# Efficiency of Selfish Investments in Network Security

Libin Jiang, Venkat Anantharam and Jean Walrand
University of California at Berkeley
Department of Electrical Engineering & Computer Science
Berkeley, CA 94720
{ljiang,ananth,wlr}@eecs.berkeley.edu *

## ABSTRACT

Internet security does not only depend on the security-related investments of individual users, but also on how these users affect each other. In a non-cooperative environment, each user chooses a level of investment to minimize its own security risk plus the cost of investment. Not surprisingly, this selfish behavior often results in undesirable security degradation of the overall system. In this paper, we first characterize the price of anarchy (POA) of network security under two models: an "Effective-investment" model, and a "Bad-traffic" model. We give insight on how the POA depends on the network topology, individual users' cost functions, and their mutual influence. We also introduce the concept of "weighted POA" to bound the region of all feasible payoffs. In a repeated game, on the other hand, users have more incentive to cooperate for their long term interests. We consider the socially best outcome that can be supported by the repeated game, and give a ratio between this outcome and the social optimum. Although the paper focuses on Internet security, many results are generally applicable to games with positive externalities.

## Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*Security and protection (e.g., firewalls)*; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

## General Terms

Security, Economics, Performance

## Keywords

Internet security, game theory, price of anarchy, positive externality

*This work is supported by NSF under Grant NeTS-FIND 0627161: Market Enabling Network Architecture


*NetEcon'08,* August 22, 2008, Seattle, Washington, USA.


## 1. INTRODUCTION

Security in a communication network depends not only on the security investment made by individual users, but also on the interdependency among them. If a careless user puts in little effort in protecting its computer system, then it is easy for viruses to infect this computer and through it continue to infect others'. On the contrary, if a user invests more to protect itself, then other users will also benefit since the chance of contagious infection is reduced. Define each user's "strategy" as its investment level, then each user's investment has a "positive externality" on other users.

Users in the Internet are heterogeneous. They have different valuations of security and different unit cost of investment. For example, government and commercial websites usually prioritize their security, since security breaches would lead to large financial losses or other consequences. They are also more willing and efficient in implementing security measures. On the other hand, an ordinary computer user may care less about security, and also may be less efficient in improving it due to the lack of awareness and expertise. There are many other users lying between these two categories. If users are selfish, some of them may choose to invest more, whereas others may choose to "free ride", that is, given that the security level is already "good" thanks to the investment of others, such users make no investment to save cost. However, if every user tends to rely on others, the resulting outcome may be far worse for all users. This is the free riding problem in game theory as studied in, for example, [1].

Besides user preferences, the network topology, which describes the (logical) interdependent relationship among different users, is also important. For example, assume that in a local network, user $A$ directly connected to the Internet. All other users are connected to $A$ and exchange a large amount of traffic with $A$. Intuitively, the security level of $A$ is particularly important for the local network since $A$ has the largest influence on other users. If $A$ has a low valuation of its own security, then it will invest little and the whole network suffers. How the network topology affects the efficiency of selfish investment in network security will be one of our focuses.

In this paper, we study how network topology, users' preference and their mutual influence affect network security in a non-cooperative setting. In a one-shot game (i.e., strategic-form game), we derive the "Price of Anarchy" (POA) [2] as a function of the above factors. Here, POA is defined as the worst-case ratio between the "social cost" at a Nash Equilibrium (NE) and Social Optimum (SO). Furthermore, we

introduce the concept of "Weighted-POA" to bound the regions of all possible vectors of payoffs. In a repeated game, users have more incentive to cooperate for their long-term interest. We study the "socially best" equilibrium in the repeated game, and compare it to the Social Optimum.

## 1.1 Related Works

Varian studied the network security problem using game theory in [1]. There, the effort of each user (or player) is assumed to be equally important to all other users, and the network topology is not taken into account. Also, [1] is not focused on the efficiency analysis (i.e., POA).

"Price of Anarchy" (POA) [2], measuring the performance of the worst-case equilibrium compared to the Social Optimum, has been studied in various games in recent years, most of them with "negative externality". These include "selfish routing game" [3], "price competition game" [4] and "resource allocation game" [5], etc. For example in the "selfish routing game", if a user sends its traffic through a link, other users sharing that link will suffer larger delays.

On the contrary, in the network security game, if a user increases his investment, the security level of other users will improve. So it falls into the category of games with positive externalities. Therefore, many results in this paper may be applicable to other similar scenarios. For example, assume that a number of service providers (SP) build networks which are interconnected. If a SP invests to upgrade its own network, the performance of the whole network improves and may bring more revenue to all SP's.

In [6], Aspnes et al. formulated an "inoculation game" and studied its POA. There, each player in the network decides whether to install anti-virus software to avoid infection. Different from our work, [6] has assumed binary decisions and the same cost function for all players.

# 2. PRICE OF ANARCHY (POA) IN THE STRATEGIC-FORM GAME

Assume there are $n$ "players". The security investment (or "effort", we use them interchangeably) of player $i$ is $x_i \geq 0$. This includes both money (e.g., for purchasing anti-virus software) and time/energy (e.g., for system scanning, patching). The cost per unit of investment is $c_i > 0$. Denote $f_i(\mathbf{x})$ as player $i$'s "security risk": the loss due to attacks or virus infections from the network, where $\mathbf{x}$ is the vector of investments by all players. $f_i(\mathbf{x})$ is decreasing in each $x_j$ (thus reflecting positive externality) and non-negative. We assume that it is convex, and that $f_i(\mathbf{x} = \mathbf{0}) > 0$ is finite. Then the "cost function" of player $i$ is

$$g_i(\mathbf{x}) := f_i(\mathbf{x}) + c_i x_i \tag{1}$$

Note that $f_i(\cdot)$ is generally different for different players.

