Contagion and Observability In Security Domains

Yoram Bachrach[*], Moez Draief [†]and Sanjeev Goyal[‡]

Abstract

We examine security domains where defenders choose their security levels in the face of a possible attack by an adversary who attempts to destroy as many of them as possible. Though the attacker only selects one target, and only has a certain probability of destroying it depending on that defender's security level, a successful attack may infect other defenders. By choosing a higher security level the defenders increase their probability of survival, but incur a higher cost of security. We assume that the adversary observes the security levels chosen by the defenders before selecting whom to attack.

We show that under this assumption the defenders over-protect themselves, exhausting all their surplus, so optimal policy requires taxing security, as opposed to the subsidies recommended by alternative models for contagious attacks which do not take into account the attacker's ability to observe the defenders' choices.

1 Introduction

A key issue in many multi-agent interactions is security — an agent's ability to counteract threats from adversaries. In many domains the adversary has many potential targets, so a defending agent cannot know for sure where the adversary may strike. Even when an adversary strikes one agent, the attack may *propagate* to another agent if the first agent is successfully attacked. For example, a successful virus attack on one node in a computer network may infect other nodes connected to it, or fire from a terrorist attack on one building may spread to neighboring buildings. To decide how to protect itself an agent must have a model of the way such attacks may propagate and the adversary's behavior.

[*]Microsoft Rresearch Cambridge
[†]Imperial College London
[‡]Univerty of Cambridge

The probability that an attack on an agent succeeds depends on the security measures employed by that agent. For example, a company may invest in security software (such as antivirus software or a firewall) to reduce the probability that its computers would be successfully attacked. However, improved security also increases costs.

Random Attacks and Contagion: Consider the case where the origin of an attack is a random event, so a defender chosen uniformly at random would be attacked. Such a defender gains some utility from her standard operations, which she stands to lose if successfully attacked. The attack strikes at random, so the defender's behavior and her investment in security measures have no influence on the probability that she would be attacked. A defender may attempt to compute the probability that she would attacked directly, or indirectly due to an attack propagating from another agent. Given this probability and her utility from standard operation and costs of security measures, the defender can select the optimal security level so as to maximize her utility.

As attacks may propagate from other agents only if they are successfully attacked (and do not propagate if an attack fails), agents benefit from other defenders choosing a high security level. Unfortunately, quantitative analysis of many such domains shows that in equilibrium agents may select a security level that is lower than the social optimum [17]. In such cases, a central authority may improve utility by subsidizing security measures.

Strategic Adversary: In some security domains, such as wars or terror attacks, the origin of attacks is not a random event, but rather the decision of a *strategic* adversary, significantly impacting the defenders' policy. For example, a strategic adversary may *observe* the defenders' strategies, and choose a target so as to maximize the damage. Many questions arise in domains with a strategic adversary and infectious attacks. What is the optimal strategy for the defenders and adversary? How are these affected by the number of targets and the probability of attack propagation? Is the behavior in equilibrium optimal for the defenders, and if not, how bad is the performance in equilibrium, and what public policy should be chosen by a central authority?

1.1 Our Model

We study the above questions by proposing a simple model of such attacks where defender agents attempt to protect themselves in settings where a *strategic* adversary decides which of them to target, where *the defense resources are costly*, and where an attack may *propagate* from one attacked

agent to another. In the basic model there are two defenders and one adversary. Each defender obtains a *status quo* utility of $Y = 1$ if she is not successfully attacked, which is eliminated if she is successfully attacked.

The adversary observes the security levels of the defenders, and selects one of them attack. A successful attack may propagate from the initially attacked defender to other defenders, and the adversary attempts to maximize the number of defenders who are successfully attacked. The probability of an attack on a defender succeeding depends on the security level employed by that defender. Each defender i selects its security level s_i, and incurs a cost $C(s_i)$ for this. The function C maps any level of security to the decrease in utility the defender incurs for employing this security level. When a defender is attacked by the adversary, the attack resources of the adversary and the security measures of the defender engage in a contest determining the attack's result: a function f_p takes the defender's security level s_i and the adversary's attack resources a_i and returns the probability of the attack succeeding. [1] One such function proposed is the Tullock contest function [34, 35]. It was shown to be the only function fulfiling reasonable axioms [31] [2]. We use the Tullock function for its desirable properties, so if a defender i with security level s_i is attacked (directly or indirectly), the attack succeeds with probability $\frac{1}{1+s_i}$. If the attack against defender i succeeds, it may propagate to every other defender with a contagion probability q. As we assume that the adversary observes the security strategy of the defenders before picking the target, we have a game of sequential moves a la Stackelberg (see [22] for a detailed examination of Stackelbeg games): the defenders (leaders) move first, and their action is observed by the adversary (follower) who then best responds to it. The defenders may choose a mixed strategy, in which case the attacker best responds to the *realized* strategy. [3]

[1] We assume the attacker has a unit attack, so $a_i = 1$.

