

Limits of social mobilization

Alex Rutherford^a, Manuel Cebrian^{b,c}, Sohan Dsouza^a, Esteban Moro^{d,e}, Alex Pentland^f, and Iyad Rahwan^{a,g,1}

^aComputing and Information Science, Masdar Institute of Science and Technology, Abu Dhabi 54224, United Arab Emirates; ^bDepartment of Computer Science and Engineering, University of California at San Diego, La Jolla, CA 92093; ^cNational Information and Communications Technology Australia, Melbourne, VIC 3010, Australia; ^dDepartamento de Matemáticas and Grupo Interdisciplinar de Sistemas Complejos, Universidad Carlos III de Madrid, 28911 Madrid, Spain; ^eInstituto de Ingeniería del Conocimiento, Universidad Autónoma de Madrid, 28049 Madrid, Spain; ^fMedia Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; and ^gSchool of Informatics, University of Edinburgh, Edinburgh EH8 9AB, United Kingdom

Edited* by Jon Kleinberg, Cornell University, Ithaca, NY, and approved March 1, 2013 (received for review September 19, 2012)

The Internet and social media have enabled the mobilization of large crowds to achieve time-critical feats, ranging from mapping crises in real time, to organizing mass rallies, to conducting search-and-rescue operations over large geographies. Despite significant success, selection bias may lead to inflated expectations of the efficacy of social mobilization for these tasks. What are the limits of social mobilization, and how reliable is it in operating at these limits? We build on recent results on the spatiotemporal structure of social and information networks to elucidate the constraints they pose on social mobilization. We use the DARPA Network Challenge as our working scenario, in which social media were used to locate 10 balloons across the United States. We conduct high-resolution simulations for referral-based crowdsourcing and obtain a statistical characterization of the population recruited, geography covered, and time to completion. Our results demonstrate that the outcome is plausible without the presence of mass media but lies at the limit of what time-critical social mobilization can achieve. Success relies critically on highly connected individuals willing to mobilize people in distant locations, overcoming the local trapping of diffusion in highly dense areas. However, even under these highly favorable conditions, the risk of unsuccessful search remains significant. These findings have implications for the design of better incentive schemes for social mobilization. They also call for caution in estimating the reliability of this capability.

social networks | human dynamics

The Internet and online social media are now credited with the unprecedented ability to coordinate the mobilization of large masses of people to achieve incredible feats that require coverage of large geographical and informational landscapes in very limited time. Social media have been used to mobilize volunteers to map natural disasters in real time (1) and to conduct large-scale search-and-rescue missions (2). Online social networks also have been an important tool in the coordination of mass political rallies (3, 4).

Endeavors such as the Defense Advanced Research Projects Agency (DARPA) Network Challenge (5) aimed to test the power of the Internet and social media in time-critical social mobilization to its absolute limits. The Network Challenge required competing teams to locate and submit the coordinates of 10 tethered weather balloons dispersed at random locations all over the continental United States. The winning team, based at Massachusetts Institute of Technology (MIT), won the challenge by locating all balloons in less than 9 h. The MIT team used an incentive scheme to kick start an information and recruitment cascade that resulted in 4,400 sign-ups to the team's Web site within 48 h. Analysis of the diffusion revealed that the recursive incentive scheme may have played an important role in maximizing the speed and branching of the diffusion to limits above what normally are observed in viral propagation schemes (6–8).

More recently, the State Department's Tag Challenge required competing teams to locate and photograph five target "thieves" (actors) in five different cities in the United States and Europe, based only on a mug shot released at 8:00 AM local time (9). The

targets were visible for only 12 h and followed normal itineraries around the cities of Stockholm, London, Bratislava, New York, and Washington DC. Our winning team located three of the five suspects using social media, without any of the team members being based in any of the target cities (10), demonstrating yet another example of time-critical social mobilization in tasks that require coverage of large geographies.

Despite these numerous successes, we still have limited understanding of the limits of technology-mediated mobilization. If we are to rely on social media to react to time-critical emergencies, it is important to understand the conditions under which they can be successful and the risks of failure associated with them. A particular case, of highly practical importance, is to understand the extent to which we can expect to cover a certain geographical area in a given amount of time. For this, we must understand the complete statistical characterization of the population recruited, geographical area covered, and completion time it takes for social mobilization to succeed in a particular task, as well as to quantify the likelihood of failure.

