Utilizing Volunteered Geographic Information to Develop a Real-Time Disaster Mapping Tool: A Prototype and Research Framework

Michael A. Erskine  
*University of Colorado Denver, michael.erskine@ucdenver.edu*

Dawn G. Gregg  
*University of Colorado Denver, dawn.gregg@ucdenver.edu*

Follow this and additional works at: [http://aisel.aisnet.org/confirm2012](http://aisel.aisnet.org/confirm2012)

Recommended Citation

[http://aisel.aisnet.org/confirm2012/27](http://aisel.aisnet.org/confirm2012/27)
Utilizing Volunteered Geographic Information to Develop a Real-Time Disaster Mapping Tool: A Prototype and Research Framework

Michael A. Erskine
University of Colorado Denver
michael.erskine@ucdenver.edu

Dawn G. Gregg
University of Colorado Denver
dawn.gregg@ucdenver.edu

Abstract
The global proliferation of personal mobile devices has provided the capability of electronic data collection to billions of people. Furthermore, recent innovations in technologies and implementation methodologies have allowed groups of people to collect and analyze large quantities of data. Examples of such systems include Wikipedia, a community-sourced digital encyclopedia, and Yelp, a directory and review tool of local businesses. The occurrence of important events or the establishment of new restaurants has motivated individuals to provide timely and accurate contributions to Wikipedia or Yelp, respectively. Considering the benefits of widespread data collection capabilities and the motivations to contribute to public knowledge, this design-science paper proposes and prototypes components of a publically driven, real-time disaster-response mapping system.

Keywords
Crowdsourcing, Design Science, Volunteered Geographic Information, Decision-Making, Geographic Information Systems and Disaster Response.

1. Introduction
Throughout history, large-scale disasters have caused enormous casualties, destruction and economic harm. But never before were communications and transportation technologies available to aid in disasters as quickly as today. However, even with such advances, often real-time data from within a disaster area are not available until days after a disaster or after responders have been deployed.
This paper posits that volunteered data, contributed by individuals from within a disaster area, can be utilized to develop real-time maps to aid disaster responders in coordinating responses more effectively. Using the data collection capability of modern mobile devices, individuals could contribute information to a centralized system that could broadcast relevant and accurate maps of the disaster area to initial emergency responders, as well as disaster responders and coordinators.
A review of existing research has revealed that practical technology and current research may be able to converge in order to solve a major obstacle for efficient and effective disaster response: providing accurate and relevant information to a central disaster command and sharing this information with disaster responders in real-time (e.g. Cheong and Cheong, 2011; Velev and Zlateva, 2011; Sutton et al., 2011).

1.1 Disaster Response

It may seem that large-scale disasters are a rare occurrence, but in the last two years alone there have been several notable disasters, including the 2010 Haiti earthquake, the 2010 Chile earthquake, the 2010 Pakistan flooding, the 2010/2011 Queensland flooding, the 2011 Tohoku earthquake and tsunami, as well as the 2011 Christchurch earthquake (UNEP, 2011). Together these six disasters alone accounted for 333,944 reported casualties and an estimated economic loss of $US 199.3 to 327.3 billion (see Table 1).

<table>
<thead>
<tr>
<th>Disaster</th>
<th>Estimated Casualties</th>
<th>Estimated Economic Impact and Damages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010 Haiti Earthquake</td>
<td>316,000 killed</td>
<td>Not available.</td>
</tr>
<tr>
<td>2010 Chile Earthquake</td>
<td>525 killed</td>
<td>$US 15-30 Billion (UNEP, 2011; USGS 2011)</td>
</tr>
<tr>
<td>2010 Pakistan Floods</td>
<td>1,500 killed</td>
<td>$US 43 Billion (Tarakzai, 2010)</td>
</tr>
</tbody>
</table>

Table 1: Notable Disasters Since 2010

Large-scale disasters will be ever-present; however, recent technology, communications and transportation systems have improved disaster response dramatically. For example, following the 2010 Haiti earthquake, the Israeli Defense Forces Medical Corps deployed an assessment team within 11 hours and installed an operational field hospital within 89 hours of the disaster (Kreiss et al., 2010).

