Utilizing Community Volunteered Information to Enhance Disaster Situational Awareness

David F. Merrick
Florida State University
dmerrick@fsu.edu

Tom Duffy
Florida State University
tduffy@fsu.edu

ABSTRACT

Social media allows the public to engage in the disaster response and recovery process in new and exciting ways. Many emergency management agencies in the United States are embracing social media as a new channel for alerts, warnings, and public outreach, but very few are mining the massive amounts of data available for use in disaster response. The research reflected in this paper strives to help emergency management practitioners harness the power of community volunteered information in a way that is still novel in most parts of the country. Field verification and research combined with survey results attempts to identify and solve many of the barriers to adoption that currently exist. By helping practitioners understand the virtues and limitations of this type of data and information, this research will encourage the use of community volunteered information in the emergency operations center.

Keywords

Social media, crowd sourcing, situational awareness, twitter, facebook, community volunteered information

DEFINING SITUATIONAL AWARENESS

In emergency and disaster situations, responders and emergency managers are faced with a myriad of decisions. These decisions often involve the application of limited resources, have serious consequences, and are time sensitive. In order to make the best decision possible in this stressful environment, an emergency manager requires the best information possible – they must have up to date situational awareness.

Situational awareness can be defined as awareness of and understand what is happening around them, be able to predict how that will change over time, and see how different factors in the environment can impact the dynamic situation. While decision makers often overlook what makes up situational awareness, they know that they require quality information and intelligence and that information must be constantly updated. In short, situational awareness, particularly in an emergency operations center (EOC) that is geographically removed from the emergency site is difficult to achieve.

SOURCES OF DATA AND INFORMATION FOR SITUATIONAL AWARENESS

Information is raw data that has been analyzed, and it contains the necessary semantics and references to be useful to decision makers. In the classic model, raw data points are analyzed and processed, creating information. The resulting information contains meaning; it is useful for decision makers and can contribute to overall situational awareness. For example, a list of house addresses is data. The same list, cross-referenced to utilities status and physical damage reports is information. Only the second example can drive situational awareness.

Traditional sources of data and information for situational awareness include reconnaissance and damage reports from first responders and other professional assets, weather reports and forecasts, general geographical information, emergency communications systems (911 call centers), population information, partner provided information and reports, and reports and information from the private sector (particularly from transportation partners).

These sources are generally viewed in the EOC as being accurate. That is, they provide data and information that accurately reflects real-world conditions. Emergency managers also view this data as consistent – the same information or conditions will be described or labeled the same way over and over.

However, these sources are not perfect. While they may be accurate and consistent, they are often slow and

cannot provide a complete picture or the emergency or disaster. This is especially true as the geographical size and scope of the disaster increases. Reconnaissance and damage assessments for large disaster areas may not be complete for days and emergency managers run a risk of missing damage reports altogether, or making decisions based on incomplete reports. In reality there is no economical way to accelerate traditional reconnaissance after a disaster. It is slow and often dangerous work, and government budgets and manpower are already stretched thin. Given these limitations, it should be possible to enhance situational awareness through non-traditional sources.

COMMUNITY VOLUNTEERED DATA AND INFORMATION

Community volunteered data (CVD) is another example of how the public at large can participate in responding to an emergency both on the ground and in the emerging "virtual space emergency" (Schneiderman & Preece, 2007; Palen et al, 2009). The pervasiveness of information and communications technologies (ICT) in recent times have given the public a medium through which they can show their altruistic tendencies in times of crisis. This isn't always through purpose built off-the-shelf technologies; much of the time this is through the re-appropriation of online technologies that suit the needs of the crisis scenario (Starbird & Palen, 2011, Starbird & Stamberger, 2010).

CVD can be found in a variety of sources and mediums. Twitter and Facebook provide a simple and familiar way for people to post data to the public, using smartphones, tablets and computers. Images are posted to Flickr and Picasa, and videos to YouTube and Vimeo. Every major news outlet on the internet provides a mechanism for users to comment on news stories, and CNN provides a news outlet strictly for CVD – CNN iReport. All of these channels can provide meaningful data for emergency managers, but identifying all the channels and separating the valid from the invalid is difficult and not intuitive.

As Tropical Storm Debby moved onshore near Jacksonville, Florida in June of 2012, a capture of Twitter traffic gathered 36,317 tweets matching a list of keywords, including "Debby", "TS Debby", "flwx", and "hurricane". This appeared to be a promising capture of potential data, but in reality was full of 'noise'. During the same week, the band One Direction tweeted a post with the word "Debby" in it (which was captured) and that single post was retweeted by almost 28,000 people. In the final analysis, over 75% of the traffic captured did not pertain to the storm. Further analysis of the remaining Debby related Twitter traffic shows that much of the data being posted relates only to references to the fact that it is currently raining, or complaints about cable TV outages. Simply put, the majority of data on Twitter during Tropical Storm Debby was noise and distraction – finding useful and actionable information was clearly a challenge.

