Securing Crisis Maps in Conflict Zones

George Chamales
Rogue Genius LLC
Oakland, USA
george@roguegenius.com

Rob Baker
Washington D.C., USA
rob@rrbaker.com

***Abstract*—Crowdsourced crisis mapping technologies aggregate information from individuals, trusted networks of informants, and publicly available media sources for use in informing decision making and response. Deploying this technology in conflict zones introduces a new set of vulnerabilities that can be exploited by hostile actors who have developed or adopted network surveillance and attack capabilities. Successful infiltration or compromise of these deployments can have a significant, negative impact on the lives of those reporting to and operating crisis maps. This paper identifies several fundamental vulnerabilities present in crisis mapping deployments and defines a set of best practices to defend those vulnerabilities from successful attack. The recommendations are adapted from established principles in computer and network security, combined with the authors' experience operating crisis mapping deployments in Afghanistan, Sudan, and Libya.**

Keywords-component; crisis mapping; crowdsourcing; conflict zones; computer security;

I. INTRODUCTION

The traditional approach to crisis response is based on monetary and material support from individuals, groups, and governments that is given to established crisis response organizations which is delivered to the crisis affected population in the form of goods and services. That aid is distributed based on information about the needs, scope, and scale of the crisis which is collected by the response organizations. That information collection process typically uses a combination of observations from the response organizations and interactions from their personnel with members of the crisis-affected population.

The rise of information technology and communication infrastructures such as the Internet and cellular phone system significantly increases the capacity to share information. In large scale crises where the communication infrastructure remains operational, these technologies can result in large spikes in messages relating the status of groups and individuals, requests for support, and the existence of threats.

Pull-based observational and polling approaches to information collection by response organizations, designed to seek information, are not designed to tap into this spontaneous, push-based, situational information flow. As a result, response organizations are often ill equipped to handle situations where they are suddenly inundated with large quantities of unfiltered, unstructured information from individuals in the field.

A. Crowdsourced Crisis Mapping

Crowdsourced crisis mapping (CCM) seeks to utilize information from those inside a crisis (the crowd on the ground) by collecting that data onto a computer-based crisis mapping platform. Once on a crisis mapping platform, a variety of tasks can be performed on the data, including: filtering, translation, categorization, structuring, geolocation, analysis, export, and alerting.

This information processing is performed using a combination of automated tools and human workers who can be located inside the crisis area or, if the platform is connected to the Internet, anywhere in the world (the crowd in the cloud). Automated analysis can provide a first pass through the data, perform several of the processing steps and prioritize messages for processing by the human workers [1]. Human workers provide a scalable workforce that is capable of approving reports and providing training data for the automated analysis components. The information processing by human workers can be enhanced by utilizing workers with specialized skills including knowledge of the local language, geography, and community.

Once processed, messages are presented to crisis response organizations and other interested parties in the forms of maps, data feeds, and alerts. Response organizations can further leverage the CCM system by providing their own datasets for information processing, giving feedback on the quality of the data they are receiving, and by directing analysis towards the information they need.

The deployment of crisis mapping platforms has been aided in the recent years by the development and adoption of professional products such as GeoIQ's Geocommons [2] as well as open source platforms such as Ushahidi and its hosted solution, Crowdmap [3]. Crisis mapping platforms allow individuals and organizations to rapidly install, deploy, and manage crowdsourced crisis maps in response to sudden onset crises ranging from natural disasters, to the ongoing instability in the Middle East and Africa.

B. Crisis Maps in Conflict Zones

Reports of abuse of information technology by hostile actors in conflict zones abound, from filtering of Internet connections to complete shutdowns by government forces in Egypt, Syria, and Tunisia. Those reporting on abuses or speaking out against these forces have found themselves targeted for attack by the forces themselves or their proxies -

with consequences ranging from harassment to imprisonment and death.

