Appendix E

Letter on Security in Crowdsourcing

By Rebecca Goolsby, Ph.D., Office of Naval Research

Published in Burns, R. and Shanley, L.A. 2013. Connecting Grassroots to Government for Disaster Management: Workshop Summary. Washington, DC: Commons Lab of the Woodrow Wilson International Center for Scholars

September 6, 2012

My article on cybersecurity, crowdsourcing, and social cyber-attack is delayed as my team begins to piece together the recent events in Assam, India that led to ethnic violence between the Bodo people and the Muslims only a few weeks ago. This is a subject that I have long had an interest in, owing to the significant implications of these new patterns of behavior, new forms of community, and new problems in crime and malicious mischief that the virtual world is experiencing everywhere we look. Consider this an informal letter/email on my thoughts on the subject that you may share as you wish, not as a publication, but as food for thought.

Information sharing has a spectrum of social impact, from the very "white," "clean," and humanitarian efforts such as disaster relief, coordination of humanitarian activities, and the promulgation of truthful information to topics more grey and even dark. This type of messaging seeks to bolster social order, relieve suffering, and promote positive social bonds. Counter-messaging, the refutation of bad information, lies, and mischief is a bit grey, colored by the propaganda that it seeks to refute. Its objectives are a bit more biased, to promote one's own "story" against the claims of others who seek to use deceit or misrepresentation to get their views across. Propaganda of every stripe—attempts to rally the base or influence others—gets a bit greyer still, with the objective of swaying others toward a particular agenda. The creation of hoaxes and scare-mongering campaigns seek to subvert public order, generate and exploit the resulting chaos so as to benefit or gain in some way. This is something of a new black art.

The use of "Photo-Shopped" images—pictures which have been altered in order to create fear and chaos—has been used before, particularly in Middle East affairs, where

one group or another alters images to suggest that police brutality, mob violence, or other acts occurred in one place and at a given time (when, in fact, the pictures were from a different place, time, and situation). Savvy social media enthusiasts know how to use "reverse image search" to find the true origins of photos—and to be skeptical of images found on the Internet. New entrants into the world of social media are not aware of these capabilities and can be readily fooled—as was the case in Assam. The use of MMS (multimedia mass texting, where images were sent to cellphones, rather than through the Internet directly) was an interesting addition. Details are scant, but it is possible that social media might have played a role, as social media enthusiasts often link their phones and emails to their accounts and often, unthinkingly, allow third-party apps (programs) to access their information—and provide links to their friends' information, which would be a good way to seed a snowball of interconnecting links. If my social media pal appeared to send me images, then I might trust that to be a true indication of what was going on near me (or near them)—when in fact it was some malefactor who poached his information and his connection to me.

It is unclear to me whether this happened in Assam. We're trying to figure that out, from a distance (myself, Dr. Huan Liu from Arizona State University, and others), but it is hard because Facebook and Twitter blocked the false content. Since they were a conduit---but not the primary conduit—for the false information, it was difficult to find this crisis at the time it broke. Discussions were in a minor language—Bengali—and thus the discussion of these images and the (false) situations they depicted did not overlap very much into English-speaking communities. The scare-mongering campaign was designed to capitalize on the social uncertainty among the Muslim community following several actual incidents in the previous weeks, leading to mass exodus to refugee camps only weeks before. That the attack was on the last day of Ramadan—a celebration of the Islam faith—was telling. Terrorists are historically interested in symbolic acts and time components figure prominently in their symbolic language.

The capability of crowdsourcing such an attack is now everywhere. Through social media, hate speech proliferates with the capability of reaching hordes of interested mischief makers who are comfortably anonymous and hard to track. Social cyber-attack as a means to bully, trick, and sow uncertainty in tense situations is not going to go away. It is not a matter of finding "the one guy behind all this" anymore, as malcontents, "trolls," and malicious actors are legion, connected in loose cyber-communities and technically capable. "Robot Twitter accounts" and other "zombie" systems can extend the reach of individuals and when these techniques are shared among like-minded anarchists and zealots, the capability of a small minority is magnified. They are thus able to pump their apparent numbers up and spread the risk of being caught around. With this capability of hiding behind dozens, even hundreds or thousands of identities, the risk of discovery is lowered, and the capability to develop an extensive cadre of cooperating "cyber-hoodlums" is growing. For those of us old enough to remember "phone phreaking" (http://en.wikipedia.org/wiki/Phreaking), this is not a new thing. The super "phreaks" can and do hide among the many, many "script-kiddies" capable of learning simple pranks and thus sowing mischief, hate, and chaos—chaos that the truly harmful

players can exploit financially, politically, or socially. Trying to find "that one guy behind it all" is to engage in a game of "whack-a-mole" with literally hundreds, thousands, and even millions of shadow puppets. It would be more profitable to try to discover clever ways of figuring out who benefits, but even then, that's a fairly small number.

There is research on this in a number of places and it would make for a useful workshop if the organizers were careful to look at the SOCIAL, as well as the TECHNICAL, aspects of this, for that is where the vulnerability is, in the connections among this shadow community. Further, the need to substantively educate the public, especially first responders, whose worlds are usually in a state of uncertainty, danger, and incipient chaos, about the need to be circumspect and savvy in information sharing, for they may be particularly at risk for exploitation, hoaxes, and scams. A crowd that carefully self-polices is the absolute best defense, as government cannot be everywhere—but the crowd certainly is.