Fusing Non-Authoritative Data to Improve Situational Awareness in Emergencies

Christopher E. Oxendine
Department of Geography & Environmental Engineering, United States Military Academy
christopher.oxendine@usma.edu

Emily Schnebele
Department of Geography & GeoInformation Science, George Mason University
eschnebe@gmu.edu

Guido Cervone
Department of Geography, The Penn State University
cervone@psu.edu

Nigel Waters
Department of Geography & GeoInformation Science, George Mason University
nwaters@gmu.edu

ABSTRACT

In order to coordinate emergency operations and evacuations, it is vital to accurately assess damage to people, property, and the environment. For decades remote sensing has been used to observe the Earth from air, space and ground based sensors. These sensors collect massive amounts of dynamic and geographically distributed spatiotemporal data every day. However, despite the immense quantity of data available, gaps are often present due to the specific limitations of the sensors or their carrier platforms. This article illustrates how non-authoritative data such as social media, news, tweets, and mobile phone data can be used to fill in these gaps. Two case studies are presented which employ non-authoritative data to fill in the gaps for improved situational awareness during damage assessments and emergency evacuations.

Keywords

Evacuation, emergency, non-authoritative data, damage assessment.

INTRODUCTION

Natural, anthropogenic, and technological hazards pose a constant threat to the development and sustainment of communities and nations. A catastrophic event can claim thousands of lives, cause billions of dollars in damage, trigger an economic depression that might directly or indirectly affect the entire world, cause tsunamis, floods, landslides, or render a region uninhabitable. Such potential catastrophic consequences are due to the emergence of megacities, the proliferation of nuclear power plants and nuclear waste storage facilities, high dams, and other facilities whose destruction poses an unacceptable risk with a global reach. Thus, the study of natural hazards and of the processes that govern their occurrence has become a fundamental challenge for the survival of our civilization (Casti, 2012).

In recent years, advances in our ability to observe the Earth and its environment through the use of air, space and ground based sensors has led to the collection of large, dynamic, and geographically distributed spatiotemporal data. During emergencies, remote sensing data from air- and space-borne platforms have become the de-facto standard for providing high-resolution information for the assessment, relief, and mitigation of the damaged areas (Jensen and Cowen, 1999). However, due to limitations in orbital revisit time, sensor characteristics, and presence of clouds, gaps in the data might occur. Real-time data collection and analysis may be limited to non-existent immediately following an emergency (Cutter, 2003). Non-authoritative data, social media in particular, can provide large, rapidly changing dynamic datasets that not only complement remote sensing observations, but also add an additional subjective view of how people perceive and react to hazards.

This article presents two applications illustrating the use of non-authoritative data to fill in the gaps and further enhance situational awareness. Non-authoritative sources include data volunteered by citizens, for example volunteered geographical information (VGI). This general class of data may contain pictures, videos, sounds, text messages, etc. which include spatial and temporal information (Goodchild, 2007). Non-authoritative data may consist of data that has been collected for purposes other than disaster assessment, including, for example, data from traffic cameras or mobile phone locations. Due to the spread of the internet to mobile devices, an unprecedented and extensive amount of data has become available, often geolocated and often in real-time. Mining this ‘big data’, it is possible to reconstruct a spatiotemporal human terrain for improved knowledge when remote sensing data are unavailable. Additionally, non-authoritative data may provide unique knowledge that is not possible to acquire from remote sensing instruments, thus improving situational awareness during disasters and emergencies.

This article discusses the fusion of remote sensing and non-authoritative sources during an emergency to assess transportation infrastructure. It is important to understand road status, so that efficient evacuations can be implemented during emergencies. We discuss two specific applications:

1. Assessment of transportation infrastructure during and after Hurricane Sandy in New York City (NYC) using crowd sourced, remote sensing imagery and social media data.

2. Identification of evacuation routes during emergencies in NYC using traffic information and mobile phone data.

TRANSPORTATION INFRASTRUCTURE ASSESSMENT

Non-authoritative data provide valuable, real-time, on the ground information during disasters when traditional sources are unavailable, lacking, or slow to respond. In addition, the utilization of data from multiple sources can help provide a more complete description of phenomena. For example, data fusion is often employed with remote sensing data to combine information of varying spatial, temporal, and spectral resolutions as well as to reduce uncertainties associated from using a single source (Zhang, 2010). The fused data provides new or better information than would be available from a single source (Pohl and Van Genderen 1998). Furthermore, redundancies in observations can increase the confidence in observations or estimates, while data from multiple sources can provide information when they might not do so if used in isolation. The fusion of multiple data sources can improve damage assessments and provide an increased understanding of the sequence of events that may lead to transportation infrastructure failure.

