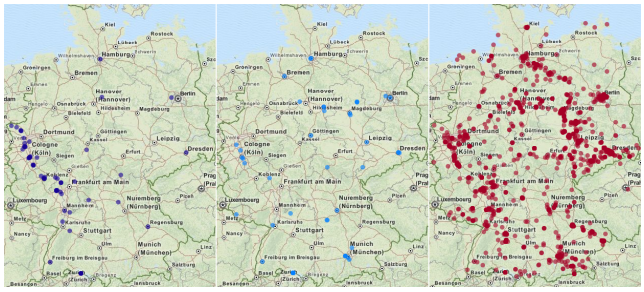

Figure 6: The spatial distributions of the flood-related tweets in the 3 time periods apparent from Figure 5, cf. its legend for colors and dates.

of the tweets in the first two periods and in the last period. In the first period, most of the tweets are aligned along the valley of the river Rhine. Most of the messages reflect the water rise that happened around Christmas (December 20-30 of 2012). The messages of the second period (winter 2013) are more scattered over the territory. The messages of the fourth period (summer 2013) cover almost the whole territory of Germany.

5.3 Spatio-temporal clustering of tweets

We applied density-based clustering (using OPTICS [5]) to the set of the flood-related tweets according to their positions in space and time, i.e., as spatio-temporal points. We chose 30 km as spatial distance threshold, 1 day (86,400 seconds) as temporal distance threshold, and 2 as minimum neighborhood size threshold. We obtained 90 clusters ranging from 3 to 783 members, together including 1,885 points (77.6% of total), and 544 points (22.4%) classified as noise. Figure 7 shows a map of all spatio-temporal tweet clusters with their convex hulls and noise points filtered out. Viewing the clusters in the STC (Figure 8) shows that only a few small clusters have been built in the first time period,

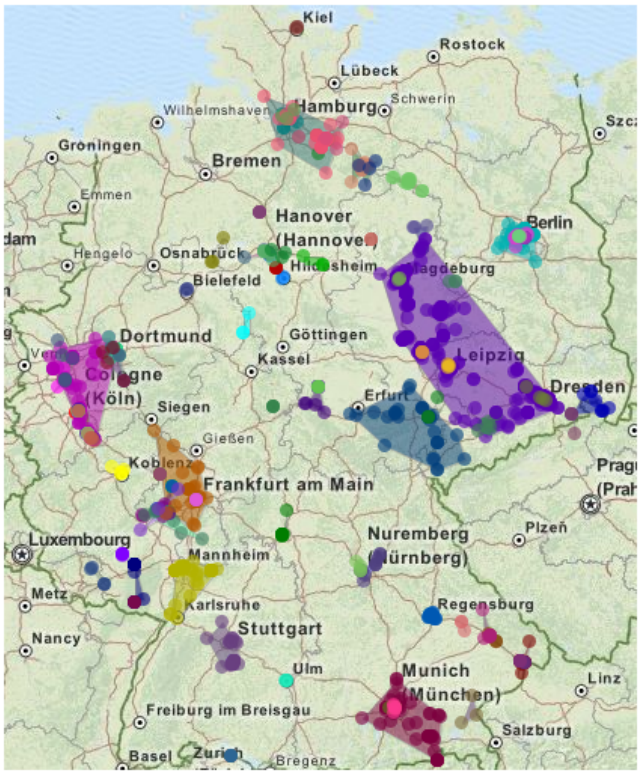

Figure 7: The spatio-temporal clusters of the flood-related tweets are shown on a map.

none in the second and third periods, and very many in the fourth period (cf. Figure 5 for time period dates – colors do not match). The largest cluster covering the area of Saxony (containing Dresden and Magdeburg; purple) began on 01 June and ended on 17 June 2013; it includes 783 tweets and its spatial extent (bounding rectangle diagonal) is 268 km.

5.4 Matching detected clusters to subsidiary sources of information

To check the validity of the obtained spatio-temporal clusters, we first used other information sources that can be found manually on the Web. This does not necessarily yield supporting evidence for all clusters. For example, in our concrete experiment no additional information could be found for the small cluster of 5 messages posted in a small town northwest of Kaiserslautern on 20 May between 15:53 and 18:31, i.e., of quite short duration.

On the other hand, for an almost equally compact cluster at Hildesheim near Hannover (3 messages, starting on 26 May at 20:10 and ending on 27 May at 6:38) we were able to find four YouTube videos (26, 27, and 28 May), several photos on platforms "Tumblr" and "FT Photo Diary" (uploaded 28 May), as well as corresponding news articles by "TheLocal Germany" online newspaper (30 May) and the international "FloodList" online portal (31 May).