In a Nash game, player $i$ chooses his investment $x_i \geq 0$ to minimize $g_i(\mathbf{x})$. First, we prove in [8] that

PROPOSITION 1. *There exists some pure-strategy Nash Equilibrium (NE) in this game.*

Denote $\bar{\mathbf{x}}$ as the vector of investments at some NE, and $\mathbf{x}^*$ as the vector of investments at Social Optimum (SO). Also denote the unit cost vector $\mathbf{c} = (c_1, c_2, \ldots, c_n)^T$.

We aim to find the POA, $Q$, which upper-bounds $\rho(\bar{\mathbf{x}})$, where

$$\rho(\bar{\mathbf{x}}) := \frac{G(\bar{\mathbf{x}})}{G^*} = \frac{\sum_i g_i(\bar{\mathbf{x}})}{\sum_i g_i(\mathbf{x}^*)}$$

is the ratio between the social cost at the NE $\bar{\mathbf{x}}$ and at the social optimum. For convenience, sometimes we simply write $\rho(\bar{\mathbf{x}})$ as $\rho$ if there is no confusion.

Before getting to the derivation, we illustrate the POA in a simple example. Assume there are 2 players, with their investments denoted as $x_1 \geq 0$ and $x_2 \geq 0$. The cost function is $g_i(\mathbf{x}) = f(y) + x_i, i = 1, 2$, where $f(y)$ is the security risk of both players, and $y = x_1 + x_2$ is the total investment. Assume that $f(y)$ is non-negative, decreasing, convex, and satisfies $f(y) \to 0$ when $y \to \infty$. The social cost is $G(\mathbf{x}) = g_1(\mathbf{x}) + g_2(\mathbf{x}) = 2 \cdot f(y) + y$.

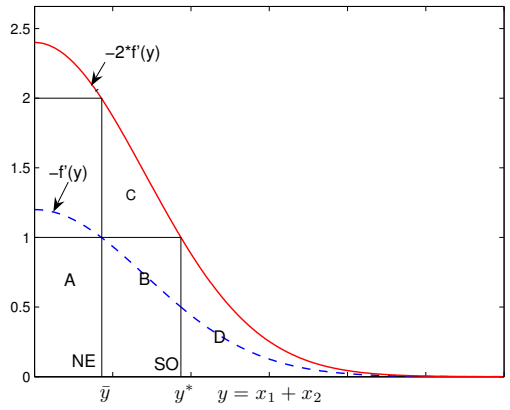


**Figure 1: POA in a simple example**

At a NE $\bar{\mathbf{x}}$, $\frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_i} = f'(\bar{x}_1 + \bar{x}_2) + 1 = 0, i = 1, 2$. Denote $\bar{y} = \bar{x}_1 + \bar{x}_2$, then $-f'(\bar{y}) = 1$. This is shown in Fig 1. Then, the social cost $\bar{G} = 2 \cdot f(\bar{y}) + \bar{y}$. Note that $\int_{\bar{y}}^{\infty}(-f'(z))dz = f(\bar{y}) - f(\infty) = f(\bar{y})$ (since $f(y) \to 0$ as $y \to \infty$), therefore in Fig 1, $2 \cdot f(\bar{y})$ is the area $B + C + D$, and $\bar{G}$ is equal to the area of $A + (B + C + D)$.

At SO (Social Optimum), on the other hand, the total investment $y^*$ satisfies $-2f'(y^*) = 1$. Using a similar argument as before, $G^* = 2f(y^*) + y^*$ is equal to the area of $(A + B) + D$.

Then, the ratio $\bar{G}/G^* = [A+(B+C+D)]/[(A+B)+D] \leq (B + C)/B \leq 2$. We will show later that this upper bound is tight. So the POA is 2.

Now we analyze the POA with the general cost function (1). In some sense, it is a generalization of the above example.

LEMMA 1. *For any NE $\bar{\mathbf{x}}$, $\rho(\bar{\mathbf{x}})$ satisfies*

$$\rho(\bar{\mathbf{x}}) \leq \max\{1, \max_k\{(-\sum_i \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_k})/c_k\}\} \tag{2}$$

*Note that $(-\sum_i \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_k})$ is the marginal "benefit" to the security of all users by increasing $x_k$ at the NE; whereas $c_k$ is the marginal cost of increasing $x_k$. The second term in the RHS (right-hand-side) of (2) is the maximal ratio between these two.*

PROOF. At NE,

$$\begin{cases} \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_i} = -c_i & \text{if } \bar{x}_i > 0 \\ \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_i} \geq -c_i & \text{if } \bar{x}_i = 0 \end{cases} \tag{3}$$

By definition,

$$\rho(\bar{\mathbf{x}}) = \frac{G(\bar{\mathbf{x}})}{G^*} = \frac{\sum_i f_i(\bar{\mathbf{x}}) + \mathbf{c}^T\bar{\mathbf{x}}}{\sum_i f_i(\mathbf{x}^*) + \mathbf{c}^T\mathbf{x}^*}$$

