Guard Your Connections: Infiltration of a Trust/Reputation Based Network

Malik Magdon-Ismail
Rensselaer Polytechnic Institute
Computer Science Department
Troy, NY 12180
magdon@cs.rpi.edu

Brian Orecchio
Rensselaer Polytechnic Institute
Computer Science Department
Troy, NY 12180
oreccb@rpi.edu

ABSTRACT

We study infiltration of a trust/reputation based network. At every time step, the agent solicits a connection request (friend request). The goal of the agent is to amass as many such connections as possible to further its goals. Our model for such an infiltration of a network relies on two properties of the actors in the network. They desire more links (an ego effect); they are more likely to connect to trusted or credible nodes (the trust by reference effect). We demonstrate the following properties of this infiltration.

(i) The trust by reference effect is critical. If agents are not trusting enough, then the network is robust to infiltration; however, with logarithmically more trust, the process phase transitions to significant infiltration.

(ii) The network structure is important. If the trust effect is small, then well clustered networks (typical social networks) are *easier* to infiltrate; when the trust effect is larger, then networks with large expansion (for example Erdös-Rényi random graphs) are easier to infiltrate.

(iii) The algorithm used by the agent plays a significant role in success of the infiltration. Random connection requests are much less successful than even simple greedy strategies, even if those greedy strategies are restricted to only using local information.

Author Keywords

Cascade, phase transition, social networks, LinkedIn

INTRODUCTION

The motivation for this paper came from the following anecdote. A student asked one of the authors for advice regarding a company that contacted him claiming to be hiring in "stealth mode"; we'll call this company SC for the "stealth company". SC doesn't have a web presence, nobody knows they are hiring, and he (the student) is extremely attractive to SC because of his skills. SC wishes to take the interview further. The student seeks advice from the author as to whether to pursue further. The author's response: "Don't waste your time with this thing, it's just spam." The retort from the student was interesting:

"I don't think it is spam because SC has a presence on LinkedIn® *and is distance three from me. So SC is linked to a connection of one of my connections."*

SC is trying to infiltrate a trust based network, and leverage this trust to exploit some scam. For example, they will ask for a $200 application fee. SC is relying on the fact that since it has links to people you trust, it has gained some credibility in your eyes, increasing their chances of success. Further, if they may link to you during this process, then they have gained more credibility in the eyes of other network actors.

Clearly, trust based infiltration is real, and being exploited by deviants. Further, this type of trust infiltration is not limited to networks like LinkedIn®, nor to such malicious endeavors as scams. Consider the academic collaboration network in which there are authors of varying status. It may be desirable to collaborate with high stature authors, to build up one's own stature. To accomplish this, authors may try to collaborate with other more prominent authors, thereby gradually increasing their credibility and hence likelihood of collaborating with an even higher stature author (the ultimate goal). A less well known author tries to "infiltrate" into the trust network of academic collaboration. Similarly, it is easier to market a product using well connected nodes than less well connected nodes, therefore it is desirable to have under ones control a set of well connected nodes.

We study infiltration of a trust based network. We introduce a natural model for studying this process. Our model is based on two main social forces driving an actor to form links: *ego* and *trust*. Actors have ego which is serviced by being more popular. Actors desire more links; that's the ego effect. Links confer a concrete status that can be used to satisfy an actors ego-need for being popular. SC gains credibility by being connected to more nodes in the network. As SC gets connected to more and more nodes, the likelihood that any particular node will desire to connect to SC increases. That is the trust effect. The ego effect is in some sense what is responsible for the infiltration starting. The trust effect is what is responsible for the infiltration gaining steam.

Though our model is relatively simple, its analysis is surprisingly difficult. Nevertheless within this simple model,

WebSci 2012, June 22–24, 2012, Evanston, Illinois, USA.

many interesting behaviors appear. The goal of this paper is to build the case in support of the following conclusions.

(i) **There are critical trust thresholds.** As n, the number of nodes in the network, gets large, we ask what fraction of the network can the infiltrator ultimately link to. We show the existence of a narrow transition regime where if the trust effect is small, below a threshold c/n, then asymptotically as $n \to \infty$ the expected fraction of the network that gets infiltrated is zero - the network is resistant to infiltration. Conversely, if the trust effect is larger than $c\log(n)/n$, then a constant fraction of the network can be infiltrated – the network succumbs to the infiltrator. We demonstrate this effect analytically for the Erdös-Rényi random graph and a random infiltration strategy where the analysis is tractable, though non-trivial. This random graph model is not necessarily a good model of a social network, but it displays all the qualitative effects which we demonstrate empirically using real social networks (a small portion of a LinkedIn network and the Enron email network).

(ii) **The network structure is important.** We consider the infiltration process on the Enron network, a typical social network, and compare it to two random graph models: an Erdös-Rényi random graph with the same number of edges as the social network and a random graph with the same degree distribution as the social network. In general, for small values of the parameter governing the trust effect, a uniform random graph is more resistant to infiltration than the social networks we tested, indicating that structural properties of social networks (such as strong triadic closure) are of benefit to the infiltrator. However, when the trust effect is large, random graphs are much easier to infiltrate, owing to their large expansion: if the trust effect is large, then it is easy to infiltrate a neighbor of an infiltrated node, and so what is desired is that the neighborhood of the infiltrated set be large (at all times) to obtain the most efficient infiltration - i.e. graphs with larger expansion are easier to infiltrate.

(iii) **The strategy used by the infiltrator is important.** We study analytically the random strategy in which SC attempts to link to a randomly selected node at every time step. This is the simplest strategy. It is most amenable to mathematical analysis and also requires least knowledge for SC to implement. Empirically, we compare the random strategy with a simple greedy strategy in which SC selects at every time step the node who is most likely to accept the link. Though both strategies have similar qualitative features (existence of critical trust thresholds, network structure dependence), quantitatively the greedy strategy is much more effective at infiltrating. We also consider a strategy that takes advantage of only local knowledge. This strategy also has critical trust thresholds but is less sensitive to network structure.