Crowdsourced crisis mapping is a relatively new development in the field of information technology, but one that can pose a threat to those hostile organizations. This was certainly the case in Egypt where an election monitoring deployment of Ushahidi encountered regular harassment by members of the Egyptian Security Services. Among others, this harassment included the security services demanding their own administrative account on the Ushahidi platform. It has been reported that documents describe close surveillance of activists including the apparent compromise of communications sent over Skype, which the deployment coordinators relied on for confidential communications [4].

Adoption and integration of CCM technology by response organizations, such as the UN's role as the initial requester for and eventual operator of the LibyaCrisisMap.net deployment [5], open the possibility that successful compromise could lead to disclosure of information regarding that organization's operations or the manipulation of the response operations. CCM technology can offer response groups with an unparalleled opportunity for widespread situational awareness, but those deployments must be carried out with a solid understanding of the risks involved and appropriate steps taken to eliminate those vulnerabilities or avoid attacks against them.

C. *Caveats to the Following*

Crowdsourced crisis mapping is a rapidly evolving field with new lessons learned from each deployment and a growing sophistication of tools and techniques being employed. At the same time, new computer and network vulnerabilities are regularly being discovered and the corresponding attacks published and utilized around the world. The ongoing development of these fields makes it difficult to know what new technologies will be available in the next months and years and what types of new attacks may render existing defenses useless.

As a result, recommendations for specific security tools, products, and protocols are purposefully avoided since there is no guarantee that those recommendations will be reliable in the future. Similarly, recommendations for general secure application deployment such as server configuration, intrusion detection, and incident response (also the subject of ongoing updates and revisions) are beyond the scope of this paper. Those interested in deploying the systems are encouraged to seek out the current best practices for secure application deployment and networked communication.

The following vulnerabilities and best practices are based on the current state of crowdsourced crisis mapping technology and the threats, both physical and online, to individuals in a crisis and to crisis mapping platforms. The recommendations are based on fundamental security concepts as they apply to the specific use case of information technology in crisis mapping deployments for conflict zones. These findings and recommendations are not intended to be a universal list applicable to all deployments or to cover the full range of potential vulnerabilities and best practices. Instead they are designed to identify and address some of the most critical issues with this technology and serve as a starting point for future discussions.

II. VULNERABILITIES OF CROWDSOURCED CRISIS

CCM technology utilizes existing technologies such as web services and mobile communication, and applies it to a new domain with a unique set of deployment and security challenges. This approach combines the vulnerabilities of those technologies with vulnerabilities specific to the use cases where it is being deployed. Understanding these new vulnerabilities is a necessary first step towards developing best practices to eliminate those vulnerabilities and prevent successful attacks.

The following are five types of vulnerabilities introduced by crisis mapping technology used in conflict zones:

A. *Identification of Reporters and Vulnerable Groups*

Hostile organizations such as oppressive governments do not necessarily need a reason to target a specific individual or group, however individuals who report on the activities by these organizations can make themselves a target for attack and retribution. In a conflict, those reporters may be citizens communicating over social media, submitting text messages to a crisis mapping platform, or professional journalists. The information related by those reports can also be used to identify vulnerable groups such as refugees, those acting against the hostile powers, or response organizations - as was the case with the Taliban's threat to target foreign aid workers responding to the 2010 floods in Pakistan [6].

B. *Control of Communications Networks*

Oppressive governments with nationalized communications infrastructure or private infrastructure that can be influenced by hostile groups can allow those organizations to monitor communications, filter access, or selectively disable communications. Recent examples of Internet disruption in Syria and Egypt Highlight this capability as do ongoing struggles with the Taliban's ability to disable the cellular infrastructure in Afghanistan [7].

C. *Programming Flaws in Crisis Mapping Platforms*

The operating systems, servers (web, email, etc.), and crisis mapping platform applications may contain vulnerabilities that could be leveraged by hostile organizations to disrupt or compromise the deployment. The risk of these flaws is increased by the use of platforms that have not received thorough security reviews to identify and eliminate exploitable flaws and adequate documentation necessary to avoid insecure configurations.