This lack of understanding is especially prone to selection bias over few successful social mobilization strategies and may lead to inflated expectations of the reliability and efficacy of these techniques (11, 12). However, it is beyond experimental capabilities to perform randomized experimentation with large crowdsourcing challenges [with notable exceptions emerging recently (13, 14)].

Modeling efforts in the wake of the H1N1 and other global pandemics also have provided a valuable insight into time-sensitive human dynamics on a large scale, via spatial simulation (15) or network-based diffusion (16, 17). In common with these efforts, we model the interaction and connection of large numbers of agents; however, we consider the propagation of a message that may be transmitted without direct physical proximity and generally shorter "incubation" times, leading to faster spreading. Thus, the mechanism of "infection" (recruitment) is independent of human mobility patterns (18–20), which, in our case, contribute only to the area searched.

In this work, we build on recent results on social network structure, information diffusion, and urban economics to elucidate the constraints they pose on social mobilization. In particular, we conduct high-resolution simulations of the DARPA Network Challenge. We obtain statistical characterizations of the population recruited, geography covered, and time to locate the 10 balloons, together with their dependencies on the instrumental variables.

Our results demonstrate, surprisingly, that the DARPA Network Challenge outcome is plausible and thus not simply a fluke that can be explained only by the role of mass media. Therefore,

Author contributions: A.R., M.C., E.M., A.P., and I.R. designed research; A.R., S.D., and I.R. performed research; A.R., M.C., E.M., A.P., and I.R. contributed new reagents/analytic tools; A.R. analyzed data; and A.R., M.C., E.M., A.P., and I.R. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: irahwan@acm.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216338110/-DCSupplemental.

the challenge lies at the limit of what time-critical social mobilization can achieve. Mobilization requires highly connected, highly active individuals to be motivated to propagate the message to many friends and to mobilize people in distant locations, overcoming the local trapping of diffusion in highly dense areas. Moreover, even under highly favorable conditions, the risk of mobilization failure remains significant. These findings have implications for the design of better incentive schemes for social mobilization. They also call for caution in estimating the reliability of this capability.

Simulation Model

In seeking to understand social mobilization, we must consider the many different dynamics that underpin such a process, particularly the branching dynamics of recruitment, the temporal dynamics of message propagation, the geographical spread of social networks, and the scales and aspects of human mobility. A full accountability of each process will render the simulation and its understanding impossible, and thus we concentrate on the main ingredients that explain the observed behavior in each of the processes.

Recruitment

Examination of the dynamics of the branching recruitment process in empirical data from the challenge (6) suggests several key features. After a large initial round of recruitment from the seed node, the reproductive number is well below the tipping point (*SI Appendix*). Thus, many subtrees first are created spreading from the root node, then steadily die out. To describe the typical branching recruitment process, we fit to the observed branching data assuming an atypical burst of recruitment when the search commences. We find a power law distribution with a mean $\langle R_0 \rangle = 0.89$. (See *Materials and Methods* for details of fit.)

Response Time

The importance of the heterogeneity in response times in viral recruitment processes has been demonstrated (21). In a study of a viral e-mail campaign, the time taken to forward a message was found to be log-normally distributed, as opposed to the commonly used Gaussian assumption, with a mean of 1.5 d with an SD of 5.5 d. This large heterogeneity has a deep impact on the propagation of information: cascade dynamics may be halted by the few individuals with very long response times, and thus recruitment events may continue up to the order of years after the seed node starts the cascade. The waiting-time distribution in a time-critical campaign such as the DARPA Network Challenge might differ fundamentally because it necessarily must end by a fixed deadline regardless of whether the campaign is successful. However, we use the distribution of ref. 21 as a reasonable approximation. We do not apply a cutoff at large times, although the tail of the distribution may be effectively truncated because a search may terminate if all balloons are found before recruits with waiting times drawn from the tail of the distribution can act. The role of burstiness in diffusion in temporal networks was investigated systematically in ref. 22 by characterizing tie strength due to both topological and temporal characteristics. Paradoxically, burstiness was found to promote efficient diffusion at small scales but to hinder it on large scales.

