Protecting Against Network Infections: A Game Theoretic Perspective

Jasmina Omic
Network Arhitecture
and Services
Delft University of Technology
Delft, Netherlands
Email: j.s.omic@ewi.tudelft.nl

Ariel Orda
Department of Electrical
Engineering Technion,
Israel Institute of Technology
Haifa, Israel 32000
Email: ariel@ee.technion.ac.il

Piet Van Mieghem
Network Arhitecture
and Services
Delft University of Technology
Delft, Netherlands
Email: P.F.A.VanMieghem@tudelft.nl

***Abstract*— Security breaches and attacks are critical problems in today's networking. A key-point is that the security of each host depends not only on the protection strategies it chooses to adopt but also on those chosen by other hosts in the network. The spread of Internet worms and viruses is only one example. This class of problems has two aspects. First, it deals with epidemic processes, and as such calls for the employment of epidemic theory. Second, the distributed and autonomous nature of decision-making in major classes of networks (e.g., P2P, ad-hoc, and most notably the Internet) call for the employment of game theoretical approaches. Accordingly, we propose a unified framework that combines the *N*-intertwined, *SIS* epidemic model with a noncooperative game model.**

We determine the existence of a Nash equilibrium of the respective game and characterize its properties. We show that its quality, in terms of overall network security, largely depends on the underlying topology. We then provide a bound on the level of system inefficiency due to the noncooperative behavior, namely, the "price of anarchy" of the game. We observe that the price of anarchy may be prohibitively high, hence we propose a scheme for steering users towards socially efficient behavior.

I. Introduction

Network security has become one of the major challenges of communication networking. Security breaches come in many forms, such as the spread of viruses and worms in the Internet, as well as social engineering compromises and direct exploitation of a host's vulnerability. In such a breach, an exposed (*infected*) host becomes a new source of infection, which attacks other unprotected machines. We shall term any such breach as a *virus*, and its spread as an *infection process*.

In order to overcome such breaches and their implied damage, network users and nodes can be equipped with protection and curing tools, to which we shall refer as *protection strategies* (or *curing strategies*). For example, a protection strategy is an antivirus software, with its signature quality and the speed of response to new virus strains. An important property of a protection strategy is the frequency with which the host is checked and secured. Several factors influence the choice of the protection strategy, most notably the significance and value of the protected information, the probability of infection, the overhead of employing the protection strategy and its (monetary) price.

A major source of complication in network security is the typically autonomous nature of decision making in the network, most notably in the Internet. Indeed, administration and policy enforcement are not possible at the inter-networking level (as opposed to intra-networking within a company), hence a majority of users is left to make independent decisions, including the choice of the protection strategy. Clearly, while such decisions are made autonomously by users and nodes, they do influence other users, through the potential infection processes. This gives rise to a *noncooperative game* [17], whose investigation is the subject of this paper.

The noncooperative nature of the process of protecting against viruses potentially has major implications on the network and its users. Indeed, the trade-offs between the damage infection and the price and overhead of a protection strategy may be vastly different across users, hence placing certain nodes in an unfair position to protect much of the network by investing more than other nodes. For example, consider an internetwork that includes a company network that has servers with vital data, as well as hosts of individual users that are divided into subnetworks. Suppose that each machine is administrated by an independent decision maker. The company's servers will seek a higher level of protection, due to the importance of the information they contain and the fact that many users' hosts will be able to connect to them. On the other hand, for individual users, the price of tools such as antivirus software and host's firewall will often be too high compared to the value of the security they provide. Moreover, a user host often has just a small number of neighbors, namely other hosts that can connect to them hence potentially endangering them. Therefore, these hosts would compromise with a lower level of protection, hence decreasing the level of security of the whole network, and, in turn, putting a higher burden on the company's servers.

Investigating such a *network security game* requires a proper model, which captures both the process of infection spread as well as the game's structure. We obtain such a model by combining game theoretic principles with epidemic theory [28].

While extensive studies have been done on spreading processes in networks, its game theoretic perspectives have hardly

been considered. Epidemics on computer networks were studied in [11]–[14]. The *SIS* (*Susceptible Infected Susceptible*) model and the influence of the topology on the spreading process were extensively studied in [7]–[10]. We shall employ the *N-intertwined model*, proposed and studied in [1], to model the spreading process under the influence of a curing process.

With the rapid growth of Internet and decentralization of services, the game theoretical framework has become an important tool for network modeling. Game theoretic models have been employed in various networking contexts, such as flow control [18], [19], routing [20], [26], and bandwidth allocation [21]. These studies mainly investigated the structure of the network operating points i.e., the Nash equilibria of the respective games. Such equilibria are inherently inefficient [27] and, in general, exhibit suboptimal network performance. As a result, the question of how much worse the quality of a Nash equilibrium is with respect to a centrally enforced optimum has received considerably attention e.g., [22]–[24]. In order to quantify this inefficiency, several conceptual measures have been proposed in the literature. Most notably, the *price of anarchy* [25] corresponds to a worst-case analysis and it is the ratio between the *worst* Nash equilibrium and the social optimum.

Recently, network security under a game theoretical setting was considered in [5]. That study addressed the interplay between protection and infection and noted the influence of the underlying topology, however it focused on the case of just two simple strategies, namely being fully protected or totally unprotected. In particular, if a node chooses the "fully protected" strategy, its security level does not depend on those of its neighbors. Somewhat similar work appears in [30], [31], where Lelarge *et al.* generalize game settings to incorporate weak security solutions. However, the problem is tractable only for sparse random graphs and trees. In [29], Jiang *et al.* consider a network security game, where the level of security is determined by weights assigned to a topology and the infection process is not modeled. A framework that is closer to the present study is that of *IDS* (Interdependent security games) [15], [16]. As opposed to [5], in *IDS* games security levels of agents are interdependent even when they choose the "protected strategy". However, the *IDS* framework does not consider the influence of the underlying topology, as it restricts its attention to the case of a complete graph.

The *N*-intertwined epidemic model takes into account the topology of the *relation network*. Each host stores IP addresses, e-mail accounts and passwords of other hosts and systems. This stored information defines a relation between hosts. If a host is compromised, then all reachable hosts can be attacked as well. The relation network is an abstraction that determines the hosts that can be infected. The relation topology is a significant aspect of the spreading process [7], [13], [14].

For given curing strategies of the individual nodes, it is possible to calculate the probability of infection and the average infection time for individual nodes [1]. With such a powerful model, we establish the existence of a Nash equilibrium point (NEP) [17] and characterize the strategies of nodes at equilibrium, as well as the overall network performance.