Since $f_i(\cdot)$ is convex for all $i$. Then $f_i(\bar{\mathbf{x}}) \le f_i(\mathbf{x}^*) + (\bar{\mathbf{x}} - \mathbf{x}^*)^T\nabla f_i(\bar{\mathbf{x}})$. So

$$\begin{aligned}\rho &\le \frac{(\bar{\mathbf{x}} - \mathbf{x}^*)^T\sum_i \nabla f_i(\bar{\mathbf{x}}) + \mathbf{c}^T\bar{\mathbf{x}} + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{c}^T\mathbf{x}^*} \\ &= \frac{-\mathbf{x}^{*T}\sum_i \nabla f_i(\bar{\mathbf{x}}) + \bar{\mathbf{x}}^T[\mathbf{c} + \sum_i \nabla f_i(\bar{\mathbf{x}})] + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{c}^T\mathbf{x}^*}\end{aligned}$$

Note that

$$\bar{\mathbf{x}}^T[\mathbf{c} + \sum_i \nabla f_i(\bar{\mathbf{x}})] = \sum_i \bar{x}_i[c_i + \sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i}]$$

There are two possibilities for every player $i$: (a) If $\bar{x}_i = 0$, then $\bar{x}_i[c_i + \sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i}] = 0$. (b) If $\bar{x}_i > 0$, then $\frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_i} = -c_i$. Since $\frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i} \le 0$ for all $k$, then $\sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i} \le -c_i$, so $\bar{x}_i[c_i + \sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i}] \le 0$.

As a result,

$$\rho(\bar{\mathbf{x}}) \le \frac{-\mathbf{x}^{*T}\sum_i \nabla f_i(\bar{\mathbf{x}}) + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{c}^T\mathbf{x}^*} \quad (4)$$

(i) If $x_i^* = 0$ for all $i$, then the RHS is 1, so $\rho(\bar{\mathbf{x}}) \le 1$. Since $\rho$ cannot be smaller than 1, we have $\rho = 1$.

(ii) If not all $x_i^* = 0$, then $\mathbf{c}^T\mathbf{x}^* > 0$. Note that the RHS of (4) is not less than 1, by the definition of $\rho(\bar{\mathbf{x}})$. So, if we subtract $\sum_i f_i(\mathbf{x}^*)$ (non-negative) from both the numerator and the denominator, the resulting ratio upper-bounds the RHS. That is,

$$\rho(\bar{\mathbf{x}}) \le \frac{-\mathbf{x}^{*T}\sum_i \nabla f_i(\bar{\mathbf{x}})}{\mathbf{c}^T\mathbf{x}^*} \le \max_k\{(-\sum_i \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_k})/c_k\}$$

where $\sum_i \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_k}$ is the $k$'th element of the vector $\sum_i \nabla f_i(\bar{\mathbf{x}})$.

Combining case (i) and (ii), the proof is completed. □

In the following, we give two models of the network security game. Each model defines a concrete form of $f_i(\cdot)$. They are formulated to capture the key parameters of the system while being amenable to mathematical analysis.

## 2.1 Effective-investment ("EI") model

Generalizing [1], we consider an "Effective-investment" (EI) model. In this model, the security risk of player $i$ depends on an "effective investment", which we assume is a linear combination of the investments of himself and other players.

Specifically, let $p_i(\sum_{j=1}^n \alpha_{ji}z_j)$ be the probability that player $i$ is infected by a virus (or suffers an attack), given the amount of efforts every player puts in. The effort of player $j$, $z_j$, is weighted by $\alpha_{ji}$, reflecting the "importance" of player $j$ to player $i$. Let $v_i$ be the cost of player $i$ if he suffers an attack; and $c_i$ be the cost per unit of effort by player $i$. Then, the total cost of player $i$ is $g_i(\mathbf{z}) = v_ip_i(\sum_{j=1}^n \alpha_{ji}z_j) + c_iz_i$.

For convenience, we "normalize" the expression in the following way. Let the normalized effort be $x_i := c_iz_i, \forall i$. Then

$$\begin{aligned}g_i(\mathbf{x}) &= v_ip_i(\sum_{j=1}^n \frac{\alpha_{ji}}{c_j}x_j) + x_i \\ &= v_ip_i(\frac{\alpha_{ii}}{c_i}\sum_{j=1}^n \beta_{ji}x_j) + x_i\end{aligned}$$

where $\beta_{ji} := \frac{c_i}{\alpha_{ii}}\frac{\alpha_{ji}}{c_j}$ (so $\beta_{ii} = 1$). We call $\beta_{ji}$ the "*relative importance*" of player $j$ to player $i$.

Define the function $V_i(y) = v_i \cdot p_i(\frac{\alpha_{ii}}{c_i}y)$, where $y$ is a dummy variable. Then $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, where

$$f_i(\mathbf{x}) = V_i(\sum_{j=1}^n \beta_{ji}x_j) \quad (5)$$

Note that $V_i(\cdot)$ is still decreasing, non-negative and convex.

PROPOSITION 2. *In the EI model defined above, $\rho \le \max_k\{1 + \sum_{i:i\neq k}\beta_{ki}\}$. Furthermore, the bound is tight.*

PROOF. Let $\bar{\mathbf{x}}$ be some NE. Denote $\mathbf{h} := \sum_i \nabla f_i(\bar{\mathbf{x}})$. Then the $k$th element of $\mathbf{h}$

$$\begin{aligned}h_k &= \sum_i \frac{\partial V_i(\sum_{j=1}^n \beta_{ji}\bar{x}_j)}{\partial x_k} \\ &= \sum_i \beta_{ki} \cdot V_i'(\sum_{j=1}^n \beta_{ji}\bar{x}_j)\end{aligned}$$