Geography of Ties

Several studies have been made of geographical scaling laws for friendship (23–25). Liben-Nowell et al. (26) analyzed a blogging network and the relationship between friendship and distance. They concluded that friendship correlates more strongly with a person's rank, a measure of the number of closer people, than simply with the distance between people:

$$P_{ij} \propto \frac{1}{\sum_{k:r_{ik} < r_{ij}} p_k}, \quad [1]$$

where P_{ij} is the probability of friendship between agents in two distinct grid cells i and j , p_k is the population in cell k , and r_{ik} is the distance from cell i to cell k . The quantity on the right-hand side is the rank of an agent in i with respect to agent in j ; it is a measure of the number of people located between i and j . Thus, the spatial distribution of a person's friends is now strongly dependent on the local population density, with the effect that two people separated by a given large distance are more likely to be friends in a rural region than in a dense, urban environment. It also was noted that friends could be classified into two distinct types. One group comprises rank-based friends chosen because of geographical proximity (e.g., a shared workplace) in accordance with rank scaling. The authors also observed a “background” probability that an agent might be friends with any other randomly chosen agent from across the country; in this case, friendships are independent of geography. Further, these two types of friends were found to exist in a ratio of 2.5 distance-independent friendships to 5.5 rank-based friendships among the user average of 8 friends. In our simulation, we apply this model of friendship to high-resolution population density data derived from census data (27, 57) (*Materials and Methods*).

Passive Recruitment

In addition to the branching, temporal, and friendship mechanisms described above, we investigated the role of two other mechanisms: passive recruitment and mobility. We describe these in turn below. The data collected during the DARPA Network Challenge recorded each person who officially registered with the MIT team, allowing them to recruit others and to report findings. However, this is only a subset of all the people who became aware of the search; the record of sign-ups gives a measure of the number of new recruits each individual successfully invites, but not the larger hidden network of individuals who search but do not sign up or recruit others. We refer to this process as passive recruitment, quantified by the number of passive recruits per individual n_{pass} . There was a considerable number of single nodes reporting findings directly, separate from any recruitment tree (5 of the 10 balloons were reported in this way). The reported traffic to the MIT team's Web site of more than 100,000 individuals given only 4,400 sign-ups is further evidence of an unreported, hidden network. This suggests that in addition to the observable chain of individuals who actively recruit others after being recruited themselves, there is a supplementary process whereby individuals become aware of the search and the associated incentives and will report any balloons in their vicinity, yet are not sufficiently motivated to recruit others. This may be the result of low affinity with the potential parent node from which they became aware of the search (28–30). The effect of mass media and possibly word of mouth would manifest itself similarly. By definition, the participation of these individuals is difficult to measure unless they report a balloon, but given the large number of submissions attributed to single nodes that were not part of a recruitment chain, we expect that a sizable number of passive recruits also were participating in the search. This process gives rise to an interesting multiplicative factor, separate from the exponential growth of recruits due to branching.

Although the number of passive recruits a person can mobilize is intrinsically hard to quantify, a good measure of this number is the number of friends of a user of a typical social networking service, such as Facebook. The average degree of the entire global network is around 200, with a large range, but it is observed to be up to 400 among the most active users (31). It is these users who have been observed to drive such viral recruitment processes (21). As discussed in *SI Appendix*, we study this parameter for a wide range of passive recruits: however, because we aim to test the

potential recruits has a greater likelihood of being searched later. However, this is exacerbated by the fact that the population is far from homogeneously distributed, demonstrating strong spatial autocorrelation (*SI Appendix*). Rather, a sparsely populated cell likely is surrounded by other sparsely populated cells; thus, there are considerably fewer opportunities for recruitment into that cell from its neighbors. Conversely, well-populated cells in urban areas experience the opposite effect. We have highlighted the extent to which a balloon becomes more easily found as it is moved to a location with higher density. In this context, it may be seen that a few of the chosen balloon locations were in challenging locations but that overall success was expected.

