How Bad are Selfish Investments in Network Security?

Libin Jiang
Venkat Anantharam
Jean Walrand

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2008-183
http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-183.html

December 20, 2008

Acknowledgement

This work is supported by the National Science Foundation under Grant NeTS-FIND 0627161: Market Enabling Network Architecture

How Bad are Selfish Investments in Network Security?

Libin Jiang, Venkat Anantharam and Jean Walrand
EECS Department, University of California, Berkeley
{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 his 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, (1) 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. (2) 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. (3) Next, we compare the benefits of improving security technology or improving incentives, and show that improving technology alone may not offset the efficiency loss due to the lack of incentives. (4) Finally, we characterize the performance of correlated equilibrium (CE) in the security game. Although the paper focuses on Internet security, many results are generally applicable to games with positive externalities.**

***Index Terms*—Internet security, game theory, price of anarchy, repeated game, correlated equilibrium, positive externality**

I. 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 his 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 himself, then other users will also benefit since the chance of contagious infection is reduced. Define each user's "strategy" as his 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 his 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.

Next, we compare the benefits of improving security technology or improving incentives, and show that improving technology alone may not offset the efficiency loss due to the lack of incentives. Finally, we consider the performance of correlated equilibrium (CE) (a more general notion than NE) in the security game and characterize the best and worst CE's. Interestingly, some performance bounds of CE coincide with the POA of NE.

A. 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

This work is supported by the National Science Foundation under Grant NeTS-FIND 0627161: Market Enabling Network Architecture

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". Roughgarden et al. shows that the POA is generally unbounded in the "selfish routing game" [3], [4], where each user chooses some link(s) to send his traffic in order to minimize his congestion delay. Ozdaglar et al. derived the POA in a "price competition game" in [5] and [6], where a number of network service providers choose their prices to attract users and maximize their own revenues. In [7], Johari et al. studied the "resource allocation game", where each user bids for the resource to maximize his payoff, and showed that the POA is 3/4 assuming concave utility functions. In all the above games, there is "negative externality" among the players: for example in the "selfish routing game", if a user sends his traffic through a link, other users sharing that link will suffer larger delays.

On the contrary, in the network security game we study here, if a user increases his investment, the security level of other users will improve. In this sense, 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 her own network, the performance of the whole network improves and may bring more revenue to all SP's.

In [8], 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, [8] has assumed binary decisions and the same cost function for all players.

II. 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). So this is not a "one-time" investment. 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 differentiable, 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 the function $f_i(\cdot)$ 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 Appendix A1 that

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

In the paper we consider pure-strategy NE. 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 ρ 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 $G(\mathbf{x}) = g_1(\mathbf{x}) + g_2(\mathbf{x}) = 2 \cdot f(y) + y$.

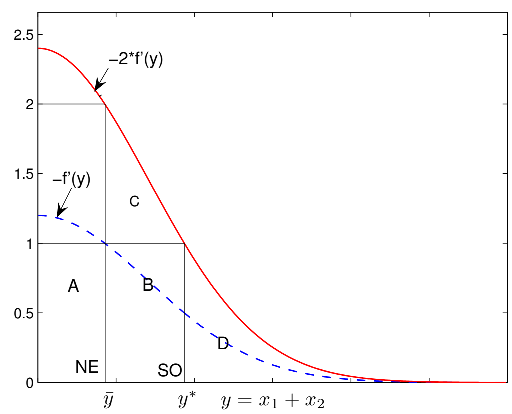

Fig. 1. POA in a simple example

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,

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

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 $f_i(\bar{\mathbf{x}}) \leq f_i(\mathbf{x}^*) + (\bar{\mathbf{x}} - \mathbf{x}^*)^T\nabla f_i(\bar{\mathbf{x}})$. So

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

Note that

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

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 $\frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i} \leq 0$ for all k, then $\sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i} \leq -c_i$, so $\bar{x}_i[c_i + \sum_k \frac{\partial f_k(\bar{\mathbf{x}})}{\partial x_i}] \leq 0$.

As a result,

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

(i) If $x_i^* = 0$ for all i, then the RHS is 1, so $\rho(\bar{\mathbf{x}}) \leq 1$. Since ρ 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,

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

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.

A. *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 α_{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

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

where $\beta_{ji} := \frac{c_i}{\alpha_{ii}}\frac{\alpha_{ji}}{c_j}$ (so $\beta_{ii} = 1$). We call β_{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

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

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

Proposition 2: In the EI model defined above, $\rho \leq \max_k\{1 + \sum_{i:i\neq k}\beta_{ki}\}$. 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 kth element of $\mathbf{h}$

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

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

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

which completes the proof. ■

(6) gives some interesting insight into the game. Since β_{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 $\rho \leq 1 + (0.6 + 0.8 + 0.8) = 3.2$. 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 $\rho \leq n$ (i.e., POA is the number of players). This also explains why the POA is 2 in the example considered in Fig 1.

The following is a worst case scenario that shows the bound is tight. Assume there are n players, $n \geq 2$. $\beta_{ki} = 1, \forall k, i$; 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.[1]

Given $\mathbf{x}_{-i} = \mathbf{0}$, $g_i(\mathbf{x}) = [(1-\epsilon)(1-x_i)]_+ + x_i = (1-\epsilon) + \epsilon \cdot x_i$ when $x_i \leq 1$, 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

[1]Although $V_i(y_i)$ is not differentiable at $y_i = 1$, it can be approximated by a differentiable function arbitrarily closely, such as the result of the example is not affected.

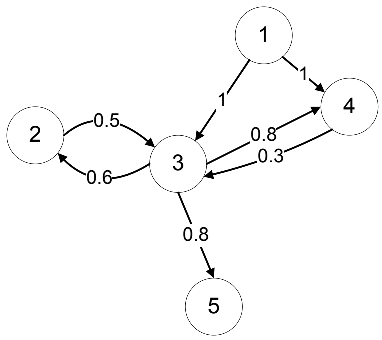

Fig. 2. Dependency Graph and the Price of Anarchy (In this figure, $\rho \leq 1 + (0.6 + 0.8 + 0.8) = 3.2$)

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 $\rho = n$.