Our contributions

The main contributions of this study can be summarized as follows.

1) Introduction of a novel framework for network security under the presence of autonomous decision makers with multiple (possibly infinite) protection strategies. The model encompasses general (arbitrary) topologies.
2) Establishment of the existence of a Nash equilibrium point and characterization of its properties.
3) Discussion of the related global (i.e., "social") optimization problem, and establishment of an upper bound on the price of anarchy.
4) Proposal of schemes for a network manager to influence the game, resulting in a potentially major improvement in the level of network security.

II. The N-intertwined model

We proceed to review the *N*-intertwined model introduced and discussed in [1]. A relation network is modeled as a connected graph with users being nodes. By separately observing each node, the security breach (virus) spread is modeled in a bidirectional network specified by a symmetric adjacency matrix A. A node i at time t can be in one out of two states: *infected*, with probability $v_i(t) = \Pr[X_i = 1]$ or *healthy*, with probability $1 - v_i(t)$. The sum of the probabilities of being infected and susceptible are equal to 1 because a node can only be in one of these two states. The state of a node i is specified by a Bernoulli random variable $X_i \in \{0,1\}$: $X_i = 0$ for a susceptible node and $X_i = 1$ for an infected node. We assume that the curing process per node i is a Poisson process with rate δ_i, and that the infection rate per link is a Poisson process with rate β which is imminent for all nodes and thus constant in the network. For a node i, we can formulate the following differential equation

$$\frac{dv_i(t)}{dt} = \beta(1 - v_i(t)) \sum_{j=1}^{N} a_{ij} v_j(t) - \delta_i v_i(t)$$

where a_{ij} is the element of the adjacency matrix A and it is equal to 1 if the nodes i and j are connected, otherwise it is 0. A node is not considered connected to itself, i.e., $a_{ii} = 0$. The probability of a node being infected depends on the probability that it is not infected $(1 - v_i(t))$ multiplied with the probability that a neighbor j is infected $a_{ij} v_j(t)$ and that it tries to infect the node i with the rate β. We denote the set of curing rates for a network by the vector $\vec{\delta} = [\ \delta_1 \quad \delta_2 \quad .. \quad \delta_N\ ]^T$. Detailed derivations are given in [1] and [2].

In the steady state $\frac{dv_i(t)}{dt} = 0$, $v_i(t) = v_{i\infty}$, for each node $1 \leq i \leq N$, we have that

$$v_{i\infty} = \frac{\beta \sum_{j=1}^{N} a_{ij} v_{j\infty}}{\beta \sum_{j=1}^{N} a_{ij} v_{j\infty} + \delta_i} \tag{1}$$

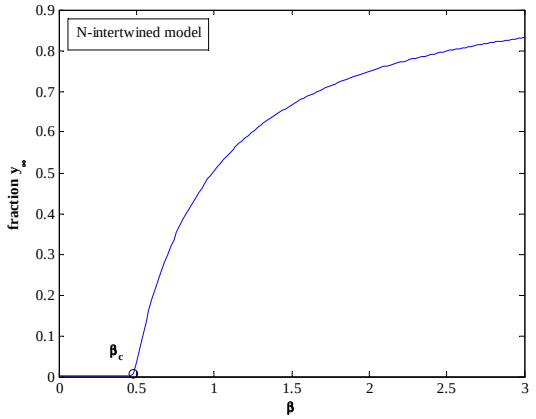

Fig. 1. Fraction of infected nodes as a function of infection rate β. The epidemic threshold is denoted by β_c.

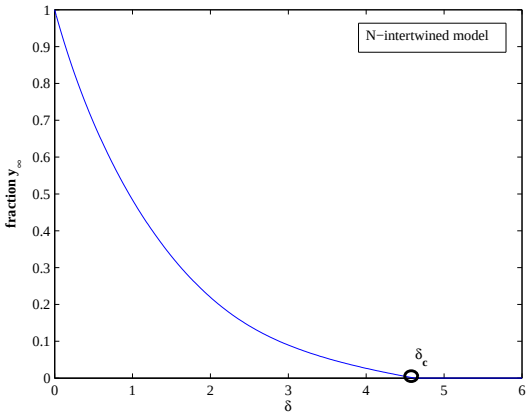

Fig. 2. Fraction of infected nodes as a function of curing rate for all curing rates equal.

This system of equations has $2N$ solutions with one positive solution and one solution equal to 0 [1]. The positive solution gives the probability of nodes being in the infected state. In the general case, the positive solution is differentiable to the threshold, where two solutions meet. At the threshold, the solution is not differentiable in general. We shall explain the meaning of the threshold later.

The fraction of infected nodes at any given time t can be calculated as a sum of probabilities that the nodes are infected $y(t) = \frac{1}{N}\sum_{j=1}^{N} v_j(t)$ and in the steady state $y_\infty = \frac{1}{N}\sum_{j=1}^{N} v_{j\infty}$.

For a Bernoulli random variable with infection probability $v_i(t)$, we have $E[X_i(t)] = v_i(t)$. For a fixed set of curing rates per nodes, the fraction of infected nodes as a function of the spreading rate per link is given in Figure 1 . The model has a threshold value $\beta = \beta_c$ below which the epidemic extinguishes and the number of infected nodes in the steady state is 0. The threshold [2] is equal to

$$\beta_c = \frac{1}{\lambda_{\max}(A_\delta)} \tag{2}$$

where $\lambda_{\max}(A_\delta)$ is a maximal eigenvalue of the matrix $A_\delta = diag(\frac{1}{\delta_i})A$. There are many different matrices A_δ with the same largest eigenvalue, and as a consequence, there are many different curing rate vectors $\overrightarrow{\delta}$ that result with the same threshold β_c. The epidemic threshold is defined as follows: for $\beta < \beta_c$, the infection dies out - the mean epidemic lifetime is of order $log(n)$ and for $\beta > \beta_c$ the epidemic persists with the average number of infected nodes y_∞. If all the curing rates are the same $\delta_1 = \delta_2 = .. = \delta_N = \delta$, the threshold is given by $\frac{\beta_c}{\delta_c} = \frac{1}{\lambda_{\max}(A)}$. For $\beta_c = 1$, critical curing rate [1] is

$$\delta_c = \lambda_{\max}(A) \tag{3}$$

For example, the largest eigenvalue of a line graph is $\lambda_{\max}(A) \simeq 2 = \delta_c$, while that of a star topology is $\lambda_{\max}(A) = \sqrt{N-1} = \delta_c$. These two graphs are interesting examples, as both have the same number of links $L = N-1$. Thus, in the homogenous case, the level of protection required for a star topology is significantly higher than for a line topology with the same number of nodes.

