Applications of Social Networks and Crowdsourcing for Disaster Management Improvement

Liliya I. Besaleva
Department of Computer Science
University of Virginia
Charlottesville, VA, 22901
besaleva@virginia.edu

Alfred C. Weaver
Department of Computer Science
University of Virginia
Charlottesville, VA, 22901
weaver@virginia.edu

Abstract— **Emergency resources are often insufficient to satisfy fully the demands for professional help and supplies after a public disaster. Furthermore, in a mass casualty situation, the emphasis shifts from ensuring the best possible outcome for each individual patient to ensuring the best possible outcome for the greatest number of patients. Historically, various manual and electronic medical triage systems have been used both under civil and military conditions to determine the order and priority of emergency treatment, transport, and best possible destination for the patients [12][13][15][16][17][18]. Unfortunately, none of those solutions has proven flexible, accurate, scalable or unobtrusive enough to meet the public's expectations [7]. In this paper, we provide insights into the trends, innovations, and challenges of contemporary crowdsourced e-Health and medical informatics applications in the context of emergency preparedness and response. Additionally, we demonstrate a system, called CrowdHelp, for real-time patient assessment which uses mobile electronic triaging accomplished via crowdsourced information. With the use of our system, emergency management professionals receive most of the information they need for preparing themselves to provide timely and accurate treatments of their patients even before dispatching a response team to the event.**

Keywords—emergency preparedness; triage; crowdsourcing; e-health; mobile technologies;

I. INTRODUCTION

Online social networking is undoubtedly one of the fastest growing phenomena seen thus far in the 21st century. Facebook announced that it serves one billion monthly active users in 2012 [14]. Consequently, social media websites are quickly becoming one of the largest identity and reputation management systems in the world where we are empowered to take on new roles in content creation, peer support and service delivery. And while the overall growth rate of content creation and active usage is starting to slow down as a part of the natural process of maturation of the web, one exception to the rule is smaller, more personalized social networks. They are often seamlessly incorporated as part of mobile applications and provide a highly customizable user experience with help from various sensors and information banks surrounding us in our everyday lives [9]. Consequently, the paradigm of information availability is shifting from primarily desktop devices to mobile and cyber physical technologies.

Simultaneously, another trend has arisen—the use of social networking and crowdsourcing for disaster management and improving public health [3][4][5]. Earthquakes, tsunamis, tornadoes, violent storms, and influenza epidemics have proven difficult to manage despite all of our technological advancements. When a disaster occurs, sometimes the best (or indeed only) way to save numerous lives is to generate a timely reaction to its outcomes. That puts extra emphasis on the importance of high-quality information management for emergency communications. Crisis management using crowdsourcing has been attempted by numerous organizations from all around the world and of all sizes [9] [12]. Following that trend, and taking into consideration the latest movements toward mobility in the technology world, we now present our design for a practical and highly capable application called CrowdHelp. CrowdHelp is a versatile application (accessible on smartphones, tablets and computers alike) which collects direct feedback from its users about their medical condition, in combination with data coming from sensors in smart devices (smartphones, tablets, laptops, etc.) and is to be used for enabling fast response of emergency professionals to devastating events. CrowdHelp is based squarely on the idea of social networks, and its functionality depends upon gathering the "wisdom of the crowd." It is an advance beyond currently existing technologies because of its centralized approach for data availability, fast data analysis, innovative combination of techniques (machine learning, cyber-physical systems, and crowdsourcing) and its focus on public safety operations. CrowdHelp was developed in collaboration with emergency response and medical specialists with experience from large natural disasters such as the 2010 Haiti earthquake [20].

II. APPLICATIONS OF WEB 2.0 TECHNOLOGIES

Vivid examples of the use of crowdsourcing can be seen in the fields of health and social welfare. Websites providing first-person experiences and advice make it easier for patients and survivors to open up about their conditions and ask for help. A popular one is the *patientslikeme* website [10] where over 150,000 people look for answers to unusual, embarrassing, or simply hard-to-talk-about health-related questions. It is a safe community where people feel protected behind their online

978-0-7695-5137-1/13 $26.00 © 2013 IEEE
DOI 10.1109/SocialCom.2013.38

personas, and no longer experience discomfort when revealing their real-life stories to others because it is a place of shared sympathy and understanding. Additionally, many social networks for survivors or friends and families of victims of disastrous events exist; they are used for connecting people needing that kind of support to cope with their grief or share their inspirational survival stories. The social importance of such communal social networks has been recognized by various authorities, including the Red Cross, and the Association for Computing Machinery (ACM) [8] [12].