[2] The axioms are monotonicity (the probability of the attack succeeding increases in the attacker resources a_i and decreasing in the security level s_i), anonymitiy (multiple players engaged in a contest are treated equally), consistency (a sub-contest between a subset of players is qualitatively similar to a contest of more players) and independence of irrelevant alternatives (the results are not affected by players not participating in the contest).

[3] A mixed-strategy of a defender is a lottery among different security levels. To be more precise, by a Stackelberg equilibrium we mean a strategy profile composed of a (mixed) strategy for each defender, where no defender can improve upon assuming the other defnder sticks to her strategy in the profile and assuming the adversary best responds to the *realized* strategies of the defenders (i.e. the attacker knows the results of the lotteries). We believe this is justified as for the defenders security levels are a matter of long term policy, while the attacker can quickly and cheaply examine the level of security of the defenders. For example, in a cyber-security context, the attacker can run a simple

Our Contribution: We examine the above domain, analyzing equilibria and optimal joint defender outcomes:

- There is an equilibrium in which defenders choose mixed strategies, with security levels in the support $[0, \bar{s}]$, for some real $\bar{s}$ while the adversary targets the lowest security defender. Changes in the number of players leaves the support of the distribution $[0, \bar{s}]$ unaltered.

- An increase in the contagion probability q across nodes leads to a first order shift in security distribution. Let q be the probability of contagion across nodes and denote by $F_q(.)$ the mixed strategy probability distribution function under contagion probability q. For a fixed number of players n, we show that $F_{q'}(s) \geq F_q(s)$, for $q' \geq q$. This implies that an increase in the contagion probability yields a decrease in the expected security level for each defender.

- Faced with a strategic adversary, defenders exhaust all surplus. The price of anarchy is unbounded as collective provision of security has positive surplus to the defenders.

Thus, in the presence of a strategic adversary, optimal public intervention could take the form of a tax on security investments while in the case of random attacks traditional wisdom recommends a subsidy on such investments.

We also examine $n \geq 2$ defenders, showing that:

- Changes in the number of defenders leads to a first order stochastic shift in the distribution of security: the distribution function for $n+1$ players first order dominates the distribution function for n players.

- An increase in the number of players leads to an increase in the expected security level for each defender.

2 Related Work

Many game theoretic models of security were proposed, with domains including security against terrorism, network security and arms races [29, 17, 12, 10, 14, 20, 6]. Many security domains are characterized by uncertainty regarding the actions taken by the adversaries. One way to model this is a Bayesian game, where the distribution of the adversaries' types

off-the-shelf penetration test software to gauge a defnder's realized security level.

is known [5]. However, such Bayesian models fail to capture cases where an attacker may observe the security policies of the defenders and use this to her advantage. Extensive literature studies conflict between two players across multiple "battle sites" with fixed budgets. One example is the Blotto games and its derivatives [30, 11, 4, 26, 27]. Similarly, work on all-pay auctions examined players with similar goals who are in conflict with each other [34, 35, 31, 9, 7, 1]. In contest domains with *asymmetric* adversaries it was shown that pre-commitment makes the better endowed player to increase effort and the less endowed player to decrease effort over the effort in a Nash equilibrium [8]. In contrast to Blotto games or contests, in our game the *defenders* act with a view to protecting individual interests while the adversary seeks to maximize damage.

The defenders are therefore engaged in competition to avoid attacks, but also provide collective goods for each other, since greater individual security lowers the risk of spread of infection. This structure seems natural in many contexts, such as computer security where security measures are decentralized and adversaries seek to exploit inter-connections between individual entities to maximize harm.

Our assumption that the adversary observes the defenders' strategy before choosing her move yields a Stackelberg model, and is realistic in various domains. One example is computer security, where a hacker can infer security settings of different computers. Another example is physical security, where an adversary may observe the way the defending forces are allocated before launching an attack. Several other authors have proposed Stackelberg models for various security domains. For a overview of such work and its applications, see [13].

One interesting Stackelberg model examines patrolling agents defending against an attacker, who must decide which areas to patrol when only observing their close environment [24]. That paper uncovers good patrolling policies, for example for UAV's (unmanned aerial vehicle) and security robots [3, 2]. Such Stackelberg security models have formed the basis for deployed security policy optimization tool used by various agencies. One such a system assists in generating security policies for the Los Angeles International Airport [25] and another is used by the Federal Air Marshals Service [32]. Stackelnerg security games may be hard to solve due to the large strategy space, however various techniques allow overcoming this difficulty and finding good strategies [36, 21, 23, 37, 33, 15, 16].