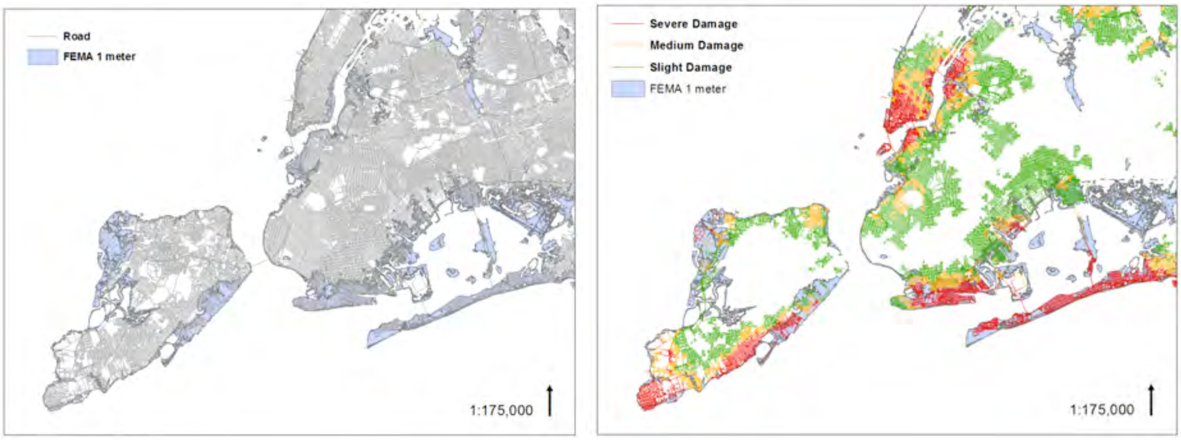
Using Hurricane Sandy as a case study, our first example illustrates how fusing non-authoritative data provides improved knowledge of flooding and road damage. Two sources of data are utilized to create a damage assessment for NYC following Hurricane Sandy. The first source was collected by the Civil Air Patrol which collected thousands of aerial photos following the impact of Hurricane Sandy. The photos were placed on a Google Crisis Map website for the public to assess visible damage through a crowdsourcing portal supported by MapMill (Google, 2012; MapMill, 2012). This yielded a damage assessment, a large data set generated from crowdsourced, non-authoritative, non-traditional sources. The crowdsourced photos (dated October 31-November 11, 2012) for the area between 33N to 26N latitude and 90W to 84W longitude were downloaded from MapMill. Because of the large number of photos and the scale of the study area, the photos were aggregated into a 500m grid. The value for each grid point is a function of the number of images present in each grid and their average crowdsourced damage assessment. As a result, each grid has a value from 1 to 10, with 1 representing no damage and 10 severe damage/flooding.

A second dataset, VGI available from YouTube videos, was collected from the Hurricane Sandy Google Earth website where videos supplied by Storyful (2012) could be accessed. These videos, which were voluntarily contributed by the public and contained geolocation information, documented flooding in New York City during and after Hurricane Sandy. The small number of videos (n=15) did not require any crowdsourcing or automated assessment and were visually assessed by the author. Furthermore, it is shown that even a small number of properly located VGI can help improve flood assessment (Schnebele and Cervone, 2013). Each location corresponding to a video point was assigned a value of 10 (severe damage/flooding).

The authors fused the data sets together using a kriging interpolation, resulting in a damage assessment surface generated solely from non-authoritative data. Kriging allows for spatial correlation between values (i.e.

locations/severity of flooding) to be considered and is often used with Earth science data (Olea and Olea, 1999; Oliver and Webster, 1990; Waters, 2008). Ordinary kriging generated a strong interpolation model. Cross-validation statistics yielded a standardized mean prediction error of 0.0008 and a standardized root-mean-squared prediction error of 0.9967 indicating that the interpolated surface well represents the data points A high-resolution road network was then layered over the damage assessment surface.