In the last couple of years local authorities have also embraced social media to enable their staff to connect more effectively with the local communities. This is a direct way for them to engage with residents, community groups and partners using blogs, Twitter, Facebook, YouTube, Flickr and LinkedIn. These social networks allow local authorities to get involved in local conversations or collect "the wisdom of the crowd" about events occurring in their locality. These organizations are slowly increasing their level of recognition of social media as an invaluable source of information.

A. State-of-the-art Solutions

State-of-the-art solutions for disaster response are based in part upon data retrieval through crowdsourcing methods. Many crisis response systems turn to popular social networks (Twitter, Facebook, etc.) or specialized mobile applications (Red Cross Mobile Apps, Help Call etc.), to both locate and provide the latest news relating to a critical situation. Growing demand makes the development of successful and effective tools even more vital and urgent. An extremely well-known and widespread example of a successful e-Health system is *Ushahidi* [5]. This is an open source crisis map platform first created in 2007, and deployed in multiple locations, including Kenya, Mexico, Afghanistan, and Haiti. It leverages Web 2.0 technologies for multiple sources data integration (phones, e-mail, social media sites such as Twitter and Facebook) and provides an up-to-date publicly available crisis map that is in turn available to relief organizations. The platform uses crowdsourcing for information collecting, and supports cooperation enabling among various organizations.

How geeks responded to a catastrophic disaster of a high-tech country: rapid development of counter-disaster systems for the great east Japan earthquake of March 2011 [2] provides a different perspective on the use of social media for natural disaster relief. Instead of only exploring the occurrence or aftereffects of such events, the authors conducted a comprehensive survey of disaster management systems, in order to record how geeks in the high-tech Japan responded to a national crisis. It is an interesting work, as it focuses on the social aspects of disaster outcome alleviation rather than technology alone.

Similarly, *Harnessing the Crowdsourcing Power of Social Media for Disaster Relief* [7] talks about different variations of similar solutions for disaster recovery. In this article the authors express a concern about the general lack of well-organized and centralized sources of information. So, while gathering opinions and reading e-mails, tweets, Facebook posts, etc., could be fast, suitable for categorization according to urgency, and easily applicable to specific cases, it also has its shortcomings. One such shortfall is the non-existent centralized source system (a one-stop-shop where we find relevant information at the same spot without having to browse through different media websites/applications). Also, there is no way to assess the credibility of the data reported because the credibility of the sources is unknown.

B. Issues of Existing Systems

Sadly, most of the information coming from crowdsourcing sources cannot be considered entirely trustworthy. These publicly available information domains normally do not perform any data integrity checks. Therefore, the entries could be malicious, or sent for the purpose of creating confusion. That is a problem some have attempted to investigate and deal with but without affirmative success. An exemplary work is *Seeking the Trustworthy Tweet: Can Microblogged Data Fit the Information Needs of Disaster Response and Humanitarian Relief Organizations* [1] which discusses the untrustworthy nature of data coming from social media, as well as the different types of issues arising from it. As stated in that paper, currently no solutions exist but what looks promising as the most scalable and effective approach is the use of machine learning, entity extraction, and text classification techniques for data analysis.

Another side of the social networking approach is its decentralized nature. Post, tweets, and blog entries are spread throughout the entire Web. While the distributed way of collecting and presenting information could be considered good for data richness, it also means people often need to visit several different data centers to locate all the information they seek. Also, that information is usually not in any purified state (cleansed of "noise"), but instead comes in a variety of forms, each of which requires additional effort on the part of the user.

Additionally, we should note that real-world medical applications typically handle a plethora of privacy-sensitive data regarding their patients' diagnoses, symptoms and medical history. Because using Web 2.0 methods for collecting and disseminating data about victims in mass disasters is still a fairly new and developing concept, not much has been done for defining security specifications and requirements for the existing applications. Therefore, most of the existing solutions do not provide the same high standards of privacy protection as traditional methods.