Searchability, Blendability, and Findability. To draw more general conclusions about the probability of searching a location, we move away from the specific balloon locations. We now can measure the ease with which each cell may be searched over the course of many different search realizations. With this in mind, we map the searchability (s) of each cell i as

$$s_i = \frac{n_i^{\text{searched}}}{N}, \quad [2]$$

where n_i^{searched} is the number of instances in which someone is recruited in cell i out of N searches ($N = 10,000$ for the following results). We see (Fig. 3) that cells located in dense metropolitan areas are easily searchable because there are many more potential searchers to recruit in those cells, whereas the opposite is true for sparsely populated areas. Fig. 3 (black points) demonstrates this saturating trend above cell population $\approx 10^4 \text{ km}^{-2}$. This is far from a linear mapping, as some places are highly searchable despite having only intermediate population. Adding more people to a cell located in a small town increases the searchability a great deal; however, the payoff for adding more people to a cell in a large city is negligible.

Intuitively, we also might expect an added difficulty in locating a target in a region of high population density, such as Manhattan, despite its density-driven high searchability. We model this difficulty to locate a target at a given place successfully by the blendability b_i of a cell i . There are (at least) two distinct sources for this difficulty. One is a characteristic of the city itself: the increased density leads to increased complexity of the physical urban environment (37–39), providing more possibilities for a target to be concealed (e.g., an adobe house in Santa Fe, NM vs. a skyscraper

in Manhattan). The other contributor to the degree of blendability of a location comes from the individual perspective: sensory overload in busy places, leading to inattentive blindness (40–42), diminished feelings of individual responsibility to report sightings in large crowds (43, 44), and/or reduced cognitive processing ability due to stress (45). In all the above cases, we may infer safely that the larger the population of a cell p_i , the larger its blendability b_i . We assume that $b_i \sim p_i^\beta$, similar to how other urban indicators scale with population (46, 47), e.g., wages and crime with $\beta = 1.25$. We also consider walking speed with $\beta = 1$ in *SI Appendix*.

Thus, if we define the blendability per person, we obtain

$$\hat{b}_i = \frac{p_i^\beta}{p_i} = p_i^{\beta-1}, \quad [3]$$

and we rescale $\{\hat{b}_i\}$ to lie in the range $[0, 1]$.

The tension between the searchability and blendability of places is modeled by the findability per cell i :

$$f_i = \frac{s_i}{b_i}. \quad [4]$$

Again we scale $\{f_i\}$ to be in the range $[0, 1]$ and plot Eq. 4 as the red points in Fig. 3 for $\beta = 1.25$ (we repeat the analysis using $\beta = 1$ in *SI Appendix*). We isolate a regime of high findability defined by a value greater than 0.8, which corresponds to the gray-shaded region, with a population density in the range of 1,100–13,500 km^{-2} . We emphasize that the exact findability threshold is not important, as the blendability is defined only up to a constant. Comparing midtown Manhattan (population density, 36,627 km^{-2}) with nearby Asbury Park, NJ (population density, 4,975 km^{-2}), we see (Fig. 3) that, counterintuitively, it may be easier to hide in the former than in the latter. The origin of this result is that if $\beta \geq 1$, then for a large p_i , the rate at which searchability increases with population is insufficient to overcome the rate at which blendability increases with population; thus, the findability is maximized in places of intermediate density (which happens when b is an increasing function of p_i ; a detailed derivation of the condition for β to display this behavior may be found in *SI Appendix*).

Finally, in Fig. 4 we visualize the variation of the searchability, blendability, and findability in the vicinity of Manhattan and Asbury Park (see *SI Appendix* for a full map of the continental United States). Manhattan has extremely high population density (strong red shading in upper circle), leading to high searchability. However, this again is counteracted by a very high blendability, resulting in a relatively lower findability than intuitively expected (medium shading in findability map). In contrast, the intermediate population density in Asbury Park leads to a fairly high searchability (medium shading in lower circle). However, because the blendability is very low (blue shading of blendability map), the findability is very high. In general, it can be seen that areas of intermediate population have high findabilities.

Discussion

Our goal is to understand the practical limits of time-critical social mobilization, and to do so in light of contemporary wisdom about the factors that may affect it: the structure and geographical distribution of social ties, the branching and temporal dynamics of information diffusion via social media, and urban economics. Where possible, we used parameters measured from large-scale empirical results to create a realistic, high-resolution simulation of a mobilization scenario akin to the DARPA Network Challenge.