From (3), we have $\frac{\partial V_i(\sum_{j=1}^n \beta_{ji}\bar{x}_j)}{\partial x_i} = \beta_{ii}\cdot V_i'(\sum_{j=1}^n \beta_{ji}\bar{x}_j) = V_i'(\sum_{j=1}^n \beta_{ji}\bar{x}_j) \ge -1$. So $h_k \ge -\sum_i \beta_{ki}$. Plug this into (2), we obtain an upper bound of $\rho$:

$$\rho \le \max\{1, \max_k\{-h_k\}\} \le Q := \max_k\{1 + \sum_{i:i\neq k}\beta_{ki}\} \quad (6)$$

□

(6) gives some interesting insight into the game. Since $\beta_{ki}$ is player $k$'s "relative importance" to player $i$, then $1 + \sum_{i:i\neq k}\beta_{ki} = \sum_i \beta_{ki}$ is player $k$'s relative importance to the society. (6) shows that the POA is bounded by the maximal social "importance" among the players. Interestingly, the bound does not depend on the specific form of $V_i(\cdot)$ as long as it's convex, decreasing and non-negative.

It also provides a simple way to compute POA under the model. We define a "dependency graph" as in Fig. 2, where each vertex stands for a player, and there is a directed edge from $k$ to $i$ if $\beta_{ki} > 0$. In Fig. 2, player 3 has the highest social importance, and $\rho \le 1 + (0.6 + 0.8 + 0.8) = 3.2$. In another special case, if for each pair $(k, i)$, either $\beta_{ki} = 1$ or $\beta_{ki} = 0$, then the POA is bounded by the maximum out-degree of the graph plus 1. If all players are equally important to each other, i.e., $\beta_{ki} = 1, \forall k, i$, then $\rho \le n$ (i.e., POA is the number of players). This also explains why the POA is 2 in the example considered in Fig 1.

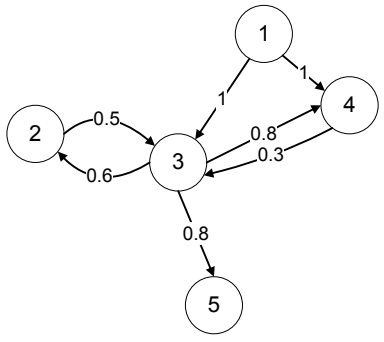


**Figure 2: Dependency Graph and the Price of Anarchy (In this figure, $\rho \le 1 + (0.6 + 0.8 + 0.8) = 3.2$)**

The following is a worst case scenario that shows the bound is tight. Assume there are $n$ players, $n \ge 2$. $\beta_{ki} = 1, \forall k, i$; and for all $i$, $V_i(y_i) = [(1 - \epsilon)(1 - y_i)]_+$, where $[\cdot]_+$

means positive part, $y_i = \sum_{j=1}^n \beta_{ji} x_j = \sum_{j=1}^n x_j$, $\epsilon > 0$ but is very small.

Given $\mathbf{x}_{-i} = \mathbf{0}$, $g_i(\mathbf{x}) = [(1-\epsilon)(1-x_i)]_+ + x_i = (1-\epsilon) + \epsilon \cdot x_i$ when $x_i \le 1$, so the best response for player $i$ is to let $x_i = 0$. Therefore, $\bar{x}_i = 0, \forall i$ is a NE, and the resulting social cost $G(\bar{\mathbf{x}}) = \sum_i [V_i(0) + \bar{x}_i] = (1-\epsilon)n$. Since the social cost is $G(\mathbf{x}) = n \cdot [(1-\epsilon)(1-\sum_i x_i)]_+ + \sum_i x_i$, the social optimum is attained when $\sum_i x_i^* = 1$ (since $n(1-\epsilon) > 1$). Then, $G(\mathbf{x}^*) = 1$. Therefore $\rho = (1-\epsilon)n \to n$ when $\epsilon \to 0$. When $\epsilon = 0$, $\bar{x}_i = 0, \forall i$ is still a NE. In that case $\rho = n$.

## 2.2 Bad-traffic ("BT") Model

Next, we consider a model which is based on the amount of "bad traffic" (e.g., traffic that causes virus infection) from one player to another. Let $r_{ki}$ be the total rate of traffic from $k$ to $i$. How much traffic in $r_{ki}$ will do harm to player $i$ depends on the investments of both $k$ and $i$. So denote $\phi_{k,i}(x_k, x_i)$ as the probability that player $k$'s traffic does harm to player $i$. Clearly $\phi_{k,i}(\cdot,\cdot)$ is a decreasing function. We also assume it is convex. Then, the rate at which player $i$ is infected by the traffic from player $k$ is $r_{ki}\phi_{k,i}(x_k, x_i)$. Let $v_i$ be player $i$'s loss when it's infected by a virus, then $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, where the investment $x_i$ has been normalized such that its coefficient (the unit cost) is 1, and

$$f_i(\mathbf{x}) = v_i \sum_{k \ne i} r_{ki}\phi_{k,i}(x_k, x_i)$$

If the "firewall" of each player is symmetric (i.e., it treats the incoming and outgoing traffic in the same way), then it's reasonable to assume that $\phi_{k,i}(x_k, x_i) = \phi_{i,k}(x_i, x_k)$.