The take home message is that the network's structure and SC's strategy play an important role in how successful SC will be. However, the tendency of actors to trust by reference plays an especially important role. If actors are just slightly more trusting than a critical threshold, the network transitions from impenetrable to penetrable. It is thus important to guard ones connections because they not only affect your own (local) status but how other's will view a potential deviant in the network, leading to a potentially global cascade. *A little misplaced trust can go a long way to aid such deviants as SC.*

Related Work

We study a sequential cascading process on a social network. Cascading processes in social networks have a large history. The simplest models are the linear threshold model where a node gets infected if a fraction of its neighbors are infected, which is appropriate for studying (for example) the diffusion of innovations [13]. In the independent cascade model, in which every infected node has one chance to infect its neighbors with a probability is the simplest of a sequence of network infection models that have its roots in the mathematical epidemiology literature (SI, SIR, SIRS, etc.) [8, 2]. Percolation models have been useful in studying the cascading of failures in networks [4, 3, 6]. Percolation and independent cascade infection are related, and the trick used in [10, 7] converts the infectious epidemic problem in the simple independent cascade model to the study of the connected components in a percolation problem.

Within all these models of cascading processes, there are two typical goals. Assuming some nodel of the network and on how an initial node set is infected, determine the extent of the spread. Given a network and model for the infectious process, determine an initial set of nodes to infect so as to maximize the spread [9]. A related problem is that of immunization: given network and a set of initially infected nodes, determine who to immunize (within a budget) so that the spread is minimized [1].

Our model can be viewed as a model of infection, where we say that a node is infected if it is linked to SC. Our model then captures several parts of different models in the literature. We are concerned with similar issues here in a sequential model in which the agent (SC) may dynamically update its strategy for infection. The instance of our model that we analyze is similar to the independent cascade model (SI model) due to the link by reference part of the model where each node that is your neighbor has a probability p_t to "infect" you by causing you to link to SC. In addition, our process also has the ego effect. Further, and more importantly, is the strategic aspect of the agent. The "infection" does not proceed on the network merely according to an endogenous process; there is an exogenous agent who may specify at each time step which node is to be requested. This introduces a strategic aspect to the agent, which changes the dynamics. We study three particular strategies for the agent (random, greedy and local strategies) but the setting offers interesting algorithmic challenges depending on how much information the agent is allowed to have.

A related, though different topic, is the infiltration of a net-

work, typically a sensor network, using multiple identities. This is usually known as a Sybil attack, and properties of such attacks, behavior of networks under such attacks and defense against such attacks have been studied (see, for example, [11, 14]). Other than the fact that the Sybil attacker attempts to infiltrate using trust in multiple identities, there is little similarity between the processes studied there and our process for infiltration of a social network.

TRUST BASED INFILTRATION

We first describe some useful notation, and then present the basic model.

Notation

The social network is a graph $G = (V, E)$ where the vertex set is $V = \{v_1, \ldots, v_n\}$ (there are the n nodes of the network) and the edge set is $E = \{e_1, \ldots, e_m\}$ (there are the m (undirected) edges in the network). We use $[n]$ for the set $\{1, \ldots, n\}$. The stealth company who is attempting to infiltrate the network will be referred to as SC. In the process of the infiltration, SC will have infected a subset S_t of nodes at time-step t. We assume that $S_0 = \emptyset$. $\delta(v)$ is the degree of node v and we will use $K_t(v)$ to denote the number of v's neighbors are also SC's neighbors (the number of common neighbors); $K_t(v) = |N(v) \cap S_t|$, where $N(v)$ is the neighborhood of v.

Model

We now present a general model for infiltration of a network. We assume for simplicity that the base network (the subgraph induced by $v_1, \ldots, v_n$) is static and the only edges that are added are connections to SC. This assumption is made since in most of the networks where this social process can happen, most nodes will not be destroying their relationships because there is a benefit in keeping them. Also, the infiltration process would most likely occur in a certain time-frame in which the structure of a well developed network would not change dramatically.

The process starts out at time 0 with no nodes linked to SC. At time-step t, there is a subset $S_t \subseteq V$ of nodes that are linked to SC. At time-step $t + 1$, SC will send connection requests to a set of nodes $\{v_j\}$ based on a certain budget of requests and a strategy for which nodes to actually request. We will simplify further to one connection request at each time step.[1] Each node requested, v_j where $j \in [n]$, will accept the connection with probability $P(v_j|S_t)$. This process continues for specified number of time-steps or until every node has been requested once. If a node rejects a connection request, no further connection requests may be sent to this node. Let $P(v_j|S_t)$ be the probability that node v_j accepts the connection request given the "infected set" S_t. We assume that $P(v_j|S_t)$ is made up of two independent processes; a probability P_T to accept based on the credibility of SC (the trust effect) and a probability P_E to accept based on an ego-factor that drives nodes to desire more links. Since these two processes are independent, the overall probability to accept the link is given by

$$P(v_j|S_t) = 1 - (1 - P_E)(1 - P_T). \quad (1)$$

We postulate that the trust effect P_T is a function of v_j's relationship to SC through S_t. The "closer" SC is to v_j, the higher P_T should be. The ego term P_E is independent of S_t and only depends on v_j's desire to accumulate more links. We anticipate that this desire to accumulate more links should be a function of $\delta(v_j)$, v_j's degree. The following two properties are natural.

(i) Monotonicity of P_T: Fix v_j. If $S_t \subseteq S'_t$ then $P_T(v_j|S'_t) \geq P_T(v_j|S_t)$. (SC's credibility can only increase as it infects more and more nodes.)

(ii) Monotonicity of P_E: $P_E(v_j|S_t)$ is decreasing in $\delta(v_j)$. (A node's ego-desire to accumulate more links is decreasing or not-increasing as it accumulates more and more links.)

The model allows us to specify the ego and trust effects on a node-by-node basis. Such generality is not needed to understand the qualitative features of the model. In our theoretical analysis we will simplify to all nodes having the same ego effect. For our experiments, we will use a simple, intuitive functional form for the ego effect that takes into account degree inhomogeneity in the network - lower degree nodes are more succeptible to the ego effect. For P_E, we used a simple functional dependence of the form

$$P_E(\delta(v_j)) = \frac{p_e}{(1 + \delta(v_j))^\alpha} \quad (2)$$

where p_e and α are parameters. It is clear that P_E is decreasing in $\delta(v_j)$. For our experiments, we set $p_e = \frac{1}{2}$ and considered different choices of α. With the exponent $\alpha > 0$, P_E decreases with the node's degree. For the trust effect P_T, we also use a simple model based on independent effects using a "connection by reference" model. For each node that is a common neighbor of v_j and SC, there is a probability p_t for v_j to accept the connection to SC. We assume that each such connection by reference is independent so

$$P_T(v_j|S_t) = 1 - (1 - p_t)^{K_t(v_j)} \quad (3)$$

The monotonicity of P_T easily follows because if $S_t \subseteq S'_t$ then $K'_t \geq K_t$. Our model is perhaps the simplest that one could construct; the closeness of SC to v_j and hence SC's credibility in the eyes of v_j is simply a function of the number of common neighbors. A more complex model might consider, for example, the shortest paths from SC to v_j through each member of S_t and so on. However, this simple model already leads to very interesting mathematics and non-trivial conclusions. Further, this model for P_T based on the number of common neighbors is leveraging the fact that triangles tend to close in social networks (triadic closure [12]); we are relying on the trust rationale for triadic closure in this paper. Finally, we arrive at the probability that a connection request to v_j is successful,

$$P(v_j|S_t) = 1 - \left(1 - \frac{p_e}{(1 + \delta(v_j))^\alpha}\right)(1 - p_t)^{K_t(v_j)} \quad (4)$$

