Editor: **Daniel Zeng,** University of Arizona, zeng@email.arizona.edu

Harnessing the Crowdsourcing Power of Social Media for Disaster Relief

Huiji Gao and Geoffrey Barbier, *Arizona State University*
Rebecca Goolsby, *US Office of Naval Research*

Social media has recently played a critical role in natural disasters as an information propagator that can be leveraged for disaster relief. After the catastrophic Haiti earthquake on 12 January 2010, people published numerous texts and photos about their personal experiences during the earthquake via social media sites such as Twitter, Flickr, Facebook, and blogs, and videos were posted on YouTube. In just 48 hours, the Red Cross received US$8 million in donations directly from texts,[1] which exemplifies one benefit of the powerful propagation capability of social media sites.

Survivors also use social media sites to keep in touch with the world after a disaster. The jammed mobile phone network in Japan caused by the recent tsunami and earthquake made it difficult for people to communicate with each other. In response, they used Twitter, Facebook, Skype, and local Japanese social networks to communicate and keep in touch with their loved ones.[2]

Although social media can positively impact disaster relief efforts, it does not provide an inherent coordination capability for easily coordinating and sharing information, resources, and plans among disparate relief organizations. Nevertheless, crowdsourcing applications based on social media applications such as Twitter and Ushahidi offer a powerful capability for collecting information from disaster scenes and visualizing data for relief decision making. This article briefly describes the advantages and disadvantages of crowdsourcing applications applied to disaster relief coordination. It also discusses several challenges that must be addressed to make crowdsourcing a useful tool that can effectively facilitate the relief progress in coordination, accuracy, and security.

Disaster Relief Systems and Crowdsourcing

Coordination is a central, challenging issue in agent-oriented distributed systems.[3] Disaster relief systems focus primarily on designing coordination protocols and mechanisms to manage government and non-governmental organization (NGO) activities. Research has shown that it is possible to leverage social media to generate community crisis maps and introduce an interagency map (see Figure 1) to allow organizations to share information as well as collaborate, plan, and execute shared missions.[4] The interagency map is designed to allow organizations to share information if they operate on the same platform or use similar data-representation formats.

Recently, the efficiency at which government and NGOs are able to respond to a crisis and provide relief to victims has gained increased attention. Researchers are exploring and developing applications to leverage crowdsourcing for disaster relief.[5] Crowdsourcing allows capable crowds to participate in various tasks, from simply "validating" a piece of information or photograph as worthwhile to complicated editing and management, such as those found in virtual communities that provide information—from Wikipedia to Digg. This is a form of collective wisdom[6] information sharing that strongly leverages participatory social media services and tools.

1541-1672/11/$26.00 © 2011 IEEE
Published by the IEEE Computer Society

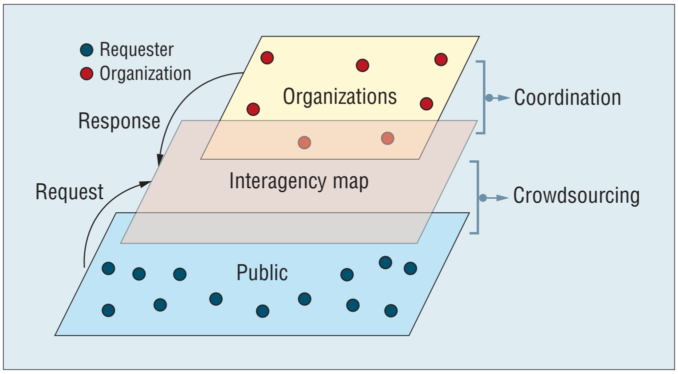

Figure 1. Interagency map. The map works as an intermediary between the public and relief organizations. Requests are collected via social media crowdsourcing. Organizations can then take actions, share information, and coordinate with each other using the information on the map.

The rapid growth of crowdsourcing applications for disaster relief profits from the development of online social media that provides an open, convenient platform that can collect data from various sources in a short time.

As an example of an attempt to adopt crowdsourcing to help the relief community to enhance cooperation, Ushahidi (www.ushahidi.com) is an open source crisis map platform created in 2007 and deployed in locations such as Kenya,[7] Mexico, Afghanistan, and Haiti. It leverages Web 2.0 technologies to integrate data from multiple sources—phones, Web applications, email, and social media sites such as Twitter and Facebook—to provide an up-to-date, publicly available crisis map that is in turn available to relief organizations. This platform uses crowdsourcing for social activism and public accountability to collectively contribute information, visualize incidents, and enable cooperation among various organizations.