Figure 2 illustrates the behavior of the fraction of infected nodes as a function of the curing rate δ.

To simplify the analysis, we scale all the rates such that the spreading rate per link is $\beta = 1$.

For a fixed $\beta = 1$ we can vary the curing rates such that when $\lambda_{\max}(A_{\delta_c}) = 1$, we have reached the threshold point. The critical curing rates are denoted by the vector $\overrightarrow{\delta_c}$. As shown in [2], there are three possibilities:

$$\begin{cases} \lambda_{\max}(A_\delta) < 1 & \text{uninfected network} \\ \lambda_{\max}(A_\delta) = 1 & \text{critical threshold} \\ \lambda_{\max}(A_\delta) > 1 & \text{infected network} \end{cases} \tag{4}$$

In Figure 1, we can observe the threshold behavior of an infected network.

A. A bound on the largest eigenvalue of the matrix A_δ

We shall employ the following lower bound, derived in [2]. The bound on the largest eigenvalue of the matrix A_δ is

$$\lambda_{\max}(A_\delta) \geq \frac{2L}{\sum_{i=1}^{N} \delta_i} \tag{5}$$

where L is the number of links in the network. The inequality is equivalent to $\lambda_{\max}(A_\delta) \geq \frac{E[D]}{E[\delta]}$, where $E[D] = \frac{1}{N}\sum_{i=1}^{N} d_i$ is the average degree and $E[\delta] = \frac{1}{N}\sum_{i=1}^{N} \delta_i$ is the average curing rate. The inequalities become equalities if $\delta_i = d_i$.

III. The virus protection game

Consider a network with N nodes defined by an adjacency matrix A. This is an underlying topology over which a virus can spread with an infection rate $\beta = 1$ per link. Each node i chooses its curing rate among an infinite number of strategies from the interval $\delta_i \in [0, \delta_{max}]$, where $\delta_{max} > 1/c_{min}$, so as to minimize its *cost function* $J^{(i)} = c_i\delta_i + v_{i\infty}$, where c_i is a

positive value that stands for the *relative price of protection* and quantifies the trade-off of the user between the money (and any overhead) invested in protection and the penalty of being infected. c_{min} is the minimum of all c_i. The curing rate state space can be bounded by $\delta_{max} > 1/c_{min}$ without loss of generality as shown in Lemma 5. For example, a firm may give much importance to security, hence its relative price of protection would be smaller than that of a private Internet user. Thus, the cost function of a node i is a weighted sum of the curing rate per node, δ_i, and the probability of infection in the steady-state, $v_{i\infty}$.

To sum up, the game has N players, corresponding to the nodes of a graph. Each node i chooses a curing strategy δ_i so as to minimize its cost function $J^{(i)}$. The strategies chosen by all nodes result in a certain steady-sate infection probability for each node, $v_{i\infty}$. The latter is also the percentage of time that the node is in the infected state. We term this game as *the virus protection game*.

A *Nash equilibrium point* (*NEP*) is a strategy profile such that no user can benefit from unilaterally changing its strategy. We shall denote an NEP by a vector $\overrightarrow{\delta} = [\ \delta_1 \quad \delta_2 \quad .. \quad \delta_N\ ]^T$ and a corresponding vector of individual probabilities of infection $V_\infty = [v_{1\infty}, v_{2\infty}, ..v_{N\infty}]^T$. The probability of infection $v_{i\infty}$ depends on the states of other nodes as in equation (1) and, therefore, the cost function $J^{(i)}(\overrightarrow{\delta}, A)$ depends on the vector of curing strategies and the system (network) parameters.

In the Appendix, a simple case with just two nodes and one link illustrates the cost function behavior and the optimization process of individual nodes. An example of a cost function for a network with two nodes is given in Figure 3. The cost function of the second node $J^{(2)}$ is calculated for different values of the constants c_1 and c_2. The cost function can only increase due to the fact that the curing rate price is much larger than the corresponding security it offers as in the case $c_1 > 1, c_2 > 1$. It can decrease due to the decrease of the infection probability if the curing rate price is not that significant as in the case $c_1 < 1, c_2 < 1$. A network is clean of viruses if the curing rates of all nodes satisfy the threshold relation (2). Whether the network is able to reach the threshold depends on the price each node is prepared to pay.

Clearly, it is of interest to establish the existence of an NEP and characterize it. We shall show that a Nash equilibrium always exists. We shall also show that the NEP's quality, in terms of overall network security and protection against viruses, largely depends on the properties of the underlying topology.

A. Nash Equilibrium

First we indicate that the individual probabilities of infection $v_{i\infty}$ are strictly convex in δ_i. This will be later used to establish the quasi-convexity of the cost function, with which we shall prove the existence of a Nash equilibrium.

The following result is taken from [2].

Lemma 1: For fixed curing rates of other nodes, the probability of infection $v_{i\infty}(\delta_i)$ is a strictly convex function in δ_i.

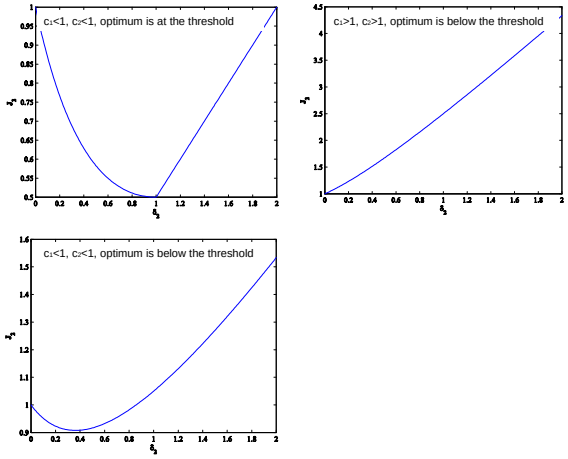

Fig. 3. Cost function for a network with two nodes and different parameters c_1 and c_2. The curing rate δ_1 of the first node is optimal.

Lemma 2: For the cost function defined as $J^{(i)}(\delta_i, \delta_{-i}) = c_i\delta_i + v_{i\infty}(\delta_i, \delta_{-i})$, $c_i > 0$, the function is quasi-convex in each $\delta_j,\ j = 1..N$.