B. 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

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

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, $\rho \leq 1 + \max_{(i,j):i\neq j} \frac{v_i r_{ji}}{v_j r_{ij}}$. 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

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

We have

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

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 $\frac{\partial f_j(\bar{\mathbf{x}})}{\partial x_j} \geq -1$. So

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

According to (2), it follows that

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

which completes the proof. ■

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. 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 $\rho \leq 2$!

To show the bound is tight, we can use a similar example as in section II-A. Let there be two players, and assume $v_1 r_{21} = v_1 r_{12} = 1$; $\phi_{1,2}(x_1, x_2) = (1-\epsilon)(1-x_1-x_2)_+$. Then it becomes the same as the previous example when $n = 2$. Therefore $\rho \to 2$ as $\epsilon \to 0$. And $\rho = 2$ when $\epsilon = 0$.

Note that when the network becomes larger, the imbalance between a certain pair of players becomes less important. Thus ρ may be much less than the worst case bound in large networks due to the averaging effect.

III. 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

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

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

Proposition 4: In the EI model,

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

In the BT model,

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

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 β_{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.

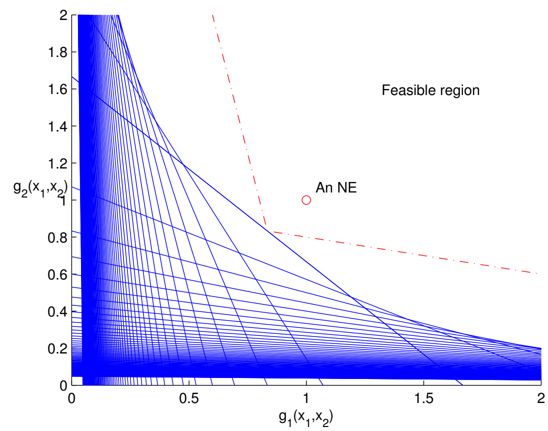

Fig. 3. Bounding the feasible region using weighted POA

IV. Repeated Game

The Folk Theorem [9] 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 ith 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 [9], 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

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

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

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

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

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

and pose the problem $\max_{\lambda'\geq \mathbf{0}} \min_{\mathbf{x}\geq \mathbf{0}} \mathcal{L}(\mathbf{x},\lambda')$.

Let λ be the vector of dual variables when the problem is solved (i.e., when the optimal solution $\mathbf{x}_E$ is reached). Then differentiating $\mathcal{L}(\mathbf{x},\lambda')$ in terms of x_i, we have the optimality condition

$$\begin{cases}\sum_k (1+\lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] = 1+\lambda_i & \text{if } x_{E,i} > 0 \\ \sum_k (1+\lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] \leq 1+\lambda_i & \text{if } x_{E,i} = 0\end{cases} \tag{13}$$

Proposition 6: The performance ratio γ 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 λ_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 $V(\mathbf{t})$:

$$V(\mathbf{t}) = \min_{\mathbf{x}\geq \mathbf{0}} \sum_i g_i(\mathbf{x}) \quad \text{s.t.} \quad g_i(\mathbf{x}) \leq t_i, \forall i \tag{14}$$

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 ([15], page 250), the "value function" $V(\mathbf{t})$ 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 λ is the vector of dual variables when the problem with $\mathbf{t} = \underline{\mathbf{g}}$ is solved. So,

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

Then

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

which completes the proof. ∎

Proposition 6 gives an upper bound on γ 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 $g_i(\mathbf{x}) = V_i(\sum_{j=1}^n \beta_{ji} x_j) + x_i$, γ 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 Appendix A2).

Note that the inequality $\gamma \leq \max_{i,j,k} \frac{\beta_{ik}}{\beta_{jk}}$ may not give a tight bound, especially when β_{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.

V. Improvement of Technology

Recall that the general cost function of player i is

$$g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i. \tag{15}$$

.

Now assume that the security technology has improved. We would like to study how effective is technology improvement compared to the improvement of incentives. Assume that the new cost function of player i is

$$\tilde{g}_i(\mathbf{x}) = f_i(a \cdot \mathbf{x}) + x_i, a > 1. \tag{16}$$

This means that the effectiveness of the investment vector $\mathbf{x}$ has improved by a times (i.e., the risk decreases faster with $\mathbf{x}$ than before). Equivalently, if we define $\mathbf{x}' = a \cdot \mathbf{x}$, then (16) is $\tilde{g}_i(\mathbf{x}) = f_i(\mathbf{x}') + x'_i/a$, which means a decrease of unit cost if we regard $\mathbf{x}'$ as the investment.

Proposition 8: Denote by G^* the optimal social cost with cost functions (15), and by $\tilde{G}^*$ the optimal social cost with cost functions (16). Then, $G^* \geq \tilde{G}^* \geq G^*/a$. That is, the optimal social cost decreases but cannot decrease more than a times.

Proof: First, for all $\mathbf{x}$, $\tilde{g}_i(\mathbf{x}) \leq g_i(\mathbf{x})$. Therefore $\tilde{G}^* \leq G^*$.

Let the optimal investment vector for the improved cost functions be $\tilde{x}^*$. Then, $g_i(a \cdot \tilde{x}^*) = f_i(a \cdot \tilde{x}^*) + a \cdot \tilde{x}^*_i$. Also, $\tilde{g}_i(\tilde{x}^*) = f_i(a \cdot \tilde{x}^*) + \tilde{x}^*_i$. Then, $a \cdot \tilde{g}_i(\tilde{x}^*) = a \cdot f_i(a \cdot \tilde{x}^*) + a \cdot \tilde{x}^*_i \geq g_i(a \cdot \tilde{x}^*)$, because $f_i(\cdot)$ is non-negative and $a > 1$.