III. CROWDHELP OVERVIEW

A. Approach

Preventing natural disasters is beyond our capabilities, but providing better information to disaster management professionals and affected persons (citizens) is not. Earthquakes, tsunamis, tornadoes, violent storms, and influenza epidemics have proven difficult to manage despite all of our technological advancements.

When a disaster occurs, sometimes the best (or indeed only) way to save numerous lives is to generate a timely reaction to its outcomes. That puts extra emphasis on the importance of high-quality information management for emergency communications. This is particularly visible in critical situations because internal and cluster communication within relief groups, non-governmental organizations, civil societies and other agencies is crucial to getting their work accomplished in a timely manner.

One of the most important steps to be taken during mass-casualty emergencies is to quickly and accurately triage (assess, sort and count) those who are in need of help. A standard medical triage scheme typically helps in the process of determining the priority of patients' treatments based on the severity of their condition (see Figure 1).

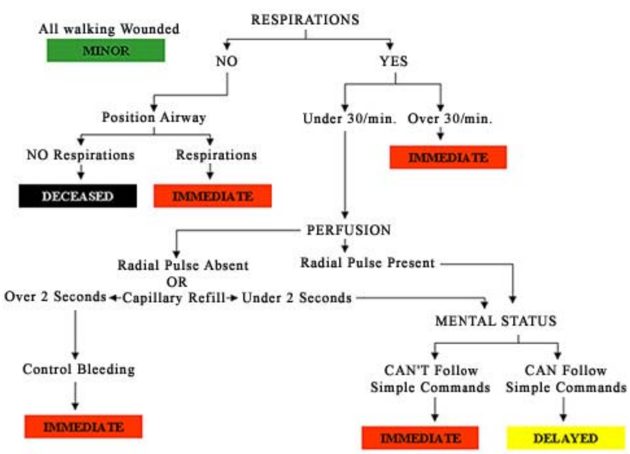

Fig. 1. Simplified schematic representation of medical triage system.

As a result of this triaging, numerous vital decisions about the best resource division and accommodation are made within the emergency response organizations. Traditionally, first aiders perform these critical tasks either by manually evaluating and color paper-tagging their patients on the spot or by using specialized electronic systems for remote monitoring. These solutions, however, have proven both labor intensive, slow, and prone to error when done by hand [6][11], or simply too obtrusive and unavailable for mass use when in the form of specialized systems.

B. System Design

As witnessed in the past few years, the newest trend among disaster response organizations is the use of social networks and crowdsourced information. Applications vary from simple manual monitoring of entries (tweets, online posts, comments, etc.) to automated versions of the same process, and intelligent systems with prediction capabilities and a ubiquitous nature. Unfortunately, none of those works have taken into consideration the obvious social issues that arise from interacting with actual human beings. Some of those problems are the lack of data integrity (often marred by ill-meaning users [15]), inaccuracy, and the need for output availability and coherency. Therefore, we have chosen to build our system using a combination of techniques for data collection, pre-processing and analysis.

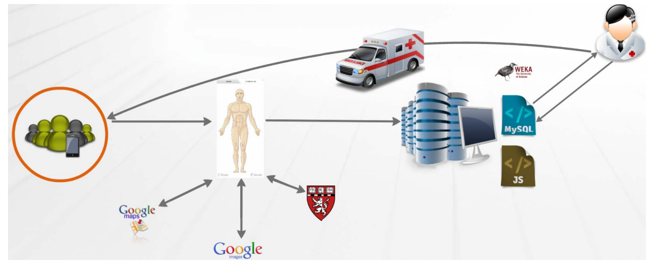

Fig.2. Schematic presentation of the workflow of the *CrowdHelp* platform

CrowdHelp is a software platform which combines two major components into its workflow (see Figure 2)—an application serving as a gateway to CrowdHelp's functionalities for ordinary users and victims, and a server layer providing the backend features of the system to be used by disaster management professionals (DMPs). From a user's point of view CrowdHelp is a downloadable application which is also accessible as a website online. CrowdHelp could be used equally well by numerous people reporting an event, as well as by a single individual who wants to determine what his/her physical condition might be and what are the possible symptoms and treatments of that condition. When used for emergency reporting our application which has several main features: it allows its users to submit information relevant to the event; provides the users with information about their possible conditions, as well as the possible causes for their symptoms; dynamically populates a list of places capable of treating the victims; and maintains a user profile with information about the user's location and reported symptoms. The data submission is done through an intuitive user interface which uses a simplistic visual reporting technique (see Figure 3). Information about the possible symptoms and causes is generated based on the user's report of their injured body area(s) and gender. As a result, a number of related imagery and videos which describe the symptoms, as well as a list of emergency medical treatment centers (which could be sorted based on their features or location) is populated dynamically (see Figure 4).