Figure 1 (a) is an example of the authoritative flood extent (1 m horizontal resolution) estimated by FEMA from field verified high water marks and storm surge sensor data (FEMA, 2012). The road network is illustrated by grey lines while the blue areas indicate flooded areas. This binary assessment is compared to the road damage map created from the fusion of non-authoritative data (Figure 1 (b)). By overlaying a high-resolution road network layer on top of the damage assessment surface, severely damaged roads can be identified at the street level. This allows authorities to prioritize site inspections, task additional aerial data collection, or identify routes that may be compromised. The application of non-authoritative data enhances the binary assessment by providing a range of damage values.



(a) FEMA flood extent **(b) Road damage classification**

Figure 1. FEMA flood extent map (a) and road damage classification (b) generated from the fusion of non-authoritative data.

By using multiple data sources, flood or damage details not captured by one source can be provided by another. For example, the geolocated videos enhanced the assessment by identifying flooding not conveyed in the Civil Air Patrol photos. Reasons for this disparity may include flooding captured on video had receded before the Civil Air Patrol flights or were captured at night, or flooding may have occurred in areas which were not in a flight path or were unable to be seen from aerial platforms (i.e. flooding in tunnels, under overpasses).

Sources of error in non-authoritative data, such as incorrect information (false positive/negative) or improper geolocation need to be considered. Incorrect information can be mitigated by including visually verified photos/videos and the application of multiple sources. Crowdsourcing, in particular, can increase accuracy and enhance information reliability compared to single source observations (Giles, 2005). Geolocation errors can be reduced with automation. In addition, sparse data or data skewed in favor of densely populated or landmark areas makes the use of non-authoritative data sources especially challenging. Increasing data volume and integrating authoritative data can yield increased confidence and include underrepresented areas. Although non-authoritative data can provide timely, local information often in large volume, they are often viewed with uncertainty. Conversely, the verification and authentication of authoritative data yield trusted results at the cost of time.

INCREASED SITUATIONAL AWARENESS DURING EVACUATIONS

Fusing non-authoritative and authoritative information provides increased situational awareness for crisis response personnel. This information enables them to better understand where incidents are occurring, how many people are at risk, and where to allocate resources. The ability to reduce response time is critical to minimizing loss of life and limb. Using mobile phone data, our second case study presents another application for using non-authoritative data to enhance situational awareness of emergency personnel.

Emergency Evacuations

No-notice emergency evacuations occur with limited or no planning. The common operating picture, enabled by Geographic Information Systems (GIS), provides a platform that the emergency operations center (EOC) uses to share collected data and information with first responders and evacuees. Planners analyze collected information to determine incident location, where to allocate resources, and how to evacuate citizens.

When an evacuation order is issued, there are many questions that are spatial in nature: What impact has the incident had on transportation? What non-profit organizations are in the region and can they provide support? How many people can we safely evacuate? Understanding where people are located during an evacuation is a major challenge. The location of individuals is constantly changing and unpredictable. The allocation of resources is usually based on information that has been collected about the situation and the situational awareness of emergency planners. As shown in Figure 2, near real-time mobile location data can be used to increase the situational awareness of emergency personnel and to redirect resources as needed within their jurisdiction.

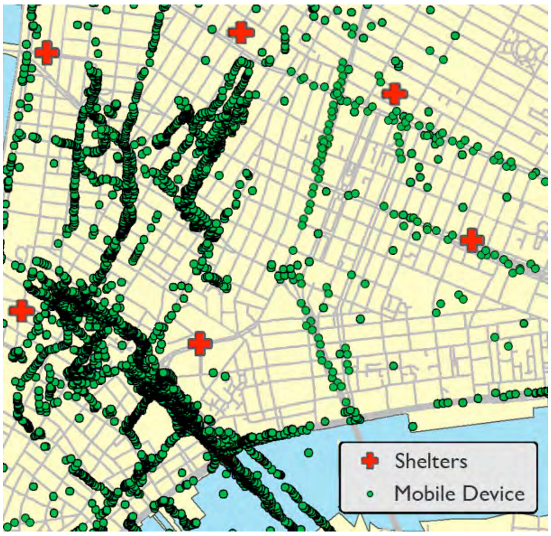


Figure 2. Notice the increased densities of mobiles phones approaching the Brooklyn Bridge in NYC.