Therefore, we have $a \cdot \sum_i \tilde{g}_i(\tilde{x}^*) = a \cdot \tilde{G}^* \geq G(a \cdot \tilde{x}^*) \geq G(\mathbf{x}^*) = G^*$, where $\mathbf{x}^*$ minimizes $G(\mathbf{x}) = \sum_i g_i(\mathbf{x})$. This completes the proof. ■

Here we have seen that the optimal social cost (after technology improved a times) is at least a fraction of $1/a$ of the social optimum before. On the other hand, we have the following about the POA after technology improvement.

Proposition 9: The POA of the network security game with improved technology does not change in the EI model and the BT model.

Proof: The POA in the EI model only depends on the values of β_{ji}'s, which does not change witht the new cost functions. To see this, note that

$$\begin{aligned} \tilde{g}_i(\mathbf{x}) &= f_i(a \cdot \mathbf{x}) + x_i \\ &= V_i(a \cdot \sum_j \beta_{ji} x_j) + x_i. \end{aligned}$$

Define the function $\tilde{V}_i(y) = V_i(a \cdot y), \forall i$, where y is a dummy variable, then $\tilde{g}_i(\mathbf{x}) = \tilde{V}_i(\sum_j \beta_{ji} x_j) + x_i$, where $\tilde{V}_i(\cdot)$ is still convex, decreasing and non-negative. So the β_{ji} values do not change. By Proposition 2, the POA remains the same.

In the BT model, define $\tilde{\phi}_{k,i}(x_k, x_i) := \phi_{k,i}(a \cdot x_k, a \cdot x_i)$, we still have $\tilde{\phi}_{k,i}(x_k, x_i) = \tilde{\phi}_{i,k}(x_i, x_k)$. So by Proposition 3, the POA has the same expression as before. ■

To compare the effect of incentive improvement and technology improvement, consider the following two options to improve the network security.

1) With the current technology, deploy proper incentivizing mechanisms (e.g., reward and punishment) to achieve the social optimum.
2) All players upgrade to the new technology, without solving the incentive problem.

With option 1, the resulting social cost is G^*. With option 2, the social cost is $\tilde{G}(\tilde{x}_{NE})$, where $\tilde{G}(\cdot) = \sum_i \tilde{g}_i(\cdot)$ is the social cost function after technology improvement, with $\tilde{g}_i(\cdot)$ defined in (16), and $\tilde{x}_{NE}$ is a NE in the new game. Define $\rho(\tilde{x}_{NE}) := \tilde{G}(\tilde{x}_{NE})/\tilde{G}^*$, then the ratio between the social costs with option 2 and option 1 is

$$\tilde{G}(\tilde{x}_{NE})/G^* = \rho(\tilde{x}_{NE}) \cdot \tilde{G}^*/G^* \geq \rho(\tilde{x}_{NE})/a$$

where the last step follows from Proposition 8. Also, by Proposition 9, in the EI or BT model, $\rho(\tilde{x}_{NE})$ is equal to the POA shown in Prop. 2 and 3 in the worst case. For example, assume the EI model with $\beta_{ij} = 1, \forall i, j$. Then in the worst case, $\rho(\tilde{x}_{NE}) = n$. When the number of players n is large, $\tilde{G}(\tilde{x}_{NE})/G^*$ may be much larger than 1.

From this discussion, we see that the technology improvement may not offset the negative effect of the lack of incentives, and solving the incentive problem may be more important than merely counting on new technologies.

VI. Correlated Equilibrium (CE)

Correlated equilibrium (CE) [10] is a more general notion of equilibrium which includes the set of NE. In this section we consider the performance bounds of CE.

Conceptually, one may think of a CE as being implemented with the help of a mediator [11]. First the mediator selects a recommended strategy profile $\mathbf{x}$ with probability $\mu(\mathbf{x})$. Then the mediator confidentially tells each player i the component x_i of this strategy profile that is recommended for him. Each player i is free to choose whether to obey the mediator's recommendations. $\mu(\mathbf{x})$ is a CE iff it would be a Nash equilibrium for all players to obey the mediator's recommendations. Note that given a recommended x_i, player i only knows $\mu(\mathbf{x}_{-i}|x_i)$ (i.e., a partial knowledge of other players' recommended strategies). Then in a CE, x_i should be a best response to the randomized stratgies of other players with distribution $\mu(\mathbf{x}_{-i}|x_i)$. CE can also be implemented with a pre-play meeting of the players [9], where they decide the CE they will play. Later they use a device which generates a signal with the distribution μ, and separately gives each player their partial information.

Interestingly, CE can also arise from simple and natural dynamics (without coordination via a mediator or a pre-play meeting). References [12] and [13] showed that in an infinite repeated game, if each player observes the history of other players' actions, and decides his action in each period based on a "regret-minimizing" criterion, then the empirical frequency of the players' actions converge to some CE. Note that in these dynamics, each player does not need to know other players' cost functions, but only their actions in the past. (Specifically in the network security game, observing the actions of his neighbors is sufficient.) This is very natural since in practice, different players tend to adjust their investments based on their observation of others' investments.

For simplicity, in this paper we focus on CE whose support is on a discrete set of strategy profiles. We call such a CE a *discrete CE*. More formally, μ is a discrete CE iff (1) it is a CE; and (2) the probability $\mu(\mathbf{x}) > 0$ only for $\mathbf{x} \in S_\mu$, where S_μ, the support of the distribution μ, is a discrete set of strategy profiles. That is, $S_\mu = \{\mathbf{x}^i \in \mathcal{R}_+^n, i = 1, 2, \ldots, M_\mu\}$, where $\mathbf{x}^i$ denotes a strategy profile, $M_\mu < \infty$ is the cardinality of S_μ and $\sum_{\mathbf{x} \in S_\mu} \mu(\mathbf{x}) = 1$. (But each player can still choose his investment from $\mathcal{R}_+$.) Discrete CE exist in the security game since pure-strategy NE exists (Proposition 1), and a pure-strategy NE is clearly a discrete CE. Appendix A3 gives an example of discrete CE which is not a pure-strategy NE.