[1] One can show that for trust based infiltration within our model, it is always better to request one connection at a time, provided there is no time constraint.

Under this model, SC tries to select a sequence of connection requests to achieve some objective. Several interesting objectives are possible.

(i) Maximize $\mathbb{E}[|S_n|]$, the expected number of nodes that will be linked to.

(ii) Given a target set C, maximize $\mathbb{E}[|C \cap S_n|]$, the expected number of nodes in C that will be linked to.

(iii) Given a target set C, maximize $\mathbb{P}[|C \cap S_n| > 0]$, the probability to infect some node in the target set.

Determining the optimal strategy for SC to attain either of these objectives given the initial social network and the parameters p_e, α, p_t is an interesting and challenging problem. For the remainder of this paper we will focus on $\mathbb{E}[|S_n|]$ and three simple strategies for SC (random, greedy and local connection request strategies). From the actor point of view, we will focus on the parameter p_t, the tendency for an actor to accept a connection via reference. From this perspective, our main goal is to show the existence of a critical threshold within a narrow range such that on the one hand if p_t is small (below this range), $\mathbb{E}[|S_n|]$ is essentially zero; if p_t is large (above this range) then $\mathbb{E}[|S_n|]$ is proportional to n, i.e. a constant fraction of the network gets infected.

RANDOM AND GREEDY STRATEGIES

There are many strategies the SC can utilize when making connection requests. In particular, we will consider three of these possible strategies. For each strategy, SC sends out a single connection request at each time-step. The random strategy selects a node uniformly at random from among all nodes that have not yet received a connection request. The advantage of this strategy is that it requires minimal information - only what nodes are in the network.

The greedy strategy selects the node $v \in V \setminus S_t$ that maximizes the probability to accept the connection $P(v|S_t)$. This greedy strategy can be efficiently implemented in simulation because it only requires one to maintain and update $K_t(v)$ for all nodes v. If a node accepts a link, then we only need to update $K_t(u)$ for nodes in $N(v)$, incrementing by 1. A priority queue with two way pointers to the node indices can be used to efficiently have access to the node that has highest probability to accept a connection request. The problem with the greedy strategy is that SC needs some way to estimate the network structure. This may be an unreasonable assumption in a practical setting but can happen in some cases.

Another strategy is more practical by taking advantage of local knowledge of the infiltrator's connections. In particular, it selects a node $v \in V$ which is a connection of the nodes in SC's neighborhood (second level connection). The implementation of this strategy requests the latest new second level connection. In a network like LinkedIn, this information is readily available and does not require SC to compute anything or use tricky heuristics.

These three algorithms are sufficient for our purposes to illustrate a reasonably powerful adversary SC who uses the greedy strategy with full network knowledge, the other extreme of a random adversary and finally an adversary who uses local knowledge of the connections of nodes in its neighborhood. However, several interesting algorithmic problems from the perspective of SC are interesting, with respect to the infiltration metrics discussed in the previous section.

(i) What is the optimal strategy given the network, and the actor parameters p_e, α, p_t and full network knowledge.

(ii) What is the optimal strategy given a network model and local information: SC only knows about the connections of the nodes in its neighborhood who it has already infected. This type of information is typically available in many social network applications.

THE STAR NETWORK

Trees are a good representation of organizational social structure. We consider the simplest such network, a star, which consists only of a root and n children. Assume that the root has ego probability P_0 and for simplicity the leaves all have ego probability P_E. Even for just this simple network, it is already an interesting problem to determine the optimal strategy for SC.

Once SC solicits the root, the network is effectively broken up into isoloted disjoint nodes. If the root accepts, the probability to convert one of the leaves is $1-(1-P_E)(1-p_t)$; if the root rejects, the probability is P_E. The only decision to be made is whether the to request the root or one of the leaf nodes. Upon requesting a leaf, the leaf may accept, in which case we have linked to 1 node and reduced the problem to a star with $n-1$ leaves, where the root probability P_0 has increased to $P_0' = 1-(1-P_0)(1-p_t)$; if the leaf rejects, again, we have reduced the problem to an identical problem with $n-1$ nodes. It is non-trivial to identify the conditions on n, P_0, P_E, p_t under which the SC requests the root.

For simplicity, we will consider a family of strategies defined by a parameter k, which corresponds to the number of leaf nodes SC requests (fixed in the strategy) followed by requesting the root node. The aspect of the optimal strategy that these strategies lack is the ability of SC to adapt to the result of each of the leaf node requests. Nevertheless, this is still an interesting class of strategies to illustrate the complexity of infiltrating even just this simple network. Let $E(k)$ be the expected number of nodes converted by strategy k. Then, after some algebra and manipulation of binomial summations,

$$\begin{aligned} E(k) &= nP_E + (1-\gamma_k)(1+np_t(1-P_E)) \\ &\quad -kp_t(1-P_E)(1-(1-p_t)\gamma_{k-1}), \end{aligned} \tag{5}$$

where $\gamma_k = (1-P_0)(1-p_tP_E)^k$. It is sub-optimal to request the root if $E(0) \leq E(1)$. We consider the strategy which requests the root if $E(0) > E(1)$, a condition that reduces to

$$\frac{P_0}{1-P_0} > \frac{P_E}{1-P_E} + p_t(nP_E - 1). \tag{6}$$

The decision on whether to request the root even for this simple network and simple strategy is already quite com-

plex. Intuitively, if P_0 is not large enough, *compared to* P_E then it is not advisable to go for the root. On the RHS, P_E and p_t are constant, and only n varies as SC requests more nodes. The strategy is to request the leaf nodes until (6) is met, where if k leaf nodes have been requested and i of them accepted, then n is replaced by $n-k$ and P_0 is replaced by $P_0' = 1-(1-P_0)(1-p_t)^i$. Only when (6) is satisfied do you go for the root node, from which point the optimal strategy is trivial.