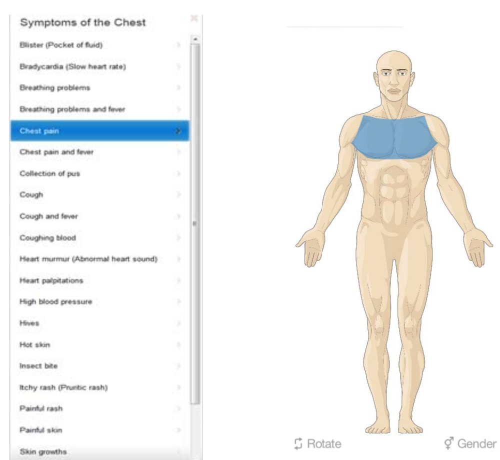

Fig. 3. CrowdHelp's interface with body part selection indicating injured area and the most likely symptoms

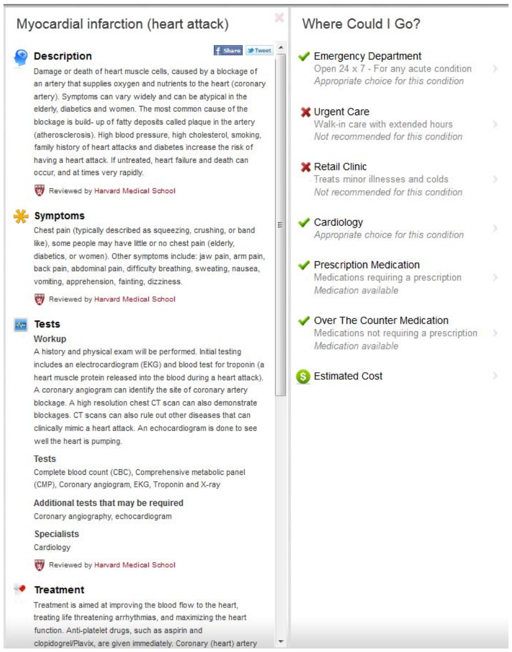

Fig. 4. Symptoms and medical centers lists

From the point of view of emergency professionals who have server access, CrowdHelp is a service for data analysis (see Figure 5) which works together with the machine learning software WEKA [19]. It provides a number of machine learning algorithms and computational formulas for clarifying, deciphering, summarizing and clustering all inputs into easily comprehensible visual images representing the geographical location, urgency and association of each entry to a specific cluster (see Figure 6). CrowdHelp provides its backend users with the ability to fully customize all testing, analytical and clusterization methods by choosing their preferred machine learning algorithms, visual characteristics of the representation and data attributes. The data that is collected through that application could be used by the responsible organizations both for immediate response and for future performance improvement analysis and event model building.

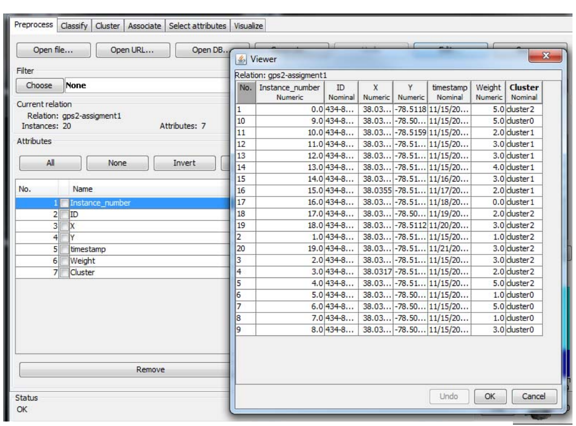

Fig. 5. CrowdHelp's backend

The server layer is designed with a user management module which provides its users with a hierarchical system of access rights, and an easy to operate interface for performing tasks such as data manipulation and analysis configuration customization (selecting the clusterization parameters).