We first write down the conditions for a discrete CE with the general cost function $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i, \forall i$. If $\mu(\mathbf{x})$ is a discrete CE, then for any x_i with a positive marginal probability, x_i is a best response to the conditional distribution $\mu(\mathbf{x}_{-i}|x_i)$, i.e., $x_i \in \arg\min_{x_i' \in \mathcal{R}_+} \sum_{\mathbf{x}_{-i}} [f_i(x_i', \mathbf{x}_{-i}) + x_i'] \mu(\mathbf{x}_{-i}|x_i)$. (Recall that player i can choose his investment from $\mathcal{R}_+$.) Since the objective function in the right-hand-side is convex and differentiable in x_i', the first-order condition is

$$\begin{cases} \sum_{\mathbf{x}_{-i}} \frac{\partial f_i(x_i, x_{-i})}{\partial x_i} \mu(x_{-i}|x_i) + 1 = 0 & \text{if } x_i > 0 \\ \sum_{\mathbf{x}_{-i}} \frac{\partial f_i(x_i, x_{-i})}{\partial x_i} \mu(x_{-i}|x_i) + 1 \geq 0 & \text{if } x_i = 0 \end{cases} \tag{17}$$

where $\sum_{\mathbf{x}_{-i}} \frac{\partial f_i(x_i, x_{-i})}{\partial x_i} \mu(x_{-i}|x_i)$ can also be simply written as $E(\frac{\partial f_i(x_i, x_{-i})}{\partial x_i}|x_i)$.

A. How good can a CE get?

The first question we would like to investigate is: does there always exist a CE that achieves the social optimum (SO) in the security game? The answer is generally not. If a CE achieves SO, then the CE should have probability 1 on the set of $\mathbf{x}$ that minimizes the social cost. For convenience, assume there is a unique $\mathbf{x}^*$ that minimizes the social cost. In other words, each time, the mediator chooses $\mathbf{x}^*$ and recommends x_i^* to player i. If $x_i^* > 0$, then it satisfies

$$\sum_k \frac{\partial f_k(\mathbf{x}^*)}{\partial x_i} = -1$$

Since $\sum_k \frac{\partial f_k(\mathbf{x}^*)}{\partial x_i} \leq \frac{\partial f_i(\mathbf{x}^*)}{\partial x_i}$, we have $\frac{\partial g_i(\mathbf{x}^*)}{\partial x_i} = \frac{\partial f_i(\mathbf{x}^*)}{\partial x_i} + 1 \geq 0$. If the inequality is strict, then player i has incentive to invest less than x_i^*. Therefore in general, CE cannot achieve SO in this game.

But, a CE can be better than all NE's in this game. Appendix A3 gives an example. The example is different in nature from that in [10] since each player can choose his investment from $\mathcal{R}_+$.

B. The worst-case discrete CE

As mentioned before, CE can result from simple and natural dynamics in an infinitely repeated game (without a mediator or pre-play meeting). But like NE's, the resulting CE may not be efficient. In this section, we consider the POA of discrete CE, which is defined as the performance ratio of the worst discrete CE compared to the SO. In the Weighted-sum model and the Bad-traffic model, we show that the POA of discrete CE is the identical to that of pure-strategy NE derived before, although the set of discrete CE's is larger than the set of pure-strategy NE's in general.

Lemma 2: The POA of discrete CE, denoted as ρ_{CE}, satisfies

$$\rho_{CE} \leq \max_{\mu(\mathbf{x}) \in \mathcal{C}_D} \{\max\{1, \max_k [E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_k})]\}\}$$

where $\mathcal{C}_D$ is the set of discrete CE's, the distribution $\mu(\mathbf{x})$ defines a discrete CE, and the expectation is taken over the distribution $\mu(\mathbf{x})$.
Although the distribution $\mu(\mathbf{x})$ seems quite complicated, the proof of Lemma 2 (shown in Appendix A4) is similar to that of Lemma 1.

Proposition 10: In the EI model and the BT model, the POA of discrete CE is the same as the POA of NE. That is, in the EI model,

$$\rho_{CE} \leq \max_k \{1 + \sum_{i: i \neq k} \beta_{ki}\},$$

and in the BT model,

$$\rho_{CE} \leq (1 + \max_{(i,j): i \neq j} \frac{v_i r_{ji}}{v_j r_{ij}}).$$

The proof is included in Appendix A5.

VII. 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 games 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.

We have compared the benefits of improving security technology and improving incentives. In particular, we show that the POA of NE's is invariant with the improvement of technology, under the EI model and the BT model. So, improving technology alone may not offset the efficiency loss due to the lack of incentives. Finally, we have studied the performance of correlated equilibrium (CE). We have shown that although CE cannot achieve SO in general, it can be much better than all pure-strategy NE's. In terms of the worst-case bounds, the POA's of discrete CE's are the same as the POA's of NE's under our two models.

Given that the POA is large in certain scenarios, a natural question is how to design schemes or mechanisms to improve the investment incentives for better network security. This has not been a focus of this paper, and we would like to study it more in the future. As discussed above, repeated games and correlated equilibriums can yield better outcomes. Another conceptually simple scheme, with the coordination of a social planner, is "due care" (see, for example, [1]). In this scheme, each player i is required to invest no less than x_i^*, the investment in the socially optimal configuration. Otherwise, he is punished according to the negative effect he causes to other players. Although in theory, "dual care" scheme can achieve the social optimum, in practice, however, it is not easy to implement. Firstly, to find the optimal level of investment by each user, a large amount of information needs to be collected. Secondly, to enforce the scheme, the social planner needs to monitor the players' actual efforts and investments, which may be hindered by the privacy concern of the players. In the future, we would like to further explore effective and practical schemes to improve the efficiency of security investments.