This strategy gives us some intuition. As the network gets larger, the required hurdle for P_0' increases. Heuristically, don't go after the well connected nodes in your network unless you are sure you will get them. Rather, go after their less connected neighbors until they have brought the probability to capture the well connected node up. Only then go after the well connected node. Needless to say, this is but a heuristic intuition for the general network, for which the analysis of SC's optimal strategy is formidable. To gain further insight, we consider random graphs, not because they are a good representation of social networks, but because certain qualitative features may appear that we might empirically verify on real networks.

INFILTRATION OF RANDOM GRAPHS

The main goal here is to study the random strategy on an Erdös-Rényi random graph with edge probability γ and identify the critical threshold p_t^* in this setting. To begin we consider $\gamma = 1$, so G is the complete graph on n nodes. In this case all node degrees are equal, and so P_E is a constant over the nodes. Let $a = 1 - P_E$ and $b = 1 - p_t$. The complete graph is easy to analyze because every node is equivalent and so all connection request strategies are equivalent. We consider the random strategy. We may assume that SC sends connection requests in the order $v_1, v_2, \ldots, v_n$. At step t, let x_t be a random variable that denotes the number of nodes that link to SC after step t. Then node v_{t+1} will accept a connection with probability

$$P_{t+1} = 1 - ab^{x_t}.$$

So, x_t is the following stochastic dynamical process:

$$x_{t+1} = \begin{cases} x_t + 1 & \text{with probability } 1 - ab^{x_t}; \\ x_t & \text{with probability } ab^{x_t}. \end{cases} \qquad (7)$$

We are mostly interested in the dependence of $\mathbb{E}[x_n]$ on b. Lets now consider $\gamma < 1$. Practically, social networks are sparse in which case $\gamma = O(d/n)$ for some constant d. Since random graphs are almost regular, we may assume that P_E is once again approximately constant over vertices; alternatively, we may set $\alpha = 0$ and choose p_e to suit our needs. Again, let x_t be the number of nodes linked to SC after step t. At step $t+1$, consider the probability that node v_{t+1} accepts the connection. We may write

$$\begin{aligned} P_{t+1} &= \sum_{K=0}^{x_t} \mathbb{P}[\text{link}|K \text{ common}] \cdot \mathbb{P}[K \text{ common}] \\ &= \sum_{K=0}^{x_t} (1 - ab^K) \binom{x_t}{K} \gamma^K (1-\gamma)^{x_t - K} \\ &= 1 - a(1 - \gamma + \gamma b)^{x_t}. \end{aligned}$$

For $\gamma < 1$, the stochastic process is identical to the complete graph with parameter $b' = (1 - \gamma + \gamma b)$; this corresponds to $p_t' = \gamma p_t$. So the stochastic process with $\gamma < 1$ and parameter p_t is equivalent to the stochastic process on the complete graph with $p_t' = \gamma p_t$. Thus, it suffices to analyze the process for the complete graph and we can infer the dynamics for general γ.

Exact Analysis via Dynamic Programming

Fix b. Let $X(n, y) = \mathbb{E}[x_n | 1 - P_E = y]$. We are interested in $X(n, a)$. We derive a recursion for $X(n, y)$ as follows. The first infection occurs at time t with probability $(1-y)y^{t-1}$ – this is the probability that the first $t-1$ fail (each failure is independent with probability y), and the tth is a success. We can write

$$X(n, y) = \sum_{t=1}^{n} \mathbb{E}[x_n | \text{first success at } t](1-y)y^{t-1}.$$

Note that $\mathbb{E}[x_n|\text{first success at } t]$ is 1 plus the expected number of further successes. After the first success, each subsequent success occurs with probability $1 - yb^{1+K}$, where $K \geq 0$ is the number of current successes. Equivalently, this probability can be written $1 - (yb)b^K$, from which we observe that the ensuing process for the next $n-t$ steps can equivalently be viewed as the original process for $n-t$ steps, but starting with $1 - P_E = by$. Thus,

$$\mathbb{E}[x_n|\text{first success at } t] = 1 + X(n-t, by),$$

and we have the recursion

$$X(n, y) = \sum_{t=1}^{n} (1 + X(n-t, by))(1-y)y^{t-1}.$$

This recursion does not serve to give much intuition for the behavior of $X(n, a)$, but since the only values of y that are needed to compute $X(n, a)$ are $a, ba, b^2a, \ldots, b^n a$, this recursion immediately gives an efficient $O(n^2)$ dynamic program to compute $X(n, a)$ exactly.

Asymptotic Analysis

The following theorem shows that there is a thin critical region for p_t, the tendency to link via reference, that separates an impenetrable network from one that is "easy" to infiltrate.

THEOREM 1. *Suppose $P_E = \frac{1}{n}$ and let c be a constant.*

(i) For $p_t \leq c/n$, $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[x_n] = 0$.

(ii) For $p_t \geq \frac{c\log(n)}{n}$ with $c > 1$, $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[x_n] > 0$.

The theorem says that as long as p_t is small enough SC cannot infiltrate the network. If p_t is just a logarithmic factor larger, then SC can essentially infiltrate a constant fraction of the network, i.e. the infiltration takes place on a global scale. Though the theorem as stated applies to the complete graph, we get the analogous result for $\gamma < 1$. The corresponding thresholds are $p_t < c/\gamma n$ for impenetrability and $p_t > c\log(n)/\gamma n$ for large scale infiltration. In passing, we note that part (i) of the theorem can be strengthened to $c \leq \log n$.

In the interests of space, we will only give a sketch of the proof, leaving the full proof for a more complete version. Recall that P_t is the probability that node v_t accepts the connection. We will need the following technical lemma.

LEMMA 2. $P_t \leq P_{t-1} + P_{t-1}(1 - P_{t-1})p_t$.