On 1 June, at the height of the flodding crisis, 9 different spatio-temporal clusters began in different regions of Germany: on rivers Main, Rhine, and Neckar on the west of the country, in Bavaria on river Danube (Passau and Regensburg) and in Saxony on rivers Spree, Elbe, Saale, and

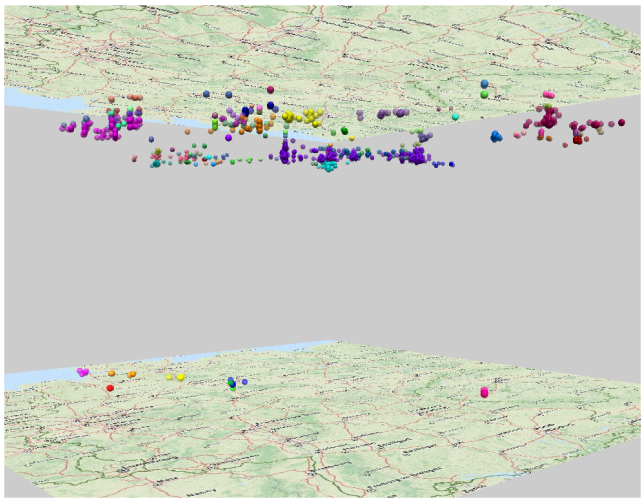

Figure 8: Temporal distribution of the clusters from Figure 7 shown in the space-time cube.

Mulde. During this period press coverage in particular increased sharply, with news articles from 1 June reporting, among others, on rising water levels on Rhine, Danube, and Neckar rivers and mentioning cities Passau and Gera; in particular, several news articles refer to multiple places from different clusters.

These findings confirm that subsidiary data sources may indeed provide valuable supporting evidence for the significance especially of small tweet clusters. However, the heterogeneous and often unstructured content – especially with respect to spatial references – as well as different delays between the actual event and the availability of corresponding artifacts mean such verifications are largely manual tasks difficult to automate.

As another promising data source and an example for non-textual spatio-temporal data, we have available hourly Erlang value measurements for for each of the approximately 50,000 cellphone towers in Germany from May 2010 until July 2013. These are numeric values roughly indicating the product of count and duration of phone calls within a cell and time interval.

Based on the assumption that disruptive events like floodings also cause a detectable change in mobile phone utilization patterns, we wanted to test if we could detect similar spatio-temporal clusters as those identified in Section 5.3. In particular, we try to detect deviations from the representative daily time series of Erlang values during an event.

As a preliminary approach, we model each of phone cell's time series as result of a Gaussian process with $\mathcal{N}(\mu, \sigma)$. To account for the circadian and weekly cycles of human activities, each cell's model comprises of 24 hours $\times$ 7 weekdays = 168 individual distributions $\mathcal{N}_i$. For the sake of confirmation of suspected events (based on the Twitter stream), we assume each cell exhibits its normal behavior during time intervals not belonging to any such detected event, which are thus used as input to the model.

To identify the cells with significant deviations we compare the actual Erlang values with the model-predicted mean values μ. If the difference exceeds 2σ we flag the cell and time slot as an anomaly. Due to their small spatial extent

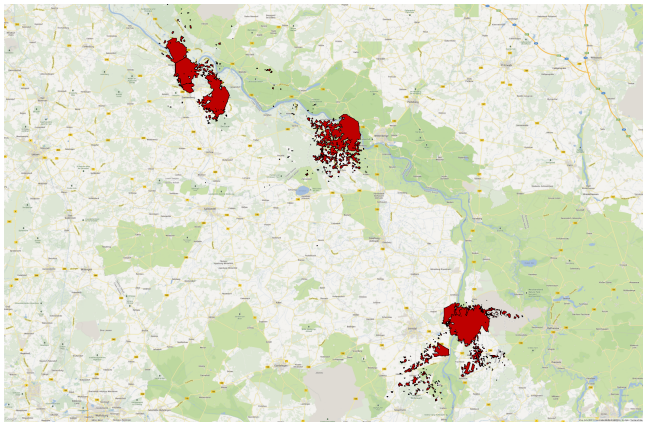

Figure 9: Affected cell towers that have abnormal readings corresponding to the purple cluster from Figures 7 and 8.

flagged cells may provide more fine-grained indication of potential event locations. Figure 9 shows hotspots of significant deviations along the Elbe river, corresponding to the time interval of the rather large purple cluster in Figures 7 and 8.

It should be noted that this rather simple approach is intended as a proof-of-concept for integration of subsidiary data rather different from social media. Section 6.2 provides more details on a more general "round table" approach to data fusion that will ultimately incorporate highly heterogeneous data sources including social media.