Advantages of Crowdsourcing for Disaster Relief

Compared to traditional relief methods, leveraging crowdsourcing for disaster relief has three advantages. First, crowdsourced data including user requests and status reports are collected almost immediately after a disaster using social media. Ushahidi-Haiti was set up two hours after the 12 January earthquake by volunteers from Tufts University in Medford, Massachusetts.[8] Soon after, organizations were able to borrow a short message service (SMS) short code phone number (Mission 4636) to send free SMS texts.[9] News of this free emergency number was spread through local and national radio stations.

As of 25 January, the Haiti crisis map had more than 2,500 incident reports, with more reports being added every day. The large amount of nearly real-time reports allows relief organizations to identify and respond to urgent cases in time.

Second, crowdsourcing tools can collect data from emails, forms, tweets, and other unstructured methods and then do rudimentary analysis and summaries, such as by creating tag clouds, trends, and other filters. These can help partition the data into bins (such as most-frequently requested resources) and requests into predetermined, most-urgent categories (such as medical help, food, shelter, or people trapped). Relief agencies can then concentrate on the issues and events that are most important to the relief effort.

Figure 2 illustrates the food requests on Ushahidi-Haiti, and Figure 3 shows the most affected locations during the Japanese tsunami based on the number of reports mapping on Ushahidi's crisis map. Using these maps, relief organizations can coordinate resource distribution and make better decisions based on their analysis of crowdsourced data. Fallback plans can be further developed for the top events or to cover the majority of events.

Third, providers can include geo-tag information for messages sent from some platforms (such as Twitter) and devices (including handheld smart phones). Such crowdsourced data can help relief organizations accurately locate specific requests for help. Furthermore, visualizing this type of data on a crisis map offers a common disaster view and helps organizations intuitively ascertain the current status.

Shortfalls of Crowdsourcing for Disaster Relief

Although crowdsourcing applications can provide accurate and timely information about a crisis, current crowdsourcing applications still fall short in supporting disaster relief efforts.[10] Most importantly, current applications do not provide a common mechanism specifically designed for collaboration and coordination between disparate relief organizations. For example, microblogs and crisis maps do not provide a mechanism for apportioning response resources, so multiple organizations might respond to an individual request at the same time.

A second shortcoming is that data from crowdsourcing applications, while useful, do not always provide all the right information needed for disaster relief efforts. The accuracy of the report's geo-tag and content is not guaranteed, although relief workers greatly need the ability to automatically and accurately locate crowdsourced data on the crisis map. That is, a geo-located tweet does not necessarily refer to the