Our model is novel in combining a *Stackelberg* game similar to [13, 16] with *contagious* attacks similar to [17, 10, 14]. In contrast to [13, 33, 15, 6], we do not focus on algorithms for solving the security game, but examine

the impact of contagion and the strategic attacker's ability to observe the the defenders' security on the choices and surplus of the defenders and public policy issues. We believe our model is more realistic in assuming the attacker observes the realized security levels, while our multiple defenders make their decisions simultaneously, without observing each other's strategy, as opposed to models such as [17, 13, 6]. Also, the clear cut results on equilibrium strategies and the price of anarchy seem to be novel relative to the existing literature.[4]

3 The Security Game

Consider a three player game: two defenders choose a security level to protect themselves against an attack from the third player. The two defenders are denoted players 1 and 2 while the adversary is denoted by the letter $\mathcal{A}$. The defenders have 'status quo' utility of $Y = 1$, which they may lose if successfully attacked. To prevent this, defender i chooses security investment $s_i \geq 0$, $i = 1, 2$ at cost $C(s_i)$. We assume that the cost function is twice continuously differentiable, $C(0) = 0$, $C'(0) = 1$, and that $C''(s) \geq 0$, for all $s \geq 0$.[5]

The security efforts of the players are observed by player $\mathcal{A}$, who then chooses to attack the nodes with the resources at his disposal. The adversary has one unit of resource and allocations take integer values: $a_i \in \{0, 1\}$. Let $a = \{a_1, a_2\}$ be the allocation of attack resources across the 2 defenders where $\sum_{i=1}^{2} a_i = 1$. First consider direct attack: suppose $\mathcal{A}$ attacks player i who has chosen s_i. Followign the Tullock function [31, 34, 35], the probability that i survives is: $\frac{s_i}{s_i + a_i}$.

Successful attacks may spread from one player to another. The probability of infection is given by $q \in [0, 1]$. If the attack does spread from i to j then it engages with defense resources at j and the probability of survival for node j is given by: $\frac{s_j}{s_j + a_i}$. We assume an immunity property with regard to attacks: if node i survives a direct or indirect attack, it is henceforth immune to further attacks.

[4] The possibility of excessive security investments and the need for collective action is discussed by [29], but as far as we are aware formal results on the structure of mixed equilibria and comparative statics have not been obtained before.

[5] Our results carry over to a setting with discrete security levels; see the section on discrete security for an example.

3.1 The attacker

Let $\sigma(i) \in [0,1]$ be the probability that player $\mathcal{A}$ attacks node i. We first focus on the optimal strategy of the attacker.

Observation 1. *A strategic attacker attacks the player with the lowest security level, and it is indifferent between attacking two defenders who choose the same level of security.*

Proof. Suppose players choose (s_1, s_2). We solve the game backwards, starting with $\mathcal{A}$'s optimal action. The expected number of successfully attacked nodes with $\sigma(1) = 1$ is:

$$\frac{1-q}{s_1+1} + \frac{q}{1+s_1}\frac{s_2}{s_2+1} + \frac{2q}{s_1+1}\frac{1}{s_2+1}$$

Similarly, the expected number of successfully attacked nodes with $\sigma(2) = 1$ is given by:

$$\frac{1-q}{s_2+1} + \frac{q}{1+s_2}\frac{s_1}{s_1+1} + \frac{2q}{s_2+1}\frac{1}{s_1+1}$$

It is easily verified that attacking 1 yields strictly greater utility than attacking 2 if and only if $s_1 < s_2$; Targeting 1 and 2 yields $\mathcal{A}$ equal utility when $s_1 = s_2$. □

3.2 The defenders

Now consider the defenders. The attacker targets the least defended agent, so if $s_1 < s_2$ then player 1's payoff is: $\left[\frac{s_1}{s_1+1}\right] - C(s_1)$. If $s_1 > s_2$ then player 1's payoff is:

$$\left[\frac{s_2}{s_2+1} + \frac{1}{s_2+1}(1-q) + \frac{1}{1+s_2}q\frac{s_1}{s_1+1}\right] - C(s_1).$$

If $s_1 = s_2$ then we assume $\mathcal{A}$ chooses to attack each of the nodes with equal probability so player 1's payoff is:

$$\frac{1}{2}\left[\frac{s_1}{s_1+1} + \frac{s_2}{s_2+1} + \frac{1-q}{s_2+1} + \frac{q}{1+s_2}\frac{s_1}{s_1+1}\right] - C(s_1).$$

Consider the incentives at work here. First, raising s_1 renders a direct and an indirect attack less successful. Second, and perhaps more interestingly, raising s_1 has the potential of shifting the threat to the other player: as s_1 moves across the s_2 threshold, the optimal response of the adversary is

to switch from player 1 to player 2. This makes security an *implicitly* "competitive" phenomenon. The competitive effect is entirely a consequence of the strategic character of the adversary: in domains where the attack is invariant with respect to security investments, matters are different (as the section on random attack below clarifies). We first show that there only exists a mixed equilibrium.