PROPOSITION 3. *In the BT model,* $\rho \le 1 + \max_{(i,j):i\ne j} \frac{v_i r_{ji}}{v_j r_{ij}}$. *The bound is also tight.*

PROOF. Let $\mathbf{h} := \sum_i \nabla f_i(\bar{\mathbf{x}})$ for some NE $\bar{\mathbf{x}}$. Then the $j$-th element

$$\begin{aligned} h_j &= \sum_i \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_j} = \sum_{i\ne j} \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_j} + \frac{\partial f_j(\bar{\mathbf{x}})}{\partial x_j} \\ &= \sum_{i\ne j} v_i r_{ji} \frac{\partial \phi_{j,i}(\bar{x}_j, \bar{x}_i)}{\partial x_j} + v_j \sum_{i\ne j} r_{ij} \frac{\partial \phi_{i,j}(\bar{x}_i, \bar{x}_j)}{\partial x_j} \end{aligned}$$

We have

$$\begin{aligned} q_j &:= \frac{\sum_{i\ne j} \frac{\partial f_i(\bar{\mathbf{x}})}{\partial x_j}}{\frac{\partial f_j(\bar{\mathbf{x}})}{\partial x_j}} = \frac{\sum_{i\ne j} v_i r_{ji} \frac{\partial \phi_{j,i}(\bar{x}_j, \bar{x}_i)}{\partial x_j}}{v_j \sum_{i\ne j} r_{ij} \frac{\partial \phi_{i,j}(\bar{x}_i, \bar{x}_j)}{\partial x_j}} \\ &= \frac{\sum_{i\ne j} v_i r_{ji} \frac{\partial \phi_{j,i}(\bar{x}_j, \bar{x}_i)}{\partial x_j}}{\sum_{i\ne j} v_j r_{ij} \frac{\partial \phi_{j,i}(\bar{x}_j, \bar{x}_i)}{\partial x_j}} \le \max_{i:i\ne j} \frac{v_i r_{ji}}{v_j r_{ij}} \end{aligned}$$

where the 3rd equality holds because $\phi_{i,j}(x_i, x_j) = \phi_{j,i}(x_j, x_i)$ by assumption.

From (3), we know that $\frac{\partial f_j(\bar{\mathbf{x}})}{\partial x_j} \ge -1$. So

$$h_j = (1 + q_j)\frac{\partial f_j(\bar{\mathbf{x}})}{\partial x_j} \ge -(1 + \max_{i:i\ne j} \frac{v_i r_{ji}}{v_j r_{ij}})$$

According to (2), it follows that

$$\rho \le \max\{1, \max_j\{-h_j\}\} \le Q := 1 + \max_{(i,j):i\ne j} \frac{v_i r_{ji}}{v_j r_{ij}} \qquad (7)$$

□

Note that $v_i r_{ji}$ is the damage to player $i$ caused by player $j$ if player $i$ is infected by all the traffic sent by $j$, and $v_j r_{ij}$ is the damage to player $j$ caused by player $i$ if player $j$ is infected by all the traffic sent by $i$. Therefore, (7) means that the POA is upper-bounded by the "maximum imbalance" of the network. (Also, one can find an example where the bound is tight [8].) As a special case, if each pair of the network is "balanced", i.e., $v_i r_{ji} = v_j r_{ij}, \forall i, j$, then $\rho \le 2$!

# 3. BOUNDING THE PAYOFF REGIONS USING "WEIGHTED POA"

So far, the research on POA in various games has largely focused on the worst-case ratio between the social cost (or welfare) achieved at the Nash Equilibria and Social Optimum. Given one of them, the range of the other is bounded. However, this is only one-dimensional information. In any multi-player game, the players' payoffs form a vector which is multi-dimensional. If an observer observes a NE payoff vector, it would be interesting to characterize or bound the region of all feasible vectors of individual payoffs, sometimes even without knowing the exact cost functions. This region gives much more information than solely the social optimum, because it characterizes the tradeoff of efficiency and fairness among different players. Conversely, given any feasible payoff vector, it is also interesting to bound the region of the possible payoff vectors at all Nash Equilibria.

We show that this can be done by generalizing POA to the concept of "Weighted POA", $Q_{\mathbf{w}}$, which is an upper bound of $\rho_{\mathbf{w}}(\bar{\mathbf{x}})$, where

$$\rho_{\mathbf{w}}(\bar{\mathbf{x}}) := \frac{G_{\mathbf{w}}(\bar{\mathbf{x}})}{G^*_{\mathbf{w}}} = \frac{\sum_i w_i \cdot g_i(\bar{\mathbf{x}})}{\sum_i w_i \cdot g_i(\mathbf{x}^*_{\mathbf{w}})}$$

Here, $\mathbf{w} \in \mathcal{R}^n_{++}$ is a weight vector, $\bar{\mathbf{x}}$ is the vector of investments at a NE of the original game; whereas $\mathbf{x}^*_{\mathbf{w}}$ minimizes a weighted social cost $G_{\mathbf{w}}(\mathbf{x}) := \sum_i w_i \cdot g_i(\mathbf{x})$.

To obtain $Q_{\mathbf{w}}$, consider a modified game where the cost function of player $i$ is

$$\hat{g}_i(\mathbf{x}) := \hat{f}_i(\mathbf{x}) + \hat{c}_i x_i = w_i \cdot g_i(\mathbf{x}) = w_i f_i(\mathbf{x}) + w_i \cdot c_i x_i$$

Note that in this game, the NE strategies are the same as the original game: given any $\mathbf{x}_{-i}$, player $i$'s best response remains the same (since his cost function is only multiplied by a constant). So the two games are strategically equivalent, and thus have the same NE's. As a result, the weighted POA $Q_{\mathbf{w}}$ of the original game is exactly the POA in the modified game (Note the definition of $\mathbf{x}^*_{\mathbf{w}}$). Applying (2) to the modified game, we have

$$\begin{aligned} \rho_{\mathbf{w}}(\bar{\mathbf{x}}) &\le \max\{1, \max_k\{(-\sum_i \frac{\partial \hat{f}_i(\bar{\mathbf{x}})}{\partial x_k})/\hat{c}_k\}\} \\ &= \max\{1, \max_k\{(-\sum_i \frac{w_i \partial f_i(\bar{\mathbf{x}})}{\partial x_k})/(w_k c_k)\}\} \quad (8) \end{aligned}$$

Then, one can easily obtain the weighted POA for the two models in the last section.