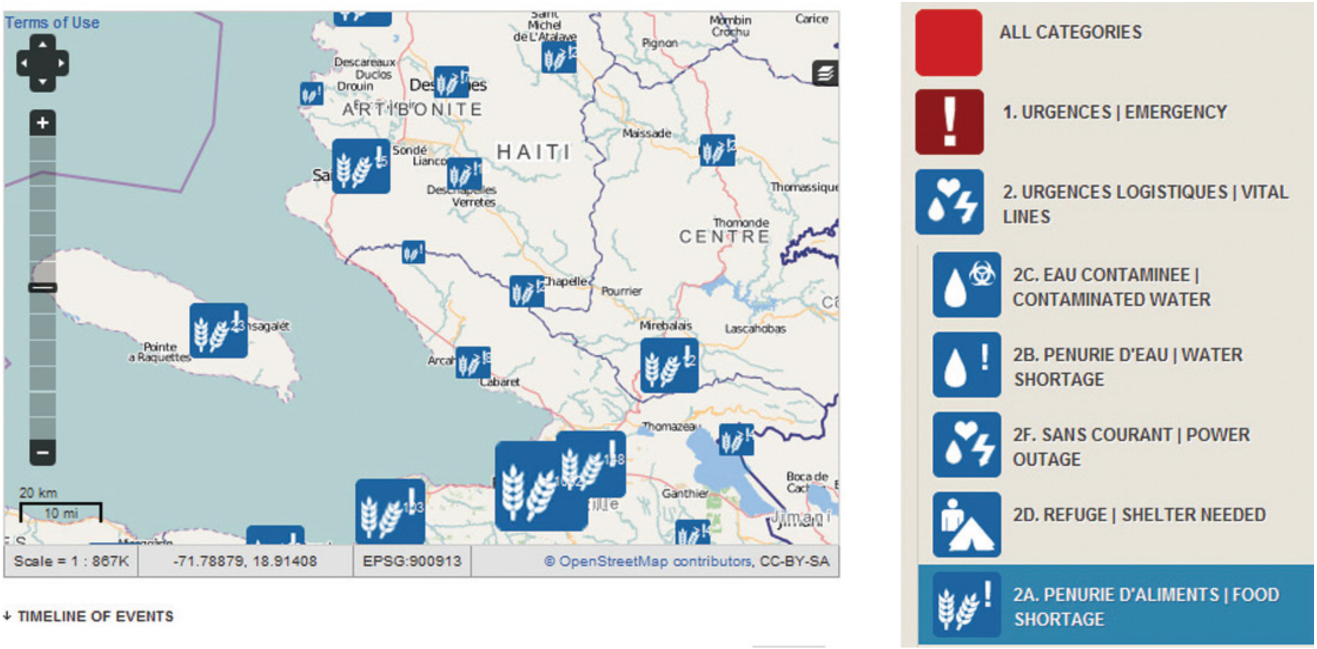

Figure 2. Food requests after the Haiti earthquake. The Ushahisi-Haiti crisis map helps organizations intuitively ascertain where supplies are most needed.

geo-location point. Someone might text the crisis map's phone number to report something they saw earlier, possibly texting from a shelter about a bridge that collapsed 10 miles down the road. Furthermore, fraud reports from malicious persons might appear as normal requests on crisis map. In addition, there are often duplicate reports, and information essential for relief coordination is not readily available or easily accessible, such as lists of relief resources or communication procedures and relief organization contact information.

Lastly, current crowdsourcing applications do not have adequate security features for relief organizations and relief operations. For example, crowdsourcing applications that are publicly available for reporting are also publicly available for viewing. Although it is important to provide information to the public, this can create conflicts when decisions must be made about where and when relief resources are needed. In some circumstances, relief workers themselves might be targeted by nefarious groups, so publicizing details about relief efforts could endanger them.

Challenges and Research Issues

Social media as a crowdsourcing mechanism provides aggregate situational awareness, important and new communications pathways, and some opportunities for assistance on an individual level. However, to make crowdsourcing a useful tool, we must address several challenges to leverage both the data and communications capabilities, including sensemaking, security, and coordination. To tackle these problems, we can employ text mining and social computing technologies by managing social knowledge and modeling social behavior during the disaster relief.[11]

Geo-tag Determination

As we mentioned earlier, crowdsourced data might include inaccurate geo-tags, or might not include them at all. Thus, using social mining to consider both a reporter's social and physical networks could help locate a report's actual spatial features.

In addition, research has shown the need to leverage the information available from the response group. For example, to supplement crowdsourcing information, groupsourcing[10] collects disaster information such as a relief situation summary, validated geo-tag information, and transportation conditions provided by a sanctioned group consisting of individuals with disparate resources, goals, and capabilities working at the disaster scene.

Report Verification

To some degree, crowdsourcing can automatically filter reports through photos, videos, and comments from other reports. Ushahidi, for example, lets users verify a report by clicking a verification button. It leaves the verification problem to the crowds to get collective feedback. However, this strategy—an open self-adjusting method, in technical terms—depends on a large number of people validating the reports. For Ushahidi maps, however, generally only a few people

