Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence

Equilibria in Epidemic Containment Games

Sudip Saha, Abhijin Adiga and **Anil Kumar S. Vullikanti**
Department of Computer Science
Network Dynamics and Simulation Science Laboratory
Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA 24060
Email: {ssaha,abhijin,akumar}@vbi.vt.edu

Abstract

The spread of epidemics and malware is commonly modeled by diffusion processes on networks. Protective interventions such as vaccinations or installing anti-virus software are used to contain their spread. Typically, each node in the network has to decide its own strategy of securing itself, and its benefit depends on which other nodes are secure, making this a natural game-theoretic setting. There has been a lot of work on network security game models, but most of the focus has been either on simplified epidemic models or homogeneous network structure.

We develop a new formulation for an epidemic containment game, which relies on the characterization of the SIS model in terms of the spectral radius of the network. We show in this model that pure Nash equilibria (NE) always exist, and can be found by a best response strategy. We analyze the complexity of finding NE, and derive rigorous bounds on their costs and the Price of Anarchy or PoA (the ratio of the cost of the worst NE to the optimum social cost) in general graphs as well as in random graph models. In particular, for arbitrary power-law graphs with exponent $\beta > 2$, we show that the PoA is bounded by $O(T^{2(\beta-1)})$, where $T = \gamma/\alpha$ is the ratio of the recovery rate to the transmission rate in the SIS model. We prove that this bound is tight up to a constant factor for the Chung-Lu random power-law graph model. We study the characteristics of Nash equilibria empirically in different real communication and infrastructure networks, and find that our analytical results can help explain some of the empirical observations.

1 Introduction

The spread of epidemics and malware is commonly modeled by diffusion processes on networks, such as the SIS or SIR models (Newman 2003; Grassly and Fraser 2008). They are typically controlled by vaccinating nodes or installing antivirus software patches. This involves a certain cost for the individual (e.g., the economic cost of the vaccine or the patch). On the other hand, an individual has no incentive to protect itself if enough of its neighbors are protected (referred to as *herd immunity*). This is a natural setting for a game-theoretical analysis, and is an active area of research e.g. (Aspnes, Rustagi, and Saia 2007; Aspnes, Chang, and Yampolskiy 2006; Kumar et al. 2010; Bauch and Earn 2004; Omic, Orda, and Mieghem 2009; Lelarge and Bolot 2009; Grossklags, Christin, and Chuang 2008; Khouzani, Sarkar, and Altman 2012; Khouzani, Altman, and Sarkar 2012; Khouzani, Sarkar, and Altman 2011). A large part of this research has been focused on simplistic assumptions about either (i) the diffusion process (e.g., a very high transmission probability), or (ii) the network (e.g., homogeneity, allowing for differential equation models). All these models involve individual utility functions with some notion of the cost of getting infected; these are generally difficult to compute and analyze in heterogeneous networks, which accounts for the limited understanding of the dynamics of epidemic games in realistic scenarios.

(Ganesh, Massoulie, and Towsley 2005) developed a spectral characterization for the dynamics of the SIS model – they show that an epidemic dies out soon if $\lambda_1(G) < \gamma/\alpha$, where $\lambda_1(G)$ is the *spectral radius* or the *largest eigenvalue* of the adjacency matrix of the contact graph G, while γ and α are the recovery rate and transmission rate of the SIS model, respectively. We use T to denote the threshold γ/α. This result is extended to other models in (Prakash et al. 2012) – this is analogous to the characterization of differential equation based epidemic models in terms of the reproductive number, R_0 (Newman 2003). While $\lambda_1(G) > T$ implies the epidemic lasts "long" in some families of graphs (and in practice, in many networks, this holds if $\lambda_1(G)$ is much larger than T), the precise converse derived by (Ganesh, Massoulie, and Towsley 2005) is in terms of the spectral gap of the Laplacian of G. A natural approach for containing an epidemic, motivated by this characterization, is to design interventions (e.g., vaccination) so that λ_1 becomes smaller than T, as discussed in (Tong et al. 2012; Mieghem et al. 2011).

Our contributions. We develop a novel formulation of an epidemic containment game motivated by the above characterization of the SIS model (Ganesh, Massoulie, and Towsley 2005; Prakash et al. 2012), and approaches to contain it by reducing the spectral radius (Tong et al. 2012; Mieghem et al. 2011). In this way, our formulation uniquely incorporates a realistic infection model in general heterogeneous networks. We characterize the structure of the equilibria in the resulting game in different graphs, and then study their properties empirically using real world networks. We discuss our contributions in greater detail below.

1. A game formulation based on spectral properties. We introduce the *Epidemic Containment (EC)* game on a network $G(V,E)$ of players, for the SIS model of epidemic spread. The individual actions of a player are either to be secured (which has a fixed cost, corresponding to the price of a vaccination/anti-virus software) or not. If a player is not secured and if $\lambda_1(G') > T$, where G' is the graph induced by the insecure nodes, he or she incurs a high cost of infection, since the epidemic is likely to last long (this is formally defined later). We show that pure Nash equilibria (NE) always exist in an EC game, and can be found by a best response strategy. Further, a minimum cost Nash equilibrium is also a social optimum, and we show that finding the social optimum is NP-complete, in general.

