Social Media Visual Analytics for Emergency Management: A Systematic Mapping

Mahshid Marbouti, Theodore D. Hellmann, and Frank Maurer
University Of Calgary, Calgary, Alberta, Canada
{mmarbout, tdhellma, fmaurer}@ucalgary.ca

Abstract—Social media visual analytics are becoming important in helping emergency managers gain situation awareness and make better decisions. In this paper, we present a systematic mapping to understand how the field is structured, find out what research topics exist in social media visual analytics for emergencies, and understand what the visual analytics application categories in this area are. This systematic mapping demonstrates that the work in this area has been spread over a large number of research communities. There is limited visual analytics work regarding information diffusion, rumor analysis, and location inference for emergency management, and Twitter is still among the most popular social media types.

Index Terms—Systematic Mapping, Social Media, Crisis Mapping, Emergency Management, Visual Analytics

INTRODUCTION

When an emergency happens, people involved in response need to gain situational awareness in order to work effectively [59]. Social media is considered a critical resource for effective response because of its large userbase, fast communication, and ease of use [60]. Data from social media can be used as one of the major sources from which emergency managers can draw information to help identify and prioritize new issues or provide more detail about known issues. However, extracting meaningful information from social media is challenging due to the sheer volume of data and its high noise to signal ratio. Visual analytics [61] can help decision-makers have better awareness about on-going situations and make better decisions during critical situations. In order to better understand the current state of the art of social media visual analytics for emergency management, we performed a systematic mapping study of this field.

A systematic mapping study takes a large number of papers and classifies and categorizes them based on their titles, abstracts and keywords. The study draws a global picture of the field by providing an overview about the topics and application areas from existing publications and determines top publication venues.

1 RELATED WORK

Literature reviews have been conducted in this area to discuss the impacts of using social media in emergency management. Flávio et al. [62] provide a systematic literature review about using volunteered geographic information (VGI) for disaster management. Veil et al. [63] performed a literature review of using social media in risk and crisis communication. Boulos et al. [64] performed an in depth review on the use of crowdsourcing and sensor technologies in crisis management. Some major challenges they highlight are regarding dealing with information overload, extract the unrelated information (distil the noise) and identify unreliable information. Hiltz et al. [65] also provides a literature review about the studies who address information overload when using social media for emergency management. Most of these studies does not follow a systematic approach which makes it impossible to repeat their work and it also shows they could have missed some relevant papers. In this work, we limit our scope to social media visual analytics to extract visualization and analysis aspects of the studies as well as emergency management aspects and we also follow a systematic approach.

2 METHODOLOGY

To reduce expert bias while gaining an understanding of this field of research, we wanted to follow a repeatable and replicable search process that has a high precision and recall. We started the systematic mapping process proposed by Peterson et al. [66] and added a step to it. The main steps of Peterson’s process are: (1) defining the research questions, (2) deriving query terms from the research questions, (3) searching databases for relevant papers, (4) screening out irrelevant papers, (5) keywording papers based on titles and abstracts, and (6) data extraction and mapping. A key concern with the process is whether the query will be able to find all relevant papers (recall), but not more than those papers (precision). As the set of relevant results is not known, we followed an iterative process in order to try to estimate how good the recall of our search process is. After screening irrelevant papers out of a set of papers – the result of either an initial search against a database or of a previous iteration of the search – we randomly choose one or more relevant papers and used them for backwards snowballing and forwards snowballing. For backwards snowballing, we looked at the reference list for a set of relevant papers and added those to the set of papers used as input for step (4) – in essence, looking backwards from the time of publication to make sure we are including older relevant papers. For forward snowballing, we use a database search engine to look for papers that cite a randomly-selected paper that made it past step (4) previously – in essence, we are looking forward from the time of publication to make sure we are including newer relevant papers. If a relevant paper is identified, we change the query so that it actually is found when we re-execute the search. We iterate this process until the snowballing does not find additional relevant papers. The final query is a conjunction of disjunctions. Each disjunction covers a specific theme: Domain, Dataset, and Technique: [Domain: Emergency (OR) Disaster (OR) Crisis (OR) Catastrophe (OR) crowdsourc*] (AND) [Dataset: Social media (OR) Twitter (OR) Facebook (OR) Flickr (OR) Instagram] (AND) [Techniques: Visual* (OR) Map*]. We used * in some terms

---

Social Media Visual Analytics for Emergency Management: A Systematic Mapping

Mahshid Marbouti, Theodore D. Hellmann, and Frank Maurer
University Of Calgary, Calgary, Alberta, Canada
{mmarbout, tdhellma, fmaurer}@ucalgary.ca

Abstract—Social media visual analytics are becoming important in helping emergency managers gain situation awareness and make better decisions. In this paper, we present a systematic mapping to understand how the field is structured, find out what research topics exist in social media visual analytics for emergencies, and understand what the visual analytics application categories in this area are. This systematic mapping demonstrates that the work in this area has been spread over a large number of research communities. There is limited visual analytics work regarding information diffusion, rumor analysis, and location inference for emergency management, and Twitter is still among the most popular social media types.

