Budget Balance or Voluntary Participation? Incentivizing Investments in Interdependent Security Games

Parinaz Naghizadeh and Mingyan Liu
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor
Email: {naghizad, mingyan}@umich.edu

***Abstract*— In a system of interdependent users, all entities are affected by the security decisions of one another. These users benefit from the improved health of the system when their neighbors invest in security measures; an effect known as *positive externality*. The externality of these decisions make security a *public good*, the optimal provision of which in a system of self-interested users requires regulation/incentives through an external mechanism. In this paper we first show that due to the non-excludable nature of security, no mechanism can achieve social optimality and ensure voluntary participation, while maintaining a balanced budget, for all instances of a security game. We then compare two incentive mechanisms in this context for improving security investment among users, namely the Pivotal mechanism and the Externality mechanism. We show that even though both mechanisms incentivize socially optimal investments, they differ in terms of budget requirements and participation. The Pivotal mechanism guarantees users' participation; however, although (weakly) budget balanced in many game environments, it runs a budget deficit in security games. The Externality mechanism on the other hand is a budget balanced mechanism, but fails to incentivize voluntary participation. We further study the effects of the information available to the mechanism designer on the budget deficit of the Pivotal mechanism.**

I. Introduction

Improving the state of security in a collection of interdependent users, such as the Internet, requires not only the study of security threats and the design of strategies to thwart the efforts of malicious entities, but also the collective effort of users in adopting such mitigating strategies. The importance of implementing protective measures becomes more prominent when we note that users' security in interconnected systems is dependent on *all* users' efforts toward securing their systems. That is, by investing in security measures, users not only protect their own systems, but improve the overall health of the environment, as they are less likely to become a host for adversaries, thus preventing further spread of attacks. Consequently, the investments of users in security can be viewed as a *public good* with *positive externalities*. In particular, the interaction among strategic users when deciding on the provision of security as a public good in an interconnected system is referred to as an *interdependent security (IDS)* game.

It is well known [1] that the provision of public goods in a system of autonomous, self-interested users is in general inefficient. The optimal level of public good in such environments is the one maximizing social welfare, and is referred to as the *socially optimal* solution. In an unregulated system, users' levels of effort are far from this optimal, as entities do not consider the externality of their actions on others, and further choose to free-ride on the externality of others' efforts. Therefore, improving the levels of effort, and ideally, achieving the socially optimal solution, requires the design and implementation of incentive mechanisms.

The design of mechanisms for optimal provision of security, i.e. improving security investments in interdependent security games, has received considerable attention in the literature, see e.g. [2], [3], [4], [5]. In general, these methods propose to either *incentivize* or *dictate* improved security behavior by users [2]. Our focus in the current paper is on the design of incentive mechanisms that use monetary payments/rewards to incentivize improved security investments. Examples of existing mechanisms in the literature include introducing subsidies and fines based on security investments [6], [3], assessing rebates and penalties based on security outcomes [3], imposing a level of due care and establishing liability rules [6], [7], etc. In this paper, we will examine two incentive mechanisms, namely the *Pivotal* [8] and *Externality* [9] mechanisms, both of which induce socially optimal user behavior by levying a monetary tax on each user participating in the proposed mechanism.

In addition to inducing socially optimal investments, an incentive mechanism is often required to satisfy other desirable properties. In particular, when dealing with monetary taxation, the mechanism designer attempts to achieve either weak or strong *budget balance (BB)*; i.e., collect enough taxes so that there is either a surplus or full redistribution, but not a budget deficit. Furthermore, it is of interest to design the taxation in a way that ensures users' *voluntary participation (VP)*; i.e., users prefer cooperating in the mechanism to opting out. Users' participation in a given mechanism itself is dependent on (1) the design of the mechanism, and (2) the options available to users when staying out. The latter is what sets the study of incentive mechanisms for security games apart from other public good problems where similar Pivotal and Externality mechanisms have been applied, e.g., [10], [11].

To illustrate the underlying difference, we note that security is a *non-excludable* public good. As a result, when staying out of the mechanism, a user can still benefit from the externalities of improved security actions by other participating users. The availability of these spill-overs in turn limits users' willingness to pay for the good or their interest in improving their actions. In contrast, with excludable public goods, e.g. transmission power allocated in a communication system [11], users' willingness to participate is determined by the change in their utilities when contributing and receiving the good, as compared to receiving no allocation at all. This means that the designer has the ability to collect more taxes and require a higher level of contribution when providing an excludable good. Consequently, tax transfer mechanisms such as the Externality mechanism (e.g. [11]) and the Pivotal mechanism (e.g. [10]) can incentivize the socially optimal solution, guarantee voluntary participation, and have (weak) budget balance.

However, in this paper we show that given the non-excludable nature of security, there is no tax transfer mechanism that can achieve social optimality, voluntary participation, and (weak) budget balance simultaneously in all instances of the interdependent security game. We then illustrate this impossibility result by studying the Pivotal and Externality mechanisms. We show that the Pivotal mechanism on one hand guarantees users' participation, but runs a budget deficit. The Externality mechanism on the other hand is budget balanced, but fails to incentivize voluntary participation. We further study the effects of the information available to the mechanism designer on the budget deficit of the Pivotal mechanism.

The rest of this paper is organized as follows. In Section II, we introduce a model for interdependent security games, as well as the corresponding impossibility result. We analyze the Pivotal and Externality mechanisms in Sections III and IV, respectively. Section V concludes the paper.