Proof: Let us first show that $J^{(i)}$ is quasi-convex in δ_j. For any $j \neq i$, the cost function $J^{(i)}$ is quasi-convex in j

$$\begin{aligned} & c_i\delta_i + v_{i\infty}(\delta_i, \delta_{-i,\lambda\delta_j+(1-\lambda)\delta'_j}) \\ \leq\ & \max\{c_i\delta_i + v_{i\infty}(\delta_i, \delta_{-i,\delta'_j}), c_i\delta_i + v_{i\infty}(\delta_i, \delta_{-i\delta_j})\} \end{aligned}$$

which holds for any c_i. The probability of a node being infected $v_{i\infty}(\delta_i, \delta_{-i})$ is convex function in δ_i. When δ_i reaches the threshold value for the curing rate δ_{ic}, the infection probability becomes zero. The cost function $J^{(i)}$ is a sum of a convex function and a linear - strictly increasing function and, therefore, it is quasi-convex in the domain of interest. ■

Theorem 3: For a set of strategies $\forall i\ \delta_i \in [0, \delta_{max}]$ which is non-empty, compact and convex, and for the continuous and quasi-convex cost function in each δ_i, the game has at least one Nash equilibrium.

Proof: The set of minimizers of a quasi-convex function on a convex set is convex. Continuity of the cost function implies upper-hemicontinuity of the point to set correspondence [3]. ■

The existence of an NEP means that the protection game has at least one stable point. We proceed to explore the properties of the Nash equilibria, which indicate the ability of a network to protect itself from epidemics.

B. Characterization of equilibrium

An NEP can be in two very different regions, namely above or at the threshold (4). The NEP does not exist below the threshold: in order to realize why, consider the following example. A node i has a curing rate δ_i and the curing rates of other nodes are fixed. If inequality holds $\lambda_{\max}(A_{\delta_c}) < 1$, the cost function of a node i is $J^{(i)} = c_i\delta_i$. Therefore, a node i is able to reduce the curing rate such that its cost function decreases, because the probability of infection $v_{i\infty} = 0$ will not change.

If the optimum of the cost function is reached at the threshold point, we can have multiple Nash equilibria. At the critical point $\overrightarrow{\delta}_c$, we have that

$$\lambda_{\max}(A_{\delta_c}) = 1 \tag{6}$$

This can be easily shown on a simple example of a two-nodes network (see the Appendix). This infinite set of NEPs is bounded and we will establish the worst case scenario.

When the network is in the regime above the threshold, numerical calculations suggest that only one equilibrium exists. However, this was not established formally, and the uniqueness of the NEP above the threshold remains an open problem.

In the case of two nodes (see the Appendix), a unique Nash equilibrium exists if $\sqrt[3]{c_1c_2^2} + \sqrt[3]{c_1^2c_2} > 1$, $(\delta_1\delta_2 < 1)$. For $\sqrt[3]{c_1c_2^2} + \sqrt[3]{c_1^2c_2} \leq 1$, $(\delta_1\delta_2 \geq 1)$, the example with two nodes shows multiple NEPs.

Next, we determine the influence of the relative price of protection vector $\overrightarrow{c}$ on the Nash equilibria. This shall be later used to bound the equilibrium value of the cost function.

In some cases, all the nodes of a network decide not to protect themselves against infection, leaving the overall network unsecured. If a node is unprotected, i.e., $\delta_i = 0$, the infection probability is always equal to 1 and it does not depend on the curing rates of other nodes.

The next theorem makes a distinction between networks with a vector $\overrightarrow{c}$ such that every node chooses not to be protected at all and networks where the equilibrium point is reached with curing rates larger than 0.

Theorem 4: In a virus protection game, for a network with N nodes and with cost function for a node i defined as

$$J^{(i)} = c_i\delta_i + v_{i\infty}$$

the following hold:

1) If $\forall i\ c_i \geq 1$, the only Nash equilibrium is defined by the curing rate vector $\overrightarrow{\delta} = [\ 0\quad 0\quad ..\quad 0\ ]^T$.
2) If $c_i < \frac{1}{d_i}$, where d_i is the degree of a node i, the curing rate of a node i in the Nash equilibrium is different from zero, $\delta_i \neq 0$.

Proof: Point 1.

Consider any two nodes in the network i, j. Since the network is a connected network with N nodes, at least one node from the pair will be connected to at least one other node. For the same δ_i, δ_j,we can compare the infection probability of neighboring nodes in a network with N nodes $(v_{i\infty}, v_{j\infty})$, with the case of a network with only two connected nodes $v_{i\infty}, v_{j\infty}^{(2)}$. The infection probability of a connected node will increase due to possible connections to infectious nodes and its neighbor will also feel this effect. It holds that

$$v_{i\infty} \geq v_{i\infty}^{(2)}, v_{i\infty} \geq \frac{1-\delta_i\delta_j}{1+\delta_i} \tag{7}$$

Similarly, we have

$$v_{j\infty} \geq v_{j\infty}^{(2)}, v_{j\infty} \geq \frac{1-\delta_i\delta_j}{1+\delta_j}$$

Equality holds for $N = 2$.

For a node i to increase δ_i from zero it has to hold that for some $\delta_i > 0$

$$\begin{aligned} J^{(i,\delta_i=0)} &> J^{(i,\delta_i>0)} \\ 1 &> c_i\delta_i + v_{i\infty} \\ v_{i\infty} &< 1 - c_i\delta_i \end{aligned} \tag{8}$$

and similarly for node j.

From (7) and (8) we have

$$\delta_j > c_i\delta_i + c_i - 1 \tag{9}$$

and similarly, for node j it holds

$$\delta_i > c_j\delta_j + c_j - 1 \tag{10}$$

from (9) and (10) we have

$$(1 - c_ic_j)\delta_j > -1 + c_ic_j$$

which gives for positive c_i, c_j, δ_j

$$\begin{aligned} \delta_j &> -1,\ c_ic_j < 1 \\ \delta_j &< -1,\ c_ic_j > 1 \end{aligned}$$

And similarly we have for δ_i. We can conclude that for $c_i > 1, c_j > 1$ for nodes i and j there is no other solution than $\delta_i = \delta_j = 0$. We can continue the process for any other two nodes in the network concluding that the only solution is $\overrightarrow{\delta} = [\ 0\quad 0\quad ..\quad 0\ ]$, which proves the first point of the theorem.

Point 2.