Many emergency managers understand there is a large amount of CVD available at their fingertips in a modern emergency operations center. However, the perception is this data is inherently flawed. Using the same four metrics applied to traditional situational awareness data and information sources – accuracy, timeliness, consistency and completeness - the only positive aspect of CVD is timeliness. CVD flows at an astonishing and often overwhelming rate. However, the general assumption in the emergency management community is that this data is unreliable and useless professionally. (Koon, 2012) The data lacks accuracy, consistency and completeness. This is a 'wisdom of crowds' (Surowiecki, 2004) approach to harvesting data. However where the crowd is useful for collecting quantitative data, it is more challenging to mine out more qualitative data and turn this into reliable information. More recent research is making attempts at improving the quality of information obtained through CVD by identifying the reliability of data providers through collaborative filtering techniques (Starbird et al, 2012). There are also more technical approaches being taken such as Rogstadius et al (2011) work on real-time social media aggregation tools that has led to the development of the CrisisTracker system (Rogstadius et al (in press), Crisis Tracker Website).

Processing CVD through a model or workflow, and removing the distracting and useless data will yield usable, meaningful information. The output of the process is called community volunteered information, or CVI. While traditional situational data sources are purpose built and immediately applicable to emergency management, creating CVI from CVD requires (at this juncture), human analysis and processing to create useful information applicable to the current situation.

CREATING CVI

As Hurricane Isaac approached the Florida panhandle and northern coast of the Gulf of Mexico, FSU researchers again considered the validity of social media as it may apply to situational awareness. To gather data on CVD accuracy and usability, the FSU Virtual Operation Support Team (VOST) partially activated in Tallahassee, Florida, and a separate field research team moved westward to follow the landfall of the storm. Isaac made landfall near New Orleans, Louisiana as a Category 1 hurricane on August 28th, 2012. The slow

forward progress of the storm caused inland flooding due to heavy rainfall, and storm surge combined with tides broke several levees southeast of New Orleans resulting in heavy flooding of residential areas. (Brown, 2012)

The FSU VOST is a developing team of digital volunteers using ideas and concepts first defined by Jeff Phillips in 2011 and tested later that year in the Shadow Lake Fire in August and September 2011. A VOST is designed to monitor media and social media channels and accomplish the task of CVD to CVI analysis. (St. Denis, 2012) This data is then provided to response and coordination agencies through an established and trusted method. The FSU VOST used this research deployment as a test of communications tools, workflow and policy.

The goal of the field research team was straightforward – measure the accuracy and consistency of social media reporting in and around New Orleans following landfall. On August 29th, 2012, the VOST in Tallahassee began forwarding social media data (CVD) to the field research team via email, with a focus on reports of damage, flooding, utility and power disruption, and available services. The VOST was not tracking the 'noise' – data of no use to situational awareness – but focusing on finding the 'signal' – data that could build situational awareness. They were, in effect, creating CVI from CVD. During this event, this process was a manual one and resource intensive. Four to five VOST researchers worked eight to ten hours per day to sort and identify potential data points.

As the FSU VOST compiled this data, they forwarded it to research personnel in New Orleans for follow-up and verification. The field research team took the data and drove to the areas indicated in the CVI and compared the reports to the current conditions on the ground. Due to limited time available to the research team in New Orleans, the sample size of data points is small (n=28). Given that fact, these results are considered to be preliminary until follow up research trips can be designed and implemented. Further, each data point represents a single post on social media. The VOST ignored re-posts of other people's reports (re-tweets), though multiple independent reports of damage or conditions were included. In other words, only original posts were forwarded to the field team for follow-up.

Of the reports investigated by the research team, 77% were confirmed to be accurate depictions on the ground. However, there was no way to gauge consistency of the reports, or completeness of the information. Research team observations while in New Orleans showed much damage and disaster impacts that were not reported on social media, making the completeness of the CVI suspect. Further, professional emergency managers would often choose a different description of an event than the person who provided the CVI. "Flooding" was referred to often in social media posts, when in reality most of the events showed only evidence of minor ponding or standing water in the roadway.

Two interesting results appeared. First, data posted by government agencies and local neighborhood associations had the same general accuracy rate as the data posted by members of the general public. This may indicate that during a disaster, these organizations are simply repeating what they are told by the public, and not performing any independent verification of public reports. Secondly, 100% of data provided (from any source) with an attached photo was shown to be accurate – the posted social media report matched the conditions on the ground.

Overall, the data presented to the field team for verification did a decent job of conveying damage to trees and utility service, as well as where supplies and food was available (typically through restaurant limited menus). Three neighborhoods of New Orleans created their own mashups of social media data and provided online tools for residents to report issues and damage. A high level of post-disaster social media usage by the community illustrates the public's increasing dependence on the medium. Government disaster management officials need a method to integrate these new data points into existing situational awareness workflows.

BARRIERS TO ADOPTION OF CVI FOR SITUATIONAL AWARENESS

There are many barriers to adopting CVI as a part of disaster situational awareness. Despite an overall observed accuracy rate of 77%, processing CVI from CVD is time consuming, and therefore expensive, to create during a disaster or crisis. While emergency managers would like to integrate CVI into the EOC environment, current doubts about accuracy and reliability hinder this effort (Koon, 2012).