II. Problem and Motivation

A. The IDS game model

Consider a collection of N interconnected users. Each user decides on a level $x_i \geq 0$ of investment on security, with increasing effectiveness. User i will incur a cost of $h_i(x_i)$ to implement this action. The *cost function* $h_i(\cdot)$ can be different among users, and is assumed to be continuous, differentiable, increasing, and convex. The convexity assumption models the fact that security actions become increasingly costly as their effectiveness increases.

Let $\mathbf{x} := \{x_1, x_2, \dots, x_N\}$ denote the vector of investments of the N users. Given this vector, the security risk of a user i is determined by the *risk function* $f_i(\mathbf{x})$. This function quantifies the amount user i has subject to loss given the investments $\mathbf{x}$. The dependence of user i's risk on *all* users' actions reflects the externality of users' investments on one another. We take $f_i(\cdot)$ to be continuous, differentiable, decreasing, and convex, in all arguments x_j, $\forall i, j$. The fact that $f_i(\cdot)$ is decreasing in $x_j, \forall j \neq i$, reflects the *positive* externality of user j's action on user i's security. The convexity assumption models the fact that users' risks decrease considerably as a result of initial investment in security; however, the rate of risk reduction slows down as the investment levels increase, as no security investment will be effective against all attacks.

The security related costs for user i, referred to as the *total cost function*, is therefore given by:

$$g_i(\mathbf{x}) = f_i(\mathbf{x}) + h_i(x_i) \ , \tag{1}$$

with the utility of user i given by $u_i(\mathbf{x}) = -g_i(\mathbf{x})$.

We assume that all users are rational, and choose their actions strategically so as to minimize their total cost function in (1). The one-shot, simultaneous move game among these strategic users will be referred to as the (unregulated) *interdependent security (IDS)* game. These games, their equilibria, and the inefficiency of the security investments in the state of anarchy when compared to the socially optimal levels of investment have been extensively studied in the literature [2], [5]. The socially optimal investment levels in these IDS game are those maximizing the sum of users' utilities, or equivalently, minimizing the *social costs* $\sum_i g_i(\mathbf{x})$, and are given by the solution to the following optimization problem:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\succeq 0} \sum_{i=1}^{N} g_i(\mathbf{x}) \ . \tag{2}$$

The inefficiency of the Nash equilibria of an unregulated IDS game, as compared to the solution $\mathbf{x}^*$ of (2), stems from the fact that self-interested users do not consider the externality of their investments when choosing a level of effort, or choose to free-ride on the externality of others' actions.

Our focus in the current paper is on the design of incentive mechanisms that use monetary taxation to incentivize improved security investments. We will examine two such incentive mechanisms, namely the Pivotal and the Externality mechanisms, both of which attempt to shape users' behavior by levying a tax t_i on each user i participating in the proposed mechanism. The tax t_i may be positive, negative, or zero, reflecting payments, rewards, or no transaction, respectively. The users' utilities are assumed to be quasi-linear, i.e., linearly decreasing in the taxation term t_i. Therefore, when participating in the mechanism, a user i's utility is given by:

$$u_i(\mathbf{x}) = -g_i(\mathbf{x}) - t_i = -f_i(\mathbf{x}) - h_i(x_i) - t_i \ . \tag{3}$$

When dealing with monetary taxation, a desirable property of a mechanism is to achieve either weak or strong *budget balance (BB)*, i.e., a surplus of resources ($\sum_i t_i > 0$) or a complete redistribution of taxes ($\sum_i t_i = 0$), respectively. In particular, the designer finds a budget deficit ($\sum_i t_i < 0$) undesirable, as this requires injection of external resources into the system. Furthermore, it is of interest to design the taxation in a way that the users' *voluntary participation* (VP) constraints are satisfied, i.e., users prefer staying in the mechanism to opting out.

Before proceeding to the analysis of the Pivotal and Externality mechanisms, we discuss an impossibility result concerning incentive mechanisms using taxation in IDS games, and show that no mechanism can simultaneously satisfy the aforementioned desirable properties in all instances of the proposed security games.

B. An impossibility result

Ideally, the goal of an incentive mechanism designed for an interdependent security game is to induce socially optimal behavior, while guaranteeing users' voluntary participation and maintaining a (weakly) balanced budget. Nevertheless, in the IDS games described in this section, there is no tax transfer mechanism that can achieve the three aforementioned goals simultaneously in all instances of the game. The following two example illustrate this inconsistency in simple instances of the IDS game. Consider a user i, referred to as the *loner*, who unilaterally opts out of the mechanism.

Our first example assumes a full information IDS game.[1] In a game of full information, when a user i opts out, the remaining participating users choose their welfare maximizing level of investment, while taking into account that the loner will best-respond to their collective action. As a result, the equilibrium investment profile is the Nash equilibrium of the game between the $N-1$ participating users and the loner.

Example 1: Consider a collection of N users, with risk functions given by the total effort model $f_i(\mathbf{x}) = \exp(-\sum_{i=1}^N x_i), \forall i$, and linear cost functions $h_i(x_i) = c_i x_i$. Let $c_1 < c_2 < c_3 < \ldots < c_N$. Let $c_1 < \frac{c_2}{N-1} < 1$.