5.5 Experiment 2: Conclusions

Spatio-temporal clustering of pre-filtered disaster-related tweets may allow detection of locations and times of disaster events. However, it should be borne in mind that

- it is not guaranteed that any event is always represented by a tweet cluster;
- some tweet clusters occur in places and times where no disaster events happen since tweets may refer to events occurring elsewhere; hence, it is necessary to check the content of the messages in each cluster;
- some tweet clusters may refer not to disaster events themselves but to consequences of disaster events, such as traffic problems; this also shows a need of checking the tweet content.

6. SUMMARY

Related work on utilizing Twitter as a form of highly distributed 'social sensors' to detect and localize events and extreme situations are based on two assumptions: first, an appropriate number of both event-related and geolocated (either directly or indirectly) messages are available; second, the given or derived geolocations are reasonably correct in representing both place and time of observed spatio-temporal events and situations. The extend to which these assumptions hold highly depends on the tweeting behavior of the users. One motivation of our examination therefore has been to assess to what extend freely available Twitter data from Germany – where personal privacy and data protection issues are perceived significantly more acutely than

in many other countries – can be utilized for this purpose. In addition, our goal has been to design a visual analytics workflow that helps the analyst in detecting significant spatio-temporal events from tweet streams as well as in interpreting and evaluating corresponding findings.

In this paper we reported on the initial results from visual exploration and analysis of a set of geolocated tweets from Germany over an eight-month period (22 November 2012 – 24 July 2013), with "ground truth" being available for the data – in particular, news and other media coverage on the severe floodings throughout Germany in the summer of 2013 as well as on several conventions in major towns.

6.1 Recommendations

Overall, it can be concluded that even with a limited amount of accurately georeferenced tweets in the case of Germany (compared to e.g. the US or Netherlands), spatio-temporal visual analysis of the data still allows to detect significant events with reasonable accuracy. However, an analyst has to exercise due care in interpreting the results. Our lessons learnt can thus be summarized as a list of recommendations:

1) To detect disasters and other significant events from the stream of tweets, it is reasonable to filter the set of incoming messages based on a predefined vocabulary of relevant terms. To find emerging topics, it is necessary to analyze a sample of all tweets, not only georeferenced ones.

2) It is therefore reasonable to combine the tweets with space- and time-referenced artifacts from other sources (e.g., YouTube, Flickr, ...), which can be filtered based on their titles and/or tags. This facilitates artifact matching based on approximate topic categories.

3) Events may be detected by spatio-temporal clustering of pre-filtered objects (tweets and, possibly, posts from other media). For this purpose, a clustering algorithm working in real time within a distributed computing architecture needs to be developed. The algorithm must be able to attach new incoming objects to appropriate existing clusters and store the history of detected clusters, i.e., how they evolve over time: move in space, expand or shrink, become denser or sparser, or keep stable.

4) For each new cluster appearing in the Twitter stream, an analyst needs to check

- if it really refers to an event occurring in the same place and time as the cluster or to an event occurring elsewhere;
- if it refers to locally experienced consequences of an event that occurred elsewhere (i.e., direct vs. collateral effects).

5) Besides specifically looking for predefined types of events using vocabulary-based filtering, it may be reasonable also to pay attention to unusual concentrations of tweets in space and time. This can be done by aggregating all tweets by suitable spatial cells and time intervals into place-related time series [4]. For each place, the current level of Twitter activities computed in real time needs to be compared with the usual level for the respective day of the week and time of the day derived from the historical data [6]. Where significant deviations from the usual levels are detected, the most frequent keywords and phrases may be analyzed to interpret what is going on.

6) Subsidiary sources of data containing neither text nor content tags may still be used to provide additional evidence for the significance of spatio-temporal clusters of tweets by building a suitable prediction model of expected behavior. Subtracting observed from predicted values yields time series of residuals from which spatio-temporal events can be extracted and clustered for correlation with tweet clusters.

6.2 Future work

The above results and recommendations represent work in progress. Future refinement of the analysis methodology and workflow would integrate relevant aspects of cited related works. In addition, we want to address the open challenges of applying space partitioning and spatio-temporal clustering approaches efficiently in streaming settings, of algorithmically detecting and classifying unknown events, as well as general scalability issues with regard to massive volume of social media data such as Twitter.

Another important aspect of future work will be on the integration of multiple data sources for verification and refinement of potential events. In this paper, a focus has been on the viability of Twitter as a primary source in general, with subsidiary information being matched largely manually. For a long-term perspective, we strive to developed a more principled approach to the fusion of multiple sources of information with the objectives of

- achieving a better spatial and/or temporal resolution,
- allowing a more specific classification of the type of event occurring,
- increasing the confidence in the validity of detected events, and
- annotating confirmed events with additional semantic information.