PROPOSITION 4. *In the EI model,*

$$\rho_{\mathbf{w}} \le Q_{\mathbf{w}} := \max_k\{1 + \frac{\sum_{i:i\ne k} w_i \beta_{ki}}{w_k}\} \qquad (9)$$

*In the BT model,*

$$\rho_{\mathbf{w}} \le Q_{\mathbf{w}} := 1 + \max_{(i,j):i\ne j} \frac{w_i v_i r_{ji}}{w_j v_j r_{ij}} \qquad (10)$$

Since $\rho_{\mathbf{w}}(\bar{\mathbf{x}}) = \frac{G_{\mathbf{w}}(\bar{\mathbf{x}})}{G^*_{\mathbf{w}}} = \frac{\sum_i w_i \cdot g_i(\bar{\mathbf{x}})}{\sum_i w_i \cdot g_i(\mathbf{x}^*_{\mathbf{w}})} \leq Q_{\mathbf{w}}$, we have $\sum_i w_i \cdot g_i(\mathbf{x}^*_{\mathbf{w}}) \geq \sum_i w_i \cdot g_i(\bar{\mathbf{x}})/Q_{\mathbf{w}}$. Notice that $\mathbf{x}^*_{\mathbf{w}}$ minimizes $G_{\mathbf{w}}(\mathbf{x}) = \sum_i w_i \cdot g_i(\mathbf{x})$, so for any feasible $\mathbf{x}$,

$$\sum_i w_i \cdot g_i(\mathbf{x}) \geq \sum_i w_i \cdot g_i(\mathbf{x}^*_{\mathbf{w}}) \geq \sum_i w_i \cdot g_i(\bar{\mathbf{x}})/Q_{\mathbf{w}}$$

Then we have

PROPOSITION 5. *Given any NE payoff vector $\bar{\mathbf{g}}$, then any feasible payoff vector $\mathbf{g}$ must be within the region*

$$\mathcal{B} := \{\mathbf{g} | \mathbf{w}^T\mathbf{g} \geq \mathbf{w}^T\bar{\mathbf{g}}/Q_{\mathbf{w}}, \forall \mathbf{w} \in \mathcal{R}^n_{++}\}$$

*Conversely, given any feasible payoff vector $\mathbf{g}$, any possible NE payoff vector $\bar{\mathbf{g}}$ is in the region*

$$\bar{\mathcal{B}} := \{\bar{\mathbf{g}} | \mathbf{w}^T\bar{\mathbf{g}} \leq \mathbf{w}^T\mathbf{g} \cdot Q_{\mathbf{w}}, \forall \mathbf{w} \in \mathcal{R}^n_{++}\}$$

In other words, the Pareto frontier of $\mathcal{B}$ lower-bounds the Pareto frontier of the feasible region of $\mathbf{g}$. (A similar statement can be said for $\bar{\mathcal{B}}$.) As an illustrating example, consider the EI model, where the cost function of player $i$ is in the form of $g_i(\mathbf{x}) = V_i(\sum_{j=1}^n \beta_{ji}x_j) + x_i$. Assume there are two players in the game, and $\beta_{11} = \beta_{22} = 1$, $\beta_{12} = \beta_{21} = 0.2$. Also assume that $g_i(\mathbf{x}) = (1 - \sum_{j=1}^2 \beta_{ji}x_i)_+ + x_i$, for $i = 1, 2$. It is easy to verify that $\bar{x}_i = 0, i = 1, 2$ is a NE, and $g_1(\bar{\mathbf{x}}) = g_2(\bar{\mathbf{x}}) = 1$. One can further find that the boundary (Pareto frontier) of the feasible payoff region in this example is composed of the two axes and the following line segments (the computation is omitted):

$$\begin{cases} g_2 = -5 \cdot (g_1 - \frac{1}{1.2}) + \frac{1}{1.2} & g_1 \in [0, \frac{5}{6}] \\ g_2 = -0.2 \cdot (g_1 - \frac{1}{1.2}) + \frac{1}{1.2} & g_1 \in [0, 5] \end{cases}$$

which is the dashed line in Fig. 3.

By Proposition 5, for every weight vector $\mathbf{w}$, there is a straight line that lower-bounds the feasible payoff region. After plotting the lower bounds for many different $\mathbf{w}$'s, we obtain a bound for the feasible payoff region (Fig 3). Note that the bound only depends on the coefficients $\beta_{ji}$'s, but not the specific form of $V_1(\cdot)$ and $V_2(\cdot)$. We see that the feasible region is indeed within the bound.