For scaled rates such that $\beta = 1$, the infection probability of a node i is

$$v_{i\infty} = \frac{\sum_{j=1}^{N} a_{ij}v_{j\infty}}{\sum_{j=1}^{N} a_{ij}v_{j\infty} + \delta_i}$$

The first derivative of the cost function for a node i for $\delta_i = 0$ is

$$\begin{aligned} \left.\frac{dJ^{(i)}}{d\delta_i}\right|_{\delta_i=0} &= c_i + \left.\frac{\delta_i \sum_{j=1}^{N} a_{ij}\frac{\partial v_{j\infty}}{\partial \delta_i} - \sum_{j=1}^{N} a_{ij}v_{j\infty}}{(\sum_{j=1}^{N} a_{ij}v_{j\infty} + \delta_i)^2}\right|_{\delta_i=0} \\ &= c_i - \frac{1}{\sum_{j=1}^{N} a_{ij}v_{j\infty}} \end{aligned} \tag{11}$$

which achieves its maximum for $\sum_{j=1}^{N} a_{ij}v_{ji} = d_i$. If $c_i < \frac{1}{d_i}$, the first derivative of the cost function is smaller than zero for any set of curing rates of other nodes. This proves the second point of the theorem. ■

Theorem 4 shows that if antivirus software or other means of protection against viruses are too expensive, such that $\forall i\ c_i \geq 1$, the NEP is unique and the network will end up in the completely infected state. In order to steer a decision maker i to chose protection over infection, the relative price should

satisfy the inequality $c_i < \frac{1}{d_i}$. The higher the degree of a node, the more it is exposed to infection, hence the required relative price is lower. For example, a large firm typically has many interactions over the Internet and thus its degree is higher. Therefore, its required relative price of antivirus software is lower than that of a smaller firm, which has less opportunities to get infected.

In order to determine the global optimum and the worst case scenario that can happen in a virus protection game, we establish an upper bound on the minimum of the cost function $J_{\min}^{(i)}$.

Lemma 5: The minimum of the cost function $J_{\min}^{(i)}$ is bounded by $J_{\min}^{(i)} \leq 1$.

Proof: For curing rate $\delta_i > 0$, $v_{i\infty}$ is bounded. The cost function for $\delta_i = 0$ is $J^{(i)}(\delta_i = 0) = 1$ and the minimum cannot be larger than this value. Therefore, we have

$$\begin{aligned} J_{\min}^{(i)} &= c_i \delta_{iopt} + v_{iopt} \leq 1 \qquad (12) \\ \delta_{iopt} &\leq \frac{1 - v_{iopt}}{c_i} \leq \frac{1}{c_i} \end{aligned}$$

In the case of a network above the threshold, inequality (12) holds because the function's minimum cannot be larger than 1 ($J^{(i)}(\delta_i = 0) = 1$). We have

$$\begin{aligned} J^{(i)} &= c_i \delta_{iopt} \leq 1 \\ \delta_{iopt} &\leq \frac{1}{c_i} \end{aligned}$$

■

If multiple Nash equilibria exist, the curing vector $\overrightarrow{\delta_c}$ is bounded as in Lemma 5. If the relative price of a protection strategy for a node *i* is too high, the other nodes in the network will have to pay more for the security of the whole network.

IV. Global optimum and price of anarchy

Clearly, if we could dictate the security strategy of the whole network, we would be able to obtain a better solution. However, as mentioned, the Internet is a decentralized system, and it is challenged by persistent virus infections. Therefore, security of the whole network depends on the decisions of independent users. Yet, is it possible and feasible to completely cure the Internet? How far is the Internet from the global optimal point in the presence of a virus protection game?

In an attempt to address these questions, in this section we discuss the global (social) optimization problem and the price of anarchy for a network with N nodes.

Assume that a "network manager" has the same relative price of security C for all the nodes. The corresponding (global) optimization problem is

Minimize

$$J_M = \sum_{j=1}^{N} v_{j\infty} + C \sum_{j=1}^{N} \delta_j$$

For some C, a network will be in the regime at the threshold and $\sum_{j=1}^{N} v_{j\infty} = 0$. The global cost function becomes $J_M = C \sum_{j=1}^{N} \delta_j$. A manager can be interested in optimizing the overall protection, such that $\sum_{j=1}^{N} v_{j\infty} = 0$, which reduces the problem to $J_M = C \sum_{j=1}^{N} \delta_j$.

In Section II, we have seen that a network can be in two significantly different states, namely above or at the threshold. These two states have to be discussed separately. Thus, we split the optimization problem into two different problems, namely: optimization of the network at the threshold and above the threshold. We assume that a manager optimizes in the regime where the network NEP is, i.e.: if a network NEP reaches the threshold, the manager will optimize with the constraint $\sum_{j=1}^{N} v_{j\infty} = 0$ while if a network NEP is above the threshold, the network manager will optimize the function $J_M = \sum_{j=1}^{N} v_{j\infty} + C \sum_{j=1}^{N} \delta_j$ with the constraint $\sum_{j=1}^{N} v_{j\infty} > 0$. In the case of multiple NEPs, where some are above and others are at the threshold, the network manager optimizes at the threshold.

The network is below the threshold if the curing rates of individual nodes satisfy the inequality

$$\lambda_{\max}(diag(\frac{1}{\delta_i})A) \leq 1$$

If the strict inequality $\lambda_{\max}(diag(\frac{1}{\delta_i})A) < 1$ holds, the vector $\overrightarrow{\delta}$ cannot be a Nash equilibrium point, because there is a point $\delta_j^* < \delta_j$ such that $\lambda_{\max}(diag(\frac{1}{\delta_i^*})A) = 1$. The equality $\lambda_{\max}(diag(\frac{1}{\delta_i})A) = 1$ can be a Nash equilibrium point, if the condition $(\forall i)\frac{\partial J^{(i)}}{\partial \delta_i} = 0$ is satisfied.

Above the critical threshold, as in Figure 1, the probabilities of infection $v_{i\infty}$ are larger than zero and interesting parameters for the optimization are the sum of infection probabilities and the sum of curing rates $J_M = \sum_{j=1}^{N} v_{j\infty} + C \sum_{j=1}^{N} \delta_j$.

A. Optimization at the threshold

As shown in section III-B, there can be an infinite number of NEPs in this regime. What is the best possible strategy a network manager can apply with the respect to the total protection used? What is the lowest total price for the complete Internet security and what is the worst state of the network? The set of Nash equilibria is bounded in this regime as $(\forall i)\delta_i < \frac{1}{c_i}$, thus the worst NEP is also bounded.

We proceed to determine the worst possible case of an NEP and the global optimal point.

