RELIEF 11-03

Report on the RELIEF 11-03 Experiments at NPS/Camp Roberts

Humanitarian Technologies for
Domestic and International HADR Operations

30 May 2011

Mr. John Crowley (contractor)
Samuel Bendett
Dr. Linton Wells II

STAR-TIDES
Center for Technology and National Security Policy
National Defense University

In Partnership with the
Naval Postgraduate School

RELIEF 11-03

Summary

From 2-5 May 2011, RELIEF convened its eighth session of field experiments for humanitarian information management and crisis mapping at the Naval Postgraduate School's (NPS) facility at Camp Roberts (Paso Robles, CA). The RELIEF experiments occurred within a partnership of the National Defense University's Center for Technology and National Security Policy and the Naval Postgraduate School.

PROBLEM DOMAINS: RELIEF 11-03 focused on three problem domains:

1. **Microtasking:** Building a microtasking management plugin for Ushahidi, which enables a workflow for a large number of volunteers to triage individual social media reports through translation, initial categorization, and geolocation, with the goal of filtering large number of messages into a small number of actionable report for responders. This work occurred in partnership with the Joint Staff QuickNets project, which is providing a deployable platform for supporting social media during crises.

2. **Exploring the use of natural language processing algorithm to identify actionable messages from large-volume social media streams**. Testing and training a natural language processing (NLP) algorithm to perform an initial filter on incoming messages that might accelerate the microtasking work from the first experiment.

3. **Improving the performance of an alpha version of OpenAerialMap**. OpenAerialMap is an open-source project with the goal of creating a free and open archive of overhead imagery, akin to way that OpenStreetMap provides a free and open map of the world. The experimentation examined ways to improve OAM's performance and to install the alpha software on an NPS server.

4. **Exploring the composition of a flyaway kit for HADR responders**. To support smaller NGOs that are on the site of a crisis within 24 hours of the onset of an emergency, Hosted Labs explored the composition of a lightweight and low-cost flyaway kit to support power and communications for a 2-person team.

APPROACH. Unlike the hackathons and code sprints which have become common among other crisis organizations (like Crisis Camps and the Random Hacks of Kindness) which generally invite technologists to invent new software platforms to solve a range of problems), the experiments at RELIEF extend existing applications. The intent is to gather the inventors of the open source software commonly used by responders and the large organizations that deploy to humanitarian emergencies. In

this case, RELIEF convened a panel of top humanitarian technologists from industry and the open-source domain. This team included the following software developers and engineers:

- **George Chamales**, core collaborator to Ushahidi and founder of Rogue Genius LLC.
- **Todd Huffman,** Synergy Strike Force
- **Rob Munro,** computational linguist, Stanford University
- **Dan Ratner,** freelance developer assisting with QuickNets
- **Jason Rexilius**, Hosted Labs/QuickNets.
- **Quicknets engineering team**
    - Mark Bradshaw, project lead
    - Dori Sewell
    - Carl Lynn
    - Tom Crow
    - Carlunle Moyet-Bruno
    - George Gallop
    - Stan VanDruff
    - Thomas "Alex" Prudencio
    - Clint Small

The RELIEF team also convened SMEs who focused on the social and policy problems around the use of open-source software in the field, including:

- **Rosa Akbari**, NPS RELIEF
- **Tristan Allen**, NPS RELIEF
- **Samuel Bendett**, NDU - TIDES
- **John Crowley**, experimentation lead and crisis mapping coordinator at both the Harvard Humanitarian Initiative and National Defense University Center for Technology and National Security Policy.
- **Christine Thompson,** Humanity Road (virtual)
- **Catherine Graham**, Humanity Road (virtual)
- **Nelly Turley**, NPS RELIEF

Over the course of four days, the team created viable solutions to problems observed in the crowdsourcing operations during the response to incidents in 2010-11, including Haiti, Pakistan, and Libya.