D. *Identification and Infiltration of Crowdsourced Workforce*

Crowdsourced workers taking part from inside the conflict zone are susceptible to the same identification and targeting as reporters to the platform. In addition, the use of unvetted crowdsourced workers either inside the conflict zone or using workers on the Internet opens up the possibility that hostile groups could pose as helpful workers - allowing them access to

internal information and sensitive systems, and the ability to manipulate workflows.

E. Use of Unverified Reports

Gathering conclusive proof of an individual report's veracity can be difficult or impossible when the report is being gathered without direct, first-hand observation from a trusted party. The use of inaccurate or false reports (either intentional or unintentional) can result in manipulation of those consuming the reports such as response organizations and individuals inside and outside the crisis zone. This vulnerability is exploitable by organizations with established propaganda and misinformation capabilities.

III. Proposed Best Practices

The following best practices are derived from the vulnerabilities listed above. They are designed to eliminate or mitigate those vulnerabilities or, where that is not possible, provide opportunities to prevent successful attacks that target those vulnerabilities. Each recommendation is presented as a high-level concept followed by sample implementations and examples from real-world experiences.

A. Enumeration of Risks and Mitigation Procedures

No two crisis mapping deployments are the same. The different types of hostile actors in a conflict zone and their varying capabilities make it necessary to identify the risks involved in each crisis mapping deployment. The enumeration process would take into account the vulnerabilities listed above and the attack capabilities of hostile groups associated with the conflict. Once identified, those risks should drive the adoption of mitigation procedures to reduce or eliminate the vulnerabilities and prevent successful attacks on those vulnerabilities that are unavoidable.

The risks and mitigation procedures should then be shared in as direct and simple a manner possible to everyone involved in the reporting, operation, and usage of the platform. This can be challenging in situations with limited communication ability to those on the ground and the potential for that information to be misinterpreted, however this step is vital to allow others to make an informed decision about their level of involvement and methods of interacting with the deployment.

In a recent CCM deployment tracking protests in Sudan, the deployment coordinators posted a descriptive set of precautions on the front page of the site warning users that SMS messages could be monitored and that reports should not be submitted over unencrypted web connections. The warnings were posted in both English and Arabic and done after thoughtful analysis of the risks posed by monitoring of those communication networks. Despite the concise nature of the warnings rumors reached Twitter that the deployment had been compromised by the Sudanese security forces [8].

B. Need to Know

Knowledge is power, and the information collected in CCM deployments can be a powerful tool for both helpful and hostile actors. In situations where individuals are being targeted for attacks, information identifying specific individuals such as names, email addresses, phone numbers, and online usernames, should be made available only to those individuals and organizations with a specific need to know that information.

If personally identifiable information must be stored, that information can be redacted on public sites, deleted, or encrypted. One-way hashing algorithms can create unique identifiers that can be used to identify reports from the same individual without exposing their actual identity (although it is possible for attackers with a known identity and a detailed understanding of how the one-way hashing is implemented to recreate that identity's unique identifier). Reversible encryption processes such as public key cryptography can generate unique identifiers that can also be decrypted to view the original value. This process relies on the protection of the encryption key [9], the security of the encryption algorithm, and the strength of the encryption algorithm relative to the adversary's capability to perform brute force decryption attacks.

Ideally, this anonymization / encryption process should be performed as early as possible when received by the platform, reducing the opportunity for sensitive information to be disclosed by a compromise of systems that utilize that sensitive information. Access provided to organizations and individuals should be regularly reviewed to ensure that they are still necessary and, if technically possible, the systems should be monitored for indications that users are attempting to abuse their permitted access.

During the LibyaCrisisMap.net deployment, the description and analysis of each report was redacted, media links to specific twitter accounts deleted, and the information was posted publicly online after a 24 hour holding period. This process allowed information from the deployment, primarily public media links and social networking posts, to be filtered, categorized, and shared, while protecting the analysis provided by the crowdsourced workforce and from on-the-ground sources.