2. Structure of equilibria in arbitrary graphs. We derive bounds on the cost of the worst NE and the Price of Anarchy in terms of the maximum degree in general graphs. There can be an exponential number of Nash equilibria, and the ratio of the maximum cost of any NE to that of the social optimum (also referred to as the price of anarchy, and denoted by PoA) can be $\Omega(\Delta(G))$, where $\Delta(G)$ denotes the maximum node degree in G. When G has a power law degree sequence with exponent $\beta > 2$, we show that the cost of the social optimum is in the interval $\left[c_1 n/T^{2(\beta-1)}, c_2 n/T^{(\beta-1)}\right]$, for constants c_1, c_2; this implies that the PoA is $O\left(T^{2(\beta-1)}\right)$. Further, we show that a $\Theta\left(T^{(\beta-1)}\right)$-approximation to the social optimum can be computed in polynomial time.

3. Structure of equilibria in random graph models. We study the structure of NE in Erdős-Rényi and Chung-Lu random graph models; the latter has been shown to be relevant for a broad class of real world networks. We prove that in Erdős-Rényi graphs $G(n,p)$, if $p = \Omega(\log n/n)$ (which is needed for the graph to be connected), and if $T^2 = O(np)$, every NE has cost $\Omega(n)$, and the PoA is $\Theta(1)$. We consider the Chung-Lu model (Chung and Lu 2006b) defined by a power-law weight vector with exponent β (defined formally later), which gives a random graph with degree sequence close to a power law. We prove that when $\beta > 2$, and the weight sequence and T satisfy additional weak assumptions, the social optimum has cost $\Theta\left(\frac{n}{T^{2(\beta-1)}}\right)$, and can be approximated by picking high weight nodes. Note that this corresponds to the upper bound we prove for general power law graphs. In contrast, the worst cost NE is $\Omega(n)$, which can be obtained by favoring low weight nodes, which leads to a PoA of $\Theta\left(T^{2(\beta-1)}\right)$ in this model.

4. Empirical analysis of the properties of the equilibria. We study the structure of equilibria in EC games in seven different real and random networks, on which malware could spread. Our main observations are summarized below: (i) We find that estimates of the minimum cost of NE scale as $\Theta\left(\frac{n}{T^{2c(\beta-1)}}\right)$ for the scale free networks, where c is a small constant close to 1, which is close to our bounds for general graphs and the Chung-Lu model; (ii) We compute a lower bound on the PoA and find that it scales roughly as $\Theta\left(T^{2c'(\beta-1)}\right)$ where c' is a constant close to 1, which is again close to our theoretical bounds; (iii) An interesting observation is that the degree distribution in the graph induced by the insecure nodes in random NE is very close to the degree distribution in the graph G in most networks; (iv) We study the community structure in the networks, and find that generally larger communities have a disproportionately high fraction of secure nodes, in contrast to their size. This might be useful in understanding what kinds of nodes have the greatest incentive to secure themselves; (v) Finally, we consider Stackelberg strategies, to explore how to mitigate the effects of distributed control and the PoA by influencing a small set of nodes. We find that the cost of random NE reduces quite a bit for a small fraction of high degree nodes secured initially.

Organization. We first describe the formal model and definitions and then discuss the results for general graphs. We then discuss the results for random graphs, and the simulation results. We summarize the related work and then conclude. Many proofs and other details are omitted because of the space limitations, and are available in (Saha, Adiga, and Vullikanti 2014).

2 Preliminaries and Model

The *Epidemic Containment* (EC) game involves an undirected graph $G(V,E)$ on the set V of players or nodes. The neighbor set and degree of a node v are denoted by $N(v)$ and $d(v)$ respectively. Let $\Delta = \Delta(G)$ denote the maximum degree and let $\lambda_1 = \lambda_1(G)$ denote the first eigenvalue of the adjacency matrix of G. We assume the SIS model of epidemic spread (Newman 2003; Ganesh, Massoulie, and Towsley 2005), in which nodes are in states *Susceptible* (S) or *Infected* (I). Initially, some source node gets infected and all other nodes are susceptible. Each infected node v infects each of its neighbors u currently in state S at rate α (the transmission rate); if neighbor u gets infected, it switches to state I. Also, each infected node v switches back to state S at rate γ. $T = \gamma/\alpha$ is referred to as the threshold. The characterization of (Ganesh, Massoulie, and Towsley 2005; Prakash et al. 2012) implies that the epidemic dies out quickly (in $o(n)$ time) if $\lambda_1(G) < T$. Each node x decides independently whether to become secured/vaccinated (denoted by $a_x = 1$) or not (denoted by $a_x = 0$); a_x is the strategy selected by node x, and $\mathbf{a} = (a_1, a_2, ..., a_n)$ denotes the strategy profile of all the nodes. We use a_{-x} to denote the strategy profile of the players other than x.