Lemma 6: The worst case Nash equilibrium, when the network is at the threshold, is bounded by

$$J_M < C \sum_{j=1}^{N} \frac{1}{c_j} \qquad (13)$$

Proof: Each curing rate is bounded by the constant c_i as in Lemma 5 and the set of Nash equilibria is therefore bounded as in (13). ■

The minimum price that has to be paid for a network which is clean of viruses is determined by the number of links.

Theorem 7: The minimum global price for a network at the threshold is

$$J_M^{(\min)} = C\sum_{j=1}^{N} d_j$$

and it is reached for each $\delta_i = d_i$.

Proof: The lower bound (5) on the largest eigenvalue of matrix A_δ, we have that if $\frac{2L}{\sum_{j=1}^{N}\delta_j} \geq 1$, the largest eigenvalue obeys $\lambda_{\max}(A_\delta) \geq \frac{2L}{\sum_{j=1}^{N}\delta_j} \geq 1$. If $\lambda_{\max}(A_\delta) \geq 1$ the network is above the threshold. Therefore, if $2L \geq \sum_{j=1}^{N}\delta_j$ which is equivalent to $\sum_{j=1}^{N} d_j \geq \sum_{j=1}^{N}\delta_j$ the network is infected. The equality $\lambda_{\max}(A_\delta) = \frac{2L}{\sum_{j=1}^{N}\delta_j} = 1$ holds if $\delta_i = d_i$ and the epidemic threshold is reached. ■

The above result is in agreement with the results obtained from subcritical branching process theory [6].

It is possible that other curing distributions satisfy $\sum_{j=1}^{N} d_j = \sum_{j=1}^{N}\delta_j$, however the minimum of the sum of curing rates cannot be lower than $\sum_{j=1}^{N} d_j = 2L$. At the threshold, the minimum of the global cost function is a linear function of the number of links in the network $J_M^{(\min)} = 2LC$. The protection's efficiency depends on topological properties. In particular, for a complete graph, the minimum of the cost function is largest among all the graphs $J_M^{(\min)} = N(N-1)C$.

For an NEP such that $\delta_1 = \delta_2 = ..\delta_N = \delta_c$ (homogenous case), the price is $J_M = N\lambda_{\max}(A)$. In section II, we compared the largest eigenvalues of two topologies, namely line and star topologies, both with the same number of links $L = N-1$. The minimal global price is the same for these two topologies; however, in the homogenous case, the NEP price is significantly higher for a star topology $(J_M = CN\sqrt{N-1})$, than for a line topology $(J_M = 2CN)$.

B. Optimization above the threshold

The network is above the threshold if the curing rates satisfy the inequality $\lambda_{\max}(diag(\frac{1}{\delta_i})A) > 1$. In general, the optimization function is $J_M = C\sum_{j=1}^{N}\delta_j + \sum_{j=1}^{N} v_{j\infty}$. Due to the complexity of the general problem, we will not discuss it here. Instead, we will consider a simpler case where $\delta_i = (1-\alpha)d_i, \alpha < 1$.

In theorem 7, the minimu of the cost function on the threshold is reached for $\delta_i = d_i$; it is of interest to proceed with the same tactic above the threshold and determine the cost function for this strategy.

Lemma 8: If $\delta_i = (1-\alpha)d_i, \alpha < 1$, the probabilities of infection are all equal, namely $(\forall i)v_{i\infty} = \alpha$.

Proof: If $\delta_i = (1-\alpha)d_i, \alpha < 1$, we have that

$$\sum_{j=1}^{N}\delta_j = (1-\alpha)\sum_{j=1}^{N} d_j \leq \sum_{j=1}^{N} d_j$$

we are certainly above the threshold. Let us assume that $v_{i\infty} = \alpha$, from 1 we have that $\delta_i = (1-\alpha)d_i$. Because $V_\infty(\overrightarrow{\delta})$ is a bijective function [1], the solution is unique and $\delta_i = (1-\alpha)d_i$ implies that $(\forall i)v_{i\infty} = \alpha$. ■

The optimization function reduces to $J_M = C(1-\alpha)\sum_{j=1}^{N} d_j + N\alpha$. This function shows threshold behavior around the point $C = \frac{N}{\sum_i d_i}$.

$$J_M = \begin{cases} N, & C \geq \frac{N}{\sum_i d_i} \\ C\sum_i d_i & C < \frac{N}{\sum_i d_i} \end{cases} \tag{14}$$

For $C < \frac{N}{\sum_i d_i}$, the threshold is reached and the cost function is equal to the sum of degrees $J_M = C\sum_i d_i < N$. In the case $C \geq \frac{N}{\sum_i d_i}$, the optimum is reached for curing rates equal to zero and $J_M = N$.

Compared with the optimization at the threshold, where the cost function minimum can be $J_M^{(\min)} = O(N^2)$ for the complete graph, the cost function minimum cannot be larger than the size of the network $J_M^{(\min)} = O(N)$.

C. Price of Anarchy

In a noncooperative networking game, it is important to know the social welfare attained at the operating points, namely the Nash equilibria. A Nash equilibrium typically exhibits nonoptimal social welfare. This penalty of selfish behavior is quantified by the price of anarchy (*PoA*), which is defined as:

$$PoA = \frac{Cost\ of\ worst\ NEP}{Social\ optimum}$$

For the virus protection game, we have two significantly different regimes. At the threshold, the cost of the social optimum is $J_M^{(\min)} = C\sum_{j=1}^{N} d_j$. The cost of the worst Nash equilibrium is upper bounded as in lemma 6, under the constraint that the network's NEP is at the threshold, which depends on the vector $\overrightarrow{c}$.

Theorem 9: The price of anarchy for a network that reaches an NEP at the threshold is bounded by

$$PoA \leq \frac{\sum_{j=1}^{N}\frac{1}{c_j}}{C\sum_{j=1}^{N} d_j}$$

Proof: Follows from Lemma 6 and Theorem 7. ■

It is interesting to note that if nodes regard security as an important issue ($c_j \ll 1$, $\sum_{j=1}^{N}\frac{1}{c_j}$ is large), the price of anarchy

can be very high. It is necessary to help the network reach a more efficient NEP, by starting the system from a point close to the optimal.

If a network is above the threshold ($v_{i\infty} > 0$), we considered a special case where curing rates are proportional to degrees with the same factor $1-\alpha$. For this special case, we can estimate the price of anarchy.