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] T. Roughgarden, "The price of anarchy is independent of the network topology", *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, 2002, pp. 428 - 437.

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

[6] A. Ozdaglar, "Price Competition with Elastic Traffic", *LIDS report*, MIT, 2006.

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

[8] 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.

[9] D. Fudenberg, J. Tirole, "Game Theory", MIT Press, Cambridge, 1991.

[10] R. J. Aumann, "Subjectivity and Correlation in Randomized strategies," *Journal of Mathematical Economics*, 1:67-96, 1974.

[11] R. B. Myerson, "Dual Reduction and Elementary Games," *Games and Economic Behavior*, vol. 21, no. 1-2, pp. 183-202, 1997.

[12] D. Foster, R. Vohra, "Calibrated Learning and Correlated Equilibrium," *Games and Economic Behavior*, 21:40-55, 1997.

[13] G. Stoltz, G. Lugosi, "Learning Correlated Equilibria in Games with Compact Sets of Strategies," *Games and Economic Behavior,* vol. 59, no. 1, pp. 187-208, April 2007.

[14] J. B. Rosen, "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games," *Econometrica*, 33, 520-534, July 1965.

[15] S. Boyd and L. Vandenberghe, "Convex Optimization", Cambridge University Press, 2004.

APPENDIX

A1. Proof of Proposition 1

Consider player i's set of best responses, $BR_i(\mathbf{x}_{-i})$, to $\mathbf{x}_{-i} \geq \mathbf{0}$. Define $x_{i,max} := [f_i(\mathbf{0}) + \epsilon]/c_i$ where $\epsilon > 0$, then due to convexity of $f_i(\mathbf{x})$ in x_i, we have

$$\begin{aligned} & f_i(x_i = 0, \mathbf{x}_{-i}) - f_i(x_i = x_{i,max}, \mathbf{x}_{-i}) \\ \geq \quad & x_{i,max} \cdot \left(-\frac{\partial f_i(x_{i,max}, \mathbf{x}_{-i})}{\partial x_i}\right) \\ = \quad & \frac{f_i(\mathbf{0}) + \epsilon}{c_i}\left(-\frac{\partial f_i(x_{i,max}, \mathbf{x}_{-i})}{\partial x_i}\right) \end{aligned}$$

. Since $f_i(x_i = 0, \mathbf{x}_{-i}) \leq f_i(\mathbf{0})$, and $f_i(x_i = x_{i,max}, \mathbf{x}_{-i}) \geq 0$, it follows that

$$f_i(\mathbf{0}) \geq \frac{f_i(\mathbf{0}) + \epsilon}{c_i}\left(-\frac{\partial f_i(x_{i,max}, \mathbf{x}_{-i})}{\partial x_i}\right)$$

which means that $\frac{\partial f_i(x_{i,max}, \mathbf{x}_{-i})}{\partial x_i} + c_i > 0$. So, $BR_i(\mathbf{x}_{-i}) \subseteq [0, x_{i,max}]$.

Let $x_{max} = \max_i x_{i,max}$. Consider a modified game where the strategy set of each player is restricted to $[0, x_{max}]$. Since the set is compact and convex, and the cost function is convex, therefore this is a convex game and has some pure-strategy NE [14], denoted as $\bar{\mathbf{x}}$.

Given $\bar{\mathbf{x}}_{-i}$, $\bar{x}_i$ is also a best response in the strategy set $[0, \infty)$, because the best response cannot be larger than x_{max} as shown above. Therefore, $\bar{\mathbf{x}}$ is also a pure-strategy NE in the original game.

A2. Proof of Proposition 7

It is useful to first give a sketch of the proof before going to the details. Roughly, the KKT condition (for the best SPE), as in equation (13), is $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] = 1 + \lambda_i, \forall i$ (except for some "corner cases" which will be taken care of by Lemma 4). Without considering the corner cases, we have the following by inequality (18):

$$\begin{aligned} \gamma \quad \leq \quad & \max_{i,j} \frac{1 + \lambda_i}{1 + \lambda_j} = \max_{i,j} \frac{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}]}{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}]} \\ \leq \quad & \max_{i,j,k} \left\{\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i} / \frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}\right\} \end{aligned}$$

which is Proposition 11. Then by plugging in $f_k(\cdot)$ of the EI model, Proposition 7 immediately follows.

Now we begin the detailed proof.

As assumed in section 4, $\mathbf{g}(\mathbf{x}) < \underline{g}$ is feasible.

Lemma 3: If $\mathbf{g}(\mathbf{x}) < \underline{\mathbf{g}}$ is feasible, then at the optimal solution of problem (12), at least one dual variable is 0. That is, $\exists i_0$ such that $\lambda_{i_0} = 0$.

Proof: Suppose $\lambda_i > 0, \forall i$. Then all constraints in (12) are active. As a result, $G_E = \sum_k \underline{g_k}$.

Since $\exists \mathbf{x}$ such that $\mathbf{g}(\mathbf{x}) < \underline{\mathbf{g}}$, then for this $\mathbf{x}$, $\sum_k g_k(\mathbf{x}) < \sum_k \underline{g_k}$. $\mathbf{x}$ is a feasible point for (12), so $G_E \leq \sum_k g_k(\mathbf{x}) < \sum_k \underline{g_k}$, which contradicts $G_E = \sum_k \underline{g_k}$. ∎

From Proposition 6, we need to bound $\max_k\{1 + \lambda_k\}$. Since $1 + \lambda_i \geq 1, \forall i$, and $1 + \lambda_{i_0} = 1$ (by Lemma 3), it is easy to see that

$$\gamma \leq \max_k\{1 + \lambda_k\} = \max_{i,j} \frac{1 + \lambda_i}{1 + \lambda_j} \qquad (18)$$

Before moving to Proposition 11, we need another observation:

Lemma 4: If for some i, $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] < 1 + \lambda_i$, then $\lambda_i = 0$.