The socially optimal level of investment in this game is the solution to (2), and is such that only the user with the lowest cost invests in security. This optimal solution is:

$$\exp(-x_1^*) = \frac{c_1}{N}, \quad x_2^* = \ldots = x_N^* = 0 \ .$$

If user 1 chooses to stay out, but his investment cost is sufficiently low, while user 2's cost is relatively high, then user 1 will continue being the only one investing. The equilibrium levels of investment $\tilde{\mathbf{x}}^1$ will be given by:

$$\exp(-\tilde{x}_1^1) = c_1, \quad \tilde{x}_j^1 = 0, \forall j = 2, \ldots, N \ .$$

If any user $j \neq 1$ decides to stay out, the equilibrium levels of security $\tilde{\mathbf{x}}^j$ will be:

$$\exp(-\tilde{x}_1^j) = \frac{c_1}{N-1}, \quad \tilde{x}_j = 0, \forall j = 2, \ldots, N \ .$$

We can now use the above to determine the voluntary participation conditions of all users. For user 1 to voluntarily participate in the mechanism, we need $g_1(\mathbf{x}^*) + t_1 \leq g_1(\tilde{\mathbf{x}}^1)$, which leads to:

$$\exp(-x_1^*) + c_1 x_1^* + t_1 \leq \exp(-\tilde{x}_1^1) + c_1 \tilde{x}_1^1 \ .$$

For any other user $j = 2, \ldots, N$, the voluntary participation condition is:

$$\exp(-x_1^*) + t_j \leq \exp(-\tilde{x}_1^j) \ .$$

To satisfy the designer's weak budget balance condition, we need $\sum_j t_j \geq 0$. However, the above VP conditions yield:

$$\sum_j t_j \leq c_1(1 - \log N) \ . \tag{4}$$

From (4), we conclude that with any number of users $N \geq 3$ and a set of costs satisfying the initial conditions $c_1 < \frac{c_2}{N-1} < 1$, the total sum of collected taxes will be negative, implying a budget deficit. We have thus found an IDS game for which regardless of how the taxes are designed, no mechanism can implement the socially optimal solution and incentivize participation, unless it injects external resources into the system. ■

Our second example presents the impossibility of achieving social optimality, voluntary participation, and budget balance, in games where users' total cost functions are their private information. Consequently, neither the mechanism designer, nor the remaining players, can reasonably assume the non-participating user's response to that of the participating users. We assume that in this case, the participating users opt for a conservative decision of assuming a zero investment by the loner user; i.e., they fully disregard the presence of the loner in the system.

Example 2: Consider the IDS game stated in Example 1, with the one difference that users' total cost functions $g_i(\cdot)$ are their private information. Assume that the unit costs of investment are such that $\frac{c_2}{N-1} < c_1 < c_2 < N-1$.

Let $\tilde{\mathbf{x}}^i$ denote the profile of investments when user i stays out of the mechanism. Here, $\tilde{\mathbf{x}}_{-i}^i = \arg\min_{\mathbf{x}_{-i} \succeq 0} \sum_{j \neq i} g_j(\mathbf{x}_{-i}, 0)$, and $\tilde{x}_i^i = \arg\min_{x \geq 0} g_i(\tilde{\mathbf{x}}_{-i}^i, x)$.[2] For the current instance of the IDS game, these profiles are given by:

$$\exp(-\tilde{x}_2^1) = \frac{c_2}{N-1}, \quad \tilde{x}_k^1 = 0, \text{ for } k \neq 2 \ ,$$

when user 1 stays out, and,

$$\exp(-\tilde{x}_1^j) = \frac{c_1}{N-1}, \quad \tilde{x}_k^j = 0, \text{ for } k \neq 1 \ ,$$

when any user $j = 2, 3, \ldots, N$ opts out.

For user 1 to participate in the mechanism, we need $g_1(\mathbf{x}^*) + t_1 \leq g_1(\tilde{\mathbf{x}}^1)$, which reduces to:

$$t_1 < \frac{c_2}{N-1} - \frac{c_1}{N} - c_1 \log \frac{N}{c_1} \ .$$

Similarly, for all other users $j \neq 1$ to participate we need:

$$t_j < \frac{c_1}{N(N-1)} \ .$$

Therefore, the total amount of taxes collected by the designer is upper-bounded by:

$$\begin{aligned} \sum_i t_i = t_1 + (N-1)t_2 &\leq \frac{c_2}{N-1} - c_1 \log \frac{N}{c_1} \\ &\leq c_1(1 - \log \frac{N}{c_1}) \ . \end{aligned} \tag{5}$$

From (5), we conclude that with any set of costs satisfying $c_1 < N \exp(-1)$ and the initial conditions $\frac{c_2}{N-1} < c_1 < c_2 < N-1$, the total sum of collected taxes will be negative, implying a budget deficit. ■

[1] We refer the interested reader to [12] for further discussion of this scenario.

[2] We assume the profile $\tilde{\mathbf{x}}_{-i}^i$ is correctly and truthfully announced by the mechanism. It is in the best interest of the loner i to use this information and best-respond accordingly.

Intuitively, this impossibility result is a consequence of the non-excludability of security as a public good, and the positive externality of users' investments on one another. In particular, there are two main types of users' in such incentive mechanisms: (1) *main investors*, who are required to increase their level of investment, often receiving a monetary reward in return (e.g. user 1 in the previous examples), and (2) *free-riders*, who are required to pay a monetary taxation and are promised an increased level of security in the system (e.g. users $j = 2, \ldots, N$ in the previous examples). Consequently, a free-rider who is considering to opt out of the mechanism does not assess his willingness to pay against the previous state of anarchy in the absence of the mechanism, but does so as compared to the scenario in which he unilaterally stays out, while others participate in the mechanism and improve the health of the system. Although this level of security is likely to be lower than when he participated, it is generally higher than that in the state of anarchy. Therefore, the free-rider is enjoying a lower risk while avoiding payment. Similarly, a main investor's compensation may not be high enough, leading this user to step out, and requiring other users to step up and improve the security instead. Again, the externality of these investments may be sufficiently high, so that a main investor opts out of the mechanism.