However, while such response times by governmental and non-governmental disaster response organizations are impressive, in the case of an earthquake, numerous studies have shown that the chance of survival is reduced significantly when trapped under rubble for more than six hours (Tanaka, 1998; Mirhashemi, 2007). Furthermore, Schultz et al. (1996) highlight the importance of immediate medical response following a natural disaster. The greatest need for medical care is between 24 to 48 hours after the disaster indicating the importance of timely and efficient emergency response. Furthermore, Schultz et al. note that the deployment of field hospitals, usually about a week after the incident, is often too late to mitigate initial fatalities. Since large-scale disaster response is frequently not initially available, local emergency responders usually provide immediate assistance. For example, Roy et al. (2003) state that local health care providers were first to respond to a major earthquake in Gujarat, India, with field hospitals and relief doctors arriving in the week after the disaster. Thus, to ensure survival, it is imperative that local emergency responders and volunteers have the necessary information and tools to respond immediately and effectively.
1.2 Disaster Response Coordination
Following a disaster, communication with front-line responders, civilians and command centers becomes increasingly difficult due to varying standards, languages, technologies, unexpected terrain, unreliable energy distribution, as well as unclear management. For example, Costa Rica developed an emergency response network that proved problematic after a major 1991 earthquake when emergency responders discovered that land-based telephone systems, essential to the communication system, were no longer operational and that they did not have the capability to transmit radio signals to and from the affected areas due to the terrain (Comfort, 1994). Similarly, after the 1999 Chi-Chi earthquake, central command underestimated the significance of the disaster because a lack of communication from military and fire radio systems gave the perception that the incident was not as massive as it actually was (Chan et al., 2006). Communications problems are not always due to technological or physical limitations. For example, responders to the 2010 Haiti earthquake made certain decisions without having a clear overall picture that a commanding agency such as the United Nations Disaster Assessment and Coordination body may have possessed (Kreiss et al., 2010). As making such decisions can inefficiently or incorrectly distribute disaster response resources, providing accurate and meaningful information to all groups can ensure that responses are most effective. Such information can be distributed through the development of disaster maps, either during disaster planning, disaster response or post-disaster.

1.3 Disaster Mapping
One important strategy utilized in disaster planning is the creation and distribution of maps. For example, in response to the 2009 Victoria Bushfires, the Victoria State Government developed maps that reveal bushfire attack risks (Victoria Building Commission, 2011). By plotting a building on such a map, its risk of harm from a bushfire can be determined. Utilizing such maps to update building codes or mandate construction methods helps ensure that appropriate fire mitigation techniques are implemented in new construction. Another example is to highlight the locations of distinct population groups that are particularly vulnerable to disasters, such as the poor, tourists, and the elderly. To ensure that emergency responders are aware of such populations, vulnerability maps can be developed to show concentrations of such high-risk populations (Morrow, 1999).

Furthermore, while most developed countries have accurate maps, some areas of the world have not been fully mapped, either due to a lack of resources or limited return on investment. To address this limitation, volunteered geographic information (VGI), or community-sourced geographic data, can be used to develop maps, as was done during the 2010 Haiti earthquake response (Zook et al., 2010). With today’s widespread use of mobile, internet-enabled devices, there is a potential to create accurate, real-time maps to aid emergency responders during a disaster event.

1.4 Mobile Devices
Nearly 6 billion mobile-cellular subscriptions reach approximately 87 percent of the world’s population. In addition, there are nearly 1.2 billion active mobile broadband subscriptions (ICT, 2011). While the data collection capabilities of mobile-cellular users may be sufficient to perform some of the tasks necessary to create a real-time disaster map, ideally relevant data-collection would occur using advanced mobile devices such as smartphones that have the ability
to collect text, photographic, as well as geospatial data, using cameras and global positioning system (GPS) receivers.

The idea of using mobile devices to communicate with emergency responders has been explored previously. For example, a wireless device to be used by first responders to track victims, provide patient monitoring and share this information with incident commanders was proposed. For such mobile devices to communicate in disaster areas, the deployment of quickly and easily deployed ad-hoc wireless networks was suggested (Killeen et al., 2006).

Assuming mobile-cellular and mobile-broadband network service can be retained or quickly restored following a disaster, anyone with the access to a networked mobile device could begin populating data to a real-time map of the disaster area.