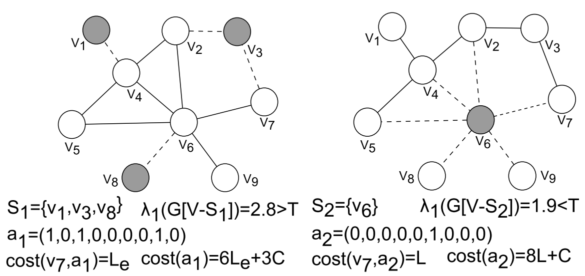
If node x decides to get secured, i.e., $a_x = 1$, it incurs a cost C (e.g., the cost of a vaccination or anti-virus software). If node x does not get secured, i.e., $a_x = 0$, its cost (denoted by $\text{cost}(x, \mathbf{a})$) depends on whether or not the epidemic dies out quickly or not in the connected component containing x (restricted to the graph induced by the insecure nodes) under the strategy profile $\mathbf{a}$ – we let $L < C$ and $L_e > C$ denote the costs in the former and latter cases, respectively. The motivation is that if the epidemic does not die out quickly in the component containing node x, then x is more likely to be infected, and incurs a higher cost than C; however, if the epidemic dies out quickly, the cost incurred is much smaller than C. Let $S = S(\mathbf{a}) = \{x \in V : a_x = 1\}$ denote the set of secure nodes in the strategy profile $\mathbf{a}$. We call the graph $G[V - S(\mathbf{a})]$ induced by the set $V - S(\mathbf{a})$ of insecure nodes as the "attack" graph. Let $G_x[V - S(\mathbf{a})]$ denote the connected component of $G[V - S(\mathbf{a})]$ that contains x. Following the characterization of (Ganesh, Massoulie, and Towsley 2005; Prakash et al. 2012), we have for any $v \in V$ and strategy profile $\mathbf{a}$:

$$\text{cost}(v, \mathbf{a}) = \begin{cases} C, & \text{if } a_v = 1, \\ L, & \text{if } a_v = 0 \text{ and } \lambda_1(G_v[V - S(\mathbf{a})]) < T, \\ L_e, & \text{if } a_v = 0 \text{ and } \lambda_1(G_v[V - S(\mathbf{a})]) \geqslant T. \end{cases}$$

An instance of the EC game is defined by the tuple (G, T, C, L, L_e). For a strategy profile $\mathbf{a}$, the social cost $\text{cost}(\mathbf{a}) = \sum_{v \in V} \text{cost}(v, \mathbf{a})$. If the epidemic dies out quickly, then $\text{cost}(\mathbf{a}) = |S|C + |V - S|L$ where $S = S(\mathbf{a})$; otherwise, $\text{cost}(\mathbf{a}) = |S|C + |V_e|L_e + |V - S - V_e|L$, where V_e is the set of insecure nodes x that are part of those components of the attack graph where the epidemic lasts long (i.e. $\lambda_1(G_x[V - S(\mathbf{a})]) \geqslant T$). The optimum social cost of an instance is denoted by C_{OPT} where, $C_{\text{OPT}} = \min_{\mathbf{a}} \text{cost}(\mathbf{a})$.

A strategy profile $\mathbf{a}$ is said to be a Nash Equilibrium (NE) if for any player i, and any alternative strategy a'_i for player i, we have $\text{cost}(i, \mathbf{a}) \leqslant \text{cost}(i, \mathbf{a}')$, where $a'_j = a_j$ for all $j \neq i$. That is, a strategy profile $\mathbf{a}$ is a NE, if no player i can benefit by switching his/her strategy, given that a_{-i} is fixed (Nisan et al. 2007).

This is illustrated in Figure 1. In Figure 1(a), $\mathbf{a}_1$ is not NE, since any of the unsecured nodes can secure itself and get its cost reduced from L_e to C. On the other hand, $\mathbf{a}_2$ in Figure 1(b) is a NE. This is because v_6 cannot benefit by switching from secure to insecure (as λ_1 of the attack graph becomes more than T), and none of the insecure nodes benefit by switching to a secured state (as that would only increase its cost from L to C). The following observation gives a simple characterization of a NE in the EC game.



(a) Not a Nash Equilibrium (b) A Nash Equilibrium

Figure 1: Example of an EC game where $\lambda_1 = 3.13$ and $T = 2$. (a) Strategy profile $\mathbf{a}_1$ where three nodes are secured but spectral radius of the attack graph is more than $T = 2$; (b) Strategy profile $\mathbf{a}_2$ where one node (v_6) is secured and spectral radius of the attack graph is below T, is a NE. For $\mathbf{a}_1$, the epidemic is likely to last long while for $\mathbf{a}_2$, the epidemic dies out quickly.

Observation 1. *For an instance (G, T, C, L, L_e) of the* EC *game, a strategy profile $\mathbf{a}$ is a NE if and only if $S(\mathbf{a})$ is a minimal set of secured nodes such that $\lambda_1(G[V - S(\mathbf{a})]) < T$.*

The proof is in (Saha, Adiga, and Vullikanti 2014). From the above observation it follows that, for this game formulation, the best NE corresponds to the social optimum. Therefore, from now on we use the terms "social optimum" (denoted by S_{OPT} and its cost C_{OPT}) and "best NE" interchangeably. To simplify the notation, for the rest of the paper, we will focus on instances (G, T, C, L, L_e) of the EC game, where $C = 1$, $L = 0$ and $L_e > 1$. Under this assumption and from Observation 1, it follows that if $\mathbf{a}$ is a NE, then $\text{cost}(\mathbf{a}) = |S(\mathbf{a})|$. All our results extend naturally to the general case. The *Price of Anarchy* (PoA) is defined as the ratio between the cost of the worst equilibrium and the social optimum cost, that is, $\text{PoA} = \frac{\max_{\mathbf{a}:\mathbf{a} \text{ is NE}} \text{cost}(\mathbf{a})}{C_{\text{OPT}}}$

We will also study Stackelberg strategies, in which a centralized authority is allowed to control the strategies of a fraction of agents, while the remainder act non-cooperatively (Roughgarden 2001).