Lemma 1. *For $q \in (0,1)$, there exists no pure strategy equilibrium for the defenders.*

Proof. First, straightforward computations show that setting zero security level is not optimal for a defender. Let us consider unequal security levels, $0 < s_1 < s_2$. Player 1 must satisfy the necessary condition for an optimum: $C'(s_1) = \frac{1}{(1+s_1)^2}$ while player 2's action s_2 must satisfy: $C'(s_2) = \frac{q}{(1+s_1)(1+s_2)^2}$. As $s_2 > s_1$ and $C(.)$ is convex, it follows that $C'(s_2) \geq C'(s_1)$. But an inspection of the (LHS) terms reveals that the (gross) marginal returns for player 2 are smaller than the marginal returns to player 1, contradicting the optimality of player actions.

Let us now examine whether the actions $s_1 = s_2 = s$ represent a pure strategy equilibrium. At such an equilibrium, however, a defender strictly gains by raising security slightly as this strictly raises his (gross) return, due to the fact that the strategic adversary switches its target. Thus, there does not exist a pure strategy equilibrium. □

Theorem 1. *There exists a symmetric mixed zero-payoff strategy with probability density function:*

$$f(s) = \frac{s+1}{1-q+s}\left((s+1)C'(s) + C(s) - 1\right) \tag{1}$$

Example 1. *Before giving the proof of Theorem 1, consider the mixed strategy profile in the theorem when the cost is linear $C(s) = s$. It is easy to see that in this case, the above symmetric mixed equilibrium has the following probability distribution function: $F(s) = s^2 + 2qs - 2q(1-q)\log\left(\frac{1+s-q}{1-q}\right)$ with support $[0, \bar{s}]$ where $\bar{s}$ is the solution of the equation: $s^2 + 2qs - 2q(1-q)\log\left(\frac{1+s-q}{1-q}\right) = 1$. In particular, if $q = 0$, we have $F(s) = s^2$ for $s \in [0,1]$.*

Proof. Assuming that the strategies of agents in equilibrium are independent drawn form a distribution with c.d.f. F, then the payoff of a given agent known that the other agent invested s in defending herself is given by:

$$\int_0^s \left[\frac{t}{1+t}+\frac{1-q}{1+t}\right] dF(t) + q\frac{s}{s+1}\int_0^s \frac{1}{1+t}dF(t) + [1-F(s)]\frac{s}{s+1} - C(s)$$

1. $F(s)$ has no mass points. This follows from the discontinuity in payoffs at $s_1 = s_2$. Suppose both players place probability mass $p > 0$ on a security level s. Observe that an individual player can strictly benefit from moving this mass from s to $s+\epsilon$, for some small $\epsilon > 0$. The cost is almost the same, but the returns register a discontinuous increase as the probability that other player has a lower security level increases by a factor of p.

2. Interval support: the support of the probability distribution is a convex set. Suppose not. Then there is an interval $[s', s'']$ such that $F(s') = F(s'')$ and $s' < s''$. From the equilibrium condition (that expected profits are zero at $s = 0$ and at the maximum security $s = \hat{s}$) the net payoffs at these two levels are also zero. For $s = s', s''$ we have:

$$\begin{aligned} C(s) &= \int_0^s \left[\frac{t}{1+t}+\frac{1-q}{1+t}\right] dF(t) \\ &\quad + q\frac{s}{s+1}\int_0^s \frac{1}{1+t}dF(t) + [1-F(s)]\frac{s}{s+1}\,. \end{aligned}$$

   Define $\tilde{s} = (s' + s'')/2$. The gross expected return at security level $\tilde{s}$ is: $\int_0^{\tilde{s}} \left[\frac{t}{1+t}+\frac{1-q}{1+t}\right] dF(t) + \frac{s}{s+1}\int_0^s \frac{q}{1+t}dF(t) + [1-F(\tilde{s})]\frac{\tilde{s}}{\tilde{s}+1}$. Denote this return by $R(\tilde{s})$. As the return function is concave in s, it is easily checked that $R(\tilde{s}) > [R(s') + R(s'')]/2$. However, $C(\tilde{s}) \leq [C(s') + C(s'')]/2$, due to the convexity of $C(.)$. So we have shown that:

$$\begin{aligned} C(\tilde{s}) &< \int_0^{\tilde{s}} \left[\tfrac{t}{1+t}+\tfrac{1-q}{1+t}\right] dF(t) + \tfrac{s}{s+1}\int_0^{\tilde{s}} \tfrac{q}{1+t}dF(t) \\ &\quad + [1-F(\tilde{s})]\tfrac{\tilde{s}}{\tilde{s}+1} \end{aligned}$$

   and the player can earn a strictly positive profit by setting a security level in (s', s''), a contradiction which shows that the support of the mixed strategy distribution must be a convex set. We thus consider a mixed equilibrium where each player uses a distribution $F(s)$ with support $[\underline{s}, \bar{s}]$. We define $\underline{s}$ and $\bar{s}$ implicitly as the points where:

$$C(\underline{s}) = \frac{\underline{s}}{\underline{s}+1}\,, \tag{2}$$

$$C(\bar{s}) = \int_0^{\bar{s}} \left[\frac{t}{1+t} + \frac{1-q}{1+t} + \frac{q}{1+t}\frac{\bar{s}}{1+\bar{s}} \right] dF(t) \tag{3}$$