1.5 Crowdsourcing

Crowdsourcing, originally defined as “the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers (Howe, 2006a, 2006b).” Furthermore, Brabham (2008, p. 87) states that, “[Crowdsourcing] is a model capable of aggregating talent, leveraging ingenuity while reducing the cost and time formerly needed to solve problems.”

Yet, while crowdsourcing originally referred to a method for recruiting labor for business tasks, more recently crowdsourcing has also referred to volunteered information. For example, Mooney et al. (2011) describe how VGI could be used to develop relevant and current maps in environments where traditional data collection methods are costly and time-consuming. Additionally, during recent disasters, such as during the 2011 Queensland flooding, Twitter, a social networking tool, was used to share real-time information (Cheong and Cheong, 2011). Sutton et al. (2011) describe how Twitter was used as an unofficial channel to communicate tsunami warning information to individuals and the media in Hawaii.

In addition to broadcasting important information, social networking tools could also be used to aggregate data in real time. Velev and Zlateva (2011) recommend the development of a social network site for emergency management purposes, as well as additional research into the use of social networking sites during disasters, as they discovered that there is limited knowledge of how such tools can effectively be deployed and utilized during disasters.

Flanagin and Metzger (2008) state that “individuals are in many cases in the best position to provide information that requires indigenous experience, esoteric understanding of a particular physical environment, and current information about local conditions.” However, such information, sourced from many, unique individuals may bring about a sense of distrust and concerns of credibility. Yet, much like Wikipedia allows for validation and discussion of data, a disaster-specific social network site could provide the capability for such validation.

Levy (1997) describes that as knowledge can be considered the source of all other wealth, it is no longer feasible to restrict knowledge to only domain-experts. Furthermore, Surowiecki (2004) suggests that in certain cases, groups can be used for decision-making and that the aggregated knowledge of a group is greater than any one individual’s knowledge. Such thinking brought about the concept of utilizing crowdsourced information, or more specifically VGI, to develop real-time disaster mapping for emergency responders.
This paper uses a design-science research methodology to address the following research questions:

- Can VGI be utilized to develop effective real-time mapping for disaster coordinators and responders?
- Is the type of information that can be provided by non-domain experts in the field adequate to create improved real-time disaster maps?

The remainder of the paper is outlined as follows: First, the design science-research model is introduced along with an overview of the prototype system and its components. This is followed by an explanation of the prototype development process along with findings produced using test data. Third, a discussion section provides implications, limitations and future research directions. Finally, a conclusion is presented.

2. Research Model
Based on the reviewed literature and the maturity and penetration of technologies, the development of a prototype crowdsource-driven, real-time disaster-mapping tool is proposed. This paper contributes to the design-science discipline of the information systems scholarship, as suggested by dimensions and ranking for the classification of design-science research developed by Gregg et al. (2001). In their work, conceptual, formal and developmental dimensions are suggested along with criteria for ranking, of which, this paper addresses all three dimensions at a medium or high rating. First, the conceptual dimension is addressed as no previous research has utilized crowdsourcing as a knowledge management input for real-time disaster mapping. Second, the formal research dimension is addressed through the description of geostatistical methods used. Finally, the developmental dimension is addresses through the development of a prototype as well as validation using test data.

2.1 Prototype Components
The proposed system will consist of several important, but independent modules. First, baseline data will be sourced from disaster planning initiatives. This baseline data will be used to develop a basemap presenting the most vulnerable areas of a disaster region. Additionally, mobile applications will allow victims, citizens and early responders to collect real-time information within the disaster area. Such information will consist of textual notes, audio notes, categorization information, photographs as well as geographic coordinates. When applicable, such information will flow through a module in which experts can re-evaluate information to ensure proper categorization and prioritization. Finally, the baseline data and the community-sourced information will be aggregated, analyzed and published. Early responders, hospitals, command centers, international relief organizations, and volunteers will be able to view such maps and respond accordingly. See Figure 1 for a model of the entire system prototype.\(^1\)

While the long-term goal of this research is to develop a fully functional prototype incorporating all components, this initial project focuses only on the analysis components of the system.

---

\(^1\) Additional details on the prototype system can be obtained from the authors.
3. Research Methods
The authors developed a working prototype of the analysis and decision-tools utilizing a design-science approach. This approach involved identifying key data that should be tracked within the system, as well as discovering appropriate geostatistical analysis methods of the baseline data and volunteered geographic information. Due to its de facto global standard and prominent use, Environmental Research Systems Incorporated’s (ESRI) ArcGIS 10, along with the ESRI Spatial Analyst, were used for the prototype development.