3 Existence and complexity of Nash Equilibria

We first discuss some properties about Nash equilibria and the complexity of computing them. It follows from observation 1 that the smallest set S such that $\lambda_1(G[V - S]) < T$ is a NE, and is also the social optimum. Further, it follows that a pure NE can be computed by iteratively choosing nodes into a minimal subset of secured nodes. Please see (Saha, Adiga, and Vullikanti 2014) for more details.

Computing the social optimum is computationally challenging, as shown below. The proof follows from a reduction of vertex cover problem to EC game problem (Saha, Adiga, and Vullikanti 2014).

Lemma 2. *Finding the social optimum of an EC game is NP complete. Moreover, the cost of social optimum cannot be approximated within a factor of* 1.3606 *unless P=NP.*

4 The structure of NE in general graphs

We consider bounds for arbitrary power law graphs (see (Saha, Adiga, and Vullikanti 2014) for additional results on other graph classes). Let n_i denote the number of nodes of degree i in G, for $i \in \{1, \ldots, d_{\max}\}$. We assume the degree sequence of G is a power law with exponent β, so that $n_i \propto 1/i^\beta$. Define $\mathcal{E}_0(x) = \sum_{i \geqslant x}^{d_{\max}} n_i$ and $\mathcal{E}_1(x) = \sum_{i \geqslant x}^{d_{\max}} i \cdot n_i$. We first observe the following useful property.

Lemma 3. *Let G be a power law graph with exponent $\beta > 2$ and let $x \leqslant cd_{\max}$ for a constant $c < 1$. Then, (1) $\mathcal{E}_0(x) = \Theta\big(n/x^{\beta-1}\big)$ and (2) $\mathcal{E}_1(x) = \Theta\big(n/x^{\beta-2}\big)$.*

See (Saha, Adiga, and Vullikanti 2014) for a proof.

Theorem 4. *Let G be a power law graph with exponent $\beta > 2$, where β is a constant and let $T^2 \leqslant cd_{\max}$ for a constant $c < 1$. Then, there exist constants c_1 and c_2 such that $c_1\big(n/T^{2(\beta-1)}\big) \leqslant C_{OPT} \leqslant c_2\big(n/T^{(\beta-1)}\big)$.*

Proof. We first consider the lower bound. Consider any strategy profile $\mathbf{a}$ that is a NE; let $I = \{v : \mathbf{a}_v = 0\}$ be the set of insecure nodes in $\mathbf{a}$. We have $\lambda_1(G[I]) \geqslant \sqrt{\Delta(G[I])}$. Let $A = \{v : d(v) \geqslant T^2\}$. It follows that for any node $v \in A$, $d_{G[I]}(v) < T^2$, for otherwise, $\lambda_1(G[I])$ would be at least T. This implies that any node $v \in A$ is either secured (i.e., has $\mathbf{a}_v = 1$) or at least $d_G(v) - T^2$ neighbors of v are secured. Let D be the smallest set such that for any $v \in A$ either: (i) $v \in D$, or (ii) at least $d(v) - T^2$ neighbors of v are in D. Then, cost($\mathbf{a}$) $\geqslant |D|$. Because of the power law degree distribution, it follows that $|A| = \Theta\big(n/T^{2(\beta-1)}\big)$. Observe that $r_e := \frac{1}{2}\sum_{v\in A}(d(v) - T^2)$ denotes a lower bound on the number of edges with end points in D which need to be removed. Since G is a power law graph with exponent β,

$$r_e = \frac{1}{2}\sum_{i\geqslant T^2}^{d_{\max}}(i - T^2)n_i = \Theta(1)\int_{T^2}^{\infty}\frac{n\big(z - T^2\big)}{z^{\beta}}dz$$
$$= \Theta\Big(n/T^{2(\beta-2)}\Big). \qquad (1)$$

Next, note that $|D|$ is minimized if the largest degree nodes are selected. Consider the largest index j such that $\sum_{i\geqslant j} i \cdot n_i = \mathcal{E}_1(j) \geqslant r_e$; then, $|D| \geqslant \sum_{i\geqslant j} n_i = \mathcal{E}_0(j)$. From Lemma 3 and (1), it follows that $j = \Theta\big(T^2\big)$ and therefore, (again from Lemma 3) $\mathcal{E}_0(j) = \Theta\big(n/T^{2(\beta-1)}\big)$.

For the upper bound, we note that $\lambda_1(G[I]) \leqslant \Delta(G[I])$. Therefore, if all the nodes in $B = \{v : d(v) \geqslant T\}$ are secured, then, $\lambda_1(G[I]) < T$. Consider any minimal set $B' \subseteq B$ such that $\lambda_1(G[V - B']) < T$. From Observation 1, B' is a NE and therefore, $C_{\text{OPT}} \leqslant |B|$. From Lemma 3, it follows that $|B| = \mathcal{E}_0(T) = \Theta\big(n/T^{(\beta-1)}\big)$. □

Corollary 5. *Let G be a power law graph with exponent $\beta > 2$, where β is a constant and let $T^2 \leqslant cd_{\max}$ for a constant $c < 1$. Then, the PoA is $O\big(T^{2(\beta-1)}\big)$.*

5 The structure of NE in random graph models

We now analyze the structure of Nash equilibria in EC games in different random graph models.