Index Terms—Systematic Mapping, Social Media, Crisis Mapping, Emergency Management, Visual Analytics

INTRODUCTION

When an emergency happens, people involved in response need to gain situational awareness in order to work effectively [59]. Social media is considered a critical resource for effective response because of its large userbase, fast communication, and ease of use [60]. Data from social media can be used as one of the major sources from which emergency managers can draw information to help identify and prioritize new issues or provide more detail about known issues. However, extracting meaningful information from social media is challenging due to the sheer volume of data and its high noise to signal ratio. Visual analytics [61] can help decision-makers have better awareness about on-going situations and make better decisions during critical situations. In order to better understand the current state of the art of social media visual analytics for emergency management, we performed a systematic mapping study of this field.

A systematic mapping study takes a large number of papers and classifies and categorizes them based on their titles, abstracts and keywords. The study draws a global picture of the field by providing an overview about the topics and application areas from existing publications and determines top publication venues.

1 RELATED WORK

Literature reviews have been conducted in this area to discuss the impacts of using social media in emergency management. Flávio et al. [62] provide a systematic literature review about using volunteered geographic information (VGI) for disaster management. Veil et al. [63] performed a literature review of using social media in risk and crisis communication. Boulos et al. [64] performed an in depth review on the use of crowdsourcing and sensor technologies in crisis management. Some major challenges they highlight are regarding dealing with information overload, extract the unrelated information (distil the noise) and identify unreliable information. Hiltz et al. [65] also provides a literature review about the studies who address information overload when using social media for emergency management. Most of these studies does not follow a systematic approach which makes it impossible to repeat their work and it also shows they could have missed some relevant papers. In this work, we limit our scope to social media visual analytics to extract visualization and analysis aspects of the studies as well as emergency management aspects and we also follow a systematic approach.

2 METHODOLOGY

To reduce expert bias while gaining an understanding of this field of research, we wanted to follow a repeatable and replicable search process that has a high precision and recall. We started the systematic mapping process proposed by Peterson et al. [66] and added a step to it. The main steps of Peterson’s process are: (1) defining the research questions, (2) deriving query terms from the research questions, (3) searching databases for relevant papers, (4) screening out irrelevant papers, (5) keywording papers based on titles and abstracts, and (6) data extraction and mapping. A key concern with the process is whether the query will be able to find all relevant papers (recall), but not more than those papers (precision). As the set of relevant results is not known, we followed an iterative process in order to try to estimate how good the recall of our search process is. After screening irrelevant papers out of a set of papers – the result of either an initial search against a database or of a previous iteration of the search – we randomly choose one or more relevant papers and used them for backwards snowballing and forwards snowballing. For backwards snowballing, we looked at the reference list for a set of relevant papers and added those to the set of papers used as input for step (4) – in essence, looking backwards from the time of publication to make sure we are including older relevant papers. For forward snowballing, we use a database search engine to look for papers that cite a randomly-selected paper that made it past step (4) previously – in essence, we are looking forward from the time of publication to make sure we are including newer relevant papers. If a relevant paper is identified, we change the query so that it actually is found when we re-execute the search. We iterate this process until the snowballing does not find additional relevant papers. The final query is a conjunction of disjunctions. Each disjunction covers a specific theme: Domain, Dataset, and Technique: [Domain: Emergency (OR) Disaster (OR) Crisis (OR) Catastrophe (OR) crowdsourc*] (AND) [Dataset: Social media (OR) Twitter (OR) Facebook (OR) Flickr (OR) Instagram] (AND) [Techniques: Visual* (OR) Map*]. We used * in some terms...
to catch different combinations of terms. For example visual* will catch visual, visualization, visualize, etc.

3 RESULTS

In this section, we describe our main findings.