Overall, the amount of tax that is collectable from the free-riders, as well as the amount to be paid out to the main investors, are greatly limited by the spill-overs of security investments, in both full information and incomplete information scenarios. In addition, as the designer maintains a budget balance, there are no external resources available to cover a possible gap among the taxes collected and those to be paid out. Thus, we observe that there is no mechanism that can simultaneously achieve social optimality, voluntary participation, and budget balance, in all instances of the interdependent security game.

Given this observation, we focus on two possible incentive mechanisms with the main goal of achieving socially optimal investments. We will show that the Pivotal mechanism will incentivize participation, but may run a budget deficit; on the other hand, the Externality mechanism is a budget balanced mechanism, but may not guarantee voluntary participation.

III. The Pivotal mechanism

A. Overview

Vickery-Clarke-Groves (VCG) mechanisms, often referred to as Groves mechanisms [1], [10], refer to a family of mechanisms in which revealing the true preference is a user's dominant strategy regardless of others' actions, and achieve allocative efficiency (i.e. implement the socially optimal solution) in games where users have quasi-linear utilities, by imposing taxes that internalize the externality of users' actions. However, the (weak) budget balance and voluntary participation conditions do not necessarily hold in these mechanisms, and are further dependent on the specifics of the design, as well as the game environment.

In general, let $u_i(k, \theta_i, t_i) = v_i(k, \theta_i) - t_i$ be user i's utility. Here, θ_i is user i's type; a user's type determines the preference of a user over the possible outcomes. In IDS games, a user i's type is his risk and cost functions $\{f_i(\cdot), h_i(\cdot)\}$, or equivalently, his total cost function $g_i(\cdot)$. Users are required to report their types to the mechanism designer, based on which the designer decides on an allocation k. In the current IDS game, an allocation is the vector of investments $\mathbf{x}$ prescribed by the mechanism.

The VCG family of mechanisms achieve truth revelation and efficiency by assigning the following taxes to users, when their reported types are $\hat{\theta}$:

$$t_i(\hat{\theta}) = \alpha_i(\hat{\theta}_{-i}) - \sum_{j \neq i} v_j(k^*(\hat{\theta}), \hat{\theta}_j) \ .$$

Here, $k^*(\hat{\theta}) = \arg\max_k \sum_i v_i(k, \hat{\theta}_i)$ is the socially optimal allocation given users' reported types, and $\alpha_i(\cdot)$ is an arbitrary function that depends on the reported types of agents other than i. Any choice of this function results in truth revelation and a socially efficient outcome, and a careful design may further result in VP and/or (W)BB.

One such choice that can achieve VP in certain environments is the *Pivotal*, or *Clarke*, mechanism [8], [10], with taxes given by:

$$t_i(\hat{\theta}) = \sum_{j \neq i} v_j(k^*_{-i}(\hat{\theta}_{-i}), \hat{\theta}_j) - \sum_{j \neq i} v_j(k^*(\hat{\theta}), \hat{\theta}_j) \ .$$

Here, $k^*_{-i}(\hat{\theta}_{-i}) = \arg\max_k \sum_{j \neq i} v_j(k, \hat{\theta}_j)$, is the outcome maximizing the social welfare in the absence of user i. This mechanism satisfies the participation constraints and achieves weak budget balance in many private and public good games [10]; however, we show that this is not necessarily the case in interdependent security games.

B. The Pivotal mechanism in IDS games

The taxes in the Pivotal mechanism for the IDS game can be set as follows:

$$t_i = \sum_{j \neq i} g_j(\mathbf{x}^*_{-i}, x^*_i) - \sum_{j \neq i} g_j(\hat{\mathbf{x}}^i_{-i}, \hat{x}^i_i) \ , \qquad (6)$$

where, $g_i(\mathbf{x}) = f_i(\mathbf{x}) + h_i(x_i)$ is user i's total cost function, $\mathbf{x}^* = \arg\min_{\mathbf{x}} \sum_i g_i(\mathbf{x})$ is the socially optimal solution, and $\hat{\mathbf{x}}^i_{-i}$ is the cost minimizing actions of users excluding user i, and is determined by $\hat{\mathbf{x}}^i_{-i} = \arg\min_{\mathbf{x}_{-i}} \sum_{j \neq i} g_j(\mathbf{x}_{-i}, \hat{x}^i_i)$.

In the latter, the choice of user i's action $\hat{x}^i_i$ is dependent on the information structure of the IDS game. In games of complete information, this action can be set to the known Nash equilibrium of the game between user i and the $N-1$ participating users. However, one of the main advantages of using VCG mechanisms is that they can lead users to reveal their true preferences in games where users have private information. In this case, in order to apply Pivotal mechanism to an IDS game, the action $\hat{x}^i_i$ has to be assumed fixed; otherwise, the function $\alpha_i(\cdot)$ would depend on i's type, which is contrary to the design principles of Pivotal mechanisms.

One possible intuitive interpretation of user participation/exclusion in incomplete information IDS games is that when user i participates in the mechanism, he not only reports his total cost function, but further allows for monitoring of his action, so that x_i becomes known. In contrast, when user i opts out of the system, and therefore this screening, other users have to assume i's action to be given. As shown later, we will set $\hat{x}_i^i = 0, \forall i$, that is, users $j \neq i$ disregard user i not only by ignoring the externality of their actions on user i (hence the optimization over $\sum_{j\neq i}$), but also by ignoring any possible externality of user i's investment on their utilities (hence setting $\hat{x}_i^i = 0$).[3]

C. Properties of the Pivotal Mechanism

We start by considering the Pivotal mechanism in a game of incomplete information, and present a set of propositions to illustrate the main properties of these mechanisms. These propositions establish two of the desirable properties of the Pivotal mechanism, namely efficiency and voluntary participation. The first proof follows directly from the classical literature on VCG mechanisms and is included for completeness; however, the second proof concerning user participation has to be somewhat modified to account for the effects of the externality of users' security actions on one another. We then show that the Pivotal mechanism will fail to achieve (weak) budget balance, and it therefore runs a budget deficit, requiring the designer to inject resources into the system to incentivize participation.