Thus for $s \in [\underline{s}, \bar{s}]$, we have:

$$\begin{aligned} C(s) &= \int_0^{s} \left[\frac{t}{t+1} + \frac{1-q}{1+t} + \frac{q}{t+1}\frac{s}{s+1} \right] dF(t) \\ &\quad + [(1-F(s)]\,\frac{s}{s+1}. \end{aligned} \tag{4}$$

3. $\underline{s} = 0$:Note that the cost to the player at $\bar{s}$ is exactly the expected return, so the net payoff is zero. Similarly at $s_1 = \underline{s}$: $F(\underline{s}) = 0$, so with probability 1, player 2 chooses a higher security. This means that $\mathcal{A}$ attacks player 1 with probability 1, and as $s_1 = \underline{s}$, the attack is successful, yielding player 1 a net payoff 0. The same reasoning applies to player 2. So each defending player earns zero in such an equilibrium. Exploiting properties of the cost function C and equation (2), one can show that $\underline{s} = 0$.

4. Probability density function for symmetric equilibrium: We now derive an expression for $f = F'$, the probability distribution function of the mixed equilibrium. Recall that the zero-payoff property is given by:

$$\begin{aligned} C(s) &= \int_0^{s} \frac{t}{1+t} f(t)dt + q\frac{s}{s+1}\int_0^{s} \frac{1}{1+t} f(t)dt \\ &\quad + \int_0^{s} \frac{1-q}{1+t} f(t)dt + \left(1 - \int_0^{s} f(t)dt\right)\frac{s}{s+1} \\ &= \frac{\int_0^{s} f(t)dt}{s+1} - \frac{q}{s+1}\int_0^{s} \frac{1}{1+t} f(t)dt + \frac{s}{s+1} \end{aligned}$$

This may be simplified as:

$$(s+1)C(s) - s = \int_0^{s} f(t)dt - q\int_0^{s} \frac{1}{1+t} f(t)dt$$

By differentiating the above expression we have:

$$(s+1)C'(s) + C(s) - 1 = f(s) - q\frac{1}{1+s} f(s)$$

which implies that

$$f(s) = \frac{s+1}{1-q+s}\left((s+1)C'(s) + C(s) - 1\right) \tag{5}$$

We thus obtained the required mixed equilibrium. □

An inspection of the equilibrium density function (5) reveals the following features of equilibrium security:

1. For $q < 1$, the density derived from the zero-payoff assumption implies that the marginal return is equal to the marginal cost. This not satisfied for $q = 1$. In fact at $q = 1$, $(0, 0)$ is a pure strategy equilibrium.

2. Note that for $q \in [0, 1)$, the density function f is increasing in q. Since $f(0) = 0$ for all $q < 1$, this implies that for $q' > q$, $F_{q'}(s) \geq F_q(s)$, for every s. This in turn implies that the maximum security level $\bar{s}$ and the average security levels are both decreasing with q.

3.3 Price of Anarchy

Let us now consider the situation where the players can coordinate their efforts and maximize joint payoffs.

Theorem 2. *The optimal social collective strategy for the two defenders when the attacker is strategic is given by*

(*i*) $s_1 = s_2 = 0$, *if* $q \leq 1/2$

(*ii*) $s_1 = s_2 = s^*$, *if* $q \geq 1/2$ *where* s^* *is the solution in* $[0, 1]$ *of the equation:* $C'(s) = \frac{1}{2}\frac{1}{(1+s)^2} + \frac{q}{(1+s)^3}$.

Proof. The social welfare function (collective payoff) for the two defenders with defense levels $s_1 > s_2$ is: $2\frac{s_2}{1+s_2} + \frac{1}{1+s_2}\left[1 - q + q\frac{s_1}{1+s_1}\right] - C(s_1) - C(s_2)$. The first order optimality conditions yield: $C'(s_1) = \frac{q}{(1+s_1)^2(1+s_2)}$ and $C'(s_2) = \frac{q+s_1+1}{(1+s_2)^2(1+s_1)}$. Using the convexity of the cost function $C(s)$, one gets a contradiction. We thus examine symmetric profiles where $s_1 = s_2 = s$, so the attacker uniformly attacks either defender, and the social welfare is: $2\frac{s}{1+s} + \frac{1}{1+s}\left[1 - q + q\frac{s}{1+s}\right] - 2C(s)$. Note that this expression is decreasing when $q \leq 1/2$, otherwise it reaches its maximum at the point s^* defined in the theorem above. □

It is easy to show that the social optimum always has a positive payoff to the defenders. This implies that security investments in face of a strategic adversary are *excessive* relative to what is collectively desirable for the defenders. One measure of the potential costs of strategic action is the Price of

Anarchy [28] (PoA). In our domain, the PoA is the ratio between the maximum feasible total utility of defenders and the sum of payoffs of defenders in the worst Nash equilibrium. We constructed a Nash equilibrium in which defenders make zero profits while our computations reveal that positive joint payoffs are feasible. It then follows that the PoA is *unbounded*!