3.1 Items of Concern and Attribute Gathering
The first step of the design-science process was to identify the key data that would need to be collected and utilized for the analysis. After a review of relevant literature and discussions with emergency responders, transportation, structures, infrastructure and natural/geological were identified as primary categories. Furthermore, specific items were identified within each of these categories. For example, critical items within the transportation category included docks, roads, railroads, and airfields/airports. Table 2 lists all of the items of concern and categories utilized for this research project.
Information pertaining to the aforementioned items provides the greatest benefits when combined with existing disaster planning information and maps. For example, knowing the population density of certain areas combined with building collapse information could allow responders to focus their attention to structures that have the greatest likelihood of victims and entrapment.
<table>
<thead>
<tr>
<th>Items</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation</td>
<td>Docks (Haas, 1970; Comfort, 1994; Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Roads (Comfort, 1994; Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Railroads (Youd et al., 1992; Byers, 2000, 2001; Day, 2002)</td>
</tr>
<tr>
<td></td>
<td>Airfields/Airports (Ghalory-Ashtiany, 2004; Kreiss et al., 2010)</td>
</tr>
<tr>
<td>Structures</td>
<td>Buildings – Structural Failure (Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Tunnels</td>
</tr>
<tr>
<td></td>
<td>Bridges (Youd et al., 1992; Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Seawalls (Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Dams (Simonoff et al., 2011)</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Water Supply (Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Wired Telephone/Electrical Systems</td>
</tr>
<tr>
<td></td>
<td>Gas Lines</td>
</tr>
<tr>
<td>Natural/Geological</td>
<td>Faults</td>
</tr>
<tr>
<td></td>
<td>Lava</td>
</tr>
<tr>
<td></td>
<td>Flooding (Greenstein, 2003)</td>
</tr>
<tr>
<td></td>
<td>Land/Mudsides (Greenstein, 2003)</td>
</tr>
</tbody>
</table>

Table 2: Items of Concern and Categorization

3.2 Basemap Prototype
Following the identification of key items of concern, the next step was to develop a prototype baseline-mapping tool. The basemap was created utilizing four geospatial inputs: the disaster area perimeter, elevation data, locations of properties that are at high risk and locations of existing emergency services. See Figure 2 for a sample basemap utilizing the four inputs. Additional basemap inputs could include existing features and attributes, such as seawall strength, drainage capacity, and forest densities.

To develop the basemap a slope analysis and a Euclidean distance analysis were conducted. Euclidean distance analysis is a commonly used geospatial analysis method to determine straight-line distances from points in space. For this research project, a Euclidean distance analysis was performed on known locations of emergency service providers, such as fire stations and hospitals, and highly vulnerable locations, such as senior care centers and low-income, high-occupancy buildings.

Additionally, a 148.57 MB file consisting of a digital elevation model (DEM) was utilized to calculate slopes within the disaster area. A slope analysis, or a method to determine the rate of change between every cell in a raster, was used to develop a new raster containing the percentage of slope. In the prototype, slopes greater than 10 percent were coded with higher values as the slope increased.

3.3 VGI Map Prototypes
Once techniques for the development of a basemap were identified, the authors focused on the analysis techniques of the volunteered geographic information. Based on the data types identified in Table 2, two analysis techniques were required. For input consisting of structures or natural/geological data the input was analyzed using Kriging, while transportation and infrastructure data was analyzed using network analysis. A test data set of volunteered geographic information was developed and consisted of point files containing disaster intensity ratings on a scale of one to three, with three being the most intense. This type of data is consistent with the type of data that could be provided by victims, citizens and early responders using a cell phone or other mobile device.
Kriging, a geostatistical interpolation method based on the regionalized variable theory, was implemented (Cressie, 1990; Oliver and Webster, 1990). As Kriging allows for optimal predictions to be made for points in space by only using measurements collected at a few known locations it was deemed appropriate for this study. As data for these locations were recorded on an intensity-scale, clear maps of the interpolated data were developed. For example, test data analyzed using Kriging revealed a clear map of a hypothetical mudslide (see Figure 3).
Ideally, a web-based decision-system would be developed in order to allow community or official experts to evaluate incoming data. A domain-expert on water supply lines may discover that a data point submitted as a critical flaw is nothing more than standard aging of a waterline. In such case, the expert may re-code the severity of the item or entirely remove the item from the map. However, as expert mediation may not be available to review incoming data, the system design emphasizes peer-validated data. As with Wikipedia, large quantities of good data will generally compensate for a few bad data points. For certain types of maps, such as those of roads or electrical grids, network analysis layers may exist. Network analysis layers are unique in that points are connected and can be used to determine all possible paths, ascertain ideal paths or locate breaks in paths between points. In cases where network layers exist, qualifying inputs were added as barriers that fully prevent traffic on the network. In the case of electrical grids, such a point may represent a collapsed electrical line or in the case of a road, large debris or damage making it impassible.