For security, as well as for performance optimization, the user management system comes with user roles—guests, operators, or administrator user by another administrator. A guest user can only view the server and cluster configurations and monitor the selected cluster. An operator can view the server and cluster configurations and perform basic administrative tasks. An administrator can view the server and cluster configurations, perform administrative tasks (e.g., switch to operations, freeze operations, offline operations), modify configuration (e.g., create new groups, add resources, delete resources and or groups), perform advanced cluster manager functions. Initially everyone begins as a guest user with the option to be promoted to an operator or administrator user by another administrator.

In CrowdHelp's workflow the following steps take place:

(a) Download and install the *CrowdHelp* application onto a smartphone/tablet (or open in a web browser).

(b) Check the user's status (e-mail login required, optional social networks login is available). If this is an established user skip to step (d).

(c) Create a new user. Confirm the newly created registration with a confirmation email (not necessary when logged in with an existing social networking registration). Ask for permission to access the user's GPS coordinates, network location and time, compass and accelerometer/gyroscope readings.

(d) Present the user with a simple interface for symptoms selection, and then send the combined sensor readings and answers to the server.

(e) In the server: store the GPS coordinates in a spatial database. Analyze the data with machine learning algorithms for clusterization (e.g., XMeans, Density Based Clustering, Expectations-maximization).

(f) In the emergency management organization: Access the stored data, choose the preferred clusterization settings and manage the stored data by marking 'jobs' (entries) as open, pending, or finished. Read automatically generated suggestions which are based on the chosen answers in the questionnaire.

(g) In the user device: Receive warning push messages for forthcoming dangerous events (e.g., blocked roads, floods, storms).

What follows next is information distribution, which is an important step of the process of crisis management with its potentially vital social importance.

IV. TESTING

A. Testing Objective

CrowdHelp is a support tool for both disaster management professionals and citizens. Its two main contributions lie in providing a previously unavailable service to emergency professionals by assisting them in processing large amounts of data more efficiently and by providing them with helpful priority suggestions. This is done through a suite of techniques for accurate data clusterization and visualization for easier data comprehension.

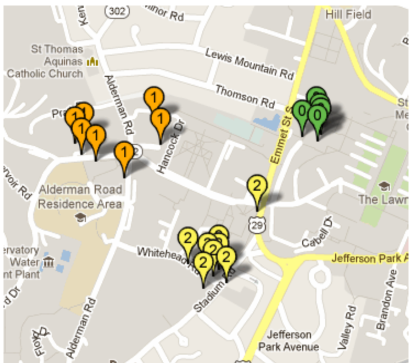

Fig. 6. Example cluster based on medical urgency

As this is a new service, not comparable to existing methods, our leading testing objectives focused on the technical feasibility of this project. Our service is designed to be highly flexible and inclusive in the types of information it provides. Clusterization of the data is done based on three different criteria – medical urgency, physical proximity to dangerous events (especially useful in the case of pollution), and geographical location with regards to neighboring entries (victims). All three clusterization maps could be visualized on separate system views (see Figures 6,7,8).

For the visualization we used the standard medical triaging conventions where all entries have levels of urgency ranging from 0-5 with zero being the lowest, therefore least urgent, and five being of the highest possible urgency. Each entry is given a number representing to which cluster is belongs, as well as a color representing the overall urgency of the cluster. The colors are selected as follows: green for urgency level of <1, yellow for urgency level of 1–2.5, orange for urgency level of >2.5 and <=3.5, red for urgency levels of >3.5.

B. Testing Setup

Priorities were assigned on two main levels – entry-wise and cluster-wise.