The Erdős-Rényi model

In Erdős-Rényi random graph model $G(n, p)$, each pair of vertices has an edge between them with probability p. The spectral radius of $G(n, p) = (1+o(1))\max\{np, \sqrt{\Delta}\}$ (Krivelevich and Sudakov 2001).

Lemma 6. *For $p \geqslant \frac{c}{n}$, where c is a suitably large constant and $np \geqslant (1 + \delta)T^2$ for any positive constant δ, for any $G \in G(n, p)$, almost surely $C_{OPT} = \Omega\Big(\frac{n^2 p}{np+\log n}\Big)$ and PoA is $O\Big(\frac{np+\log n}{np}\Big)$.*

See (Saha, Adiga, and Vullikanti 2014) for a proof. Note that the connectivity threshold in $G(n, p)$ is $p = \Omega(\log n/n)$. This implies that if $np \geqslant (1 + \delta)T^2$, any NE has cost $\Omega(n)$ above the connectivity threshold and hence, the PoA is $\Theta(1)$.

Random power law graphs

We consider a random graph model by Chung and Lu (Chung and Lu 2006b; 2006a). Given a weight sequence $\mathbf{w} = (w(v_1, V), w(v_2, V), ..., w(v_n, V))$ for nodes $v_i \in V$, the Chung-Lu model $G(\mathbf{w})$ defines the random graph $G(V, E)$ as follows: for every pair $v_j, v_k \in V$, v_j is adjacent to v_k with probability $\frac{w(v_j,V)w(v_k,V)}{\sum_{v_i\in V} w(v_i,V)}$. Extending the notion discussed in Section 4, we say that $G(\mathbf{w})$ has a power-law weight sequence with exponent $\beta > 2$ if the number of nodes with weight $i \propto 1/i^{\beta}$; for succinctness, we refer to such a G as a Chung-Lu power law graph with exponent β. It is easy to see that the expected degree of any node v_i equals its weight $w(v_i, V)$. For any $V' \subseteq V$, let $w(V')$, $w_{\max}(V')$, $w_{\min}(V')$ and $\tilde{w}(V')$ denote the expected average degree, maximum weight, minimum weight and second order average degree respectively of $G[V']$, the graph induced by V'.

We derive the following results for the bounds on the best and the worst NE in this graph.

Lemma 7. *Let $G(\mathbf{w})$ be a Chung-Lu random power law graph on n nodes with power law exponent $\beta > 2$ and $w_{\min}(V)$ a constant. Let T be the epidemic threshold such that $w_{\max}(V) \geqslant (1 + \delta)T^2$ for a constant $\delta > 0$ and $T = \Omega(\log n)$. Then, $C_{OPT} = \Omega(\frac{n}{T^{2(\beta-1)}})$, almost surely.*

The complete proof is available in (Saha, Adiga, and Vullikanti 2014). We show that if any node of weight at least $(1 + \epsilon)T^2$ is left unvaccinated, then, almost surely, the eigenvalue of the residual graph is more than T. We make use of volume arguments and concentration bounds for node degrees.

Lemma 8. *Let $G(\mathbf{w})$ be a Chung-Lu random power law graph on n nodes with power law exponent $\beta > 2$ and $w_{\min}(V)$ a constant. Let T be the epidemic threshold such that $w_{\max}(V) \geqslant (1+\delta)T^2$ and $T \geqslant (1+\gamma)\tilde{w}(G)\log^2 n$ for any positive constants δ and γ. Then, $C_{OPT} = O\big(\frac{n}{T^{2(\beta-1)}}\big)$ almost surely.*

Lemma 9. *In a Chung-Lu random power law graph $G(\mathbf{w})$ of n nodes and power law exponent $\beta \geqslant 2$ and $w_{\min}(V)$ a constant, if $w_{\max}(G) \geqslant (1 + \delta)T^2 w(V)$ for any positive constant δ and $T = \Omega(\log^2 n)$, then the cost of the worst NE is $\Theta(n)$. Therefore, the size of the largest vaccinated set corresponding to NE is $\Theta(n)$.*

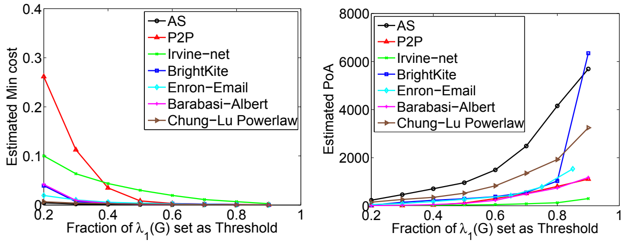
Combining Lemmas 7 and 9, we have the following tight bound for the price of anarchy of the EC game.

Theorem 10. *Consider a Chung-Lu random power law graph $G(\mathbf{w})$ of n nodes and power law exponent $\beta > 2$ and $w_{\min}(V)$ a constant, such that $w_{\max}(G) \geqslant (1 + \delta)T^2 w(V)$ for any positive constant δ and $T = \Omega(\log^2 n)$. The PoA of the* EC *game in $G(\mathbf{w})$ is $\Theta(T^{2(\beta-1)})$ almost surely.*

6 Empirical results

Table 1: Networks used in our experiments and their relevant properties: Five real (Leskovec 2011; Opsahl and Panzarasa 2009) and two synthetic graphs.