Proposition 1: In the Pivotal mechanism with taxes given by (6), reporting the true type, i.e., the true total cost function $g_i(\cdot)$, is a dominant strategy for all users i. Therefore, the socially optimal solution is implemented.

Proof: The utility of user i when reporting $\tilde{g}_i(\cdot)$, while others report $\tilde{g}_j(\cdot),\ j \neq i$, is given by:

$$u_i(\tilde{\mathbf{x}}) = -g_i(\tilde{\mathbf{x}}) - \sum_{j\neq i} \tilde{g}_j(\tilde{\mathbf{x}}) + \sum_{j\neq i} \tilde{g}_j(\hat{\mathbf{x}}^i) \ ,$$

where $\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}\succeq 0} \sum_{k=1}^N \tilde{g}_k(\mathbf{x})$ is the allocation that is optimal given the reported types $\tilde{g}_k(\cdot), \forall k$. We first note that the last term is independent of user i's report. Then, as the allocation $\tilde{\mathbf{x}}$ is chosen according to the minimization problem $\arg\min_{\mathbf{x}\succeq 0} \sum_{k=1}^N \tilde{g}_k(\mathbf{x})$ over the reported types, the sum of the first and second terms is maximized at $\tilde{g}_i(\cdot) = g_i(\cdot)$. Therefore, users will reveal their true cost functions, irrespective of other users' reports. Consequently, the socially optimal investment profile will be prescribed by the mechanism designer. ■

Proposition 2: The Pivotal mechanism with taxes given by (6) satisfies voluntary participation.

Proof: Let $\hat{x}_i^i = 0, \forall i$ in (6). We assume that the investment profile $\hat{\mathbf{x}}_{-i}^i$, which is implemented by other users in the absence of user i, is known to the non-participating user, e.g., is publicly and truthfully announced by the mechanism. It is in the best interest of a user i to consider this information and best respond accordingly when choosing an investment level. Let $\tilde{x}_i$ denote the strategy of user i when best responding to $\hat{\mathbf{x}}_{-i}^i$.

The change in the utility of a user i when staying in vs. staying out of the mechanism is given by:

$$\begin{aligned} u_i(\mathbf{x}^*) - u_i(\hat{\mathbf{x}}) &= -g_i(\mathbf{x}^*) - \sum_{j\neq i} g_j(\mathbf{x}^*) \\ &\quad + \sum_{j\neq i} g_j(\hat{\mathbf{x}}_{-i}^i, 0) + g_i(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) \\ &\geq -g_i(\mathbf{x}^*) - \sum_{j\neq i} g_j(\mathbf{x}^*) \\ &\quad + \sum_{j\neq i} g_j(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) + g_i(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) \\ &= -\sum_{j} g_j(\mathbf{x}^*) + \sum_{j} g_j(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) \geq 0 \ . \end{aligned} \tag{7}$$

In the above, the first inequality follows from the fact that $\tilde{x}_i \geq 0$, and all user $j \neq i$'s costs are decreasing in user i's investment (positive externality), and thus $g_j(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) \leq g_j(\hat{\mathbf{x}}_{-i}^i, 0)$. The second inequality is due to the fact that $\mathbf{x}^*$ is the socially optimal solution given by the minimizer of the sum of all users' costs. We conclude that it is in the best interest of users to participate in the Pivotal mechanism with the given taxes. ■

It is interesting to note the implications of choosing the assumed action $\hat{x}_i^i$ in (6). As seen in the first inequality in (7), the inequality $g_j(\hat{\mathbf{x}}_{-i}^i, \tilde{x}_i) \leq g_j(\hat{\mathbf{x}}_{-i}^i, \hat{x}_i^i)$ will fail to hold for any choice of $\hat{x}_i^i > \tilde{x}_i$, and subsequently, voluntary participation will no longer be satisfied. If user i's utility function is not necessarily known to other users or the mechanism designer, the only viable choice for $\hat{x}_i^i$ to ensure voluntary participation is to set $\hat{x}_i^i = 0, \forall i$ in (6).

Despite achieving efficiency and incentivizing participation, the Pivotal mechanism with taxes given by (6) is not necessarily budget balanced. This imbalance is predicted by the counter-example presented in Example 2. To further elaborate on this observation, the following example provides instances of the problem with and without budget deficit in a simple game.

Example 3: Consider $N = 2$ users. Let the cost functions be linear, such that $h_i(x_i) = c_i x_i$. We assume the risk functions of these users are given by:

$$\begin{aligned} f_1(x_1, x_2) &= \alpha \exp(-(x_1 + kx_2)) \ , \\ f_2(x_1, x_2) &= \exp(-(x_1 + kx_2)) \ . \end{aligned}$$

Here, α and k are given positive constants. Choose the unit costs of investment such that $c_2 \leq \min\{2, k(\alpha+1)\}$, $c_1 \leq \alpha$. The socially optimal investments in this game will be:

$$x_1^* = 0, \quad \exp(-kx_2^*) = \frac{c_2}{k(\alpha+1)} \ .$$

When user 1 stays out of the mechanism, user 2 assumes $\hat{x}_1^1 = 0$, and chooses his investment accordingly, leading to $\exp(-k\hat{x}_2^1) = \frac{c_2}{k}$. Similarly, when user 2 stays out, user 1 will assume $\hat{x}_2^2 = 0$, invest $\exp(-\hat{x}_1^2) = \frac{c_1}{\alpha}$.

