Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy

Galen Pickard
MIT Media Lab and MIT Lincoln Lab
Cambridge, MA 02139
gpickard@mit.edu

Iyad Rahwan
MIT Media Lab and Masadar Institute of Sci.&Tech.
Cambridge, MA 02139
irahwan@acm.org

Wei Pan
MIT Media Lab
Cambridge, MA 02139
panwei@media.mit.edu

Manuel Cebrian
MIT Media Lab
Cambridge, MA 02139
cebrian@mit.edu

Riley Crane
MIT Media Lab
Cambridge, MA 02139
rcrane@media.mit.edu

Anmol Madan
MIT Media Lab
Cambridge, MA 02139
anmol@media.mit.edu

Alex (Sandy) Pentland
MIT Media Lab
Cambridge, MA 02139
sandy@media.mit.edu

Abstract

It is now commonplace to see the Web as a platform that can harness the collective abilities of large numbers of people to accomplish tasks with unprecedented speed, accuracy and scale [10]. To push this idea to its limit, DARPA launched its Network Challenge, which aimed to "*explore the roles the Internet and social networking play in the timely communication, wide-area team-building, and urgent mobilization required to solve broad-scope, time-critical problems*" [5]. The challenge required teams to provide coordinates of ten red weather balloons placed at different locations in the continental United States. This large-scale mobilization required the ability to spread information about the tasks widely and quickly, and to incentivize individuals to act. We report on the winning team's strategy, which utilized a novel recursive incentive mechanism to find all balloons in under nine hours. We analyze the theoretical properties of the mechanism, and present data about its performance in the challenge.[1]

1 Time-Critical Social Mobilization

Crowdsourcing is to harness the collective abilities of large numbers of people to accomplish tasks with unprecedented speed, accuracy and scale [11, 23, 4]. A particularly challenging class of crowdsourcing problems require not only the recruitment of a very large number of participants, but also extremely fast execution. Examples include search-and-rescue operations in the aftermath of natural disasters, hunting down wanted outlaws on the run, reacting to health threats that need instant attention, and rallying supporters to vote in a political campaign. In time-critical social mobilization problems, it is often not practical, or even impossible, to create sufficient mobilization through mass media, due to the extremely high cost of reaching everybody, or due to severe infrastructure damage[14].

[1] A full version of this paper can by found at `http://arxiv.org/abs/1008.3172`

Another common characteristic of these social mobilization problems is the presence of some sort of search process. For example, search may be conducted by members of the mobilized community for survivors after a natural disaster. There is growing literature on search in social networks. It has long been established that social networks are very effective at finding target individuals through short paths [20], and various explanations of this phenomenon have been given [13, 25, 1, 22].

However, it is important to recognize that the challenge of a successful search in social mobilization depends on the *incentive mechanism*. Indeed, in an empirical study, Dodds et al conclude that "*although global social networks are, in principle, searchable, actual success depends sensitively on individual incentives*" [8]. It has also been observed that the the success of crowdsourcing mechanisms, in general, can vary depending on the details of the financial incentive scheme in place [19].

Recognizing the difficulty of time-critical social mobilization, the Defense Advanced Research Projects Agency (DARPA) announced the DARPA Network Challenge. The announcement, which coincided with the 40th anniversary of the Internet. Through this challenge, DARPA aimed to "*explore the roles the Internet and social networking play in the timely communication, wide-area team-building, and urgent mobilization required to solve broad-scope, time-critical problems*" [5]. The challenge is to provide coordinates of ten red weather balloons placed at different locations in the continental United States. According to DARPA, "*a senior analyst at the National Geospatial Intelligence Agency characterized the problem as impossible*" by conventional intelligence gathering methods [7].

2 The Recursive Incentive Mechanism

According to the DARPA report, between 50 and 100 serious teams participated in the DARPA Network Challenge, from a total of $4,000$ teams [7]. Moreover, approximately $350,000$ people participated in the DARPA Network Challenge in various ways, ranging from searching for balloons, to simply being aware of the challenge and willing to report a balloon if spotted.

The MIT Team, which won the challenge [6], completed the challenge in 8 hours and 52 minutes. In approximately 36 hours prior to the beginning of the challenge, the MIT Team was able to recruit almost $4,400$ individuals through a recursive incentive mechanism.

The MIT Team's approach was based on the idea that achieving large-scale mobilization towards a task requires (a) diffusion of information about the tasks through social networks; and (b) provision of incentives for individuals to act, *both* towards the task and towards the recruitment of other individuals. We consider the MIT Team's approach to the DARPA Network Challenge to be an instance of a more general class of mechanisms for distributed task execution. We now define this class of mechanisms. But we first need to define the setting in which such mechanisms operate. We define a *diffusion-based task environment* which consists of the following: $N = \{\alpha_1, \dots, \alpha_n\}$ is a set of *agents*; $E \subseteq N \times N$ is a set of *edges* characterizing social relationships between agents; $\Psi = \{\psi_1, \dots, \psi_m\}$ is a set of *tasks*; $P : N \times \Psi \rightarrow [0, 1]$ returns the *success probability* of a given agent in executing a given task; $B \in \mathbb{R}$ be the *budget* that can be spent by the mechanism.