Problem 1: Microtasking

Problem Domain

During operations in Haiti and Libya, crowdsourcing platforms have enabled responders to collect information from populations affected by crisis and present their reports in novel ways. However, making sense of the streams of data flowing through SMS infrastructures has been very time consuming. During each operation in 2010-11, hundreds of volunteers have cooperated to transform unstructured SMS messages into actionable, geolocated reports. The methods used by these crowds have been inefficient, error prone, and incompatible with the austere and unforgiving environments associated with most humanitarian operations.

The workflow of many previous crowdsourcing operations required an individual volunteer to shepherd a message through the entire assembly line: from initial triage and categorization to determination of the message's actionability, veracity, and geolocation. This process made it difficult for volunteers to step in and out of the process; they had to see each method through the entire workflow, which for some messages took hours. It also required organizations to train volunteers in a wide range of tasks.

Microtasking: A New Approach

The experiment at RELIEF 11-03 flipped the assembly-line metaphor: it enabled volunteers to stand at one (virtual) place in the assembly line, performing a single specialized task. The developers—George Chamales and Rob Baker of Rogue Genius and Universities for Ushahidi—created a microtasking plugin that extends the base Ushahidi code base to allow for this new experimental workflow.

Methodology

The experiment convened 30 international volunteers from Humanity Road and the Standby Task Force (SBTF), with the intent of providing those volunteers with several thousand messages to parse in bursts of activity. The volunteers were tasked with processing messages using the new workflow, moving messages between three queues: an initial filtering for actionability, categorization, and geolocation. To provide data that could be compared against previous crowdsourcing engagements, the experiment used 30,000 anonymized SMS messages from the Haiti response in a professional English translation. The QuickNets system was able to send about 9000 of those messages through the system, the volume limitation coming from the discovery that the SMS provider (Google Voice) limits any single Google Voice account to sending 250 SMS messages per day.

Results

The volunteers worked in shifts, processing over 9000 messages resent via SMS into the QuickNets platform. Over the course of the test, volunteers processing approximately 2000 messages, generating 700 actionable reports. Volunteers who had experience with previous operations reported that both the workflow and interface were far easier to use than previous methods.

The experiment provided a platform for testing the application under simulated real-world conditions. This opportunity exposed several issues for the developers as they move towards an alpha version of the code.

1. Exposing the need for web caching: The first deals with the frequency with which the browser of each volunteers refreshed the queue of messages to be processed. The pre-alpha code used in this experiment began with the code pinging the server every 3 seconds (20 times per minute). Even at 30 volunteers, the database operations necessary to support this refresh rate slowed down the web server to a noticeable lag. In the field environment of Camp Roberts, the developers simply increased the size of the Amazon EC2 instance to get around the issue. However, this scalability issue will be addressed by adding caching that should nearly eliminate this issue.

2. Refactoring queue logic: the process by which the system selected the list of messages that any individual volunteer could select from a given queue exposed a race condition that occasionally gave two volunteers the same task. The alpha version of the code will alter the logic of the code to prevent this race condition.

The experiment also generated a third dataset based on the Haiti 4636 messages: the original dataset created by Ushahidi at Tufts during the operation, a second set generated by professional translators working with Stanford University, and this new set generated at Camp Roberts. These datasets open the opportunity for natural language algorithms to reprocess the data based on three passes of humans over the SMS messages.

Next Steps

The developers would like to explore ways to bench mark volunteer throughput, so that they can advise large institutions (including DoD) on the actual capacities of crowdsourcing applications and be able to better manage workflows during future operations.

Problem 2: Natural Language Processing Algorithm

Problem Statement:

Crowdsourcing largely relies on humans to process messages. While this method can process messages in minutes, there are problems with volunteers. First, they must be mobilized, which can take valuable time during the initial hours of the response and proved difficult during the Pakistan response. Second, at high volumes, backups can occur.

Natural language processing (NLP) algorithms can augment the capacity of volunteers, creating a symbiotic relationship: the algorithms can learn from humans what messages are actionable, and can then help the humans filter, categorize, and geolocate messages more effectively.

The experiment tested a machine learning/NLP that attempts to learn (from zero initial knowledge) how to identify actionable humanitarian information from an incoming stream of SMS messages in any language. The algorithm monitors humans who manually tag messages as "actionable" or "not actionable" and iteratively learns the linguistic and spatio-temporal features that will let it automatically prioritize future incoming messages.