[3] As shown later this section, even though the amount of taxes collected will increase with a choice of $\hat{x}_i^i > 0$, the voluntary participation constraint may no longer be satisfied.

Given these profiles, users' taxes can be assessed using (6). The sum of the taxes will be given by:

$$\begin{aligned}\sum_{i=1}^{2} t_i &= (\alpha+1)\frac{c_2}{k(\alpha+1)} + c_2 x_2^* \\ &\quad - c_1 - c_1\hat{x}_1^2 - \frac{c_2}{k} - c_2\hat{x}_2^1 \\ &= \frac{c_2}{k}\log(1+\alpha) - c_1(1+\log(\frac{\alpha}{c_1}))\ . \qquad (8)\end{aligned}$$

A choice of $(k, \alpha, c_1, c_2) = (0.25, 10, 9, 2)$ makes the sum in (8) positive, indicating that the Pivotal mechanism has achieved weak budget balance. However, a choice of $(k, \alpha, c_1, c_2) = (1, 10, 9, 2)$ leads to a negative sum in (8), indicating a budget deficit in that instance of the game. ■

D. The full information scenario

As mentioned before, a main advantage of the Pivotal mechanism is that, as shown in Proposition 1, it induces truthful revelation of the total cost functions $g_i(\cdot)$ in dominant strategies. Therefore, there is not need for full information to achieve the socially optimal solution. However, a disadvantage of this lack of full information by the mechanism designer is that she will not be extracting the maximum possible amount of tax from users in the mechanism. This in turn results in a potentially higher budget deficit. In this section, we study the effect of having full information by the designer in decreasing the deficit gap.

To illustrate, consider the second term in (6), when the assumed actions of the non-participating users are $\hat{x}_i^i = 0, \forall i$. Let $\text{BR}_i(\mathbf{x}_{-i}^i)$ denote user i's best response to a given collective action $\mathbf{x}_{-i}^i$ of other users in user i's absence. Similarly, let $\text{BR}_{-i}(x_i^i)$ be the cost minimizing response of the other users given the action x_i^i by user i. We have:

$$\begin{aligned}\sum_{j\neq i} g_j(\text{BR}_{-i}(0), 0) &\geq \sum_{j\neq i} g_j(\text{BR}_{-i}(0), \text{BR}_i(\text{BR}_{-i}(0)) \\ &\geq \sum_{j\neq i} g_j(\text{BR}_{-i}(\text{BR}_i(\text{BR}_{-i}(0))), \text{BR}_i(\text{BR}_{-i}(0)) \\ &\geq \cdots \\ &\geq \sum_{j\neq i} g_j(\tilde{\mathbf{x}}_{-i}^i, \tilde{x}_i^i). \qquad (9)\end{aligned}$$

Here, the first, third, ... inequalities are due to the positive externality of i's actions on users' $j \neq i$ costs, and the second, forth, ... inequalities are due to the fact that BR_{-i} is a cost minimizing, best response function. The vector $\tilde{\mathbf{x}}^i$ is the profile of investments at the Nash equilibrium of the game between the non-participating user i, and the $N-1$ participating users.

We conclude that, by having full information, the mechanism designer can calculate $\tilde{\mathbf{x}}^i$ beforehand, and design the taxes in (6) accordingly. By (9), this choice of taxes minimizes the second sum in (6), in turn maximizing the total amount the designer is able to collect from the system. This choice further maintains voluntary participation; the proof follows similar to Proposition 2. By Example 1 however, this mechanism will still have a budget deficit for some instances of the IDS game.

IV. The Balanced Externality Mechanism

A. The game form

In this section, we examine a taxation mechanism that can achieve the socially optimal solution to the IDS game with full information, while maintaing a balanced budget. This mechanism is adapted from the work of Hurwicz in [9]. The components of the mechanism are as follows.

The message space: Each user i provides a message $m_i := (\mathbf{x}_i, \boldsymbol{\pi}_i)$ to the mechanism designer. $\mathbf{x}_i \in \mathbb{R}^N$ denotes user i's proposal on the public good, i.e., it proposes the amount of security investment to be made by everyone in the system, referred to as an *investment profile*.

$\boldsymbol{\pi}_i \in \mathbb{R}_+^N$ denotes a *pricing profile* which suggests the amount to be paid by everyone. As illustrated below, this is used by the designer to determine the taxes of all users. Therefore, the pricing profile is user i's proposal on the private good.

The outcome function: The outcome function takes the message profiles $\mathbf{m} := \{m_1, m_2, \ldots, m_N\}$ as input, and determines the security investment profile $\hat{\mathbf{x}}$ and a *tax* profile $\hat{\mathbf{t}}$ as follows:

$$\hat{\mathbf{x}}(\mathbf{m}) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\ , \qquad (10)$$

$$\begin{aligned}\hat{t}_i(\mathbf{m}) &= (\boldsymbol{\pi}_{i+1} - \boldsymbol{\pi}_{i+2})^T\hat{\mathbf{x}}(\mathbf{m}) \\ &\quad + (\mathbf{x}_i - \mathbf{x}_{i+1})^T \text{diag}(\boldsymbol{\pi}_i)(\mathbf{x}_i - \mathbf{x}_{i+1}) \\ &\quad - (\mathbf{x}_{i+1} - \mathbf{x}_{i+2})^T \text{diag}(\boldsymbol{\pi}_{i+1})(\mathbf{x}_{i+1} - \mathbf{x}_{i+2}), \forall i. \qquad (11)\end{aligned}$$

In (11), for simplicity $N+1$ and $N+2$ are treated as 1 and 2, respectively.