In a diffusion-based task environment, agents are *not* aware of the tasks a priori. Instead, they *become* aware of tasks as a result of either (1) being directly informed by the mechanism through advertising; or (2) being informed through recruitment by an acquaintance agent [26].

Another characteristic of diffusion-based task environments is that, when a task is completed, the mechanism is able to identify not only the agent who executed it, but also the information pathway that led to that agent learning about the task. The pathway leading to the successful completion of task ψ_i is captured by the sequence $\mathcal{S}(\psi_i) = \langle a_1, \dots, a_r \rangle$ of unique agents, where a_r is the agent who completed the task, a_r was informed of the task by a_{r-1} and so on up to agent a_1 who was initially informed of the task by the mechanism. By slightly overloading notation, let $|\mathcal{S}(\psi_i)|$ denote the length of the sequence (i.e. the number of agents in the chain), and let $\alpha_j \in \mathcal{S}(\psi_i)$ denote that agent α_j appears in sequence $\mathcal{S}(\psi_i)$.

We can now define a class of mechanisms that operate in the above settings. A *diffusion-based task execution mechanism* specifies the following: $I \subseteq N$ is a set of *initial nodes* to target (e.g.

via advertising); ρ_i is the *payment* made to agent α_i; such that the following constraint is satisfied: $c|I| + \sum_{\alpha_i \in N} \rho_i \leq B$.

In words, the mechanism makes two decisions. First, it decides which nodes to target initially via advertising. Second, it decides on the payment (if any) to be made each agent. The mechanism must do this within its budget B.

In the DARPA Network Challenge, each ψ_i represents finding a balloon, and $v(\psi_i) = 4,000$ for all $\psi_i \in \Psi$. Moreover, since the ten tasks are all identical (namely finding a balloon), $\forall \alpha_i \in N, \forall \psi_k, \psi_l \in \Psi$ we have $P(\alpha_i, \psi_k) = P(\alpha_i, \psi_l)$. That is, the success probability of a particular agent is the same for all balloons.

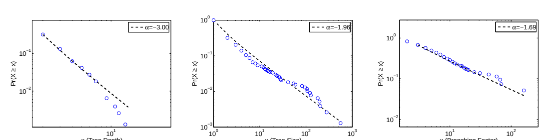
We are now ready to define the MIT Team mechanism, referred to as a *recursive incentive mechanism*. Given I initial targets, and assuming $v(\psi_i) = B/|\Psi|$, divide the budget B such that each task $\psi_i \in \Psi$ has budget $B_i = B/|\Psi|$. If agent $j \in N$ appears in position k in sequence $\mathcal{S}(\psi_i)$, then j receives the following payment: $\frac{v(\psi_i)}{2^{(|\mathcal{S}(\psi_i)|-k+1)}}$. Hence, the total payment received by agent j is the sum of payments for all sequences in which j appears: $\rho_j = \sum_{\psi_i | j \in \mathcal{S}(\psi_i)} \frac{v(\psi_i)}{2^{(|\mathcal{S}(\psi_i)|-k+1)}}$. The surplus is therefore: $S = B - \sum_{\alpha_j \in N} \rho_j$. See the figure in our full paper[21] for illustrations of this mechanism.

As demonstrated in our full paper[21], the MIT mechanism is never in deficit. Under reasonable assumptions, we also prove that the best strategy for a participant under this incentive structure is to recruit all one's peers. However, this mechanism is also limited: for instance, it is not resistant to *false name* attack.

3 Empirical Data

We explore how our mechanism's performance compares with previous studies on search and recruitment in social networks based on the diffusion data we collected during the Network Challenge.

One measure of success is the *size of the cascades*, both in terms of *number of nodes*, as well as *depth*. Results vary in existing literature. In a study of the spread of online newsletter subscriptions, in which individuals were rewarded for recommending the newsletter to their friends, the $7,188$ cascades varied in size between 2 and 146 nodes, with a maximum depth of 8 steps [12], over a time span of three months.[2] In our data, if we ignore the MIT root node, there are 845 trees recruited within only three days. The largest tree contained 602 nodes, and the deepest tree was 14 levels deep. We found that tree depth follows a power law. Furthermore, Tree size also shows a power-law distribution with exponent -1.96, as predicted by models of information avalanches on sparse networks [24]. The plots are presented in Figure 1.



(a) Tree depth (b) Tree size (c) Branching factor

Figure 1: (a) Distribution of tree depth on a log-log scale with a power law fit. (b) Distribution of tree size on a log-log scale with a power law fit. (c) Distribution of the branching factor on a log-log scale with a power law fit.

Previous empirical studies reported significant *attrition rates* (aka *discard rate*), which measures the percentage of nodes that terminate the diffusion process. For example, in a study of email-based global search for 18 target persons, attrition rate varied between $60-68\%$ in 17 out of the 18 searches performed [8]. It has been argued that the "*lack of interest or incentive, not difficulty, was the main reason for chain termination*" [8]. In another study of the diffusion of online recommendations, an attrition rate of 91.21% was repoted *despite* providing incentives to participants by offering them a

[2] Esteban Moro, personal communication.

chance in a lottery [12]. In the DARPA Network Challenge, if we ignore isolated single nodes, our mechanism achieves a significantly lower attrition rate of 56%.