Mobile Phone Data

Mobile phone companies collect and store information from individual mobile phones. After sanitizing personally identifiable information, mobile phone data are often aggregated for additional analysis, and sold for use in other applications, such as navigation, traffic services, or business development. During emergencies, citizens use their mobile phone to share information with family and friends; however they are suspicious of the government collection of this data.

Although the collection of mobile data during emergencies increases situational awareness, the government needs to evaluate the cost and benefit of using the data. The commercial sector has implemented methods to use mobile phone data to support business development while alleviating the concerns of citizens. Lessons-learned from the commercial sector may prove beneficial during the development of government policies for using mobile phone data during emergencies. While citizens may be concerned with privacy they are generally more concerned with efficient emergency response (Gruteser and Liu, 2004; Kar, B. et al., 2013).

Application of Mobile Phone Data in Evacuations

The fusion of mobile phone data with existing data introduces a method for increasing situational awareness during emergencies. Emergency response personnel often have challenges determining where people are located, even more so when there is a loss of power. Previously collected mobile data that account for diurnal population change and population change during major events, such as sporting events and festivals, provide response personnel with prediction models to use in planning.

In developing countries, mobile phone saturation rates are increasing. Mobile phone data provide information on population movements within a country enabling response organizations to prioritize their response efforts (Bengtsson, L. et al. 2011). These tools provide citizens and emergency personnel with increased information and damage estimates for their response areas.

Finally, mobile phones provide an opportunity for evacuees to provide data to emergency personnel through text messaging, social media, or phone calls. Although there may be a loss of power, mobile phones usually work on battery power for several hours after an incident. This provides an alternate means of communication and enhances situational awareness when other communication methods are limited or non-existent.

CONCLUSION

This article illustrates how non-authoritative data can fill in the gaps in traditional data sources by providing valuable information during emergencies. Non-authoritative sources offer an additional layer of information that not only tell us the severity of a potentially dangerous event, but also how citizens react to the developing danger and cope with damage. By presenting two examples, a road damage assessment for NYC following Hurricane Sandy and improved evacuation using cell phone data, the breadth and scope of possible applications of non-authoritative data for improved situational awareness are evident.

REFERENCES

1. Bengtsson, L. Lu, X., Thorson, A. Garfield, A. (2011) "Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti." *PLoS medicine* 8.8.
2. Casti, J.L. (2012) X-Events. HarperCollins New York.
3. Cutter, S. L. (2003) GI Science, Disasters, and Emergency Management. *Transactions in GIS*, 7: 439–446.
4. FEMA (2012) FEMA GeoPlatform <<http://fema.maps.arcgis.com>>
5. Giles, J. (2005) Internet Encyclopedias Go Head to Head. *Nature*, 438, 7070, 900–901.
6. Google (2012) Hurricane Sandy Google Crisis Map << http://google.org/crisismap/sandy-2012>>
7. Goodchild, M. (2007) Citizens as sensors: The world of volunteered geography. *GeoJournal,* 69, 4, 211–221.
8. Gruteser, M., Liu, X. (2004) Protecting privacy in continuous location-tracking applications. *IEEE Security and Privacy*, 2(2):28–34.
9. Jensen, J.R., Cowen, D.C. (1999) Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing,* 65, 611–622.
10. Kar, B., Crowsey, R. C., and Zale, J. J. (2013). The myth of location privacy in the United States: Surveyed attitude versus current practices. *The Professional Geographer*, 65(1):47–64.
11. MapMill (2012) MapMill: Crowdsourced Map Processing <<http://mapmill.org>>
12. Olea, R.A. (1999) Geostatistics for Engineers and Earth Scientists. Kluwer Academic, Boston.
13. Oliver, M.A., Webster, R. (1990) Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System,* 4, 3, 313–332.
14. Pohl, C., Van Genderen, J. (1998) Review article multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing* 19(5):823–854.
15. Schnebele, E., Cervone, G. (2013) Improving remote sensing flood assessment using volunteered geographical data. *Natural Hazards and Earth System Science,* 13, 669–677.
16. Storyful (2012) The World's First Social News Agency <<https://www.storyful.com>>
17. Waters, N. (2008) Representing surfaces in the natural environment: Implications for research and geographical education. *Representing, it Modeling and Visualizing the Natural Environment: Innovations in GIS* 13.
18. Zhang, J. (2010) Multi-source remote sensing data fusion: Status and trends. *International Journal of Image and Data Fusion,* 1, 1, 5–24.