Note that as $\sum_i \hat{t}_i = 0$ by (11), the budget balance condition is satisfied through this construction. What this means is that the designer will not be spending resources or making profit, as the users whose tax $\hat{t}_i$ is positive will be financing the rewards for those who have negative taxes. In other words, the mechanism proposes a tax *redistribution* scheme to incentivize improved security investments.

We refer the interested reader to our earlier work [12], for the theorems and proofs which establish the mechanism's optimality. In particular, we can show that a profile $(\hat{\mathbf{x}}(\mathbf{m}^*), \hat{\boldsymbol{t}}(\mathbf{m}^*))$, derived at any possible NE $\mathbf{m}^*$ of the Externality regulated IDS game, is the socially optimal solution. Furthermore, we can establish the converse of this statement, i.e., given an optimal investment profile, there exists an NE of the proposed game which implements this solution. In the following section, we present some intuitive interpretation for this mechanism.

B. An intuitive explanation

Intuitively, the above mechanism operates as follows. The investment profile $\hat{\mathbf{x}}$ gives the levels of investment suggested by the mechanism designer for each user, and is derived by taking the average of all users' proposals for the public good. To ensure that these proposals are consistent, and eventually match the socially optimal levels of investment, the designer sets the taxes according to (11). Equation (11) itself consists of three terms.

First, we note that a user i can only affect the first term $(\boldsymbol{\pi}_{i+1} - \boldsymbol{\pi}_{i+2})^T \hat{\mathbf{x}}(\mathbf{m})$ in its net payment by altering its proposal on the investment profile and is closely related to the Lindahl prices of the public good [9]. We will illustrate the role of this term shortly. The second term in (11) is included to punish discrepancies among users' proposals on the investment profile by increasing their net payment in case of disagreement. Lastly, the third term, which is independent of user i's message, is included to fully redistribute taxes. In fact, the last two terms will be zero at an equilibrium of this regulated IDS game. Nevertheless, the inclusion of these terms is required to ensure convergence to the socially optimal solution, and also for balancing the budget.

We now highlight the role of the first term in (11), and its close relation to the positive externality effects of users' actions. It can be shown [12, Theorem 1] that at the equilibrium $\mathbf{m}^*$ of the regulated IDS game, the tax of a user i reduces to $\hat{t}_i = \mathbf{l}_i^{*T} \hat{\mathbf{x}}(\mathbf{m}^*)$, where $\mathbf{l}_i^* := \boldsymbol{\pi}^*_{i+1} - \boldsymbol{\pi}^*_{i+2}$. If payments are determined according to these prices, the socially optimal investments $\hat{\mathbf{x}}(\mathbf{m}^*)$ will be individually optimal as well, i.e.,

$$\hat{\mathbf{x}}(\mathbf{m}^*) = \arg\min_{\mathbf{x}\succeq 0} \quad f_i(\mathbf{x}) + h_i(x_i) + \mathbf{l}_i^{*T}\mathbf{x} \ . \tag{12}$$

As a result, for all i, and all j for which $\hat{x}_j \neq 0$, the Karush-Kuhn-Tucker (KKT) conditions on (12) yield:

$$l^*_{ij} = -\frac{\partial f_i}{\partial x_j}(\hat{\mathbf{x}}(\mathbf{m}^*)) \ . \tag{13}$$

The interpretation is that by implementing this mechanism, each user i will be financing part of user $j \neq i$'s reimbursement. According to (13), this amount is proportional to the positive externality of j's investment on user i's utility.

C. An example

To close this section, we present a family of IDS games to illustrate the lack of voluntary participation under the Externality mechanism. We identify instances in which voluntary participation fails to hold, as well as instances for which the balanced Externality mechanism can incentivize voluntary participation in achieving socially optimal investments. First note that the net payments of users in the externality mechanism can be determined using (12) and (13), and are given by:

$$t_i^* = -\sum_j x_j^* \frac{\partial f_i}{\partial x_j}(x^*) - \frac{\partial h_i}{\partial x_i} x_i^* \ .$$

Example 4: Consider a collection of N users, with risk functions given by the total effort model $f_i(\mathbf{x}) = \exp(-\sum_{i=1}^N x_i), \forall i$, and linear cost functions $h_i(x_i) = c_i x_i$. Let $c_1 < c_2 < c_3 < \ldots < c_N$. Let $\frac{c_2}{N-1} < c_1 < c_2 < 1$.

The socially optimal level of investment in this game is the solution to (2), and is given by:

$$\exp(-x_1^*) = \frac{c_1}{N}, \quad x_2^* = \ldots = x_N^* = 0 \ .$$

The taxes assessed to users in the Externality mechanism are given by:

$$t_1^* = -c_1 x_1^* (1 - \frac{1}{N}), \quad t_j^* = c_1 x_1^* \frac{1}{N}, \forall j = 2, \ldots, N \ .$$

If user 1 chooses to stay out, the equilibrium levels of investment $\tilde{\mathbf{x}}^1$ will be given by:

$$\exp(-\tilde{x}_2^1) = \frac{c_2}{N-1}, \quad \tilde{x}_j^1 = 0, \forall j \neq 2 \ .$$