3.1 What are the current topics in social media visual analytics for emergency management?

We did open coding to identify what types of problems are being solved with social media visual analytics. We mapped each paper to at least one topic. Figure 1 shows the distribution of number of publications by topic and year. The topics we identified were:

Crisis Mapping: Crisis mapping publications [1, 2, 4-6, 8, 11-16, 18, 42] cover different technological aspects of gathering, integrating, and visualizing crisis related information from different sources including social media publicly available sources, maps, satellite images, etc. Crisis mapping allows the public to contribute real-time information during emergencies, like images, which can be used as supplementary information by responders for a more effective response. One of the major examples of using crisis mapping was the 2010 Haiti earthquake in which many volunteers gathered to help in mapping the basic infrastructure and to translate text messages and add them to live web-based mapping services like Crisis Camp Haiti, OpenStreetMap, Ushahidi, and GeoCommons [18]. Social media can also be used as a crisis platform to generate crisis maps. Microblogging applications such as Twitter provide a powerful capability for collecting disaster-related information from affected areas and visualize them on maps. This could be very useful during response in order to respond to people’s urgent needs.

Event exploration: Social media visual analytics can provide a rich mechanism for exploring events before, during, and after emergencies. Before an emergency, social media data can be monitored and analysed to identify events [49]. During a disaster, small scale emergencies and help requests can be identified [32]. In this publication subset, interactive visual interfaces via data mining techniques have been introduced to help users detect anomalies [23], events [22, 27, 35, 39, 51], or topics [21, 24, 29, 30, 33, 53] from different social media sources (usually textual data from microblogs) in a real-time manner. Some of the studies in this group apply visual analysis techniques to filter social media data. While some of them provide simple filtering techniques such as keyword based filtering or filtering by location [7, 67], the others has incorporated machine-learning classifiers into the visual analytics process. These studies engage the user into the classification process [32, 37, 40].

Others: The remaining publications discuss various topics: Information Diffusion is about propagation of information through social media. Visualization of information diffusion allows tracking when and where information is spreading, how different topics compete with each other, and what the public’s reaction is to particular events [43, 44, 48]. Being able to track information diffusion can be useful in the response phase for understanding how and when information spreads. The reliability of social media information was another topic we identified. Rumours can spread quickly and lead to wrong decisions and actions. Hashimoto et al. [57] propose a framework to extract rumours from social media data. The framework clarified topics on the basis of time and extracts rumours. Xu et al. [55] propose a graph based approach to detecting the location of users based on social media data. This is interesting because only a tiny portion of users shares their locations in social media websites. That’s why being able to extract users’ locations based on their social, visual, and textual information is important. Integrating heterogeneous information can help emergency managers gain situation awareness [45]. We identified two publications that focus on integrating and visualizing various information sources (maps, satellite images, government databases, and social media) during a disaster in order to help users gain information faster and have a more effective response. Lanfranchi et al. [56] discuss design implications for gaining situation awareness during a disaster. Their paper proposes visual design for various problems like topic discovery and

Figure 1 Number of Publications by topic and year

---

[Image of a graph showing the number of publications by topic and year, with categories such as Crisis Mapping, Event Exploration, and Others, and years from 2006 to 2016.]
Table 2. Publications by Topic and Application Category

<table>
<thead>
<tr>
<th>Application Category</th>
<th>Topics</th>
<th>Time &amp; Space</th>
<th>Graph</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crisis Mapping</td>
<td>[7, 11]</td>
<td>[1, 2, 4-9, 11-18]</td>
<td></td>
<td>[3, 10]</td>
</tr>
<tr>
<td>Event Exploration</td>
<td>[19, 20, 24, 25, 28, 29, 32-35, 39, 49, 50]</td>
<td>[19-24, 26, 27, 31, 36-39, 41, 50-53]</td>
<td>[19, 31, 33, 40, 52, 53]</td>
<td>[54]</td>
</tr>
<tr>
<td>Others</td>
<td>[44, 48]</td>
<td>[43, 46, 47, 55]</td>
<td>[45, 46, 56-58]</td>
<td></td>
</tr>
</tbody>
</table>

Finally, the last two publications discuss collecting, visualizing, and analysing social media data for disaster management [47] and animal bio-preparedness [46].

3.2 Which application categories do these topics relate to?

From a visual analysis perspective, we classified papers based on application categories. Sun [68] and Liu et al. [69] classifies recent visual analytics works into a set of application categories including Space and Time, Multivariate, Text, Graph and Network, Map and Other Applications. Inspired by the categories used in these publications, we performed closed coding to extract application categories from our paper set. We ended up with the following application categories:

- **Text:** A tremendous amount of text data can be found in most social media services. This group of studies have focused on extracting the relevant information from ambiguous, high volume, and usually unstructured textual data [69].

