Computers & Graphics 38 (2014) 51–60

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on Visual Analytics

Public behavior response analysis in disaster events utilizing visual analytics of microblog data

Junghoon Chae [a], Dennis Thom [b], Yun Jang [c,*], SungYe Kim [d], Thomas Ertl [b], David S. Ebert [a]

[a] Purdue University, USA
[b] University of Stuttgart, Germany
[c] Sejong University, South Korea
[d] Intel Corporation, USA

ARTICLE INFO

Article history:
Received 22 June 2013
Received in revised form
7 October 2013
Accepted 8 October 2013
Available online 21 October 2013

Keywords:
Visual analytics
Social media analysis
Disaster management

ABSTRACT

Analysis of public behavior plays an important role in crisis management, disaster response, and evacuation planning. Unfortunately, collecting relevant data can be costly and finding meaningful information for analysis is challenging. A growing number of Location-based Social Network services provides time-stamped, geo-located data that opens new opportunities and solutions to a wide range of challenges. Such spatiotemporal data has substantial potential to increase situational awareness of local events and improve both planning and investigation. However, the large volume of unstructured social media data hinders exploration and examination. To analyze such social media data, our system provides the analysts with an interactive visual spatiotemporal analysis and spatial decision support environment that assists in evacuation planning and disaster management. We demonstrate how to improve investigation by analyzing the extracted public behavior responses from social media before, during and after natural disasters, such as hurricanes and tornadoes.

1. Introduction

For emergency and disaster management, analysis of public behavior, such as how people prepare and respond to disasters, is important for evacuation planning. As social media has played a pervasive role in the way people think, act, and react to the world (more than 40 million Americans use social media web sites multiple times a day [1]), social media is changing the way people communicate not only in their daily lives, but also during abnormal events, such as natural disasters. In emergency situations, people even seek social confirmation before acting in response to a situation, where they interact with others to confirm information and develop a better informed view of the risk [2]. A study commissioned by the American Red Cross, found that roughly half of the respondents would mention emergencies and events on their social media channels, and more than two-thirds agree that response agencies should regularly monitor postings on their websites [3]. Moreover, a growing number of people are using Location-based Social Network services, such as microblogs, where they create time-stamped, geo-located data and share this information about their immediate surroundings using smart phones with GPS. Such spatiotemporal data has great potential for enhancing situational awareness during crisis situations and providing insight into the evolving event, the public response, and potential courses of action.

For public behavior analysis in disasters, however, finding meaningful information from social media is challenging. It is almost impossible to perform a straightforward qualitative analysis of the data, since the volume of the data exceeds the boundaries of human evaluation capabilities and normal computing performance. Even though we could extract certain information from the data, it is not always easy to determine whether the analysis result of the extracted information is meaningful and helpful. Thus, there is a need for advanced tools to handle such big data and aid in examining the results in order to understand situations and glean investigative insights. Given the incomplete, complex, context-dependent information, a human in this analysis and decision-making loop is crucial. Therefore, a visual analytics approach offers great potential through interactive, scalable, and verifiable techniques, helping analysts to extract, isolate, and examine the results interactively. In this paper, we present an interactive visual analytics approach for spatiotemporal microblog data analysis to improve emergency management, disaster preparedness, and evacuation planning. We demonstrate the ability to identify spatiotemporal differences in patterns between emergency and normal situations, and analyze spatial relationships among spatial distributions of microblog users, locations of multiple types of infrastructure, and severe weather conditions.

* Corresponding author. Tel.: +82 1025483148.
E-mail address: jangy@sejong.edu (Y. Jang).

0097-8493/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cag.2013.10.008

Furthermore, we show how both spatiotemporal microblog and disaster event data can help the analysts to understand and examine emergent situations, and evaluate courses of action.

This study is performed using Twitter messages called Tweets, as Twitter has been the most popular microblog service in the United States. In this paper, we extend our previous work [4] with additional features of our system and examine their capabilities with several expanded examples in Section 4.2. We also add a discussion section for comparisons and analysis of the case studies.

Our system evaluates visual analytics of spatiotemporal distribution of Tweets to identify public behavior patterns during natural disasters. The main features of our approach are as follows:

- *Spatial analysis and decision support*: The system provides effective analysis for exploring and examining the spatial distribution of Twitter users and supporting spatial decision-making using a large volume of geo-located Tweets and multiple types of supplementary information during specific time periods (i.e., disaster events).
- *Temporal pattern analysis*: Our visualization system enables the analysts to analyze the temporal distribution of the number of Twitter users posting Tweets in a given location and time.
- *Spatiotemporal visualization*: We provide a visualization that allows the analysts to simultaneously analyze both aspects: space and time in a single view.