The popular reaction to the DARPA Network Challenge was that it would be impossible without mass media. Our main finding is that success actually is expected only with social media and under realistic parameters. Assuming an initial burst of motivated

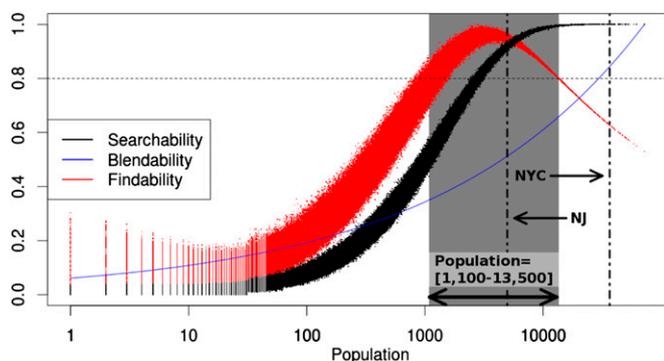


Fig. 3. Scatter plot of searchability (black), population-scaled blendability function (blue), and their ratio, defined as findability (red), as a function of population for all 5,060,288 cells. The shaded region marks the range of population density for which cells have a findability greater than 0.8. The vertical dashed lines represent midtown Manhattan, NY, and Asbury Park, NJ. [The cells within 15 km of the starting cell at MIT have been removed because they are extraordinarily searchable due to their privileged position close to the source of the search (*SI Appendix*).]

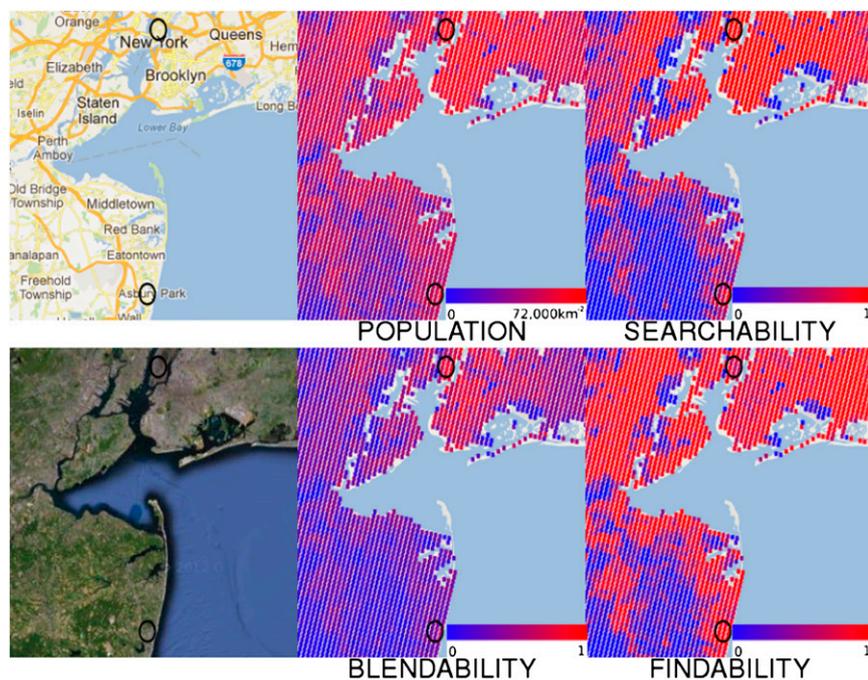


Fig. 4. Maps of Manhattan and Asbury Park showing variation in population density, searchability, blendability, and findability, as well as underlying satellite and road maps. Black circles indicate locations of Manhattan, NY (*Upper*), and Asbury Park, NJ (*Lower*). Population density is on a logarithmic scale; all others are on a linear scale. Unpopulated cells are not shown.

individuals, success takes place despite the branching factor being lower than the critical point.

That being said, we found two sobering and instructive qualifiers. First, despite the average completion time coinciding with the experience in the actual challenge, the long tail distribution of completion time suggests that the risk of failing to locate the targets within a short time frame also is significant. Second, the challenge lies at the limits of what social mobilization can achieve. Success relies on all parameters being at their practical limits: highly connected individuals must be motivated to propagate the message to many friends and to mobilize people in distant locations, overcoming the local trapping of diffusion in highly dense areas.