We envision to employ an iterative process schema known from industrial manufacturing called Kanban. The process may look like a "round table meeting of experts," each with her own area of expertise (here: the capability of a given data source to refine or extent one or more of the above aspects). The idea is to consult additional sources for complementary information, refining the current findings until the constraints on the four aspects above are satisfied or no further sources are available. This may specifically include input from a human analyst as "Expert-in-the-loop" to guide algorithmic methods for intermediate steps, as well as reaching a final assessment in case of ambiguous end results.

7. ACKNOWLEDGMENTS

This work has been supported by the *EU 7th framework programme* (EU-FP7) as part of the INSIGHT project (`http://www.insight-ict.eu/`).

8. REFERENCES

[1] Washington post: Twitter turns 7: Users send over 400 million tweets per day. `http://articles.washingtonpost.com/2013-03-21/business/37889387_1_tweets-jack-dorsey-twitter`. Accessed: 2013-08-27.

[2] G. Andrienko, N. Andrienko, M. Mladenov, M. Mock, and C. Poelitz. Discovering bits of place histories from people's activity traces. In *Proceedings IEEE Visual Analytics Science and Technology (VAST)*, pages 59–66. IEEE Computer Society Press, 2010.

[3] N. Andrienko and G. Andrienko. Spatial Generalization and Aggregation of Massive Movement Data. *IEEE Transactions on Visualization and Computer Graphics*, 17(2):205–19, 2011.

[4] N. Andrienko and G. Andrienko. A visual analytics framework for spatio-temporal analysis and modelling. *Data Mining and Knowledge Discovery*, 27(1):55–83, 2013.

[5] M. Ankerst, M. Breunig, H. Kriegel, and J. Sander. Optics: Ordering points to identify the clustering structure. In *ACM SIGMOD'99 Int'l Conf. on Management of Data*, pages 49–60. ACM Press, 1999.

[6] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. Ebert, and T. Ertl. Spatiotemporal social media analytics for abnormal event detection using seasonal-trend decomposition. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2012.

[7] A. Crooks, A. Croitoru, A. Stefanidis, and J. Radzikowski. #earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 17:124âĂŞ147, 2013.

[8] K. Gimpel, N. Schneider, B. O'Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, and N. Smith. Part-of-speech tagging for twitter: annotation, features, and experiments. In *Proceedings of 49th Annual Meeting of the ACL: Human Language Technologies (HLT'11)*, pages 42–47. Association for Computational Linguistics, 2011.

[9] K. Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–21, 1972.

[10] X. Liu, S. Zhang, F. Wei, and M. Zhou. Recognizing named entities in tweets. In *Proceedings of 49th Annual Meeting of the ACL: Human Language Technologies (HLT'11)*, pages 359–367. Association for Computational Linguistics, 2011.

[11] F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley. Is the sample good enough? comparing data from twitterâĂŹs streaming api with twitterâĂŹs firehose. In *Proceedings of ICWSM*, 2013.

[12] S. Roller, M. Speriosu, S. Rallapalli, B. Wing, and J. Baldridge. Supervised text-based geolocation using language models on an adaptive grid. In *Conference on Empirical Methods on Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL'12)*, pages 1500–1510. ACL, 2012.

[13] H. Saif, Y. He, and H. Alani. Alleviating data sparsity for twitter sentiment analysis. In *Proceedings of Making Sense of Microposts (MSM2012)*, 2012.

[14] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World Wide Web*, pages 851–860. ACM, 2010.

[15] H. Stange and S. Bothe. Reality monitoring. *Crisis Prevention*, 2/2013(1):25–27, 2013.

[16] D. Thom, H. Bosch, and T. Ertl. Inverse document density: A smooth measure for location-dependent term irregularities. In *Proceedings of International Conference on Computational Linguistics (COLING)*, 2012.

[17] D. Thom, H. Bosch, S. Koch, M. Wörner, and T. Ertl. Spatiotemporal anomaly detection through visual analysis of geolocated twitter messages. In *Pacific Visualization 2012*, 2012.

[18] B. Wing and J. Baldridge. Simple supervised document geolocation with geodesic grids. In *Proceedings of the 49th Annual Meeting of the ACL: Human Language Technologies (HLT '11)*, pages 955–964, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

[19] D. Zhang, B. Ooi, and A. Tung. Locating mapped resources in web 2.0. In *26th IEEE International Conference on Data Engineering (ICDE'10)*, pages 521–532. IEEE, 2010.

[20] W. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li. Comparing twitter and traditional media using topic models. In *Proceedings of the 33rd European conference on Advances in information retrieval (ECIR'11)*, pages 338–349, 2011.