4. Findings
The identification of necessary items, the creation of test data and the development of prototype components revealed that developing clear and concise maps for use by emergency responders or emergency commanders can be accomplished quickly and simply when sourced from VGI using only three levels of intensity. Furthermore, it was demonstrated that a basemap could easily be developed using data collected prior to the disaster. Overall, the development of a real-time, crowdsourced disaster area mapping tool appears to be viable, warranting further research.

5. Discussion
While this initial research demonstrates the feasibility of utilizing VGI for the development of real-time disaster mapping, there are several limitations.

5.1 Limitations
First, this design assumes that cellular or mobile broadband coverage will be available in the disaster area, allowing participants to submit data to the system. Additionally, it is assumed that mobile device users will learn about the necessary VGI application during a time of other essential and competing communication messages. Second, an important consideration is to determine if data collected by non-experts provides sufficient information for further classification by domain-experts. If not, perhaps the secondary classification should be eliminated for certain data or specific instructions could be provided to the non-expert data collector. Third, error rates of bad data, or missing data, must be evaluated to determine what percentage of such data will completely invalidate a map to an emergency responder. During a disaster there may be a propensity to only collect data that is ranked highly, potentially reducing the utility of the map. Additionally, data collection at the disaster epicenter may be limited, causing the epicenter to appear as an area of no concern to responders.
Usability studies will need to be conducted to determine how many data points are required to create a meaningful map. Furthermore, it must be better understood if trust of non-authoritative information as well as if geospatial reasoning ability affect usability (Erskine and Gregg, 2011).

5.2 Benefits
These initial findings provide a benefit to both practice and scholarship. Benefits to the information systems scholarship include a clear overview of the necessary geostatistical analysis techniques to analyze crowdsourced disaster data as well as a presentation of an initial framework for the development of a crowdsourced disaster-mapping prototype. Practice can benefit from a better understanding of how crowdsourced data can be used for mapping purposes.

5.3 Future Research
While this research explored the analysis components of a real-time disaster mapping system, additional research will need to explore the data collection and information output components, along with usability. Furthermore, once a full system prototype exists all combined modules should be tested in a simulated disaster environment to test technical limitations as well as acceptance by emergency responders.

5.3.1 Crowdsourced Data Input Prototype
Crowdsourced data are essential to the functionality of this system. Ideally, a mobile application that can function on as many devices as possible, and collect data and store information locally during times of no network connectivity, should be developed. Furthermore, this application should be easy to use and function internationally without the need to redevelop or redeploy based on language or region.

5.3.2 Data Output Prototype
Upon completion of the data processing a raster image displaying intensity is created, which can be overlaid with satellite imagery or other maps for emergency responders or command centers to quickly interpret the information provided. Ideally, such information will be publically available for relief organizations as well as responders in the field. A prototype of this component will be developed in the near future.

5.4 Conclusion
This paper presents a framework for the development of a crowdsourced, real-time disaster mapping system, consisting of four distinct modules: mobile applications for VGI collection, analysis tools for basemap data, analysis tools for VGI and a visualization component to present data to commanders and field responders. Results of this design-science research demonstrate that crowdsourcing concepts can be used to develop effective real-time mapping for disaster coordinators and responders. In addition, evaluation of the prototype system shows that the type of data that could be provided by victims, citizens and early responders using a cell phone or other mobile device is sufficient to create improved real-time disaster maps. This suggests that real-time VGI disaster mapping systems have the potential to be more effective than current disaster response systems.
References