Moreover, in the newsletter subscription experiment [12], the dynamics of diffusion were slow, which was attributed to a heterogeneous, non-Poissonian distribution of individuals' response time. Interestingly, we observe an exponential distribution of inter-signup time[21]. This contrasts with the empirically observed power-law distribution of inter-response time in human activity [2, 17] and information cascades [12]. Ongoing initiatives that utilize our approach could determine whether this deviation is due to the incentive mechanism.

4 Conclusion

From an *observational perspective*, previous studies have shown that the success of information cascades on social networks is affected by various factors[26, 3, 12, 16, 8, 24, 16]. However, from the *perspective of an incentive designer* seeking large-scale social mobilization, the problem boils down to two questions: (i) which individuals to target directly? (ii) what incentives to provide in order to encourage participation? While others have addressed the first question (i.e. [9]), here we addressed the question of incentives, which has not received much attention in the literature, as pointed out by Dodds et al [8], until the DARPA Network Challenge. In particular, our mechanism *simultaneously* provides incentives for participation and for recruiting more individuals to the cause. This mechanism is already being used in different contexts, such as social mobilization to fight world hunger, in games of cooperation and prediction, and for marketing campaigns.

We believe that it is not a coincidence that the winning strategy in the DARPA Network Challenge was one that exploited ideas from both incentive design [18] and computational social science [15]. After all, people are self-interested individuals, but also embedded within social networks. It is hoped that this paper will stimulate theoretical and empirical efforts to devise incentive mechanisms for a variety of challenging, time-critical social mobilization problems.

References

[1] L. A. Adamic and E. Adar. How to search a social network. *Social Networks*, 27:187–203, 2005.

[2] A. Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 2005.

[3] A.-L. Barabási. *Bursts : the hidden pattern behind everything we do*. Dutton, New York, N.Y., 2010.

[4] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, et al. Predicting protein structures with a multiplayer online game. *Nature*, 466(7307):756–760, 2010.

[5] Defense Advanced Research Projects Agency. DARPA Network Challenge, Accessed May 2010. http://networkchallenge.darpa.mil/.

[6] Defense Advanced Research Projects Agency. MIT Red Balloon Team wins DARPA Network Challenge, December 5, 2009. Press Release.

[7] Defense Advanced Research Projects Agency. DARPA Network Challenge Project Report, February 16, 2010.

[8] P. S. Dodds, R. Muhamad, and D. J. Watts. An Experimental Study of Search in Global Social Networks. *Science*, 301(5634):827–829, 2003.

[9] P. Domingos and M. Richardson. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66, New York, NY, USA, 2001. ACM.

[10] E. Hand. Citizen science: People power. *Nature*, 466(7307):685, 2010.

[11] J. Howe. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. Three Rivers Press, New York, NY, USA, 2009.

[12] J. Iribarren and E. Moro. Impact of human activity patterns on the dynamics of information diffusion. *Physical Review Letters*, 103(3), 2009.

[13] J. Kleinberg. Navigation in a small world. *Nature*, 406:845, August 2000.

[14] G. Krakow. Ham radio operators to the rescue after Katrina: Amateur radio networks help victims of the hurricane. `http://www.msnbc.msn.com/id/9228945/`, September 6 2005.

[15] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al. Social science: computational social science. *Science*, 323(5915):721, 2009.

[16] D. Liben-Nowell and J. Kleinberg. Tracing information flow on a global scale using internet chain-letter data. *Proceedings of the National Academy of Sciences*, 105(12):4633–4638, March 2008.

[17] R. Malmgren, D. Stouffer, A. Campanharo, and L. Amaral. On universality in human correspondence activity. *science*, 325(5948):1696, 2009.

[18] E. Maskin. Mechanism design: How to implement social goals. *American Economic Review*, 98(3):567–576, 2008.

[19] W. Mason and D. J. Watts. Financial incentives and the "performance of crowds". In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, pages 77–85, New York, NY, USA, 2009. ACM.

[20] S. Milgram. The small world problem. *Psychology Today*, 1(1):6067, 1967.

[21] G. Pickard, I. Rahwan, W. Pan, M. Cebrian, R. Crane, A. Madan, and A. Pentland. Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy. *Arxiv preprint arXiv:1008.3172*, 2010.

[22] M. Rosvall, P. Minnhagen, and K. Sneppen. Navigating networks with limited information. *Physical Review E*, 71(6):66111, 2005.

[23] L. von Ahn. Games with a purpose. *Computer*, 39:92–94, 2006.

[24] D. J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, April 2002.

[25] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and Search in Social Networks. *Science*, 296(5571):1302–1305, 2002.

[26] D. J. Watts and J. Peretti. Viral marketing in the real world. *Harvard Business Review*, May 2007.