Our results have implications for the use of social mobilization to achieve time-critical tasks, such as mapping crises in real time or conducting search-and-rescue operations over large geographies. Novel mobilization mechanisms need to focus on incentivizing the elements of the network that are most conducive to successful mobilization: highly connected people with distant friends, and rapid reaction time. These characteristics may be exploited in a new measure of influence. One may envisage variants of the winning team's recursive incentive strategy that provide network centrality, distance, and/or time-sensitive rewards to recruit such influentials.

We studied the tension between the benefits and difficulties of searching for physical objects in highly populated areas by defining measures of searchability, blendability, and findability. On one hand, hiding in a sparsely populated town makes it less likely for someone from that town to be recruited to find the target. However, as soon as someone gets recruited, identification becomes trivial. On the other hand, in a city with high density, one might be able to “blend into the crowd.” Our results show that, short of hiding “in the middle of nowhere,” one's best bet is to hide “in plain sight.” The role of human mobility in the context of blendability is not completely clear and certainly warrants further investigation using a more detailed treatment. Models of geographical ties and mobility should explicitly account for variations in density, particularly the deviation from pure rank

scaling (48) demonstrating the increased likelihood of city-based users to have longer-range ties.

It is worthwhile to put our work in the context of search in social networks. Milgram's landmark “small-world” experiment showed that, in principle, people are findable using six hops on the global social network (49), a result that has been reaffirmed in the Internet age (29). However, Milgram searchability relies on people's ability to form a reliable estimate of distance to the target to exploit the large jumps afforded by small-world networks (50–52). For example, if the target is known to be a professor who lives in Kyoto, one might try to reach her through a friend who lives in Tokyo, Japan, as he or she is more likely to know someone who lives in Kyoto, who in turn may know someone in academia, and so on. However, if information about a target is scarce (e.g., searching for a person in an entire country based only on a mug shot), we cannot rely on distance estimates. In other words, the problem becomes that of “uninformed” (i.e., “blind”) search (53) and thus requires large-scale social mobilization. Endeavors such as the Tag Challenge (54), in which search may benefit from partial knowledge of target location, require elements of both uninformed and heuristic search, a topic that deserves further study in the context of social mobilization.

Our work is not without limitations. First, we focused on mobilization processes that are fully driven by social ties. In reality, however, mobilization often also benefits from the use of mass media [e.g., AMBER (America's Missing: Broadcast Emergency Response) Alerts distributed via radio stations and cable television] and social media hubs (e.g., highly followed blogs or Twitter accounts). Surely, such media can accelerate social mobilization, as they complement the social diffusion process and seed it over large areas (55, 56). Second, our work is limited by our use of a simple model of human mobility. For a task like the Network Challenge, this is unlikely to be a problem. However, for scenarios that involve searching for mobile targets, as was the case in the Tag Challenge (54), more sophisticated models of human mobility should be incorporated (32).

Materials and Methods

Materials. High-resolution population data were taken from publicly available sources (57) based on US census data (27). These data comprise 7,820,528 cells, each with an area of 1 km², of which 5,060,288 are populated (i.e., 2,760,240 are empty). Empirical data from ref. 6 were used to parameterize the branching factor power law distribution as follows. We exclude the first generation of recruitment directly from the MIT team, as this is anomalously high (164 child recruits) and, because of the team's own unique personal association with the task, likely to be atypical. We also exclude 611 single nodes who signed up directly and did not recruit any child nodes; we assume these are examples of passive recruits who signed up independently. The distribution of the branching factor among a subset of the remaining nodes is described by a power law with mean $\langle R_o \rangle = 0.89$ (SI Appendix).

Methods. A set of seed nodes located at MIT is chosen, the number of which matches those initially recruited by the MIT team. All these nodes are active, i.e., they continue to recruit themselves in contrast to passive recruits, which do not continue the recruitment tree. Each newly activated node looks around in its vicinity (within a distance radius of r_{mob}) and reports any balloon it sees

within that radius. Each newly activated node also chooses an outdegree (a constant number n_{pass} of "passive" recruits and a power law-distributed number n_a of "active" recruits, where n_{pass} is drawn from the distribution seen in the MIT Red Balloon team's recruitment data). Each chosen friend, passive or active, is chosen to be rank-based with respect to geography, using 1-km² population density data across the United States, with probability 5.5/8 and uniform overpopulation with probability 2.5/8. Each active new recruit selects a delay, chosen from a log-normally distributed waiting time distribution with mean 1.5 d and SD 5.5 d (21), becomes activated, and completes its own recruitment after that time delay.