Theorem 10: For global optimum calculated for curing rates proportional to the degrees, the price of anarchy above the threshold is bounded by

$$PoA \leq \begin{cases} \frac{\sum_{j=1}^{N} \frac{1}{c_j}}{N}, & C \geq \frac{N}{\sum_i d_i} \\ \frac{\sum_{j=1}^{N} \frac{1}{c_j}}{C \sum_{j=1}^{N} d_j}, & C < \frac{N}{\sum_i d_i} \end{cases}$$

Proof: Follows from Lemma 6 and equation (14). ■

V. Managing a network by constraining the infection probabilities

We proceed to discuss how a manager can influence and control the Nash equilibria of the virus protection game. In section III-B, we have shown how a Nash equilibrium depends on the relative price of protection vector $\overrightarrow{c}$. If a network is at the epidemic threshold, more Nash equilibrium points exist. By varying the relative price of protection vector $\overrightarrow{c}$ a manager can influence the network equilibrium point. A manager may be able to do that by determining (or affecting, e.g., through subsidies) the cost of protection means, e.g., antivirus software, hence indirectly influencing c_i. Here, the "manager" may be an antivirus supplier, which gives cheaper (per unit) antivirus to entities that have many Internet interactions and are densely connected to other nodes.

In section III-B, Theorem 4, some conditions are introduced that can give guidance to the choice of the relative price of protection. If all $c_i > 1$, there is only an unprotected state, and no one will buy antivirus protection. If $c_i < \frac{1}{d_i}$, a node will always invest some money in protecting itself. These results make it possible for an antivirus supplier to estimate what price will make a network more secure. In Theorem 9, we have seen that too low relative prices can lead a network further away from the global optimum. If large firms invest in expensive security, other nodes can buy cheaper antivirus software such that the network reaches the threshold.

The other option for a manager is to set up upper bounds on infection probabilities, for all relative prices $c_i \geq 1$, which will determine the Nash point as presented in Theorem 11.

Theorem 11: If $\forall i \ v_{i\infty} \leq B_i$, $\forall i \ c_i > 1$, the only Nash equilibrium is reached for

$$(\forall i) \ \delta_{i\min} = \frac{(1-B_i)\sum_j a_{ij} B_j}{B_i}$$

Proof: The result for the unconstrained case with N nodes shows that a node will tend to decrease its curing rates till they all become terminally infected (Theorem 4). The only NEP is out of the bounded region, thus the feasible minimum will be on the bound such that $\forall i \ v_{i\infty} \leq B_j$. Nodes that are above the bound will tend to decrease their curing rates, which draws other nodes to do the same till they all reach the constraint of infection probability B_j. Thus, the minimum is reached for $\forall i \ v_{i\infty} = B_j$. The minimum point for all the nodes exists, and the corresponding curing rate can be calculated from Equation (1), for $v_{j\infty} = B_j$

$$B_i = \frac{\sum_j a_{ij} B_j}{\sum_j a_{ij} B_j + \delta_i}$$

Now, the curing rates are

$$(\forall i) \ \delta_{i\min} = \frac{(1-B_i)\sum_j a_{ij} B_j}{B_i}$$

■

For $B_i \to 0$ and B_j finite for $j \neq i$, the curing rate of node i will tend to infinity $\delta_i \to \infty$.

For $B_j = B$, $\delta_i = d_i(1-B)$, where d_i is the degree of a node i, we have the vector of curing rates $\overrightarrow{\delta}$

$$\overrightarrow{\delta_{\min}} = \begin{bmatrix} \beta d_1(1-B) & \beta d_2(1-B) & .. & \beta d_N(1-B) \end{bmatrix}^T$$

However, this is not a stable point. If there is an unfair player in the game, which reduces its security against the rules $v_i > B$, it can cause other players to pay more than what was planned. The security of the whole network is harmed.

This result suggests a strategy for steering autonomous systems (ASs), or Internet service providers, to invest money in their own security, which is proportional to the number of "links", that is, interactions they have with other ASs. The way to "force security" upon ASs is by asking a certain fixed probability of infection $v_i < B$, for all relative prices $c_i > 1$. Together with the fact that the cheapest threshold, in terms of the total security ($\sum \delta_i$), is reached when the nodes are protected proportionally to their own degrees, this seems to be a very fair way to provide overall security. Bigger ASs with more connections towards other ASs will have to protect themselves more, in order to provide the same level of security, while smaller ASs will invest proportionally to their sizes and profits.

VI. Conclusion

We presented a novel framework for network security under the presence of autonomous decision makers. We have established the existence of a Nash equilibrium point (NEP) investigated its properties. In particular, we showed that, when the price of protection is relatively high (namely, $\forall i \ c_i \geq 1$), the only equilibrium point is that of a completely unprotected network; while if this price is sufficiently low for a node (namely, $c_i < \frac{1}{d_i}$), it will always invest in protecting itself.

A network can be in two significantly different regimes, namely above or at the threshold. If a network reaches Nash equilibrium at the threshold, multiple equilibria may exist. The question of uniqueness of the Nash equilibrium above the threshold remains an open question.

We determined the global (social) optimum for the case that the network is at the threshold and for a specific case when it is above the threshold. At the threshold, the minimum of the social cost function is $O(L)$, where L is the number of links in the network. Although the optimal value of the social cost is the same for networks with the same number of links L, the non-optimal distribution of curing rates at an NEP results in much worse social welfare in some topologies (e.g., a star graph) than in other topologies (e.g., a line graph). When optimizing above the threshold, we considered a specific case, for which we showed that the global cost function is always smaller than the number of nodes in the network. This specific case provides some insight on the social performance in the general case.

Finally, we have proposed two methods for steering the network equilibrium, namely by influencing the relative prices and by imposing an upper bound on infection probabilities.

ACKNOWLEDGEMENT

This research was supported by the Netherlands Organization for Scientific Research (NWO) under project number 643.000.503 and by Next Generation Infrastructures (Bsik). We would like to thank Tom Kleiberg for helping to prepare the final version of this paper.