Using this lemma, we obtain the first part of the theorem as follows. Assume $p_t \leq c/n$.

$$\Delta P_t = P_t - P_{t-1} \leq P_{t-1}(1 - P_{t-1})p_t \Delta t.$$

We may thus obtain an upper bound for P_t by dividing both sides by $P_{t-1}(1 - P_{t-1})$ and summing. Approximating the sums by integrals (this can be made more rigorous),

$$\int_{P_E}^{P_t} dx \, \frac{1}{x(1-x)} \leq p_t \int_1^t dt \leq \frac{c(t-1)}{n}.$$

Expanding the LHS using partial fractions, integrating and then solving for P_t, we obtain

$$P_t \leq \frac{e^{c(t-1)/n}}{n - 1 + e^{c(t-1)/n}},$$

where we used the fact that $P_E = 1/n$. To complete the argument, $\mathbb{E}[x_n] = \sum_{t=1}^{n} P_t$. Again, we may upper-bound this sum with an integral, giving

$$\begin{aligned} \mathbb{E}[x_n] &= \sum_{t=1}^{n} P_t \approx \int_{t=1}^{n} dt \, \frac{e^{c(t-1)/n}}{n - 1 + e^{c(t-1)/n}} \\ &= \frac{n}{n-1} \frac{e^c - 1}{c}. \end{aligned}$$

Thus, $\mathbb{E}[x_n]$ approaches a constant, and so $\mathbb{E}[x_n]/n \to 0$. We showed more than we needed, namely that the expected total number of infiltrated nodes is bounded by a constant.

We now give the rough intuition for the second part of the theorem. Assume that $p_t \geq c\log(n)/n$. To begin, observe that since $P_E = 1/n$. A standard calculation ([5]) shows that the probability that the first infection occurs in the first $n/2$ steps is $1 - (1 - P_E)^{n/2} \approx 1 - 1/\sqrt{e}$ (asymptotically in n). We may condition on this event and conclude that

$$\mathbb{E}[x_n] \geq \left(1 - \frac{1}{\sqrt{e}}\right) \mathbb{E}[x_n | x_{\frac{n}{2}} \geq 1].$$

It thus suffices to show that $\mathbb{E}[x_n | x_{\frac{n}{2}} \geq 1]$ is proportional to n (asymptotically in n). From now on, condition on this event that $x_{\frac{n}{2}} \geq 1$. For $t \geq n/2$, we then have that

$$P_{t+1} = 1 - (1 - P_E)(1 - p_t)^{x_t} \geq 1 - (1 - p_t)^{x_t}.$$

and $x_t \geq 1$. Since P_{t+1} is the success probability for incrementing x_t, the expected waiting time is $1/P_{t+1}$. Since $p_t \approx 0$, we can approximate $P_{t+1} \approx 1 - (1 - x_t p_t) = x_t p_t = x_t \log(n)/n$. Since we have a geometric process, the expected waiting time till $x_t = 2$ is $1/P_{t+1}$ [5], which is approximately $n/\log(n)$. The additional time to wait till $x_t = 3$ is approximately $n/2\log(n)$. Thus, we can estimate the time to wait till x_t is K as

$$\frac{n}{2} + \frac{n}{\log n} + \frac{n}{2\log n} + \cdots + \frac{n}{(K-1)\log n}.$$

Since the total time we wait is n, it must be that

$$\frac{n}{2} + \frac{n}{\log n} \sum_{i=1}^{K-1} \frac{1}{i} = n.$$

We conclude that $H(K-1) = \frac{1}{2}\log n$, where $H(\ell)$ is the ℓth Harmonic number. It is well known that $H(\ell) = \Theta(\log \ell)$, from which we gather that $H(K-1) \approx \log(K) = \frac{1}{2}\log n$. We conclude that $K = \Omega(\sqrt{n})$. We can repeat this argument conditioning not on $x_{\frac{n}{2}} \geq 1$ but instead on $x_{n\epsilon} \geq 1$ for any $\epsilon > 0$. The right hand side would then become $(1 - \epsilon)\log n$ and we would conclude that $K = \Omega(n^{1-\epsilon})$. A more subtle argument gives $K = \Omega(n)$; the details are deferred.

EXPERIMENTS

We now discuss the simulations that were conducted to investigate our model of infiltration.

Data

The real data sets that we used are one author's LinkedIn InMap ® and the Enron Email Communication data set. The LinkedIn data set is a very small network that consists of the induced subgraph obtained from one author's neighborhood on LinkedIn. The network is illustrated in Figure 1.

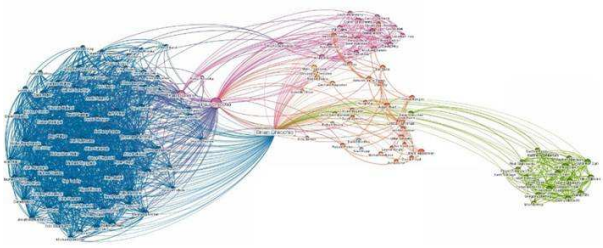
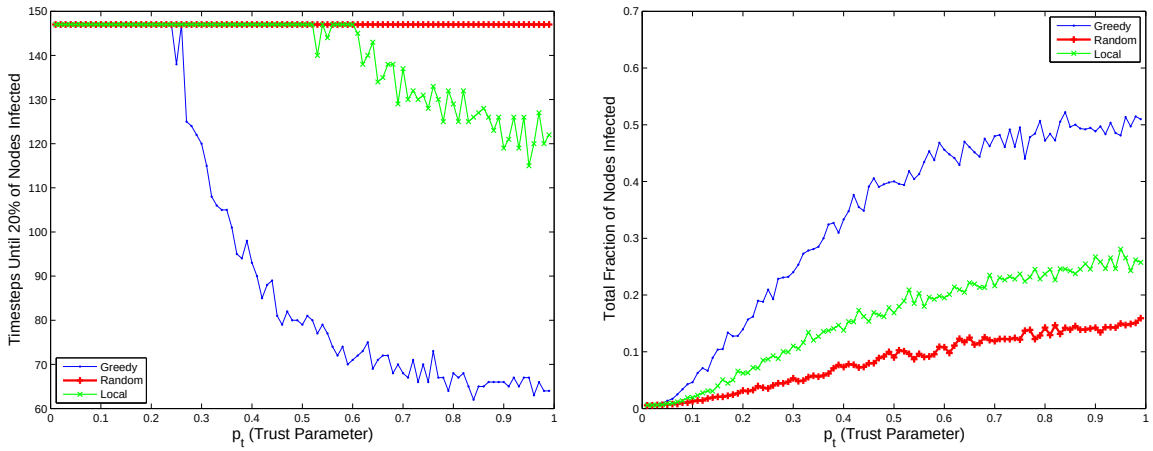