ACKNOWLEDGMENTS. We thank Wei Pan for assistance with DARPA Balloon Challenge data, Galen Pickard for useful comments, and Mohammed Mekki for support with high performance computing resources. M.C. acknowledges support from the National Science Foundation under Grant 0905645, from DARPA/Lockheed Martin Guard Dog Program under PO 4100149822, and the Army Research Office under Grant W911NF-11-1-0363. E.M. acknowledges funding from Ministerio de Educacion y Ciencia (Spain) through projects FIS2006-01485 (Modelización, Simulación y Analisis de Sistemas Complejos) and FIS2010-22047-C05-04.

- Okolloh O (2009) Ushahidi or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory Learning and Action* 59(1):65–70.
- Hellerstein J, Tennenhouse D (2011) Searching for Jim Gray: A technical overview. *Commun ACM* 54:77–87.
- González-Bailón S, Borge-Holthoefer J, Rivero A, Moreno Y (2011) The dynamics of protest recruitment through an online network. *Sci Rep*, 10.1038/srep00197.
- Ghoniim W (2012) *Revolution 2.0: The Power of the People Is Greater Than the People in Power* (Houghton Mifflin Harcourt, New York).
- Tang J, et al. (2011) Reflecting on the DARPA Red Balloon Challenge. *Commun ACM* 54:78–85.
- Pickard G, et al. (2011) Time-critical social mobilization. *Science* 334(6055):509–512.
- Cebrian M, Coviello L, Vattani A, Voulgaris P (2012) Finding red balloons with split contracts: Robustness to individuals' selfishness. *Proceedings of the ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York), pp 775–788.
- Naroditskiy V, Rahwan I, Cebrian M, Jennings NR (2012) Verification in referral-based crowdsourcing. *PLoS ONE* 7(10):e45924.
- Anonymous (September 1, 2012) Monitor: Six degrees of mobilisation. *Economist: Technology Quarterly*: 7.
- Rahwan I, et al. (2013) Global manhunt pushes the limits of social mobilization. *IEEE Computer*, 10.1109/MC.2012.295.
- Golub B, Jackson MO (2010) Using selection bias to explain the observed structure of Internet diffusions. *Proc Natl Acad Sci USA* 107(24):10833–10836.
- Bakshy E, Hofman J, Mason W, Watts D (2011) Everyone's an influencer: Quantifying influence on twitter. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining* (Association for Computing Machinery, New York), pp 65–74.
- Bond RM, et al. (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298.
- Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341.
- Epstein JM (2009) Modelling to contain pandemics. *Nature* 460(7256):687.
- Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A (2004) Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys Rev Lett* 92(17):178701.
- Wang P, González MC, Hidalgo CA, Barabási AL (2009) Understanding the spreading patterns of mobile phone viruses. *Science* 324(5930):1071–1076.
- Balcan D, et al. (2009) Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA* 106(51):21484–21489.
- Nicolaides C, Cueto-felgueroso L, Gonzalez MC, Juanes R (2012) A metric of influence spreading during contagion dynamics through the air transportation network. *PLoS One* 7(7):e40961.
- Brockmann D, Hufnagel L, Geisel T (2006) The scaling laws of human travel. *Nature* 439(7075):462–465.
- Iribarren JL, Moro E (2009) Impact of human activity patterns on the dynamics of information diffusion. *Phys Rev Lett* 103(3):038702.
- Mirtello G, Moro E, Lara R (2011) Dynamical strength of social ties in information spreading. *Phys Rev E Stat Nonlin Soft Matter Phys* 83(4 Pt 2):045102.
- Krings G, Calebrese F, Ratti C, Blondel VD (2009) Urban gravity: a model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment* (7):L07003.
- Ratti C, et al. (2010) Redrawing the map of Great Britain from a network of human interactions. *PLoS ONE* 5(12):e14248.
- Barthélemy M (2011) Spatial networks. *Phys Rep* 499:1–101.
- Liben-Nowell D, Novak J, Kumar R, Raghavan P, Tomkins A (2005) Geographic routing in social networks. *Proc Natl Acad Sci USA* 102(33):11623–11628.