3.4 Random attack

Now consider a *non-strategic* adversary. Suppose for simplicity that each player is attacked with equal probability. Then player 1 maximizes the following payoff:

$$\max_{s_1} \quad \frac{1}{2}\left[\frac{s_1}{1+s_1}+\frac{s_2}{1+s_2}+\frac{1-q}{1+s_2}+\frac{q}{1+s_2}\frac{s_1}{1+s_1}\right]-C(s_1)$$

Differentiating and using symmetry of security effort across players, we get that the optimal s must solve: $C'(s)=\frac{1}{2}\left[\frac{1}{(1+s)^2}+\frac{q}{(1+s)^3}\right]$. Moreover the security level that maximizes joint payoff is similar to the one derived for the strategic attacker setting, i.e. it solves: $C'(s)=\frac{1}{2}\frac{1}{(1+s)^2}+\frac{q}{(1+s)^3}$.

Observe that for given s, the right side expression of the collective optima is larger than that for the individual optima. Since the cost of security is weakly convex, while the returns are strictly concave, it follows that individual security choices are too small relative to what is in the players' joint interest. Recall, that, by contrast, faced with a strategic adversary, players choose too high a level of security.

Finally, we study the total security expenditure in both the mixed Nash equilibrium of Theorem 1 and the social optimum of Theorem 2 (equal to the equilibrium investment in the random attack setting). Examine the case of $q=0$ in the setting of Example 1 when $C(s)=s$. In this case the social optimum is reached at: $s^*=1-\frac{\sqrt{2}}{2}$. Simple calculations show an average security investment of $2/3$ in equilibrium. Thus the total investment of the two players in the social optimum is roughly 0.6 while it is 1.33 in equilibrium.

3.5 The n-defender game

In the benchmark model, we studied the case of two defenders and one adversary. In this section we consider $n \geq 2$ players. Suppose that infection can spread from the initially attacked node, if infected, to other nodes with probability $q \in [0,1]$. No further infections occur subsequently.

For concreteness and to get close form solutions, we only present our analysis for the linear cost function $C(s) = s$. Some key conclusions, namely that the support of the mixed strategy distribution is unaltered by a change in the number of players and that the distribution function is stochastically increasing with the number of players, are valid for the family of cost functions introduced earlier.

Theorem 3. *In the case with n players with $q \in (0,1)$ and $C(s) = s$, there exists no pure strategy equilibrium for the defenders. There exists a symmetric mixed zero-payoff strategy with probability density function:*

$$F(s) = 1 - \left(1 - s^2 - 2qs + 2q(1-q)\log\left(\frac{1+s-q}{1-q}\right)\right)^{\frac{1}{n-1}}$$

with support $[0, \bar{s}]$ where $\bar{s}$ is the solution of the equation:

$$s^2 + 2qs - 2q(1-q)\log\left(\frac{1+s-q}{1-q}\right) = 1$$

Proof. One can show that the zero-payoff condition yields:

$$\begin{aligned} s &= \int_0^s f^{(n-1)}(t)\left(\frac{t}{1+t} + q\frac{1}{1+t}\frac{s}{1+s} + \frac{(1-q)}{1+t}\right) \\ &\quad + (1-F(s))^{n-1}\frac{s}{1+s} \end{aligned} \tag{6}$$

where $f(s)$ and $F(s)$ are, respectively, the density function and the probability distribution function of the security profile, and where the density of the lowest order statistic among $n-1$ i.i.d random variables with distribution F is: $f^{(n-1)}(s) = (n-1)f(s)(1-F(s))^{n-2}$. Moreover, f is defined over the interval $[0, \bar{s}]$ where $\bar{s}$ is the solution of: $\int_0^{\bar{s}} f(t)(n-1)(1-F(t))^{n-2}\left(\frac{t}{1+t} + \frac{q\bar{s}}{(1+t)(1+\bar{s})} + \frac{(1-q)}{1+t}\right)$.

Analyzing (6) as in the two-player case yields: $f^{(n-1)}(s) = \frac{2s(1+s)}{1+s-q}$. Integrating we get: $(1-F(s))^{n-1} = 1 - \int_0^s \frac{2t(1+t)}{1+t-q}dt$. Hence, $F(s)$ is given by equation (3) and $F(s)$ is an increasing function of q as in the two player case. Note that the maximum security level does not depend on n as it is the solution to the Equation (3). Equation (3) shows that $F(s)$ declines with n, for all s. We conclude that an increase in n leads to a first order stochastic dominant shift in the distribution of security investments. □

3.6 Discrete Security Levels

Our analysis so far assumed that security is a continuous variable, while attack is a discrete variable. We provide an example which illustrates that continuous security levels are not needed: excessive security investments are obtained for even a moderately rich menu of discrete security choices.