Figure 1. Author's LinkedIn InMap Network - a strongly clustered network with clustering coefficient 0.79

The Enron Email network (obtained from the Stanford Large Network Data Collection) was picked because it represents a typical social network where this type of infiltration could occur. It is reasonable to say that if an unknown person emails you but you know they have corresponded with people you have emailed in the past then you might trust them to a larger extent. We also generated two different types of synthetic networks to run the simulation on for comparison. The first type is the Erdös-Rényi random network which is generated by two parameters, n (the number of vertices) and γ (the probability that there will be an edge between each pair



(a) Time to infiltrate 20% if nodes

(b) Expected fraction of nodes infiltrated

Figure 2. Infiltration of the small LinkedIn network using various strategies. The greedy strategy is more efficient. The local strategy is better than the random strategy. We observe demonstrates a phase transition for $p_t > p_t^* \approx 0.25$, where the time to hit 20% infection rapidly drops.

of nodes). We will refer to this type of network as an E-type graph and we will see that infiltration of this type of network is very different that the other two networks. The other synthetic network is a random graph that is generated using the degree distribution of a real network. We call this an R-type random network and it has more proporties of a social network than the E-type. We used a method for generating an R-type network that produces a network that has approximately the same degree distribution as the input network. All these networks are represented by undirected graphs with no edge weights for simplicity. Basic information about these networks along with their average clustering coefficient is shown in Table 1. As can be observed, these networks span a wide range of sizes and clustering coefficients ρ. Typically, the social networks tend to have a higher clustering coefficient. The diversity of network types that was used in this research was done to show how the infiltration process changes across different types of networks.

Network	# nodes	# edges	ρ
LinkedIn InMap	146	2032	0.79
Enron	36692	183831	0.49
E-type	36692	183831	0.0002
R-type	36692	183831	0.032

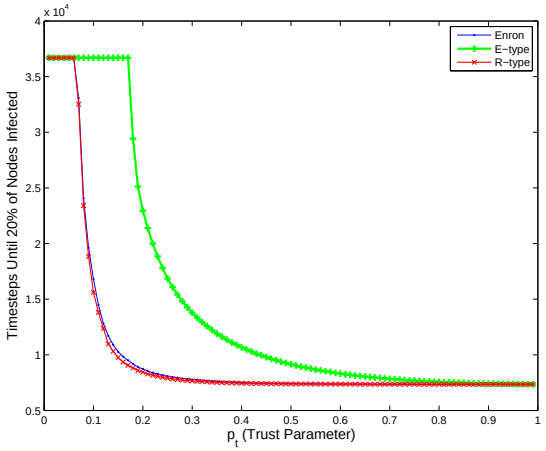
Table 1. Summary of network statistics for the four networks included in this study. ρ is the correlation coefficient, the average number of closed triangles per node.

In all of our simulations, SC sends one connection request at each time step. We set the parameter p_e, (the base ego probability) to be .5 which gives a reasonable probability to accept connections when a node has a very small degree. The α parameter (exponent of the degree in the ego effect) is set at 2 for most experiments except when it was varied to investigate the effect of α. The dependence on p_t is the parameter of our focus and our main goal is to study the extent of infiltration as we vary p_t.

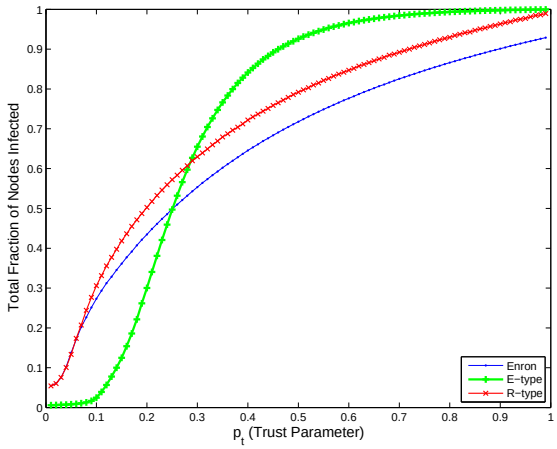
Critical Values for p_t

We will start by discussing the results of running the infiltration simulation on the small LinkedIn network. This network has 146 nodes and 2032 edges and is very well clustered as shown by the visualization of the network in Figure 1. We consider two metrics to measure the efficiency of the trust based infiltration process. The first metric is the number of time steps until the network is infiltrated to some percentage. For this paper, we chose 20% as a reasonable target infiltration percentage. The second metric is the total percentage of the network infiltrated after a certain number of time steps. Here, we chose to simulate n time steps so that all nodes are requested once. Figure 2 shows the average results over many simulations. In Figure 2a, the infiltration using the random strategy was never able to achieve 20% infiltration which you can see in Figure 2b by the fraction of infiltrated nodes never breaching 0.2. Considering the local strategy, SC only breaches 0.2 at a very high value of the trust parameter which is better than the random one but not nearly good enough to be successful. However, the greedy strategy was able to hit the 20% threshold at $p_t = 0.25$ and in the range of $p_t = [0.26, 0.35]$, the time to hit 20% rapidly dropped. This phase transition illustrates that even though Theorem 1 was established for random graphs and the random strategy, the existence of a critical p_t^* still holds for this well clustered small network and the greedy strategy. A similar yet less apparent trend is seen in the plot that shows the total fraction of nodes infected at the end of the simulation. This critical range of values where the success of the infiltration changes greatly is even smaller for larger social networks as we will show with the Enron Email network.