Graph, G	Nodes, n	$\lambda_1(G)$	$\Delta(G)$	β
AS (Oregon-1)	10670	58.72	2312	2.23
P2P (Gnutella-6)	8717	22.38	115	NA
Irvine-net	1899	48.14	255	1.34
Brightkite	58228	101.49	1134	2.01
Enron-email	36692	118.42	1383	1.86
Barabasi-Albert	5000	12.51	151	2.61
Chung-Lu power law	4235	62.31	1878	2.1



(a) Estimated Minimum Cost (b) Estimated PoA

Figure 2: (a) Minimum cost and (b) PoA of NE (on y-axis), as a function of T (x-axis). The minimum costs are normalized by network size. Results are shown for seven networks (see Table 1).

We now study characteristics of Nash equilibria of EC game in several social/communication networks and two random graph models, as summarized in Table 1. We study the costs of the cheapest and random NE, PoA, degree distributions in random NE, community structure and effects of Stackelberg strategies. We estimate the maximum and minimum NE cost by two heuristics, that we call the *High Degree* (HDG) and *Low Degree* (LDG) strategies; these are obtained by running iterative strategies for finding NE by removing nodes in decreasing or increasing order of their degrees, respectively (see (Saha, Adiga, and Vullikanti 2014) for details). From Section 4, it follows that the NE resulting from HDG has cost within a $\Theta(T^{\beta-1})$ factor of the social optimum in general power law graphs and the NE resulting from LDG is within a $\Theta(1)$ factor of the maximum cost NE in random power law graphs. Our main observations are:

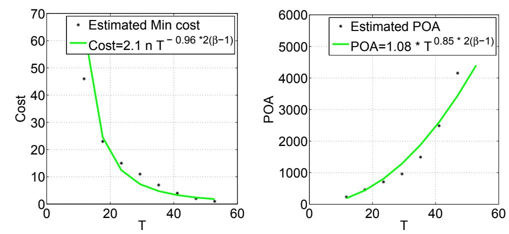
1. **Minimum and Maximum Cost of NE:** For all networks, the estimated minimum cost of NE decreases with T (see Figure 2(a)). For power law graphs, we find that the estimated minimum cost of any NE scales as $\frac{cn}{T^{2c'(\beta-1)}}$ for constants c, c' where c' is close to 1 (particularly $0.5 \leqslant c' \leqslant 1.5$) (see Figure 3(a) for results for the AS network and (Saha, Adiga, and Vullikanti 2014) for other networks), which is quite close to our theoretical bounds for general power law graphs and the Chung-Lu power law model. The estimated maximum cost of NE in most networks turns out to be close to the network size for $0 < T < \frac{1}{2}\lambda_1(G)$ (see (Saha, Adiga, and Vullikanti 2014) for results on the maximum cost of NE).

2. **PoA:** In all the networks, the PoA is an increasing function of T (see Figure 2(b)). Further, for scale-free networks, the estimated PoA scales as $cT^{2c'(\beta-1)}$ for constants c, c', with $0.5 \leqslant c' \leqslant 1.5$ (see Figure 3(b) for results for the AS network), which is also quite close to our analytical bounds.

3. **Degree distribution of secured nodes in NE:** We examine the degree distribution of secured nodes in random NE, obtained by running iterative strategies for finding NE by removing nodes in random order. Results show that, secured nodes tend to have higher degrees, compared to unsecured nodes. See (Saha, Adiga, and Vullikanti 2014) for plots.

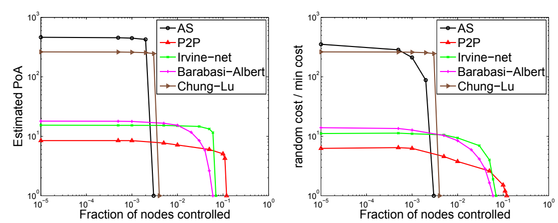
4. **Stackelberg strategies:** As defined earlier, a Stackelberg strategy involves securing a small fraction of nodes in a centralized manner by a leader (or central authority), while the remaining nodes act non-cooperatively. In our experiments, we consider a specific Stackelberg intervention strategy in which high degree nodes are secured by the leader, which works well in both arbitrary and random graphs, as discussed earlier. However, we find that such degree based Stackelberg strategies have little effect in reducing the maximum cost of NE and the random NE cost. As shown in Figure 4(a), the PoA does not change by much, though the cost of random NE reduces slightly (Figure 4(b)). The sudden drop in the plots happens when the Stackelberg strategy itself suffices.

5. **Distribution of secured nodes among communities:** We study the community structure in NE to understand relationships between the secured nodes. We find that to get small equilibria, it generally suffices to secure nodes in a few "important" communities (as shown in Figure 5 for the P2P network). This holds for different small equilibria computed with strategies *HDG* and *HEC* (*HEC* strategy is obtained by iteratively securing nodes in decreasing order of their principal eigenvector components.) One reason might be that nodes contributing to the high eigenvalue seem to be concentrated only in a few communities, because underlying communities typically have uneven degree distributions.