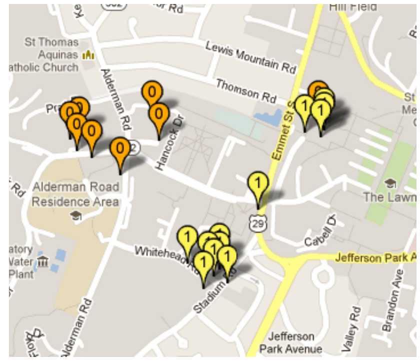

Fig. 7. Clustering based on proximity to neighboring entries

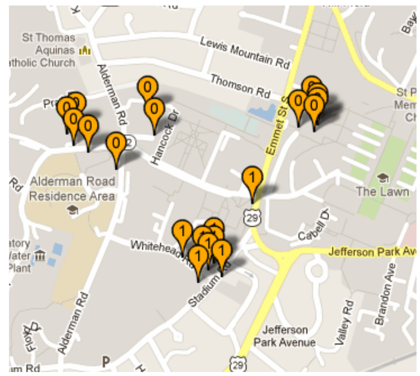

Fig. 8. Clustering based on proximity to dangerous events

a) Entries level

In the case of doing clusterization according to **medical urgency**, because *CrowdHelp* provides information about concrete medical symptoms (e.g., bleeding, numbness), we assigned priorities based on the standard medical triage system used in the United States (see Figure 1). Our testing schema consisted of the following. We designed 100+ test cases with different symptoms and manually assign each an urgency level based on the triage system. Then we randomly generated 5000 combinations of those test cases where each combination takes the highest urgency of its members. For instance, if we categorize a bleeding head trauma as an urgency level 4, and a minor leg injury as 1, a case with both a bleeding head trauma and a minor leg injury will be considered as urgency 5. After that, we customized an XMeans algorithm to use our specific data, and used the previously assigned urgency levels and color coordinated the entries according to the triage scheme.

Over 500 similar tests of different sizes of the test data were performed and all of them showed uniform results representing a neat visualization of the entries on a map. In the cases of entry overlaying some of the entries are automatically merged into a group shown through a single entry point on the map which contains meta-data about the elements of which it is made.

When doing clusterization based on **physical proximity to dangerous events**, we assigned priority levels based on the location of the victim and the anticipated size of the area to be impacted by the event. The nearest 15 % are assigned as high priority, following by medium priority for the next 35%, and green for the remainder. For instance, if we have an event which is expected to affect an area with diameter three miles (15840 feet) from its epicenter, all people within 2376 feet from it will be priority red, within 2377 – 7920 feet will be yellow, and the rest who are further away will be green.

To confirm the integrity of the resulting output from our system we performed over 350 tests and compared the actual physical location of the entries to a fixed point on the map and each time they were properly color prioritized according to our clustering rules.

Lastly, when we cluster based on **proximity to neighboring entries**, we do density based clustering which is used for finding non-linear shapes structures based on the entries' density**.** When doing density clustering we took into consideration two factors – density reachability and density connectivity.

For reachability we assumed the following rule: *a point **p** is said to be density reachable from a point **q** if point **p** is within distance **ε** from point **q** and **q** has a sufficient number of points in its neighbors which are within distance **ε**.*

For connectivity we used the following rule: *points **p** and **q** are said to be density connected if there exists a point **r** which has a sufficient number of points in its neighbors and both the points **p** and **q** are within the distance **ε**. This is a chaining process. So, if **q** is a neighbor of **r**, **r** is a neighbor of **s**, and **s** is a neighbor of **t** which in turn is a neighbor of **p**, this implies that **q** is a neighbor of **p***

b) Clusters level

All formed clusters are assigned overall priority levels which represent the median of all priorities of the entries in that cluster. This was a straightforward task which required going through the priorities of all entries and sorting them out to find the median. This is an additional feature of the system which is added for richer functionality and higher customization capabilities.

V. Conclusion

There is a plethora of articles, scientific works, reports and applications of various ubiquitous technologies for emergency response available at the moment. Naturally, by creating more accessible materials about a certain topic they only attract even further curiosity and discussions, and the popularity of the topic grows exponentially. This phenomenon is especially impressive in the field of emergency management through crowdsourcing and online media. With the addition of an application like CrowdHelp we are stepping into the next stage of crowdsourced information for crisis management. By needing virtually nothing but the smartphone in our pocket to access CrowdHelp, this lightweight application could be the solution to low-battery power, weak service signal or even circumstantial obstacles (e.g. excessive noise) to a cry for help from the heart of a disaster. Additionally, with its user-friendly design, strong crowdsourcing fundamentals and machine learning analytical powers, CrowdHelp could be considered more than a tool for collecting data but an important instrument for providing a faster and more accurate model of a situation's current state.

References

[1] A. H. Tapia, K. Bajpai, J. Jansen, J. Yen and L. Giles, "Seeking the Trustworthy Tweet: Can Microblogged Data Fit the Information Needs of Disaster Response and Humanitarian Relief Organizations", in Proceedings of the 8th International ISCRAM Conference – Lisbon, Portugal, May 2011.