Methodology

The NLP experiment was tested as part of the microtasking experiment that replayed the Haiti Mission 4636 messages within the Ushahidi platform. Based on a comparison between the data processed anew by volunteers and the NLP, the algorithm's inventor, Rob Munro of Stanford University, evaluated the accuracy of the machine learning process.

Results

On average, the algorithm identified actionable messages approximately 90% of the time (in jargon, an actionable incoming message would be in the 90th percentile for priority). This result opens the possibility of increasing the speed of prioritizing messages by approximately on order of magnitude (10x). This capability would be critical for high-volume processing. On its own, it could be a strategy for identifying actionable information in needle-in-a-haystack scenario like social media flows.

Problem #3: OpenAerialMap

Problem Statement

During a disaster, there is no single place where imagery from all providers can be aggregated and made available to responders in both the government and crisis mapping space. OpenAerialMap solves this problem by creating a distributed archive of imagery that has been processed to a common open standard.

Work Completed

OpenAerialMap has been rebuilt on volunteer time through several avenues, one of which has been multiple work sessions at RELIEF experiments. During RELIEF 11-02, Schuyler Erle created the first working version of the new OAM platform. During RELIEF 11-03, Schuyler started the process of installing the software stack on a server that NPS MOVES Institute donated. The software divides the imagery hosting problem into three components, any one of which can be hosted anywhere on the Internet:

1. Catalogue Server: an index of available imagery, including feeds from external sources.
2. Processing Server: a tile-generation service that processes all submitted imagery to a single open OAM standard.
3. Storage Server: a storage node that holds tiles processed into OAM format.

Working from the code base that is currently on oam.osgeo.org, Schuyler initiated the installation of the software onto a catalogue server and storage node at the Naval Postgraduate School's MOVES Institute. The Institute provided a 48TB storage cluster as well as small 1U Linux server that are capable of providing catalogue services.

Remaining Work

The OAM team will continue to work with the MOVES Institute to complete the installation of the code and further development to integrate the node at NPS with the node at oam.osgeo.org.

Problem #4: Lightweight, Low-cost Flyaway Kit

Problem Statement

During the initial hours of a disaster, it is common for NGOs to send small teams into the theatre of operations. While larger NGOs can afford relatively expensive kits to support these initial units, smaller NGOs are seeking more affordable solutions. This experiment evaluated a two-person, forward-deployed disaster response information gathering system. It intended to find a workable combination of geat to create a light-weight, rapid-response package for smaller NGO's that are on scene within 24 hours of the disaster. The report from Hosted Lab follows.

RELIEF Exercise May 2011

Our primary experiment was to evaluate a two-man, forward-deployed disaster response information gathering system. A light-weight, rapid-response package for smaller NGO's that are on scene within 24 hours of the disaster.

Assumptions

1) 100% COTS - low-cost readily accessible components.
2) Commercial air transport carry-on - small form factor, no security issues.
3) No existing power grid in theater - intermittent access to vehicles or generators at best for 48-96 hours.
4) Comms networks will be overwhelmed or destroyed – ad hoc networking capability and multiple channels a requirement
5) Responding team will have light IT training/expertise. Not experts but not cavemen.

Objectives

1) Validate the ability of the system to operate for 48 hours without access to power grid at full operational tempo.
2) Evaluate components, identify gaps.
3) Practice deployment procedures and operating conditions in non-sheltered environments.
4) Test connectivity and ad hoc networking procedures.