The same simulation was run on the Enron Email network as well as the E-type and R-type synthetic networks. The Enron email network (representing a typical social network) has 36692 nodes and 183831 edges and the other two graphs are generated based on this data so they have the same number of nodes and approximately the same number of edges. In Figure 3a, the plot showing the time until 20% of the net-

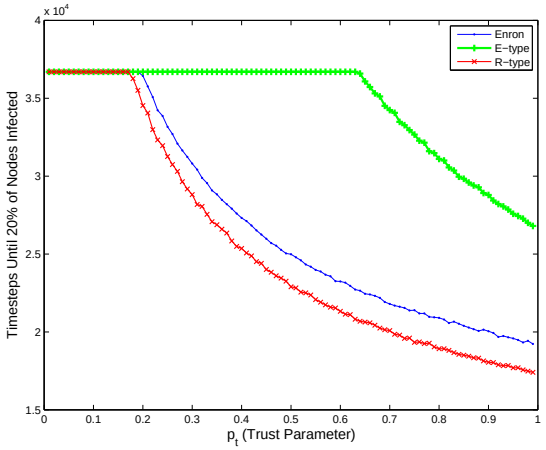


(a) Time to infiltrate 20% of nodes

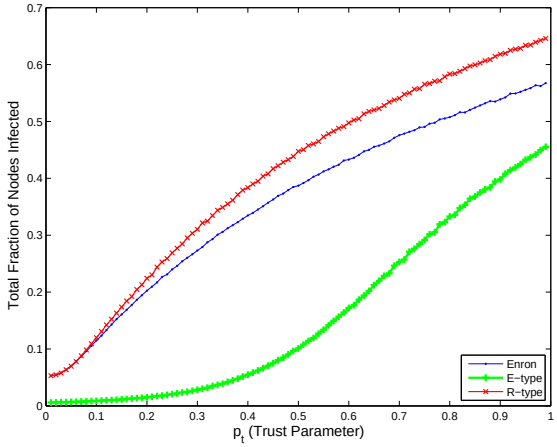


(b) Expected fraction of nodes infiltrated

Figure 3. Infiltration of the Enron network, compared to similar E-type and R-type random networks for the greedy infiltration strategy. On this larger network, the phase transitions are more pronounced. For small values of p_t, we see that the E-type random graph is resistant to infiltration, more so than the social networks and the R-type random graph – clustering helps. As p_t increases, it appears that expansion becomes more useful for infiltration and it is now easier to infiltrate the E-type random graph.



(a) Time to infiltrate 20% of nodes



(b) Expected fraction of nodes infiltrated

Figure 4. Infiltration of the Enron network, compared to similar E-type and R-type random networks for the random infiltration strategy. The phase transition is less pronounced with the random strategy, and in general, the E-type random graph is most resistant to infiltration.

work is infected using the greedy strategy, we see a clear phase transition similar to the one in the LinkedIn network but more pronounced. This is due to the larger size of the network. A similar trend is less pronounced in Figure 5a and Figure 4a for the local and random strategies. Again, the existence of a critical trust threshold is clear for all three networks in Figure 3a.

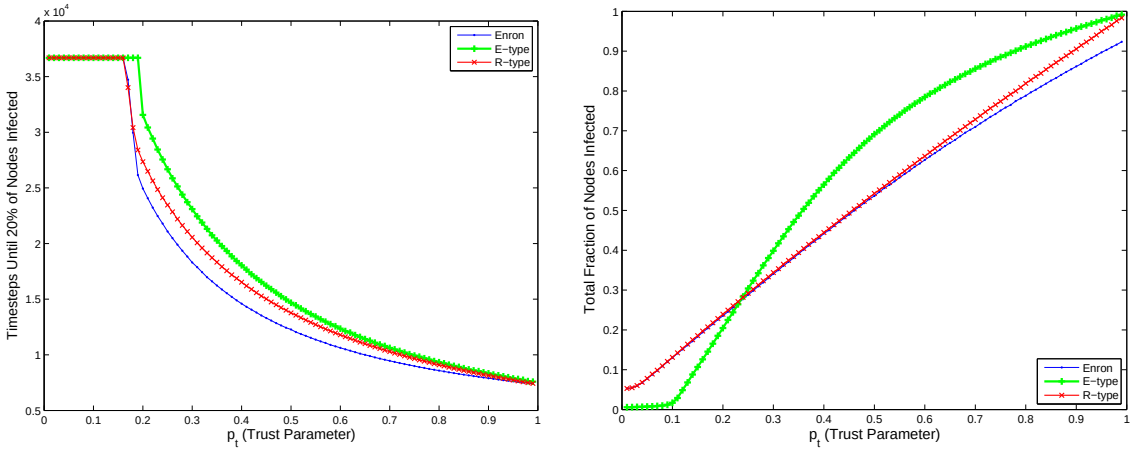
Conclusion. There is a critical value p_t^* below which infiltration is hard and above which infiltration is easy. This is the case for well clustered graphs (LinkedIn, Enron), random graphs, a greedy infiltration strategy, local or a random one. Even small graphs display this phase transition at a critical p_t^* though the phase transition is more pronounced for larger graphs as is expected.

Network Structure

We can also notice that the network structure most definitely plays a role in how successful the infiltration is. In Figure 3a, the critical p_t value is at 0.1 for the Enron network and R-type network while it is at 0.2 for the E-type network. Since the R-type network has a similar structure to that of the Enron network because it is constructed to match the degree distribution, the infiltration process is similar on these networks. We observe that the E-type network is more robust to this infiltration.

Conclusion. Random un-clustered graphs are harder to infiltrate while typical social networks which display strong clustering are more susceptible. This is especially so for the random infiltration strategy, and the greedy strategy for small p_t.

When considering Figures 3a and 4a which show the frac-



(a) Time to infiltrate 20% of nodes

(b) Expected fraction of nodes infiltrated

Figure 5. Infiltration of the Enron network, compared to similar E-type and R-type random networks for the local infiltration strategy. This strategy is more successful than random, and least sensitive to network structure changes.

tion of infected nodes that result from the infiltration using the greedy and random strategies, we see that the structure of the network not only affects the speed at which the infiltration takes place but also the resulting total infiltration success. The E-type network may be more robust to infiltration at lower values of p_t but at higher values, it ends up with more of it's nodes infected when SC is using the greedy strategy. This is because as p_t increases and the infiltration process gets started, the neighbors of infiltrated nodes will likely accept a connection. The more neighbors the better, i.e. graphs with high expansion are good when p_t is large and E-type graphs are known to have large expansions. However, when p_t is small, more clustered networks are easier to infiltrate.

In analyzing the local strategy, we notice that infiltration using this strategy is less sensitive to network structure. This robustness is hypothesized to be due to the strategy not taking advantage of a typical social network's structure. Under this algorithm, SC selects the node which is last in the list generated of second level connections so even if SC has infiltrated a community successfully, he/she will not necessarily use that advantage in selecting nodes to request. This is unlike the greedy strategy where SC will take full advantage of successful community infiltration. Thus we see this strategy resulting in similar success rates with all three networks.