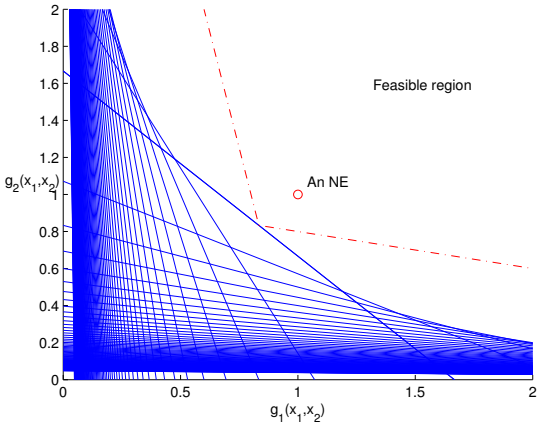


**Figure 3: Bounding the feasible region using weighted POA**

# 4. REPEATED GAME

The Folk Theorem [7] provides a Subgame Perfect Equilibrium (SPE) in a repeated game with discounted costs when the discount factor sufficiently close to 1, to support any cost vector that is Pareto-dominated by the "reservation cost" vector $\underline{\mathbf{g}}$. The $i$th element of $\underline{\mathbf{g}}$, $\underline{g_i}$, is defined as

$$\underline{g_i} := \min_{x_i \geq 0} g_i(\mathbf{x}) \text{ given that } x_j = 0, \forall j \neq i$$

and we denote $\underline{x_i}$ as a minimizer. $\underline{g_i} = g_i(x_i = \underline{x_i}, \mathbf{x}_{-i} = \mathbf{0})$ is the minimal cost achievable by player $i$ when other players are punishing him by making minimal investments 0.

Without loss of generality, we assume that $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, instead of $g_i(\mathbf{x}) = f_i(\mathbf{x}) + c_i x_i$ in (1). This can be done by normalizing the investment and re-defining the function $f_i(\mathbf{x})$.

For simplicity, we make some additional assumptions in this section:

1. $f_i(\mathbf{x})$ (and $g_i(\mathbf{x})$) is *strictly* convex in $x_i$ if $\mathbf{x}_{-i} = \mathbf{0}$. So $\underline{x_i}$ is unique.
2. $\frac{\partial g_i(\mathbf{0})}{\partial x_i} < 0$ for all $i$. So, $\underline{x_i} > 0$.
3. For each player, $f_i(\mathbf{x})$ is strictly decreasing with $x_j$ for some $j \neq i$. That is, positive externality exists.

By assumption 2 and 3, we have $g_i(\underline{\mathbf{x}}) < g_i(x_i = \underline{x_i}, \mathbf{x}_{-i} = \mathbf{0}) = \underline{g_i}, \forall i$. Therefore $\mathbf{g}(\mathbf{x}) < \underline{g}$ is feasible.

## A Performance Bound of the best SPE

According to the Folk Theorem [7], any feasible vector $\mathbf{g} < \underline{\mathbf{g}}$ can be supported by a SPE. So the set of SPE is quite large in general. By negotiating with each other, the players can agree on some SPE. In this section, we are interested in the performance of the "socially best SPE" that can be supported, that is, the SPE with the minimum social cost (denoted as $G_E$). Such a SPE is "optimal" for the society, provided that it is also rational for individual players. We will compare it to the social optimum by considering the "performance ratio" $\gamma = G_E/G^*$, where $G^*$ is the optimal social cost, and

$$\begin{array}{rll} G_E = & \inf_{\mathbf{x} \geq \mathbf{0}} & \sum_i g_i(\mathbf{x}) \\ & \text{s.t.} & g_i(\mathbf{x}) < \underline{g_i}, \forall i \end{array} \tag{11}$$

Since $g_i(\cdot)$ is convex by assumption, due to continuity,

$$\begin{array}{rll} G_E = & \min_{\mathbf{x} \geq \mathbf{0}} & \sum_i g_i(\mathbf{x}) \\ & \text{s.t.} & g_i(\mathbf{x}) \leq \underline{g_i}, \forall i \end{array} \tag{12}$$

where $g_i(\mathbf{x}) \leq \underline{g_i}$ is the rationality constraint for each player $i$. Denote by $\mathbf{x}_E$ a solution of (12). Then $\sum_i g_i(\mathbf{x}_E) = G_E$.

Recall that $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, where the investment $x_i$ has been normalized such that its coefficient (unit cost) is 1. Then, to solve (12), we form a partial Lagrangian

$$\begin{array}{rcl} \mathcal{L}(\mathbf{x}, \lambda') & := & \sum_k g_k(\mathbf{x}) + \sum_k \lambda'_k[g_k(\mathbf{x}) - \underline{g_k}] \\ & = & \sum_k (1 + \lambda'_k) g_k(\mathbf{x}) - \sum_k \lambda'_k \underline{g_k} \end{array}$$

and pose the problem $\max_{\lambda' \geq \mathbf{0}} \min_{\mathbf{x} \geq \mathbf{0}} \mathcal{L}(\mathbf{x}, \lambda')$. Let $\lambda$ be the vector of dual variables when the problem is solved (i.e., when the optimal solution $\mathbf{x}_E$ is reached).

PROPOSITION 6. *The performance ratio $\gamma$ is upper-bounded by $\gamma = G_E/G^* \leq \max_k\{1 + \lambda_k\}$.*

The result can be understood as follows: if $\lambda_k = 0$ for all $k$, then all the incentive-compatibility constraints are not active at the optimal point of (12). So, individual rationality is not a constraining factor for achieving the social optimum. In this case, $\gamma = 1$, meaning that the best SPE achieves the social optimal. But if $\lambda_k > 0$ for some $k$, the individual rationality of player $k$ prevent the system from achieving social optimum. Larger $\lambda_k$ leads to a poorer performance bound on the best SPE relative to SO.