Example 2. *Consider a model with three discrete security levels,* $(0, 1, 2)$. *The cost of each unit of security is* $c > 0$. *As before the adversary has a single discrete unit of attack, and can choose integer allocations only.*

Simple computations for Example 2, along the same lines as the computations in Section 3, reveals an interesting structure. The joint payoffs to the defenders for the various strategy profile are given in Table 1. It is

Strategies	Welfare	Strategies	Welfare
$(0,0)$	0	$(1,0)$	$1/2 - c$
$(1,1)$	$(10/8 - 2c)$	$(2,1)$	$4/3 - 3c$
$(2,2)$	$2[7/9 - 2c]$	$(2,0)$	$2/3 - 2c$

Table 1: Strategy profiles and their total defenders' welfare (joint payoffs) in the discrete game.

straightforward to deduce that joint optimal outcomes are as follows: Strategy profile $(2, 2)$ is optimal if and only if $c < 11/72$, strategy profile $(1, 1)$ is optimal if and only if $11/72 < c < 5/8$, while strategy profile $(0, 0)$ is optimal if and only if $c > 5/8$.

Consider equilibrium choices. The payoffs to player i under different strategy configurations are: $\Pi_i(0,0) = 0, \Pi_i(1,0) = 1/2 - c, \Pi_i(0,1) = 0, \Pi_i(1,1) = 5/8 - c, \Pi_i(2,0) = 2/3 - 2c, \Pi_i(0,2) = 0, \Pi_i(2,1) = 5/6 - 2c, \Pi_i(1,2) = 1/2 - c, \Pi_i(2,2) = 7/9 - 2c$. Then $(0, 0)$ is an equilibrium if and only if $c > 1/2$. Similarly, $(1, 1)$ is an equilibrium if and only if $15/72 < c < 45/72$. Finally, $(2, 2)$ is an equilibrium if and only if $0 < c < 20/72$.

A comparison of the equilibria and joint optima shows: If $c \in (0, 11/72)$ then both the joint optimum and equilibrium is (2,2); If $c \in (11/72, 15/72)$ then the joint optimum is $(1, 1)$ but $(2, 2)$ is the unique equilibrium, so there is *over-investment* in equilibrium; If $c \in (15/72, 20/72)$ then the joint optimum is $(1, 1)$ but $(2, 2)$ and $(1, 1)$ are both equilibrium, So there is a potential for over-investment; If $c \in (20/72, 36/72)$, then both the joint optimum and equilibrium is $(1, 1)$; If $c \in (36/72, 45/72)$ then the joint optimum

is $(1,1)$ but equilibria are $(1,1)$ and $(0,0)$, so there is a potential for *under-investment*; Finally, if $c \in (45/72, 1)$ both joint optimum and equilibria is $(1,1)$.

The above analysis for Example 2, with only three security levels, illustrates two points. First, competition for security generates pressures toward *over-investment* in security even with a discrete security levels when costs of security are low. Second, there is the prospect of under-investment in security when costs are high. Thus both over-investment and under-investment relative to collective optima are possible.

4 Conclusion

We examined security games focusing on the effect of a strategic adversary who observes the defenders' security level has on the equilibrium in settings with contagious attacks. We showed that in this domain the defenders over-secure and thus exhaust all their surplus. Though we relied on various assumptions, we believe that our results point to a general theme: the strategic nature of the adversary may result in players "competing" for security, causing an over-invest in security. In this case, coordinated action is attractive, and the optimal public policy is *taxing* security, in contrast to conventional wisdom which argues in favor of *subsidizing* it in contexts of contagious attacks.

Our analysis leaves many questions open for further research. First, one can consider extensions of our model where the attacker can allowed split its budget among the players, dedicating $a_i \in (0,1)$ to defender i. Second, we assumed that defenders are symmetric. Asymmetries in value of the defenders will lead to different security investments, which is interesting to analyze. Finally, we assumed that a successful attack may infect all of the remaining defenders, but in many domains such attacks can only spread in certain ways, captured by the network structure (see [19, 18]). Extending our models to a network setting is welcome.

References

[1] D. Arce, D. Kovenock, and B. Roberson. Suicide terrorism and the weakest link, 2009.

[2] N. Basilico, N. Gatti, and F. Amigoni. Leader-follower strategies for robotic patrolling in environments with arbitrary topologies. In *AAMAS*, pages 57–64, 2009.

[3] R. Beard and T. McLain. Multiple UAV cooperative search under collision avoidance and limited range communication constraints. In *Decision and Control*, 2003.

[4] V. Bier, S. Oliveros, and L. Samuelson. Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, 9:1–25, 2006.