APPENDIX

A. Unconstrained case with 2 nodes

For a network with two nodes and one link, each node chooses its strategy out of the interval $\delta_i \in [0, \delta_{max}]$. The cost function is defined as $J^{(i)} = c_i\delta_i + v_i$. The probabilities of infection follow from (1) as $v_1 = \frac{1-\delta_1\delta_2}{1+\delta_1}; v_2 = \frac{1-\delta_1\delta_2}{1+\delta_2}$. The Nash equilibrium point (NEP) is reached for $\delta_{1opt} = \sqrt[3]{\frac{1}{c_2c_1^2}} - 1, \delta_{2opt} = \sqrt[3]{\frac{1}{c_2^2c_1}} - 1$. If optimal solutions $\delta_{1opt}, \delta_{2opt}$ satisfy $\sqrt{\delta_{1opt}\delta_{2opt}} > 1$, the network NEP will be at the threshold and the cost functions reduce to $J^{(1)} = c_1\delta_1, J^{(2)} = c_2\delta_2$. In this case, both nodes will choose smaller curing rates than $\delta_{1opt}, \delta_{2opt}$ such that new values $\delta'_{1opt}\delta'_{2opt}$ satisfy $\sqrt{\delta'_{1opt}\delta'_{2opt}} = 1$. All the solutions that satisfy $\delta'_{1opt} < \sqrt[3]{\frac{1}{c_2c_1^2}} - 1, \delta'_{2opt} < \sqrt[3]{\frac{1}{c_2^2c_1}} - 1$ and $\sqrt{\delta'_{1opt}\delta'_{2opt}} = 1$ are optimal and nodes will not change their curing rates. This yields an infinite number of Nash equilibrium points.

In Figure 3, for a network with two nodes, the cost function of the second node $J^{(2)}$ is calculated for different values of constants c_1 and c_2.

REFERENCES

[1] P. Van Mieghem, J. Omic, R. Kooij, "Virus Spread in Networks", *accepted for IEEE/ACM Transactions on Networking.*

[2] P. Van Mieghem, J. Omic, "In-homogeneous Virus spread in Networks", *Technical report 20080801, http://www.nas.its.tudelft.nl/people/Piet/TUDelftReports.html*

[3] J.B. Rosen, "Existence and Uniqueness of Equilibrium Point for Concave N-person Games", *Econometrica,* Vol. 33, No.3, pp. 520-534, Jul. 1965.

[4] K. Keong Tan, J. Yu and X. Yuan, " Existence Theorems of Nash Equilibria for Non-Cooperative N-Person Games", *International Journal of Game Theory,* Vol. 24, No. 3, pp. 217-222, Sep. 1995

[5] J. Aspnes, K. Chang, A. Yampolskiy, "Inoculation Strategies for Victims of Viruses and Sum-of-squares Partition Problem", *Journal of Computer and System Sciences*, Vol. 72, pp. 1077-1093, 2006.

[6] A. Ganesh, Private communication.

[7] A. Ganesh, L. Massoulié and D. Towsley, "The Effect of Network Topology on the Spread of Epidemics", *IEEE INFOCOM,* 2005.

[8] M. Garetto, W. Gong, D. Towsley, "Modeling Malware Spreading Dynamics", *IEEE INFOCOM'03, San Francisco, CA*, Apr. 2003.

[9] G. S. Canright and K. Engo-Monsen, "Spreading on Networks: A Topographic View", *Complexus*, 2006.

[10] C. Asavathiratham, "The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains", *Ph.D. thesis, Massachusetts Institute of Technology,* Oct. 2000.

[11] Albert, R. and H. Jeong and A.-L. Barabsi, "Error and Attack Tolerance of complex networks", *Nature Vol.* 406, pp. 378-382, 27 Jul. 2000.

[12] J. O. Kephart and S. R. White, "Direct-Graph Epidemiological Models of Computer Viruses", *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy,* pp. 343-359, May 1991.

[13] R. Pastor-Satorras and A. Vespignani, "Epidemic Spreading in Scale-Free Networks", *Physical Review Letters*, Vol. 86, No. 14, 3200-3203.

[14] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint", *In 22nd International Symposium on Reliable Distributed Systems (SRDS'03), IEEE Computer,* pp. 25–34. Oct. 2003.

[15] H. Kunreuther and G. Heal, "Interdependent security", *Journal of Risk and Uncertainty (Special Issue on Terrorist Risks)*, 2003.

[16] M. Kearns and L. Ortiz. "Algorithms for Interdependent Security Games", *Advances in Neural Information Processing Systems*, 16. MIT Press, Cambridge, MA, 2004.

[17] R. B. Myerson, "Game Theory: Analysis of Conflict", Harvard University Press, Cambridge, MA, 1991.

[18] E. Altman, "Flow Control Using the Theory of Zero Sum Markov Games," *Decision and Control, 1992., Proceedings of the 31st IEEE Conference on*, pp. 1632–1637 vol.2, 1992.

[19] Y. A. Korilis and A. A. Lazar, "On the Existence of Equilibria in Noncooperative Optimal Flow Control," *Journal of the ACM*, vol. 42, no. 3, pp. 584–613, 1995.

[20] E. Altman, T. Basar, T. Jiménez, and N. Shimkin, "Competitive Routing in Networks with Polynomial Cost," in *INFOCOM*, 2000, pp. 1586–1593.

[21] A. A. Lazar, A. Orda, and D. E. Pendarakis, "Virtual Path Bandwidth Allocation in Multiuser Networks," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 861–871, 1997.

[22] E. Koutsoupias and C. H. Papadimitriou, "Worst-Case Equilibria," *Lecture Notes in Computer Science*, no. 1563, pp. 404–413, 1999.

[23] T. Roughgarden, "Designing Networks for Selfish Users is Hard," in *FOCS '01: Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 2001, p. 472.

[24] T. Roughgarden and E. Tardos, "How Bad is Selfish Routing?" *J. ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[25] C. Papadimitriou, "Algorithms, Games, and the Internet", in *STOC '01: Proceedings of the thirty-third annual ACM symposium on Theory of computing.* New York, NY, USA: ACM, 2001, pp. 749–753.

[26] A. Orda, R. Rom and N. Shimkin, "Competitive Routing in Multiuser Communication Networks", *IEEE/ACM Transactions on Networking*, vol. 1, no. 5, pp. 510–521, 1993.

[27] G. Debreu, "A Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Science*, vol. 38, pp. 886–893, Oct. 1952.

[28] D. J. Daley and J. Gani, "Epidemic Modelling: An Introduction," Cambridge University Press, 1999.

[29] L. Jiang, V. Anantharam and J. C. Walrand, "Efficiency of Selfish Investments in Network Security," *NetEcon 2008*, pp. 31–36, Aug. 2008.

[30] M. Lelarge and J. Bolot, "Network Externalities and the Deployment of Security Features and Protocols in the Internet," *SIGMETRICS 2008*, pp. 37–48, Jun. 2008.

[31] M. Lelarge and J. Bolot, "A Local Mean Field Analysis of Security Investments in Networks," *CoRR abs/0803.3455*, 2008