

NAME: Maxfiled Green
DATE: 1/15/2019
Reading reflections: Reading Group 4
STAT/CS 387

Title(s) read

- Harnessing the Crowd-sourcing Power of Social Media for Disaster Relief
- The Rise of Crowdsourcing
- Time-Critical Social Mobilization
- Error and attack tolerance of collective problem solving: The DARPA Shredder Challenge

Reflection

Disaster relief is a crucial part of recovering from natural or social disasters in heavily populated areas. Families and individuals can be displaced from homes, property destroyed while in parallel streams of communication and information cut off. This makes relief planning/strategy difficult. Using social media to leverage space and time stamped data could be a huge aid in making strategy decisions to optimize the use of resources such as funding and human power. The readings assigned this week cover multiple applications of crowd sourcing information from major social networks, both virtual and in real life. We see how crowdsourcing can provide instant access to large amounts of data with the risk of security and reliability. Applications such as disaster response and complex tasks completion are discussed and analyzed in depth.

The first article discusses the applications of crowd sourcing in disaster relief efforts. In 48 hours the red cross generated \$8,000,000. This article describes crowd sourcing applications used in disaster relief coordination. The article reports three main advantages of crowd sourcing in the context of disaster relief. The first advantage is time, crowd-sourced data can be collected immediately after a disaster has occurred, this is crucial for responders to prioritize the most urgent aid requests. The second advantage is the volume and variety of data sources to validate and confirm requests. The third advantage is that some information is inherently geo-tagged, like twitter. This provides accurate location data for relief organizations, who otherwise would potentially have to infer or use further resources to discover this type of information.

In addition to the great benefits from crowd-sourcing information about disasters comes host of downfalls. The main problem is that there is no way to validate an individual piece of information caught in the virtual fish net. The geo-tag may be incorrect, the information help in text may be incorrect, the data may be generated from a bot or other user of malicious intent. These are all problems that need to be addressed.

The rise of crowd sourcing describes how the masses in some situations are replacing professional services for a much lower cost to the consumer. The writer puts this in words nicely "Distributed labor networks are using the Internet to exploit the spare processing power of millions of human brains.". The main example given is the role of a trained stock photographer. This career used to pay well, now its easy for amateurs to upload many many images to a database and sell the license

for very cheap to an add agency. This possess a problem to professional specialists as their services are often undercut by the masses. Traditionally, normal or even hobby photographers didn't have access to the market of photo buyers, this was exclusive to professional's.

The third reading "Time Critical Social Mobilization" is a study of one solution the DARPA Network Challenge. The DARPA Network Challenge simply asked teams from all over the world to retrieve the location of 10 red weather balloons. The winning team, from MIT, crowd-sourced the location information of each of the balloons. The team was able to recruit 4400 individuals through a "recursive incentive mechanism". This way, all of the people involved in finding the balloon received a reward. A really interesting referral tree is thus created by people that continually got involved in the search process. Within the first day, 845 trees were recruited, the largest contained 602 nodes and the deepest went 14 levels down! Incredible. Additionally, the average branching factor was reported to be .93. The branching factor follows a power law distribution which indicates that there are clusters of people that recruited loads of other people, this makes for very effective diffusion on a network. However, this also makes the network very vulnerable to attack, say if a central hub had incentive to propagate false information. To compare the success of their strategy to other teams, MIT used twitter mentions as a proxy for involvement. This way they could compare over time how the different teams crowd sourcing technique's worked. Some teams got more people involved but for a shorter period of time while other teams never generated enough steam to truly go viral. The MIT team attributes its success to the monetary incentives provided. I thought this was a really great article, there are so many applications to the iterative incentive mechanism. The true genius is incentivising both recruitment and participation.

The final article title "Error and attack tolerance of collective problem solving: The Darpa Shredder Challenge" is an analysis of a solution strategy for the DARPA Shredder Challenge. This challenge involved reassembling shredded documents. This task requires high level cognitive skills to identify patterns over many disjoint observations and pieces of data. People were incentivised to participate through monetary compensation. Participants were further monetarily incentivised to recruit more workers to complete the task. Because the process and results of the crowd-sourcing were open source, competing challenge teams were able to attack the network by recruiting workers to spam the result. This allowed the UCSD team to analysis the system under two conditions, under attack and in a normal state. They found that it took very little action from the attackers to bend crowdsourced tasks out of order and beyond repair. The research team developed a method to classify users as contributors or attackers and then analyzed the behavior of both classes.