If any user $j \neq 1$ decides to stay out, the equilibrium levels of security $\tilde{\mathbf{x}}^j$ will be:

$$\exp(-\tilde{x}_1^j) = \frac{c_1}{N-1}, \quad \tilde{x}_j = 0, \forall j = 2, \ldots, N \ .$$

We can now use the above to determine the voluntary participation conditions of all users. For user 1 to voluntarily participate in the mechanism, we need $g_1(\mathbf{x}^*) + t_1 \leq g_1(\tilde{\mathbf{x}}^1)$, which leads to:

$$\exp(-x_1^*) + c_1 x_1^* - c_1 x_1^* (1 - \frac{1}{N}) \leq \exp(-\tilde{x}_2^1) \ .$$

This can be further simplified to yield:

$$\frac{c_1}{N}(1 + x_1^*) \leq \frac{c_2}{N-1} \qquad (\mathrm{VP_1}) \ .$$

For any other user $j = 2, \ldots, N$, the voluntary participation condition is:

$$\exp(-x_1^*) + c_1 x_1^* \frac{1}{N} \leq \exp(-\tilde{x}_1^j) \ .$$

With further simplification, we have:

$$\frac{c_1}{N}(1 + x_1^*) \leq \frac{c_1}{N-1} \qquad (\mathrm{VP_j}) \ .$$

Note that ($\mathrm{VP_1}$) is satisfied if ($\mathrm{VP_j}$) holds. Therefore, for voluntary participation to hold in a problem instance, we need N, c_1, c_2 to satisfy:

$$c_1 > N \exp(-\frac{1}{N-1}), \quad \frac{c_2}{N-1} < c_1 < c_2 < N - 1 \ .$$

Therefore, there indeed exist instances of the IDS game where the balanced Externality mechanism can incentivize cooperation. In problem instances where the above is not satisfied, the Externality mechanism fails to satisfy users' participation constraints. ■

V. Conclusion

In this paper, we considered two incentive mechanisms, namely the Pivotal and the Externality mechanisms, which can induce socially optimal behavior in interdependent security games by using monetary taxation. We illustrated that regardless of the mechanism used to design taxes, the non-excludable nature of security as a public good prevents taxation mechanisms from getting users to voluntarily participate while maintaining a (weakly) balanced budget. Consequently, the Pivotal mechanism can only guarantee participation, while the Externality mechanism will only maintain a balanced budget.

One of the advantages of the Pivotal mechanism is that, unlike the Externality mechanism, it does not require prior information on users' cost functions. Nevertheless, in both mechanisms, users should be able to accurately determine their own cost functions; in practice these may be costly, if not impossible, to determine. We note that although theoretically, Nash equilibria describe users' actions in a game of complete information, the Nash equilibrium in the Externality game can be interpreted as the convergence point of an iterative process, in which each user adjusts his action at each round based on his observations of other users' actions, until unilateral deviations are no longer profitable [9]. Such interpretation would eliminate the need for knowledge or reporting of the cost functions.

Another drawback of the current mechanisms, as well as prior work using taxation, is the need to accurately observe users' actions. Due to the possible costs and privacy concerns, it is of interest to design incentive mechanisms that do not require such information to operate. In addition, future work will consider the use of non-taxation based mechanisms, as well as the effects of users' risk attitudes (e.g. risk aversion), in alleviating the current impossibility result.

ACKNOWLEDGMENT

The authors would like to thank Hamidreza Tavafoghi for many useful discussions. This work is supported by the Department of Homeland Security (DHS) Science and Technology Directorate, Homeland Security Advanced Research Projects Agency (HSARPA), Cyber Security Division (DHS S&T/HSARPA/CSD), BAA 11-02 via contract number HSHQDC-13-C-B0015.

REFERENCES

[1] A. Mas-Colell, M. D. Whinston, J. R. Green *et al.*, *Microeconomic theory*. Oxford university press New York, 1995, vol. 1.

[2] A. Laszka, M. Felegyhazi, and L. Buttyán, "A survey of interdependent security games," *CrySyS*, vol. 2, 2012.

[3] J. Grossklags, S. Radosavac, A. A. Cárdenas, and J. Chuang, "Nudge: Intermediaries' role in interdependent network security," in *Trust and Trustworthy Computing*. Springer, 2010, pp. 323–336.

[4] R. Böhme and G. Schwartz, "Modeling cyber-insurance: Towards a unifying framework." in *Workshop on the Economics of Information Security (WEIS)*, 2010.

[5] L. Jiang, V. Anantharam, and J. Walrand, "How bad are selfish investments in network security?" *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 549–560, 2011.

[6] H. Kunreuther and G. Heal, "Interdependent security," *Journal of Risk and Uncertainty*, vol. 26, no. 2-3, pp. 231–249, 2003.

[7] H. Varian, "System reliability and free riding," *Economics of information security*, pp. 1–15, 2004.

[8] E. H. Clarke, "Multipart pricing of public goods," *Public choice*, vol. 11, no. 1, pp. 17–33, 1971.

[9] L. Hurwicz, "Outcome functions yielding walrasian and lindahl allocations at nash equilibrium points," *The Review of Economic Studies*, pp. 217–225, 1979.

[10] D. C. Parkes, "Iterative combinatorial auctions: Achieving economic and computational efficiency," Ph.D. dissertation, University of Pennsylvania, 2001.

[11] S. Sharma and D. Teneketzis, "A game-theoretic approach to decentralized optimal power allocation for cellular networks," *Telecommunication systems*, vol. 47, no. 1-2, pp. 65–80, 2011.

[12] P. Naghizadeh and M. Liu, "Voluntary participation in cyber-insurance markets," *Workshop on the Economics of Information Security (WEIS)*, 2014.