We first review previous work in Section 2 and describe our interactive analysis system in Section 3. We present analysis results for two natural disaster events in Section 4 and discussion in Section 5. Finally conclusion and future directions are presented in Section 6.

2. Related work

In recent years social media data has become a popular topic in a range of application domains. Several researchers have proposed and presented systems for social media analysis and important studies covering the use of social media during crisis events have been conducted. Most recent analysis environments for crisis-related social media exploration and visualization are from MacEachren et al. [5], Marcus et al. [6], and Thom et al. [7]. Their systems combine traditional spatial and geographic visualizations with means for automated location discovery, trend and outlier search, anomaly and event discovery, large scale text aggregation and highly interactive geovisual exploration. Approaches putting less focus on visualizations and more on fully automated data mining mechanisms have been proposed by Sakaki et al. [8] that use Kalman and Particle Filters to detect the location of earthquakes and typhoons based on Twitter. Various techniques for spatiotemporal data analysis and anomaly detection using visualization or machine learning techniques have been proposed by Andrienko et al. [9], Lee and Sumiya [10], and Pozdnoukhov and Kaiser [11]. Twitcident from Abel et al. [12] provides a web-based framework to search and filter crisis-related Tweets. Using the Netherlands emergency broadcast system, Twitcident automatically reacts on reported incidents and collects related information from Twitter based on semantic enrichment. In all these system the focus is primarily on individual messages and aggregated message volumes and how insight can be generated by understanding their content. In contrast, our system investigates a more user focused approach that tries to identify the whereabouts and movements of people in order to understand mass behavior.

Researchers have also examined the usage of Twitter during incidents and disasters. Terpstra et al. [13] investigate more than 90k Twitter messages that were sent during and after a storm hit the Belgium *Pukkelpop* musicfestival in 2011. They categorize Tweets into warnings about the severe weather conditions, rumors and self organization of relief measures. They show that valuable information for crisis response and decision support can be gathered from the messages. Vieweg et al. [14] investigate the differences in reaction to different crisis events. For their study they investigate eyewitness reports in Twitter from people that were affected by Oklahoma Grassfires in April 2009 and Red River Floods in March and April 2009. Their research also demonstrates the high value that the extraction of meaningful comments from crisis-related communication can have to generate insights. Furthermore, Heverin et al. [15] demonstrate that Twitter can also be a useful source of information for smaller events as they investigate the reaction to a shooting of four police officers and the subsequent search for the suspect that took place in the Seattle-Tacoma area. Based on the collection and categorization of 6000 messages they are able to show that citizens use the service to communicate and seek information related to the incident.

In this paper we also present a case study on crisis-related information gathered from Twitter data. However, in contrast to the discussed studies that harvest information directly out of the content of the messages, our method is primarily based on observing movement patterns and identifying local hotspots in order to learn about the effects of the crisis and the performance of evacuation measures.

3. Problem statement and interactive analysis process design

Analysis of public behavior, such as how people prepare and respond to disasters, plays an important role in crisis management, disaster response, and evacuation planning. Recently, social media becomes popular and people utilize it for communications not only in their daily lives, but also in abnormal disastrous situations. Thus, Location-based Social Networks services offer a new opportunity for enhancing situational awareness during disaster events. Unfortunately, collecting relevant data can be costly and finding meaningful information from the huge volume of social media data is very challenging. Therefore, there is a need for an advanced tool to analyze such massive ("big") streaming data and aid in examining the analysis results to better understand situations more efficiently.

Our proposed visual analytics approach provides multiple analysis methods: spatial analysis, spatial decision support, temporal pattern analysis, abnormal topic analysis, and interactive spatiotemporal visualization as shown in Fig. 2. In our system, all methods are tightly integrated based on a user-centered design in order to enhance the ability to analyze huge social media data (Fig. 2 (A–C)). Our Tweet collection component obtains real-time Tweets using the Twitter API—to collect about 2.2 million geo-tagged Tweets within the United States per day. In general for spatial analysis, the required accuracy of the geocoordinate depends upon the required level of location granularity. The data, however, is generated by very reliable GPS and software. We can be reasonably certain about the data accuracy as illustrated in [16]. For the temporal accuracy of Tweets, we use the time when each Tweet is created. Therefore, it is highly accurate if the time setting of the device posting a Tweet is correct. This large volume of data is stored in our database in order to maintain and track the history of the Twitter stream. Our system allows the analysts to query Tweets with a specific area and time span condition (Fig. 2(A)). The initially selected spatiotemporal context of Tweets can be represented by two different analytics components: spatial analysis and spatiotemporal visualization. Spatial analysis allows the