Results

1) Camp Roberts in May can be considered ideal solar conditions. This defines the upper bound of the possible. Given that, there was ample solar power and battery back-up to operate continuously off-grid. Extrapolation of these conditions implies that for cloud cover, rainy, higher latitude environments (worst case conditions) it will be necessary to increase solar array and battery capacity. Overall assessment is that within 35 degrees latitude of the equator, even in cloudy conditions, the kit can operate without access to external power.
2) There were a few components known missing from the kit prior to deployment that precluded testing of some capabilities. There were also a few components identified as gaps. Overall the system was functional but more tests need to be conducted. A result of this experiment is that a list of explicit tests for the hardware components is being developed for the next exercise.
3) Some component gaps were identified here that impacted operating in hot temperature, non-sheltered environments. One challenge with standard COTS equipment is operating temperature expectations. This is not a serious problem but does require some specific measures to mitigate.
4) Ad hoc networking was the largest problem area, as was expected. The typical problems of flooded unlicensed spectrum (i.e. WiFi) were encountered as well as other typical configuration and security issues. There were also interoperability problems at the application layer born of typical issues such as mismatched

versions of software between parties (one party upgrades to newer version prior to deployment, other party does not, etc.). Overall this is the known problem spot and can be solved with IT people in the field but other measures need to be developed and richer capabilities need to be added to the kit to support the assumption of lightly trained operators.

Summary

Camp Roberts is an ideal environment for this exercise. Connectivity is poor, its hot, dusty and moderately barren. Adding another venue somewhere like Washington state may be useful to test other conditions. The system performed as expected and testing identified some gaps to be filled. The biggest problem, interoperability, is a know one and there are paths to addressing it. It is expected that the next exercise in August will validate the improvements and mark the system as field-ready.

Self-Sufficient Phone Kit

A recommended set of Commercial Off The Shelf (COTS) components that will provide an individual field user a small computing and communications platform while maintaining energy self-sufficiency.

The smart phone component of the system provides the user with three radio channels for communications (GSM, WiFi and Bluetooth) along with embedded sensors such as camera, microphone, GPS and motion. It also provides removable data storage in the form of a microSD card slot. A full “qwerty” keyboard and reasonable screen size provides a moderately comfortable user interface for heavy and extended use. The 528MHz CPU and 192MB RAM along with the open source Android OS provide a robust application-computing environment on par with laptops of 10 years ago.

This core set of physical capabilities, combined with energy self-sufficiency, make this kit well suited for operations in disaster conditions, developing nations or other challenging environments. The flexibility and openness of the Android operating system enable it to serve in a variety of roles in these demanding environments.

The total package weighs 2.2lbs (1kg), costs $900 retail and fits easily in carry-on luggage. It will provide 2 to 7 days of use out of the box before solar recharge is needed. In addition to solar recharging it can scavenge power from various other sources such as wall outlets, car lighters, AA batteries and any USB power source.

Example Capabilities:

- Multi-media Information Collections (text, photo, audio, etc.), GPS tagged
- Communications (voice, SMS, email)
- Way-pointing and Mapping (offline with OpenStreetMaps)
- SMS Gateway, relay messages from field into other systems
- Information store-and-forward relay (from other field nodes, sensors)
- Field language translation services (phrases, dictionary, text-to-speech)

Self-Sufficient Base-station Kit

A recommended set of Commercial Off The Shelf (COTS) components that will provide a stationary computing and communications platform for coordinating activities while maintaining energy self-sufficiency.

The base-station component of the system provides a field user with a light "command center" capability but primarily serves as an information gathering and relay system. It can operate and provide services without an uplink, relying on local networks and communications channels if needed. Power is provided by a solar panel and set of batteries, which enable 12-hour operations in most sunlight conditions. An additional power module can be added for extend hours of operation or sub-optimal sunlight conditions. The base-station provides a local WiFi access-point as well as local Ethernet network. SMS modem, RF hand-held radios and loud speakers provide multiple communications services.

The S2B kit weighs 40lbs (20kg), costs $4,500 and can be a carry-on piece of luggage on normal airlines. It will provide up to 12 hours of heavy operation on first charge out of the box before recharge is needed. It can scavenge power from its solar panel, wall outlets and car lighters as well as other USB power sources (Solio and Fueltank batteries from S2Phone Kit). Like the S2Phone Kit external batteries can be couriered to a power source while continuing operations.

Example Capabilities:

- Information collation, tagging and forwarding
- Operations mapping and sector coordination
- SMS gateway
- Local phone support line
- Data scavenging center
- Sneaker-net forwarding center
- Common operating picture station
- Transport-layer interconnection junction
- GIS annotation