[5] J. Brynielsson and S. Arnborg. Bayesian games for threat prediction and situation analysis. In *FUSION*, 2004.

[6] H. Chan, M. Ceyko, and L. Ortiz. Interdependent defense games: Modeling interdependent security under deliberate attacks.

[7] D. Clark and K. Konrad. Asymmetric conflict. *Journal of Conflict Resolution*, 51(3):457, 2007.

[8] A. Dixit. Strategic behavior in contests. *The American Economic Review*, 77(5):891–898, 1987.

[9] J. Esteban and D. Ray. A model of ethnic conflict. *Journal of the European Economic Association*, 2007.

[10] J. Grossklags, N. Christin, and J. Chuang. Security investment (failures) in five economic environments: A comparison of homogeneous and heterogeneous user agents. 2008.

[11] S. Hart. Discrete colonel blotto and general lotto games. *International Journal of Game Theory*, 36(3):441–460, 2008.

[12] K. Hartley and T. Sandler. *Handbook of Defense Economics: Defense in a globalized world*. North Holland, 2007.

[13] M. Jain, B. An, and M. Tambe. An overview of recent application trends at the aamas conference: Security, sustainability and safety. 2011.

[14] B. Johnson, J. Grossklags, N. Christin, and J. Chuang. Uncertainty in interdependent security games. *Decision and Game Theory for Security*, pages 234–244, 2010.

[15] D. Korzhyk, V. Conitzer, and R. Parr. Complexity of computing optimal stackelberg strategies in security resource allocation games. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 805–810, 2010.

[16] D. Korzhyk, Z. Yin, C. Kiekintveld, V. Conitzer, and M. Tambe. Stackelberg vs. nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness. *Journal of Artificial Intelligence Research*, 41(2):297–327, 2011.

[17] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2):231–249, 2003.

[18] M. Lelarge. Economics of malware: Epidemic risks model, network externalities and incentives. In *Communication, Control, and Computing*, 2009.

[19] M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the internet. In *SIGMETRICS*, 2008.

[20] J. Letchford and Y. Vorobeychik. Computing optimal security strategies for interdependent assets.

[21] S. Leyffer and T. Munson. Solving multi-leader-follower games. 2005.

[22] M. J. Osborne and A. Rubenstein. *A Course in Game Theory*. The MIT Press, Cambridge, Massachusetts, 1994.

[23] P. Paruchuri, J. Pearce, J. Marecki, M. Tambe, F. Ordonez, and S. Kraus. Playing games for security: an efficient exact algorithm for solving bayesian stackelberg games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 895–902. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[24] P. Paruchuri, J. Pearce, M. Tambe, F. Ordonez, and S. Kraus. An efficient heuristic approach for security against multiple adversaries. In *AAMAS*, 2007.

[25] J. Pita, M. Jain, J. Marecki, F. Ordóñez, C. Portway, M. Tambe, C. Western, P. Paruchuri, and S. Kraus. Deployed armor protection: the application of a game theoretic model for security at the los angeles international airport. In *AAMAS*, pages 125–132, 2008.

[26] R. Powell. Sequential, non zero-sum blotto: Allocating defensive resources prior to attack. *Games and Economic Behavior*, 67(2):611–615, 2009.

[27] B. Roberson. The colonel blotto game. *Economic Theory*, 29(1):1–24, 2006.

[28] T. Roughgarden. *Selfish routing and the price of anarchy*. The MIT Press, 2005.

[29] T. Sandler. Collective action and transnational terrorism. *The World Economy*, 26(6):779–802, 2003.

[30] M. Shubik and R. Weber. Systems defense games: Colonel blotto, command and control. *Naval Research Logistics Quarterly*, 28(2):281–287, 1981.

[31] S. Skaperdas. Contest success functions. *Economic Theory*, 7(2):283–290, 1996.

[32] J. Tsai, S. Rathi, C. Kiekintveld, F. Ordonez, and M. Tambe. Iris-a tool for strategic security allocation in transportation networks. *AAMAS*, 2009.

[33] J. Tsai, Z. Yin, J. young Kwak, D. Kempe, C. Kiekintveld, and M. Tambe. Urban security: Game-theoretic resource allocation in networked physical domains. AAAI, 2010.

[34] G. Tullock. The welfare costs of tariffs, monopolies, and theft. *Economic Inquiry*, 5(3):224–232, 1967.

[35] G. Tullock. Efficient rent seeking. *Toward a theory of the rent-seeking society*, 97:112, 1980.

[36] A. Washburn and K. Wood. Two-person zero-sum games for network interdiction. *Operations Research*, 43(2):243–251, 1995.

[37] Z. Yin, D. Korzhyk, C. Kiekintveld, V. Conitzer, and M. Tambe. Stackelberg vs. nash in security games: Interchangeability, equivalence, and uniqueness. In *AAMAS*, pages 1139–1146, 2010.