In addition to concerns about accuracy, emergency managers know that creating CVI is time intensive, and EOC staffing levels and funding are not increasing. Digital volunteers such as a VOST can be (and are) used, but these must be well trained and well managed to produce useful CVI, again causing stresses on EOC personnel and processes.

Finally, no model or process exists for adoption and integration into an EOC's situational awareness workflow. Without a "plug and play" process and tool, most EOCs lack the technical and process expertise and resources to create one from scratch.

CURRENT RESEARCH AND INITIATIVE TO OVERCOME BARRIERS TO ADOPTION

Current research at Florida State University focuses on improving and expanding the field data verification started with Hurricane Isaac. In addition to performing additional CVI validation, researchers will conduct interviews with the public and emergency responders regarding the perception of social media in the context of disaster and emergency management. It is the current opinion of researchers that public perception of the value of the data in social media is significantly different than the perceptions of emergency managers. Future research surveys will confirm or deny this.

In addition to additional field research, the Center for Disaster Risk Policy at Florida State University (CDRP) is building a model for mining CVD and creating CVI for situational awareness. Using currently available tools such as Swift River by Ushahidi combined with custom applications for monitoring, alerting, and reporting, CDRP plans to offer a platform to the Florida Division of Emergency Management (FDEM) in the fall of 2013. This applied research component will include the knowledge gained from field research and CVI validation to allow operators to collaborate and create information and intelligence that will enhance the overall disaster situational awareness. Discussions with FDEM show a need for a platform with these capabilities as well as a desire to integrate quality information into situational awareness. (Koon, 2012) (Butgereit, 2012) This platform can be used by EOC's or VOST organizations in support of disaster operations.

The State of Florida's current situational awareness portal is based on ESRI GIS systems and is available to emergency managers via a web portal. According to Richard Butgereit, GIS Administrator for the agency, the portal can consume data from a variety of sources using standard protocols. (2012) Using this capability, the CDRP platform will be able to transmit CVI directly to the production situational awareness platforms already in use throughout Florida.

These research activities will address the major barriers of adoption of community volunteered information by increasing emergency manager trust in the information provided as well as simplifying the creation and integration of quality information.

REFERENCES

1. Brown, K. a. (2012, August 28). Hurricane Isaac Advisories. Retrieved December 30, 2012, from National Weather Service, National Hurricane Center: http://www.nhc.noaa.gov/archive/2012/al09/al092012.public_b.031.shtml
2. Butgereit, R. (2012, November 16). GIS Administrator, Florida Division of Emergency Management. (D. Merrick, Interviewer) Tallahassee, Florida.
3. Crisis Tracker Website, http://ufn.virtues.fi/~jakob/twitter/about.php, Accessed 15/1/2013
4. Koon, B. (2012, August 25). Director of the Florida Division of Emergency Management. (D. Merrick, Interviewer) Tallahassee, Florida.
5. Palen, P., Vieweg, S., Liu, S.B. & Hughes, A.L., (2009) Crisis in a Networked World. Features of Computer-Mediated Communication in the April 16, 2007 Virginia Tech Event, Social Science Computer Review OnlineFirst, April 8, 2009, Sage Publications
6. Rogstadius, J., Teixeira, C., Vukovic, M., and Kostakos, V., Karapanos, E., Laredo, J. CrisisTracker: Crowdsourced Social Media Curation for Disaster Awareness, IBM Journal of Research and Development, *in press*.
7. Rogstadius, J., Kostakos, V., Laredo, J. and Vukovic, M. A real-time social media aggregation tool: Reflections from five large-scale events. ECSCW 2011 *CSCWSmart? Collective Intelligence and CSCW in Crisis Situations*, Aarhus, Denmark (2011).
8. Schneiderman, B. & Preece, J., (2007), 911.gov, Science, Vol 315, pg 994, AAAS
9. St. Denis, L., Hughes, A., Palen, L. (2012). Trial by Fire: The Deployment of Trusted Digital Volunteers in the 2011 Shadow Lake Fire, Proceedings of the 9th International ISCRAM Conference – Vancouver, Canada April 2012, L.Rothkrantz, J. Ristvej and Z. Franco, eds
10. Starbird, K. & Stamberger, J. (2010) Tweak the Tweet: Leveraging Microblogging Proliferation with a Prescriptive Syntax to Support Citizen Reporting, Proceedings of the 7th International ISCRAM Conference-Seattle, USA, May 2010
11. Starbird, K. and Palen, L., (2011) Voluntweeters:" Self-Organizing by Digital Volunteers in Times of Crisis. ACM 2011 Conference on Computer Human Interaction (CHI 2011), Vancouver, BC, Canada.

12. Starbird, K., Palen, L. & Muzny, G. (2012) Learning from the Crowd: Collaborative Filtering Techniques for Identifying On-the-Ground Twitterers during Mass Disruptions, Proceedings of the 9th International ISCRAM Conference – Vancouver, Canada April 2012, L.Rothkrantz, J. Ristvej and Z. Franco, eds

13. Surowiecki, James (2004). The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations, Little, Brown