C. Isolation of Operations

Multiple, distinct operations including message collection, automated analysis, crowdsourced processing, presentation and analysis are performed in a CCM deployment. It is not necessary that these processes take place on the same server. Isolating each of the components used in the CCM system can minimize the impact from a successful compromise of one of the deployment's components.

Isolated systems must communicate between one another, and standardizing the communication between those systems can make it possible to detect irregularities that could indicate a component has been compromised or is being misused. This would be the case if a compromised system were used as a launching point to probe and attack other servers in the platform - those network connections could be identified as anomalous. Similarly, failures in an individual component could be detected if anticipated connections from that component are not received.

Distributing operations across multiple platforms - such as utilizing a third-party crowdsourcing service for message

processing as was done in Haiti [10] - can isolate hostile users who infiltrate the crowdsourced workers. That risk can be mitigated by requiring reports to be viewed by multiple members of the crowdsourced workforce, with the messages only being promoted once multiple workers agree on the messages content. This approach was taken during the Crowdflower microtasking system used in Pakreport.org to enable more accurate geolocation of messages [11].

Utilizing a public CCM server that presents publicly available data and a separate password-protected site containing sensitive reports can help ensure that sensitive information is only available to those individuals and organizations with an established need to know. This was done during the LibyaCrisisMap.net deployment - making it possible to reduce the impact of a successful compromise of the public site, and significantly reduce the attack surface of the private instance.

D. Multi-Factor Trust

Deployments that collect and relay information from disparate sources such as SMS messages, social media sites, mainstream media articles, and reports from response organizations are at least one step removed from the information. The second-hand nature of the messages being presented and the unpredictable nature of the situation on the ground can present a significant challenge for deployments seeking to provide proof of a given message's trustworthiness.

Establishing report veracity is complicated by the various levels of proof that the deployment's users may require: the verification needs of an organization tracking human rights abuses are different than those of organizations looking to identify vulnerable populations for aid delivery. The challenge is only increased by the potential for hostile actors to submit false, misleading, and manipulative reports to the system.

One approach to establishing trustworthiness that has emerged from the processing of reports from the field during CCM deployment is to divide the reasons for trusting a given reporter or message into the following factors:

- **Reporter History**: Consistent accuracy of a reporter's messages.
- **Reporter Reputation**: What other trusted sources say about the given reporter's trustworthiness.
- **Message Corroboration**: Identifying other reports and facts that lend credence to the information provided in a given message.

Determining the trustworthiness of a given report or reporter should utilize multiple factors to decrease the possibility inaccurate reports and reporters. This approach was clearly demonstrated in the recent case of a fabricated website purporting to provide the first-person experiences of a lesbian blogger in Syria. The blog, "A Gay Girl in Damascus", contained one hundred and forty five posts over the course of four months, many of which contained highly detailed accounts of life inside Syria. The site was proven to be a hoax when a journalist with a well-established reputation for accurate reporting from the Middle East received messages from a set of sources he trusted that certain facts on the site did not appear to be accurate [12].

E. Anticipate Disruptions

CCM deployments rely on the interaction of numerous components including crisis affected populations, communications infrastructures, crowdsourced workers, machine learning algorithms, software platforms, and response organizations.

There are many, many things that can go wrong, and in a rapidly evolving crisis situation, many things will. Issues encountered during past deployments include:

- Sudden restrictions on SMS message transmission prior to elections in Egypt hindered the collection of reports from the ground [13].
- During the LibyaCrisisMap.net deployment, enabling secure web communications with SSL caused the system to crash because that functionality had never been tested in the field.
- In the middle of the SudanVoteMonitor.com deployment an email glitch caused spam to the deployment's email account to be repeatedly downloaded, filling the hard drive and crashing the system.
- Processing of messages by local workers in Afghanistan was designed to take into account regular breaks for prayers and was periodically interrupted by power outages.
- Following the acceptance of a new set of volunteers to the LibyaCrisisMap.net deployment, an untrained volunteer began approving reports without following established sanitization and security procedures.