[2] A. Utani, T. Mizumoto and T. Okumara, How geeks responded to a catastrophic disaster of a high-tech country: rapid development of counter-disaster systems for the great east Japan earthquake of March 2011. In proceedings of SWID '11, ACM, article No. 5.

[3] American Red Cross, "Superstorm Sandy: One month Update", http://www.redcross.org/images/MEDIA_CustomProductCatalog/m15840096_SandyUpdate-MonthOne.pdf

[4] B.W. Hesse, D.Hansen,T. Finholt, S. Munson, W. Kellogg, J.C. Thomas, "Social Participation in Health 2.0," *Computer* , vol.43, no.11, pp.45,52, Nov. 2010

[5] C. Fugate, "Understanding the Power of Social Media as a Communication Tool in the Aftermath of Disasters", http://www.dhs.gov/ynews/testimony/testimony_1304533264361.shtm

[6] G. Asaeda, MD. The Day That the START Triage System Came to a STOP: Observations from the World Trade Center Disaster. *Academic Emergency Medicine*, vol. 9, Issue 3, pp. 255-256, March 2002. Article first published online: June 28, 2008. DOI: 10.1197/aemj.9.3.255.

[7] H. Gao, G. Barbier and R. Goolsby, Harnessing the Crowdsourcing Power of Social Media for Disaster Relief, Intelligent Systems, IEEE , vol.26, no.3, pp.10-14, May-June 2011.

[8] Klasnja, P., Hartzler, A., Powell, C., Phan, G., & Pratt, W. "HealthWeaver Mobile: Designing a Mobile Tool for Managing Personal Health Information during Cancer Care". ACM Conference on Human Factors in Computing Systems 2010.

[9] N. A. Christakis and J. H. Fowler, Connected: The Surprising Power of Our Social Networks and How they Shape Our Lives, published by Little, Brown and Company, Hachette Book Group, September 2009.

[10] Patients like Me, http://www.patientslikeme.com/ [last visit 8/9/2013].

[11] R. E. Weinerman, S. R. Rutzen, and D.A. Pearson, 1965. Effects of medical 'triage' in hospital emergency service. *Public Health Rep*., May 1965, 80(5), pp. 389-399.

[12] R. L. Briones, B. Kuch, B.F. Liu and Y. Jin, Keeping up with the digital age: How the American Red Cross uses social media to build relationships, Public Relations Review, Volume 37, Issue 1, March 2011, pp. 37-43.

[13] Regus, A global report from Regus, A global survey of business social networking. [Online]. Available: http://www.regus.com/images/Social%20Success%5B1%5D_tcm8-21559.pdf

[14] S. Haley and D. Hardawar, Facebook hits the big one: 1 billion monthly active users [STUDY]. [Online]. Available: http://venturebeat.com/2012/10/04/facebook-hits-1-billion-monthly-users/

[15] S. R. Hiltz, P. Diaz and G. Mark, "Introduction: Social media and collaborative systems for crisis management", ACM Trans. Comput.-Hum. Interact. 18, 4, Article 18 (December 2011), 6 pages.

[16] T. Gao, T. Massey, L. Selavo, D. Crawford, B. Chen, K. Lorincz, V. Shnayder, L. Hauenstein, F. Dabiri, J. Jeng, A. Chanmugam, D. White, M. Sarrafzadeh, and M. Welsh, The Advanced Health and Disaster Aid Network: A Light-Weight Wireless Medical System for Triage, *Biomedical Circuits and*

Systems, IEEE Transactions on , vol.1, no.3, pp.203-216, Sept. 2007, doi: 10.1109/TBCAS.2007.910901.

[17] T. Gao, and D. White, A next generation electronic triage to aid mass casualty emergency medical response. Proceedings of the 28th IEEE EMBS Annual International Conference, 2006.

[18] T. Massey, T. Gao, M. Welsh, J. H. Sharp, and M. Sarrafzadeh, The Design of a Decentralized Electronic Triage System. *AMIA Annual Symposium Proc*. 2006, pp. 544-548.

[19] Weka. [Online] Available at: http://www.cs.waikato.ac.nz/ml/weka/

[20] [37] Wikipedia, Haiti earthquake (2010) [Online]. Available: http://en.wikipedia.org/wiki/2010_Haiti_earthquake