- US Census Bureau (2001) Census 2000: Census Tract Cartographic Boundary Files. Available at www.census.gov/geo/www/cob/tr2000.html. Accessed March 15, 2013.
- Watts DJ, Dodds PS, Newman ME (2002) Identity and search in social networks. *Science* 296(5571):1302–1305.
- Dodds PS, Muhamad R, Watts DJ (2003) An experimental study of search in global social networks. *Science* 301(5634):827–829.
- Iribarren JL, Moro E (2011) Affinity paths and information diffusion in social networks. *Soc Networks* 33(2):134–142.
- Ugander J, Karrer B, Backstrom L, Marlow C (2011) The anatomy of the Facebook social graph. arXiv:1111.4503.
- González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782.
- Cho E, Myers S, Leskovec J (2011) Friendship and mobility: User movement in location-based social networks. *Proceedings of the 17th ACM International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York), pp 1082–1090.
- Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C (2012) A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE* 7(5):e37027.
- Bagrow JP, Lin Y-R (2012) Mesoscopic Structure and Social Aspects of Human Mobility. *PLoS ONE* 7(5):e37676.
- Phithakkittukoon S, Smoreda Z, Olivier P (2012) Socio-geography of human mobility: A study using longitudinal mobile phone data. *PLoS ONE* 7(6):e39253.
- Rapoport A, Hawkes R (1970) The perception of urban complexity. *J Am Inst Plann* 36:106–111.
- Rosvall M, Trusina A, Minnhagen P, Sneppen K (2005) Networks and cities: An information perspective. *Phys Rev Lett* 94(2):028701.
- Lee S, Holme P (2012) A greedy-navigator approach to navigable city plans. arXiv:1205.0537.
- Mack A, Rock I (1998) *Inattentive Blindness* (MIT Press, Cambridge, MA).
- Most SB, et al. (2001) How not to be seen: The contribution of similarity and selective ignoring to sustained inattentive blindness. *Psychol Sci* 12(1):9–17.
- Miller GA (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol Rev* 63(2):81–97.
- Milgram S (1970) The experience of living in cities. *Science* 167(3924):1461–1468.
- Darley J, Latane B (1968) Bystander intervention in emergencies: Diffusion of responsibilities. *J Learn Sci* 8:377–383.
- Lederbogen F, et al. (2011) City living and urban upbringing affect neural social stress processing in humans. *Nature* 474(7352):498–501.
- Bettencourt LM, Lobo J, Helbing D, Kühnert C, West GB (2007) Growth, innovation, scaling, and the pace of life in cities. *Proc Natl Acad Sci USA* 104(17):7301–7306.
- Bornstein M, Bornstein H (1976) The pace of life. *Nature* 259:557–559.
- Backstrom L, Sun E, Marlow C (2010) Find me if you can: Improving geographical prediction with social and spatial proximity. *Proceedings of the 19th International Conference on World Wide Web* (Association for Computing Machinery, New York), pp 61–70.
- Milgram S (1967) The small world problem. *Psychol Today* 61:60–67.
- Kleinberg JM (2000) Navigation in a small world. *Nature* 406(6798):845.
- Kleinberg JM (2000) The small-world phenomenon: An algorithm perspective. *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York), pp 163–170.
- Adamic LA, Adar E (2005) How to search a social network. *Soc Networks* 27:187–203.
- Knuth D, et al. (1973) *Sorting and Searching, The Art of Computer Programming* (Addison-Wesley, Reading, Massachusetts), Vol 3.
- Firth N (2012) Social media web snares 'criminals.' *New Sci* 2860:18.
- Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. *Proc Natl Acad Sci USA* 105(41):15649–15653.
- Myers SA, Zhu C, Leskovec J (2012) Information diffusion and external influence in networks. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York), pp 33–41.
- National Oceanic and Atmospheric Administration, National Climatic Data Center (2008) Gridded 1 km Population for the Conterminous United States. Available at www.ncdc.noaa.gov/oa/climate/research/population. Accessed March 15, 2013.