(a) Min cost in AS (b) PoA in AS

Figure 3: Minimum cost of NE and PoA (y-axis) estimated with *HDG* and *LDG* strategy,as a function of T (x-axis), along with the best fit function of the form $\frac{cn}{T^{2c'(\beta-1)}}$ and $kT^{2k'(\beta-1)}$ respectively, where constants c' and k' are close to 1 (see n, β values in Table 1).



(a) Effects on PoA (b) Effects on $\frac{\text{Random NE cost}}{\text{Min NE cost}}$

Figure 4: Effects of Stackelberg strategies on inefficiency metrics. Plots in (a) and (b) show the change in PoA and $\frac{\text{random NE cost}}{\text{Min NE cost}}$ (on y-axis) as different fractions (x-axis) of high degree nodes are secured as part of Stackelberg strategy. Results are shown for five networks and threshold set to $T = 0.3\lambda_1(G)$.

7 Related Work

A fundamental property about the dynamics of epidemics in many models is that of a phase transition from a small number of infections, to a large number of infections; this is characterized in terms of the reproductive number in differential equation based models (Newman 2003), and in terms

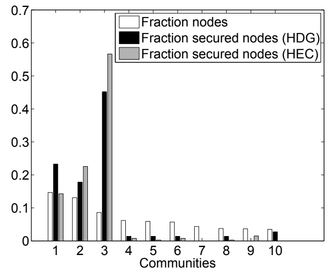


Figure 5: Secured nodes in the smallest NE (estimated with the *HDG* and *HEC* strategy) as distributed among the communities of P2P network. For each of the biggest 10 communities, the left bar shows the fraction of nodes that the community has. The other two bars show the fraction of secured nodes each community has. The middle and the right bar show them for NE's computed with the *HDG* and *HEC* strategy respectively. In to both the NE's, secured nodes are concentrated in few "important" communities.

of the spectral radius for SIS/SIR models defined on networks (Ganesh, Massoulie, and Towsley 2005; Prakash et al. 2012). This has motivated the development of interventions for controlling epidemics or malware spread that reduce the reproductive number or the spectral radius (Newman 2003; Tong et al. 2012; Mieghem et al. 2011) which is by vaccinations or anti-virus software as the case may be. However, it is commonly observed that there is limited compliance to directives to get vaccinated, or to install anti-virus software, because of the competing incentives. There is a large literature on modeling such behavior using non-cooperative game models.

There are several works in this regard based on differential equation models, e.g., (Bauch and Earn 2004; Khouzani, Sarkar, and Altman 2012; Khouzani, Altman, and Sarkar 2012; Khouzani, Sarkar, and Altman 2011; Galvani, Reluga, and Chapman 2007; Reluga 2010). These commonly rely on simplified assumptions about uniform mixing of the players in the population which greatly simplifies the problem and enables the derivation of tight analytical bounds and a detailed characterization. However, it is not easy to extend these approaches to heterogeneous networks. The work of (Aspnes, Rustagi, and Saia 2007) was among the first to study these problems on networks, especially from an algorithmic perspective and was further developed in (Kumar et al. 2010). They characterize NE in terms of the network properties, such as the maximum degree and conductance, and develop algorithms for approximating the PoA. However, both these approaches focus on an SIR model with a transmission probability of 1, so that it suffices to consider connectivity instead of percolation. (Omic, Orda, and Mieghem 2009) develop a formulation by combining a N-intertwined, SIS epidemic model with a non-cooperative game model, which simplifies the diffusion process by a mean-field approximation. Our *EC game* formulation incorporates a realistic epidemic model over a generous heterogeneous network, thus bridging both these approaches.

A common issue with all such game-theoretical formulations is that they involve utility functions that require quite a lot of non-local information to compute, and it is not clear how implementable such games might be. For instance, in most epidemic game formulations (e.g., (Aspnes, Chang, and Yampolskiy 2006; Kumar et al. 2010; Omic, Orda, and Mieghem 2009)), whether they be network-based or not, the utility function involves a term of the form "$p_v(\mathbf{a})$" (or something similar), which corresponds to the probability that node v gets infected, given the strategy vector $\mathbf{a}$. In general, this is computationally hard to estimate, and requires a lot of information about the network.

The approach of (Kumar et al. 2010) attempts to address some of these issues by limiting the amount of graph information needed in the utility function; specifically, they fix a parameter d, and consider $p_v(\mathbf{a})$ restricted to the graph induced by nodes within distance d of node v. The utility function in our *EC game* also involves such a global quantity in the form of the spectral radius, which determines the cost $\text{cost}(v, \mathbf{a})$ for node v. However, from the characterization of (Ganesh, Massoulie, and Towsley 2005), it follows that the epidemic duration is dependent on the spectral radius, which can be a good proxy for estimating this cost.

Other network security aspects have also been extensively studied via non-cooperative games, e.g., Interdependent Security games (IDS) (Kearns and Ortiz 2004) and the security game models of Grossklags et al. (Grossklags, Christin, and Chuang 2008). Another related thread is the use of Stackelberg strategies, e.g., (Jain, Conitzer, and Tambe 2013).