Strategy Comparison

When SC has no knowledge of the target network and uses a random strategy to attempt their infiltration, as shown in Figure 4, we see that the infiltration is not that successful. The local strategy improves on the random strategy due to the trust of at least one shared connection. This improvement also comes without much effort; SC just requires knowledge of the connections of the nodes in his/her neighborhood. This is realistic information to have in practice and the resulting infiltration benefit is significant. Lastly, using the greedy strategy requires SC to have full knowledge of the network and not surprisingly, it enables much quicker infiltration of the network than either of the other two strategies.

Conclusion. There is a significant performance difference between different strategies for SC.

We also observe that the phase transition is more apparent (sensitivity of the infiltration around the critical threshold p_t^*) with the greedy strategy than with the random strategy. This indicates that if defending against a powerful adversary who has network knowledge, it is all the more important to "guard your connections".

The Ego-effect

The trust parameter, p_t, is the main factor that determines the ultimate success or failure of the infiltration but changing the weight of the ego factor can also affects the process. To study this, the simulation of infiltration using the greedy strategy was conducted for different values of the parameter α. From this simulation, the critical value, p_t^* (where the phase transition occurs), was extracted and plotted versus the parameter alpha for each of the three networks as shown in Figure 6. With higher values of alpha, the probability to accept a connection request decreases based on our model so it is not surprising that p_t^* increases as α increases. However, the decrease on α is not very sensitive.

DISCUSSION

From our simulation of the infiltration process on real social networks, we have seen that the structure of the network plays an important role in the success or failure of the infiltration. In addition, the time until a certain level of infiltration reached is very sensitive around a critical p_t^* value, indicative of the phase transition. We showed that this will be true empirically for some typical social networks. For random graphs and a random infiltration strategy, we theoretically established such a phase transition for the infiltration around a narrow range of p_t, complementing the empirical results. This means actors need to be careful about how much trust they place in their connections since too much

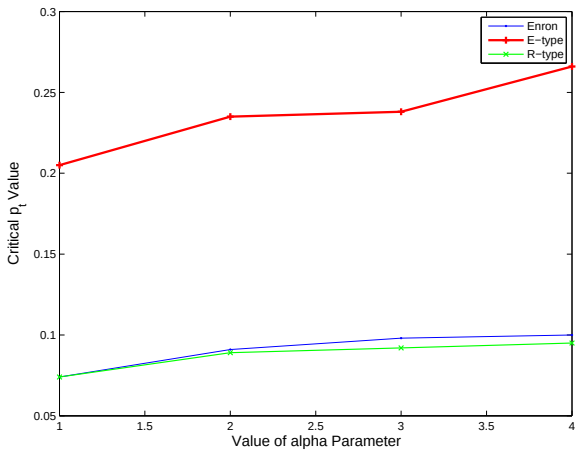


Figure 6. Effect of Ego parameter α on Infiltration (for the greedy strategy). As α increases, P_E decreases and the critical threshold p_t^* to start a cascade increases, as is expected. However, the sensitivity to α does not appear too large. Also apparent is the fact that infiltration of the E-type network is harder – a higher value of p_t is required to get significant infiltration.

trust can result in a dramatically faster infiltration of the network. Three strategies that SC can use to carry out the infiltration were also explored. The greedy approach was far more successful but requires full knowledge of the network while the random strategy requires zero knowledge of the network, but nonetheless can still be quite successful. The local strategy achieves better results than the random strategy by just using local knowledge that is available in a lot of real networks.

Realistically, SC will never have full knowledge of the target network so he can always use the random approach. However, since SC might have local knowledge around the nodes to which he is connected, a greedy local strategy could be to send connection requests to nodes who are connections of current connections using some form of greedy approach. This strategy would hopefully do much better than the local strategy presented in this paper by finding the optimal node to request at each time-step based on local information instead of arbitrarily picking a node like in the current local strategy. This will result in a more effective strategy than the random or local approach due to using local information in a greedy way instead of an arbitrary way. This type of strategy together with the optimal strategy for infection are left as avenues for future work.

Another interesting question is quantify the greedy strategy with respect to how close it is to optimal on an arbitrary graph. We can also pose the question with respect to SC's ultimate goal being to infect a specific node or community of nodes. The strategy could possibly change in that scenario. Again, these questions will be addressed in further research.

Finally, what is SC's optimal strategy for a general star with arbitrary starting ego-probabilities at each node. What other simple networks can be analyzed (for example the path, the tree, the complete d-ary tree, etc.)?

ACKNOWLEDGMENTS

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

REFERENCES

1. Anshelevich, E., Chakrabarty, D., Hate, A., and Swamy, C. Approximations for the firefighter problem: Cuts over time and submodularity. In *Proc. 20th Int. Symp. on Alg. and Comp. (ISAAC)* (2009).
2. Bailey, N. T. J. *The Mathematical Theory of Infectious Diseases and its Applications*. Hafner Press, 1975.
3. Callaway, D. S., Newman, M. E. J., Strogatz, S. H., and Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett 85* (2000), 5468–5471.
4. Cohen, R., Erez, K., ben Avraham, D., and Havlin, S. Resilience of the Internet to random breakdowns. *Phys. Rev. Lett. 85* (2000), 4626–4628.
5. Feller, W. *An Introduction to Probability Theory and its Applications*, vol. 1. John Wiley & Sons, New York, 1968.
6. Fisher, M. E., and Essam, J. W. Some cluster size and percolation problems. *J. Math. Phys.* (1961), 609–619.
7. Grassberger, P. On the critical behavior of the general epidemic process and dynamical percolation. *Math. Biosci. 63* (1982), 157–172.
8. Hethcote, H. W. The mathematics of infectious diseases. *SIAM Review 42* (2000), 599–653.
9. Kempe, D., Kleinberg, J., and Tardos, E. Maximizing the spread of influence in a social network. In *Proc. KDD 2003* (2003).
10. Mollison, D. Spatial contact models for ecological and epidemic spread. *J. Roy. Stat. Soc. B 39* (1977), 283–326.
11. Newsome, J., Shi, E., Song, D., and Perrig, A. The sybil attack in sensor networks: analysis & defenses. In *Proc. IPSN* (2004), 259 – 268.
12. Rapoport, A. Spread of information through a population with socio-structural bias: i: Assumption of transitivity. *Bulletin of Mathematical Biophysics 15*, 4 (1953), 523–533.
13. Rogers, E. *Diffusion of Innovations*. Free Press, 1995.
14. Yu, H., Kaminsky, M., Gibbons, P. B., and Flaxman, A. D. Sybilguard: defending against sybil attacks via social networks. *IEEE/ACM Trans. Netw. 16*, 3 (2008), 576–589.