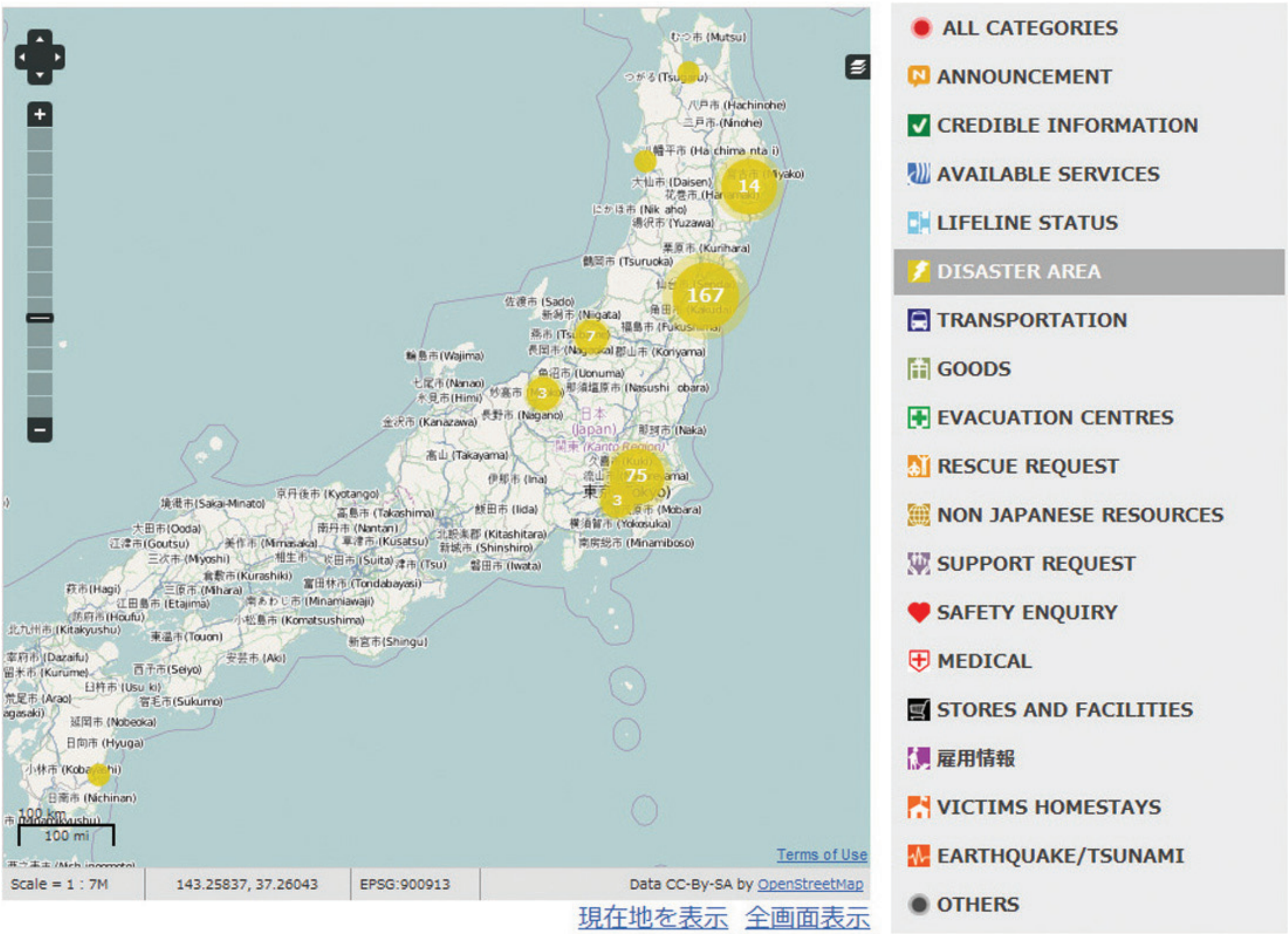

Figure 3. Ushahidi's Japanese tsunami map. By visualizing the most affected locations, the map helps relief organizations coordinate and prioritize their responses.

serve as verifiers, across the spectrum of Ushahidi map cases. A fraudulent message could conceivably propagate itself quickly with good camouflage, especially in a chaotic disaster environment, unless better incentives for verification are present.

As crisis maps and similar crowdsourcing applications are developed, it will be necessary to develop trust management systems that can recognize and report questionable postings and flag them for more scrutiny and verification.

Automated Report Summarization

Police departments and other first responders will need summarized request reports and the ability to drill down to individual tweets and texts. Because the volume and tempo of messages in a massive disaster will be so high, automated summaries and analysis to make sense of hundreds, if not thousands, of unstructured messages will be another greatly needed capability.

Spatial-Temporal Mining for Social Behavior Prediction

Currently, volunteers and victims on the ground report incidents and requests—that is, people who have a phone or other communication device. There are no applications for forecasting future requests or for filling in the likely requests that would come from blacked-out areas if such people experience limited communication ability.