8 Conclusion

Our EC game formulation allows for a tractable way to incorporate realism in both the network and disease models; this is a natural game-theoretic analogue of the approaches to reduce the spectral radius to control epidemics. The main technical contribution of our paper is the analysis of the rich network effects in the structure of equilibria, which might give further insights to understanding the incentives for individuals to secure themselves, and to affect it. We obtain tight bounds on the maximum and minimum cost NE, and the PoA in general and random graphs. We find it interesting that our empirical results on several real and random networks corroborate well with our analytical bounds. Our results show that the PoA is high in general, and degree based Stackelberg strategies do not help in mitigating it; developing more effective strategies is an interesting open problem. The spectral properties of general and random graphs that we identify would be useful in future studies of the epidemic processes in these networks.

Acknowledgments. This work has been partially supported by the following grants: DTRA Grant HDTRA1-11-1-0016, DTRA CNIMS Contract HDTRA1-11-D-0016-0010, NSF Career CNS 0845700, NSF ICES CCF-1216000, NSF NETSE Grant CNS-1011769 and DOE DE-SC0003957. Also supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D12PC000337, the US Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

References

Aspnes, J.; Chang, K.; and Yampolskiy, A. 2006. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *J. Comput. Syst. Sci.* 72(6):1077–1093.

Aspnes, J.; Rustagi, N.; and Saia, J. 2007. Worm versus alert: Who wins in a battle for control of a large-scale network? In *Proceedings of OPODIS*.

Bauch, C., and Earn, D. 2004. Vaccination and the theory of game. *PNAS* 101(36):13391–13394.

Chung, F. R. K., and Lu, L. 2006a. *Complex graphs and networks*, volume 107. American Mathematical Society.

Chung, F. R. K., and Lu, L. 2006b. The volume of the giant component of a random graph with given expected degrees. *SIAM J. Discrete Math.* 20(2):395–411.

Galvani, A.; Reluga, T.; and Chapman, G. 2007. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. *Proceedings of the National Academy of Sciences* 104(13):5692–5697.

Ganesh, A.; Massoulie, L.; and Towsley, D. 2005. The effect of network topology on the spread of epidemics. In *Proceedings of INFOCOM*.

Grassly, N., and Fraser, C. 2008. Mathematical models of infectious disease transmission. *Nature Reviews Microbiology* 6:477–487.

Grossklags, J.; Christin, N.; and Chuang, J. 2008. Secure or insure? a game-theoretic analysis of information security games. In *Proceedings of World Wide Web Conference (WWW)*.

Jain, M.; Conitzer, V.; and Tambe, M. 2013. Security scheduling for real-world networks. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Kearns, M., and Ortiz, L. 2004. Algorithms for interdependent security games. In *Proceedings of the Advances in Neural Information Processing Systems*.

Khouzani, M. H. R.; Altman, E.; and Sarkar, S. 2012. Optimal quarantining of wireless malware through reception gain control. *IEEE Trans. Automat. Contr.* 57(1):49–61.

Khouzani, M. H. R.; Sarkar, S.; and Altman, E. 2011. A dynamic game solution to malware attack. In *Proceedings of INFOCOM*, 2138–2146.

Khouzani, M. H. R.; Sarkar, S.; and Altman, E. 2012. Saddle-point strategies in malware attack. *IEEE Journal on Selected Areas in Communications* 30(1):31–43.

Krivelevich, M., and Sudakov, B. 2001. The largest eigenvalue of sparse random graphs. *ArXiv Mathematics e-prints*.

Kumar, V. S. A.; Rajaraman, R.; Sun, Z.; and Sundaram, R. 2010. Existence theorems and approximation algorithms for generalized network security games. In *Proceedings of IEEE ICDCS*.

Lelarge, M., and Bolot, J. 2009. Economic incentives to increase security in the internet: The case for insurance. In *Proceedings of IEEE Infocom*.

Leskovec, J. 2011. Snap: Stanford network analysis project.

Mieghem, P. V.; Stevanovic, D.; Kuipers, F. F.; Li, C.; van de Bovenkamp, R.; Liu, D.; and Wang, H. 2011. Decreasing the spectral radius of a graph by link removals. *Physical Review E.* 84.

Newman, M. 2003. The structure and function of complex networks. *SIAM Review* 45(2):167–256.

Nisan, N.; Roughgarden, T.; Tardos, E.; and Vazirani, V. V. 2007. *Algorithmic Game Theory*. Cambridge University Press.

Omic, J.; Orda, A.; and Mieghem, P. V. 2009. Protecting against network infections a game theoretic perspective. In *Proceedings of IEEE INFOCOM*.

Opsahl, T., and Panzarasa, P. 2009. Clustering in weighted networks. *Social Networks* 31(2).

Prakash, B. A.; Chakrabarti, D.; Valler, N.; Faloutsos, M.; and Faloutsos, C. 2012. Threshold conditions for arbitrary cascade models on arbitrary networks. *Knowledge and Information Systems* 33:549–575.

Reluga, T. 2010. Game theory of social distancing in response to an epidemic. *PLOS Computational Biology* 6(5). e1000793.

Roughgarden, T. 2001. Stackelberg scheduling strategies. In *Proceedings of the Annual ACM symposium on Theory of computing (STOC)*.

Saha, S.; Adiga, A.; and Vullikanti, A. K. S. 2014. Equilibria in epidemic containment games. Technical report, available at http://staff.vbi.vt.edu/ssaha/papers/ecgame_extended.pdf.

Tong, H.; Prakash, B. A.; Eliassi-Rad, T.; Faloutsos, M.; and Faloutsos, C. 2012. Gelling, and melting, large graphs by edge manipulation. In *Proceedings of CIKM*.