PROOF. Consider the following convex optimization problem parametrized by $\mathbf{t} = (t_1, t_2, \ldots, t_n)$, with optimal value $V(\mathbf{t})$:

$$V(\mathbf{t}) = \begin{array}{ll} \min_{\mathbf{x} \geq \mathbf{0}} & \sum_i g_i(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \leq t_i, \forall i \end{array} \quad (13)$$

When $\mathbf{t} = \underline{\mathbf{g}}$, it is the same as problem (12) that gives the social cost of the best SPE; when $\mathbf{t} = \mathbf{g}^*$, it gives the same solution as the Social Optimum. According to the theory of convex optimization, $V(\mathbf{t})$ is convex in $\mathbf{t}$. Therefore,

$$V(\underline{\mathbf{g}}) - V(\mathbf{g}^*) \leq \nabla V(\underline{\mathbf{g}})(\underline{\mathbf{g}} - \mathbf{g}^*)$$

Also, $\nabla V(\underline{\mathbf{g}}) = -\lambda$, where $\lambda$ is the vector of dual variables when the problem with $\mathbf{t} = \underline{\mathbf{g}}$ is solved. So,

$$\begin{array}{rcl} G_E & = & V(\underline{\mathbf{g}}) \\ & \leq & V(\mathbf{g}^*) + \lambda^T(\mathbf{g}^* - \underline{\mathbf{g}}) \\ & = & G^* + \lambda^T(\mathbf{g}^* - \underline{\mathbf{g}}) \\ & \leq & G^* + \lambda^T \mathbf{g}^* \end{array}$$

Then

$$\gamma = \frac{G_E}{G^*} \leq 1 + \frac{\lambda^T \mathbf{g}^*}{\mathbf{1}^T \mathbf{g}^*} \leq \max_k \{1 + \lambda_k\}$$

□

Proposition 6 gives an upper bound on $\gamma$ assuming the general cost function $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$. Although it is applicable to the two specific models introduced before, it is not explicitly related to the network parameters. In the following, we give an explicit bound for the EI model.

PROPOSITION 7. *In the EI model where $g_i(\mathbf{x}) = V_i(\sum_{j=1}^n \beta_{ji} x_j) + x_i$, $\gamma$ is bounded by*

$$\gamma \leq \min\{\max_{i,j,k} \frac{\beta_{ik}}{\beta_{jk}}, Q\}$$

*where $Q = \max_k\{1 + \sum_{i:i \neq k} \beta_{ki}\}$.*

The part $\gamma \leq Q$ is straightforward: since the set of SPE includes all NE's, the best SPE must be better than the worst NE. The other part is derived from Proposition 6 (its proof is included in [8] due to the limit on space).

Note that the inequality $\gamma \leq \max_{i,j,k} \frac{\beta_{ik}}{\beta_{jk}}$ may not give a tight bound, especially when $\beta_{jk}$ is very small for some $j, k$. But in the following simple example, it is tight and shows that the best SPE achieves the social optimum. Assume $n$ players, and $\beta_{ij} = 1, \forall i, j$. Then, the POA of the one-shot game is $\rho \leq Q = n$ according to (6). In the repeated game, however, the performance ratio $\gamma \leq \max_{i,j,m} \frac{\beta_{im}}{\beta_{jm}} = 1$ (i.e., social optimum is achieved). This illustrates the performance gain resulting from the repeated game.

# 5. CONCLUSIONS

We have studied the equilibrium performance of the network security game. Our model explicitly considered the network topology, players' different cost functions, and their relative importance to each other. We showed that in the strategic-form game, the POA can be very large and tends to increase with the network size, and the dependency and imbalance among the players. This indicates severe efficiency problems in selfish investment. Not surprisingly, the best equilibrium in the repeated game usually gives much better performance, and it's possible to achieve social optimum if that does not conflict with individual interests. Implementing the strategies supporting an SPE in a repeated game, however, needs more communications and cooperation among the players.

Given that the efficiency can be bad with selfish investment, a natural question is how to induce good or optimal performance. With a social planner, a well-known "due care" scheme can achieve social optimum theoretically (see, for example, [1]). In this scheme, each player $i$ is required to invest at least $x_i^*$, the investment in the socially optimal solution. Otherwise, he will be punished according to the amount of "damage" he causes to other players. It can be shown that the best strategy of player $i$ is to invest $x_i^*$. Although this scheme is quite simple conceptually, in practice it is not easy to implement. Firstly, the social planner needs to collect a large amount of information about the players in order to find the optimal level of investment by each user. Then, it needs to enforce this punishment scheme by monitoring the players' actual efforts/investments. Meanwhile, the privacy concern of the players can further hinder the intervention of the social planner. So, in the future, we would like to explore effective and practical schemes to improve the efficiency of investments in network security.

# 6. REFERENCES


[1] H. R. Varian, "System Reliability and Free Riding", *Workshop on Economics and Information Security*, 2002.

[2] E. Koutsoupias, C. H. Papadimitriou, "Worst-case equilibria," *Annual Symposium on Theoretical Aspects of Computer Science*, 1999.

[3] T. Roughgarden, É Tardos, "How bad is selfish routing", *Journal of the ACM*, 2002.

[4] D. Acemoglu and A. Ozdaglar, "Competition and Efficiency in Congested Markets", *Mathematics of Operations Research*, 2007.

[5] R. Johari and J.N. Tsitsiklis, "Efficiency loss in a network resource allocation game", *Mathematics of Operations Research*, 29(3): pp. 407–435, 2004.

[6] J. Aspnes, K. Chang, A. Yampolskiy, "Inoculation Strategies for Victims of Viruses and the Sum-of-Squares Partition Problem", *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pp. 43-52, 2005.

[7] Fudenberg and Tirole, "Game Theory", Massachusetts Institute of Technology, 1991.

[8] L. Jiang, V. Anantharam and J. Walrand, "Efficiency of Selfish Investments in Network Security", Technical Report, UC Berkeley, June 2008. (http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-77.html)