Scientific data on earthquakes, floods, and other phenomena could augment user data to populate models to predict the time and location of future requests and needs. Crowdsourced data plus model results could help first responders, governmental agencies, and NGOs get ahead of the demand curve and become more proactive in deploying aid and rescue capabilities. One possibility is to design spatiotemporal classifiers for models that will assist in such demand forecasts.

Scalability and Safety

Scalability is a problem for many intelligent distributed systems because the increasing number of requests, agencies, and agent behaviors significantly enlarges the dimension of the search space.[3] Large-scale data-management technology will be necessary to maintain a system's stability.

Safety is another problem that must be considered in a disaster relief management plan, especially in the design of a crowdsourcing system. While making data publically available, such systems must protect the NGOs' privacy and ensure the safety of their workers.

Crowdsourcing integrated with crisis maps has been a powerful tool in humanitarian assistance and disaster relief. Future crowdsourcing applications must provide capabilities to better manage unstructured messages and enhance streaming data. Enhancements will also include methods of fusing information with data from other sources, such as geological data about fault lines and reports of various earthquake aftershocks. Researchers are just beginning to consider trust management systems to help evaluate messages with respect to the huge troves of messages being traded in cyberspace and via cell phones.

Furthermore, metrics that gauge the success of crowdsourcing and coordination systems for disaster relief will be designed and leveraged for system evaluation and improvement. Creating an automatic recommendation system that helps improve resource distribution coordination and transportation-route suggestions is another research direction that can utilize real-time data collection and relief history.

There is no doubt that new coordination platform and crowdsourcing applications such as crowdsourced translation, filtering, and categorization will continue to become available and will significantly contribute to future disaster relief efforts. We are confident that great numbers of lives will be saved as governments and NGOs find better ways to employ the integration of social media, crowdsourcing, and collaborative tools into disaster relief systems.

References

1. J. Morgan, "Twitter and Facebook Users Respond to Haiti Crisis," BBC News, 15 Jan. 2010; http://news.bbc.co.uk/2/hi/americas/8460791.stm.
2. Digital Blogger, "Japan Tsunami and Earthquake Use of Twitter, Facebook, Skype and Other Social Networks," 12 Mar. 2011; http://bubblecube.wordpress.com/2011/03/12/japan-tsunami-earthquake.
3. A. Bedrouni et al., *Distributed Intelligent Systems: A Coordination Perspective*, Springer, 2009.
4. R. Goolsby, "Social Media as Crisis Platform: The Future of Community Maps/Crisis Maps," *ACM Trans. Intelligent Systems Technology*, vol. 1, no. 1, 2010, article no. 7.
5. J. Howe, "The Rise of Crowdsourcing," *Wired*, vol. 14, no. 6, 2006, pp. 1–4.
6. J. Surowiecki, "The Wisdom of Crowds," *Am. J. Physics*, vol. 75, no. 2, 2007, p. 190.
7. P. Meier and K. Brodock, "Crisis Mapping Kenya's Election Violence," iRevolution blog, 23 Oct. 2008; http://irevolution.net/2008/10/23/mapping-kenyas-election-violence.
8. J. Heinzelman and C. Waters, *Crowdsourcing Crisis Information in Disaster-Affected Haiti*, US Inst. of Peace, 2010.
9. CrowdFlower, "Mission 4636," 2010; www.mission4636.org.
10. H. Gao et al., "Promoting Coordination for Disaster Relief: From Crowdsourcing to Coordination," *Proc. Int'l Conf. Social Computing, Behavioral-Cultural Modeling, and Prediction*, Springer-Verlag, 2011, pp. 197–204.
11. F.-Y. Wang et al., "Social Computing: From Social Informatics to Social Intelligence," *IEEE Intelligent Systems*, vol. 22, no. 2, 2007, pp. 79–83.

Huiji Gao is a PhD student in the Data Mining and Machine Learning (DMML) lab at Arizona State University. Contact him at Huiji.Gao@asu.edu.

Geoffrey Barbier is a doctoral candidate in computer science and a student in the Data Mining and Machine Learning (DMML) lab at Arizona State University. Contact him at gbarbier@asu.edu.

Rebecca Goolsby is an anthropologist for the US Office of Naval Research. Contact her at rebecca.goolsby@navy.mil.