It is important that those deploying CCM systems be aware of the potential for disruption of service and plan accordingly to ensure that their operations are capable of continuing in reduced capacity and recovering from these unexpected events. As more and more CCM deployments are performed in the field it is expected that additional lessons learned will be published for discussion and mitigation techniques will be developed to avoid them in the future.

IV. CONCLUSION

Crowdsourced crisis mapping technology holds great promise for the tracking of ground-truth from disasters and conflict zones. As this technology grows it will likely continue to be utilized in situations where it may be seen as a threat to hostile actors. Identifying the vulnerabilities in the current state of CCM deployments is necessary to develop best practices and prevent successful exploitation of these systems.

This paper represents a first step in the definition of those vulnerabilities and best practices. The content and recommendations are intended as a starting point and it is expected that these recommendations will need to be adapted over time to fit the changing types of communication systems,

information collection platforms, processing methods, automated analysis technology and the sophistication of attacks by hostile groups. Doing so will enable those deploying CCM technology to this technology to support those impacted by crisis and those organizations seeking to provide aid.

REFERENCES

[1] R. Munro, "Subword and spatiotemporal models for identifying actionable information in Haitian Kreyol," in *Fifteenth Conference on Computational Natural Language Learning (CoNLL 2011)*, Portland, 2011 (in press).

[2] GeoCommons, available at http://geocommons.com.

[3] Ushahidi, available at http://ushahidi.com

[4] S. Stecklow, P. Sonne, M. Bradley, "Mideast Uses Western Tools to Battle the Skype Rebellion," in *The Wall Street Journal*, available at http://online.wsj.com/article/SB10001424052702304520804576345970862420038.html#ixzz1O46A44T9

[5] LibyaCrisisMap.net, available at http://libyacrisismap.net

[6] J. White, "Taliban issues threat to 'foreign horde' of aid workers in Pakistan helping flood relief effort", *The Daily Mail*, available at http://www.dailymail.co.uk/news/article-1306597/Taliban-threat-foreign-aid-workers-Pakistan.html

[7] P. Quinn, R. Faiez, "Taliban Turn Cell Phones Back On In Afghanistan", *Associated Press*, available at http://abcnews.go.com/Blotter/wireStory?id=13299404

[8] P. Meier, "Sudan, Crowdmap, Misinformation, and Repression," available at http://blog.ushahidi.com/index.php/2011/04/13/sudan-misinformation/

[9] Auguste Kerckhoffs, "La cryptographie militaire," Journal des sciences militaires, vol. IX, pp. 5–83, Jan. 1883, pp. 161–191, Feb. 1883.

[10] V. Hester, A. Shaw, and L. Biewald. 2010. "Scalable crisis relief: Crowdsourced SMS translation and categorization with Mission 4636". In *Proceedings of the First ACM Symposium on Computing for Development (ACM DEV '10*). ACM, New York, NY, USA, , Article 15 , 7 pages. DOI=10.1145/1926180.1926199 http://doi.acm.org/10.1145/1926180.1926199

[11] Pakreport.org, available at http://pakreport.org

[12] E. Addley, "Syrian lesbian blogger is revealed conclusively to be a married man", *Gualdian,* available at http://www.guardian.co.uk/world/2011/jun/13/syrian-lesbian-blogger-tom-macmaster

[13] Sherif Ashour, Mohamed Megahed and Omar el-Hadi, "SMS Messaging Restricted in Bid to Preempt Pre-Election Activism," Al-Masry al-Youm English, October 11, 2010, http://www.almasryalyoum.com/en/news/restrictions-placed-sms-messages-avert-promoting-anti-regime-incitations (accessed November 5, 2010).