Proof: From (13), it follows that $x_{E,i} = 0$. Since $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] < 1 + \lambda_i$, and every term on the left is non-negative, we have

$$(1 + \lambda_i)[-\frac{\partial f_i(\mathbf{x}_E)}{\partial x_i}] < 1 + \lambda_i$$

That is, $\frac{\partial f_i(\mathbf{x}_E)}{\partial x_i} + 1 = \frac{\partial g_i(\mathbf{x}_E)}{\partial x_i} > 0$. Since $f_i(\mathbf{x})$ is convex in x_i, and $x_{E,i} = 0$, then

$$g_i(\underline{x_i}, \mathbf{x}_{E,-i}) \geq g_i(x_{E,i}, \mathbf{x}_{E,-i}) + \frac{\partial g_i(\mathbf{x}_E)}{\partial x_i}(\underline{x_i} - 0) > g_i(\mathbf{x}_E)$$

where we have used the fact that $\underline{x_i} > 0$.

Note that $g_i(\underline{x_i}, \mathbf{x}_{E,-i}) \leq g_i(\underline{x_i}, \mathbf{0}_{-i}) = \underline{g_i}$. Therefore,

$$g_i(\mathbf{x}_E) < \underline{g_i}$$

So $\lambda_i = 0$. ∎

Proposition 11: With the general cost function $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, γ is upper-bounded by

$$\gamma \leq \min\left\{\max_{i,j,k}\left\{\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i} / \frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}\right\}, Q\right\}$$

where Q is the POA derived before for Nash Equilibria in the one-shot game (i.e., $\rho \leq Q$), and $\mathbf{x}_E$ achieves the optimal social cost in the set of SPE.

Proof: First of all, since any NE is Pareto-dominated by $\underline{\mathbf{g}}$, the best SPE is at least as good as NE. So $\gamma \leq Q$.

Consider $\pi_{i,j} := \frac{1+\lambda_i}{1+\lambda_j}$. (a) If $\lambda_i = 0$, then $\pi_{i,j} \leq 1$. (b) If $\lambda_i, \lambda_j > 0$, then according to Lemma 4, we have $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] = 1 + \lambda_i$ and $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}] = 1 + \lambda_j$. Therefore

$$\pi_{i,j} = \frac{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}]}{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}]} \leq \max_k\left\{\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i} / \frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}\right\}$$

(c) If $\lambda_i > 0$ but $\lambda_j = 0$, then from Lemma 4, $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}] = 1 + \lambda_i$ and $\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}] \leq 1 + \lambda_j$. Therefore,

$$\pi_{i,j} \leq \frac{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}]}{\sum_k (1 + \lambda_k)[-\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}]} \leq \max_k\left\{\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i} / \frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}\right\}$$

Considering the cases (a), (b) and (c), and from equation (18), we have

$$\gamma \leq \max_{i,j} \pi_{i,j} \leq \max_{i,j,k}\{\frac{\partial f_k(\mathbf{x}_E)}{\partial x_i}/\frac{\partial f_k(\mathbf{x}_E)}{\partial x_j}\}$$

which completes the proof. ■

Proposition 11 applies to any game with the cost function $g_i(\mathbf{x}) = f_i(\mathbf{x}) + x_i$, where $f_i(\mathbf{x})$ is non-negative, decreasing in each x_i, and satisfies the assumption (1)-(3) at the beginning of section 4. This includes the EI model and the BT model introduced before. It is not easy to find an explicit form of the upper bound on γ in Proposition 11 for the BT model. However, for the EI model, we have the simple expression shown in Proposition 7:

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

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

Proof: 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. Also, since $\frac{\partial f_k(\mathbf{x}_E)}{x_i} = \beta_{ik} V_k'(\sum_m \beta_{mk} x_{E,m})$, and $\frac{\partial f_k(\mathbf{x}_E)}{x_j} = \beta_{jk} V_k'(\sum_m \beta_{mk} x_{E,m})$, using Proposition 11, we have $\gamma \leq \max_{i,j,k} \frac{\beta_{ik}}{\beta_{jk}}$. ■

A3. An example where a CE is more efficient than all NE's

Assume that the cost functions are strictly convex, then all NEs must be pure-strategy NE's (because the best response of each player i to any randomization of $\mathbf{x}_{-i}$ is unique). In the following example, the CE has high probability on an efficient strategy profile, while all NE's are less efficient.

Consider the EI model with only 2 players, with cost functions $g_1(\mathbf{x}) = f(x_1 + \alpha \cdot x_2) + x_1$, and $g_2(\mathbf{x}) = f(x_2 + \alpha \cdot x_1) + x_2$, where $\alpha > 1, \mathbf{x} \geq 0$. (Note that the cost functions of the two players are symmetric.) We compute the pure NE's first. Assume that there exists $y_{NE} > 0$ such that $f'(y_{NE}) + 1 = 0$. Then the best response of player 1 to x_2 is $BR_1(x_2) = (y_{NE} - \alpha \cdot x_2)_+$, and the best response of player 2 to x_1 is $BR_2(x_1) = (y_{NE} - \alpha \cdot x_1)_+$. Then there are 3 pure-strategy NE's, shown in Fig. 4.