- **Space & Time:** Some social media services, like Twitter, provide spatio-temporal datasets. Spatial and temporal relationships within these datasets can be valuable for extracting events in emergencies.

- **Graph:** This category contains publications that use visualization of graphs to represent different elements in social media data and the connection between those elements. Graph analysis and visualization can be used for solving different problems, such as representing topics [53] and information diffusion [43].

The rest of the publications provide visual analysis design recommendations, best practices, and shortfalls in the area. Table 1 organizes papers by considering two dimensions: Their topics and their application category.

3.3 What social media data source has been used for addressing different problems?

To discover which social media sources have been studied in this field, we used closed coding. We focused on most-used social media types, which has been included in our query:

- Twitter
- Facebook
- Flickr
- Instagram
- Multiple Social Media (more than one in one publication)
- Others (papers which talk generally about social media without specifying any specific service or talk about social media types that are not included in our list e.g. Foursquare)

In Table 2, we represent which publications discussing each social media type arranged by the topic of each paper. Sometimes you can see the same publications in different rows of table because they discuss more than one type of social media. Among our social media list, Twitter has the most number of publications, while Facebook, Flickr and Instagram were only used infrequently.

3.4 What venues are most important for this field?

The studies are widely distributed among different research communities from different fields. The fact that there are many dimensions associated within this topic causes this distribution. In our paper set, publications in *Emergency and crisis management* related venues focus on research about deployment of information systems for crisis management using crowd sourcing. Publications in *Visualization and Visual Analytics* venues focus on visualization, graphics, and visual analytics techniques to enable use of social media in emergency management. Studies in *GIS* venues focus on geotagged social media data visual analysis, interacting and collaborating with social media data on maps, and raising spatial awareness for emergency response. Studies in *Intelligent systems* venues focus on computational methods for social media analysis. Publications in *Natural Language Processing* (NLP) venues focus on applying NLP techniques to extract events and stories from the text of social media data. Works in *Crowdsourcing* venues focus on crowd sourcing techniques, information extraction through social media, and volunteered geographic information retrieval. There are many other venues in different categories such as data mining, big data, health, multimedia, network, wireless communications, etc. As you can see in Figure 2 many venues has one or two publication. There are three venues with more than 3 publications. Those venues are:

- **International Conference on Information Systems for Crisis Response and Management (ISCRAM)** is an annual conference that publishes research on subjects related to information systems for all phases of emergency management.
- **IEEE Conference on Visual Analytics Science and Technology (VAST)** is an annual conference focusing on advances in visual analytics.
- **IEEE Transactions on Visualization and Computer Graphics (TVCG)** is a monthly journal focusing on computer graphic and visualization techniques.

Table 1. Publications by Topic and Social Media Type

<table>
<thead>
<tr>
<th>Social Media Type</th>
<th>Crisis Mapping</th>
<th>Event Exploration</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter</td>
<td>[1-10]</td>
<td>[19-41]</td>
<td>[43-48]</td>
</tr>
<tr>
<td>Facebook</td>
<td></td>
<td></td>
<td>[46, 47]</td>
</tr>
<tr>
<td>Flickr</td>
<td>[17]</td>
<td>[40, 41, 50, 51, 54]</td>
<td>[58]</td>
</tr>
<tr>
<td>Instagram</td>
<td>[1, 17]</td>
<td></td>
<td>[56, 58]</td>
</tr>
<tr>
<td>Multiple Social Media</td>
<td>[17]</td>
<td>[50-52, 54]</td>
<td>[56, 58]</td>
</tr>
<tr>
<td>Others</td>
<td>[11-16, 18, 42]</td>
<td>[49, 53]</td>
<td>[55, 57]</td>
</tr>
</tbody>
</table>

Table 2. Publications by Topic and Application Category
4 Conclusion

In this paper, we present the findings of a systematic mapping study for using social media visual analytics in emergency management. We addressed a variety of research questions. We found out among the identified topics, Crisis Mapping and Event Exploration had the most publications, while there are limited publications on visual analytics work in Information Diffusion, rumor detection and location inference. Among the social media sources, Twitter has been widely used compared to other social media data types. This could be because Twitter has a broad user base, has a developer-friendly API, and provides functionality for both spatio-temporal data and hashtag mechanisms that facilitate extracting spatio-temporal patterns and applying text-mining techniques. Finally, given the interdisciplinary nature of the research topic (emergency management, social media, visualization, visual analytics, etc.) publications have been distributed among very different venues.

References