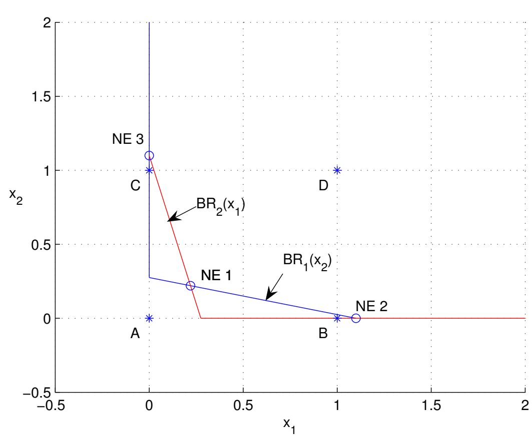

Fig. 4. Pure-strategy NE's

Denote by A, B, C, D the strategy profiles (0,0), (1,0), (0,1), (1,1) respectively (Fig. 4). We would like to construct a CE where only these profiles have positive probability and $\mu(A) : \mu(B) : \mu(C) : \mu(D) = 1 : \beta_1 : \beta_1 : \beta_1\beta_2$, where $\beta_1, \beta_2 > 1$.

Consider player 1 (the argument for player 2 is similar), we have $\frac{\partial g_1(\mathbf{x})}{\partial x_1} = f'(x_1 + \alpha x_2) + 1$. Let

$$\begin{aligned}
\frac{\partial g_1(A)}{\partial x_1} &= f'(0) + 1 = -r_1 \\
\frac{\partial g_1(B)}{\partial x_1} &= f'(1) + 1 = -r_2 \\
\frac{\partial g_1(C)}{\partial x_1} &= f'(\alpha) + 1 = r_3 \\
\frac{\partial g_1(D)}{\partial x_1} &= f'(1+\alpha) + 1 = r_4
\end{aligned} \tag{19}$$

where $r_1, r_2, r_3, r_4 > 0$, $r_1 > r_2$, $r_3 < r_4$ (consistent to the convexity of $f(\cdot)$) and satisfy

$$r_1 = \beta_1 r_3 \text{ and } r_2 = \beta_2 r_4. \tag{20}$$

Then, we have

$$\begin{aligned}
&\mu(A|x_1 = 0)\frac{\partial g_1(A)}{\partial x_1} + \mu(C|x_1 = 0)\frac{\partial g_1(C)}{\partial x_1} \\
\propto\ & \mu(A)\frac{\partial g_1(A)}{\partial x_1} + \mu(C)\frac{\partial g_1(C)}{\partial x_1} \\
\propto\ & -r_1 + \beta_1 r_3 = 0
\end{aligned}$$

and

$$\begin{aligned}
&\mu(B|x_1 = 1)\frac{\partial g_1(B)}{\partial x_1} + \mu(D|x_1 = 1)\frac{\partial g_1(D)}{\partial x_1} \\
\propto\ & \mu(B)\frac{\partial g_1(B)}{\partial x_1} + \mu(D)\frac{\partial g_1(D)}{\partial x_1} \\
\propto\ & -\beta_1 r_2 + \beta_1\beta_2 r_4 = 0.
\end{aligned}$$

Therefore, by condition (17), it is the best response of player 1 to obey the recommended actions (0 or 1) from the distribution μ. Due to symmetry of the cost functions and the distribution μ, player 2 also obeys the recommended actions. Therefore μ is a CE.

Let the function $f(y)$ satisfy the conditions in (19) and (20). For example, Fig. 5 shows $1 + f'(y)$ of such a function. For convenience, $1 + f'(y)$ is assumed to be piecewise linear, and also satisfy $1+f'(1+\epsilon) = 0$ (where $\epsilon > 0$), $1+f'(2+\alpha) = 1$, and $f(\infty) = 0$. Then $f(1 + \alpha), f(\alpha), f(1)$ and $f(0)$ can be computed according to Fig. 5.

Hence, in the CE μ, the expected cost of player 1 is

$$\begin{aligned}
&E_\mu(g_1(\mathbf{x})) \\
=\ & \frac{1}{1 + 2\beta_1 + \beta_1\beta_2}\{f(0) + \beta_1[f(1) + 1] + \\
& \beta_1 f(\alpha) + \beta_1\beta_2[f(1 + \alpha) + 1]\}
\end{aligned}$$

and by symmetry, $E_\mu(g_2(\mathbf{x})) = E_\mu(g_1(\mathbf{x}))$. Thus the expected social cost is $E_\mu(g_1(\mathbf{x}) + g_2(\mathbf{x})) = 2E_\mu(g_1(\mathbf{x}))$.

Also, since $1 + f'(1 + \epsilon) = 0$, we have $y_{NE} = 1 + \epsilon$. From here the social costs of all three pure-strategy NE's in Fig. 4 can be obtained.

To give a numerical example, let $\alpha = 5$, $\beta_1 = 8, \beta_2 = 4$, $r_1 = 2.4, r_2 = 2, r_3 = 0.3, r_4 = 0.5$, $\epsilon = 1$. Then, it

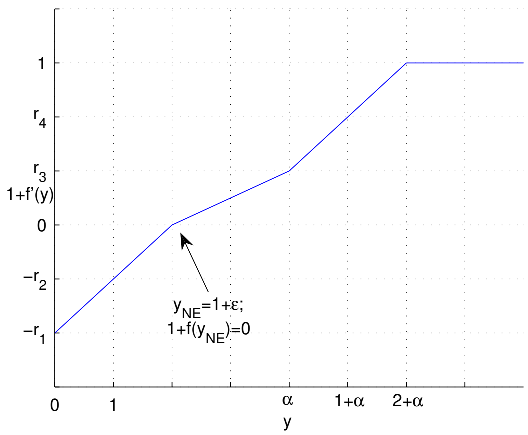

Fig. 5. The shape of $1 + f'(y)$

can be computed that the expected social cost at the CE is $E_\mu(g_1(\mathbf{x})+g_2(\mathbf{x})) = 4.351$. And the social costs at the NE 1, NE 2 and NE 3 in Fig. 4 are 7.467, 5.4 and 5.4 respectively. Therefore the expected social cost at the CE is lower than all pure-strategy NE's.

A4. Proof of Lemma 2

Proof: The performance ratio between the discrete CE $\mu(\mathbf{x})$ and the social optimal is

$$\rho(\mu) = \frac{G(\mu)}{G^*} = \frac{E[\sum_i (f_i(\mathbf{x}) + x_i)]}{\sum_i [f_i(\mathbf{x}^*) + x_i^*]}$$

where the expectation is taken over distribution $\mu(\mathbf{x})$.

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

$$\begin{aligned} &\rho(\mu) \\ \leq\ & \frac{E[(\mathbf{x}-\mathbf{x}^*)^T \sum_i \nabla f_i(\mathbf{x}) + \mathbf{1}^T\mathbf{x}] + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{1}^T\mathbf{x}^*} \\ =\ & \frac{E\{-\mathbf{x}^{*T}\sum_i \nabla f_i(\mathbf{x}) + \mathbf{x}^T[\mathbf{1} + \sum_i \nabla f_i(\mathbf{x})]\} + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{1}^T\mathbf{x}^*} \end{aligned}$$

Note that

$$\mathbf{x}^T[\mathbf{1} + \textstyle\sum_i \nabla f_i(\mathbf{x})] = \sum_i x_i[1 + \sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i}].$$

For every player i, for each x_i with positive probability, there are two possibilities: (a) If $x_i = 0$, then $x_i[1 + \sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i}] = 0, \forall \mathbf{x}$; (b) If $x_i > 0$, then by (17), $E(\frac{\partial f_i(\mathbf{x})}{\partial x_i}|x_i) = -1$. Since $\frac{\partial f_k(\mathbf{x})}{\partial x_i} \leq 0$ for all k, then $E(\sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i}|x_i) \leq -1$. Therefore for both (a) and (b), we have $E[x_i(1+\sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i})|x_i] = x_i \cdot E[1+\sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i}|x_i] \leq 0$. So,

$$\begin{aligned} & E\{\sum_i [x_i(1 + \sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i})]\} \\ =\ & \sum_i E\{E[x_i(1 + \sum_k \frac{\partial f_k(\mathbf{x})}{\partial x_i})|x_i]\} \leq 0. \end{aligned}$$

As a result,

$$\rho(\mu) \leq \frac{-E[\mathbf{x}^{*T}\sum_i \nabla f_i(\mathbf{x})] + \sum_i f_i(\mathbf{x}^*)}{\sum_i f_i(\mathbf{x}^*) + \mathbf{1}^T\mathbf{x}^*}. \quad (21)$$

Consider two cases:

(i) If $x_i^* = 0$ for all i, then the RHS is 1, so $\rho(\mu) \leq 1$. Since $\rho(\mu)$ cannot be smaller than 1, we have $\rho(\mu) = 1$.

(ii) If not all $x_i^* = 0$, then $\mathbf{1}^T\mathbf{x}^* > 0$. Note that the RHS of (21) is not less than 1, by the definition of $\rho(\mu)$. 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,

$$\rho(\mu) \leq \frac{-E[\mathbf{x}^{*T}\sum_i \nabla f_i(\mathbf{x})]}{\mathbf{1}^T\mathbf{x}^*} \leq \max_k\{E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_k})\}$$

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 cases (i) and (ii), we have

$$\rho(\mu) \leq \max\{1, \max_k E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_k})\}.$$

Then, ρ_{CE} is upper-bounded by $\max_{\mu \in \mathcal{C}_D} \rho(\mu)$. ■

A5. Proof of Proposition 10

Proof: Since $\mu(\mathbf{x})$ is a discrete CE, by (17), for any x_i with positive probability, $E(-\frac{\partial f_i(\mathbf{x})}{\partial x_i}|x_i) \leq 1$. Therefore $E(-\frac{\partial f_i(\mathbf{x})}{\partial x_i}) \leq 1$.

In the EI model, we have

$$-\frac{\partial f_i(\mathbf{x})}{\partial x_k} = \beta_{ki}[-\frac{\partial f_i(\mathbf{x})}{\partial x_i}].$$

Therefore

$$E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_k}) = E(-\sum_i \beta_{ki}\frac{\partial f_i(\mathbf{x})}{\partial x_i}) \leq \sum_i \beta_{ki}.$$

So, $\rho_{CE} \leq \max_k\{1 + \sum_{i:i\neq k} \beta_{ki}\}$, the same as the POA in NE.

In the BT model, similar to the proof in Proposition 3, it's not difficult to see that the following holds for any $\mathbf{x}$:

$$[-\sum_{i:i\neq j} \frac{\partial f_i(\mathbf{x})}{\partial x_j}]/[-\frac{\partial f_j(\mathbf{x})}{\partial x_j}] \leq \max_{i:i\neq j} \frac{v_i r_{ji}}{v_j r_{ij}}.$$

Then,

$$-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_j} \leq (1 + \max_{i:i\neq j} \frac{v_i r_{ji}}{v_j r_{ij}})[-\frac{\partial f_j(\mathbf{x})}{\partial x_j}].$$

If $\mu(\mathbf{x})$ is a discrete CE, then $E(-\frac{\partial f_j(\mathbf{x})}{\partial x_j}) \leq 1, \forall j$. Therefore $E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_j}) \leq (1 + \max_{i:i\neq j} \frac{v_i r_{ji}}{v_j r_{ij}})$. So,

$$\rho_{CE} \leq \max_j E(-\sum_i \frac{\partial f_i(\mathbf{x})}{\partial x_j}) \leq (1 + \max_{(i,j):i\neq j} \frac{v_i r_{ji}}{v_j r_{ij}}),$$

which is also the same as the POA in NE in this model. ■