Can Less Be More? A Game-Theoretic Analysis of Filtering vs. Investment

Armin Sarabi, Parinaz Naghizadeh, and Mingyan Liu

University of Michigan, Ann Arbor, USA
[arsarabi,naghizad,mingyan]@umich.edu

Abstract. In this paper we consider a single resource-constrained strategic adversary, who can arbitrarily distribute his resources over a set of nodes controlled by a single defender. The defender can (1) instruct nodes to filter incoming traffic from another node to reduce the chances of being compromised due to malicious traffic originating from that node, or (2) choose an amount of investment in security for each node in order to directly reduce loss, regardless of the origin of malicious traffic; leading to a *filtering* and an *investment* game, respectively. We shall derive and compare the Nash equilibria of both games for different resource constraints on the attacker. Our analysis and simulation results show that from either the attacker or the defender's point of view, none of the games perform uniformly better than the other, as utilities drawn at the equilibria are dependent on the costs associated with each action and the amount of resources available to the attacker. More interestingly, in games with highly resourceful attackers, not only the defender sustains higher loss, but the adversary is also at a disadvantage compared to less resourceful attackers.

1 Introduction

The continuous attempts by malicious entities to discover and exploit security vulnerabilities in networks and the ensuing efforts of network administrators at patching up such exploits have evolved into a cat and mouse game between attackers and defenders. In addition to research on mitigating security flaws and building more robust networks by analyzing specific hardware and software involved in a network, the problem has also been addressed by game theorists. Game theory provides a broad framework to model the behavior of rational parties involved in a competitive setting, where each party seeks to maximize their own net worth. For instance, the interdependent nature of cyber-security leads to numerous studies on games describing the behavior of multiple interdependent agents protecting their assets in a network [8].

In this paper we study the strategic interaction between an attacker and a defender[1], both taking actions over a set of interconnected nodes or entities. The

[1] For the remainder of the paper, to eliminate confusion, we will use *he/him* to refer to the attacker, and *she/her* to refer to the defender.

attacker is resource limited but can arbitrarily spread his resources or effort over this set of nodes. The amount of effort exerted over a node determines the attacker's likelihood of infiltrating the node and inflicting a certain amount of loss; a compromised node can go on to contaminate other connected nodes to inflict further loss. From the defender's point of view, the interactions between nodes present possible security risks, but also value derived from the communication.

We consider two types of actions the defender can take. The first is inbound filtering, whereby a certain amount of traffic from another node is blocked. This is routinely done in practice, through devices such as firewalls and spam filters, based on information provided by sources such as host reputation blacklists (RBLs) [9, 7], where traffic originating from IP addresses suspected of malicious activities (listed by the RBLs) are deemed unsafe and blocked. Ideally, if the defender could distinguish between malicious and innocuous traffic, she could block all malicious traffic and achieve perfect security. However, blocking traffic comes at a price, since no detection mechanism is without false alarms. Thus, filtering decisions leads to tradeoffs between balancing security risks and communication values. The second type of action is self-protection through investing in security. In this case the defender foregoes filtering, but instead focuses on improving its ability to resist malicious effort in the presence of tainted traffic. Self-protection is more costly than inbound filtering, but it does not put legitimate communication at risk since it carries no false alarms.

These two types of actions result in a *filtering game* and an *investment game*, respectively, which we analyze in this paper. Specifically, we derive Nash equilibria in both scenarios. We shall see that for both games, more powerful attackers, or those with larger amounts of resources, do not necessarily draw more utility at the equilibrium. By contrast, a defender always prefers to face less powerful attackers. In addition, we will compare these two games, and conclude that highly resourceful attackers favor facing a defender that invests, while less resourceful attackers' preference depends on the cost of security investments.

Most of the existing literature on interdependent security games focus on a collection of agents responding to a constant exogenous attempt to breach their systems and inflict damage, while fewer publications have addressed games with a strategic adversary. In reality, malicious sources have shown highly strategic behavior. For instance, in November 2008 the McColo ISP was effectively blocked by the rest of the Internet due to its massive operation in spam, and its takedown was estimated to have contributed to a two-thirds reduction in global spam traffic in the immediate aftermath [2]. However, by the second half of March, the seven-day average spam volume was back at the same volume seen prior to the blocking of McColo ISP [1]. In other words, if a defender decides to completely secure her assets, then the attacker will likely respond strategically by redirecting resources.

Studies on strategic attackers and most relevant to the present paper include [3, 10, 4–6]. Specifically, in [3] Fultz and Grossklags propose a complete information game consisting of a single attacker and N defenders, where defenders can decide to protect their systems through security measures and/or self-insurance. The attacker is assumed to decide only on the number of targeted nodes, with the intensity being equal among all. In [10] Nochenson and Heimann consider a single attacker competing with a single defender in a game of incomplete information, where the players can only choose from a set of action classes (e.g. protecting the highest value node, protecting proportional to nodes' values, etc). In [4–6], Hausken considers one-shot and sequential attacker-defender games, under different assumptions on independent and interdependent security models, attacker income and substitution effects, and so on.

Compared to the above references, the present paper examines a network with a large number of nodes, where the attacker can spread his efforts over the network arbitrarily. Moreover, the utility models studied herein differ from those in [4–6]. Perhaps most importantly, our study complements existing literature by considering filtering actions, in addition to the security investment actions, in order to evaluate and compare the effectiveness of security measures and blacklisting against strategic attackers from a game theoretical point of view.

In the remainder of the paper we present our model, provide intuition on how it relates to current cyber-security practices, and derive the Nash equilibria of games under discussion. We will then simulate, discuss and compare the games and their respective equilibria. The proof of the theorems are omitted for brevity.

2 Filtering

We consider a network consisting of N inter-connected nodes. There is a single attacker and a single defender both acting over these nodes. The attacker has a fixed amount of resources he can use toward compromising any subset of the N nodes. A compromised node sustains a certain amount of (direct) loss; a compromised node is also assumed to inflict further (indirect) losses on nodes it communicates with, thus modeling interdependence. On the defender's side, one mitigating option is inbound/outbound filtering over these nodes. Filtering traffic can effectively reduce the amount of malicious traffic received by a node, thereby reducing its probability of being compromised, or the incurred losses. The extreme form of filtering is *takedown*, whereby traffic from a node is completely blocked, effectively isolating this node from the rest of the network. An advantage of filtering is low cost; it takes relatively little to perform inbound filtering, and we will assume its cost is zero in our analysis. The downside of filtering is false positives, which reduce the *value* represented by communication between two nodes; this aspect is explicitly modeled in this case.

Following the discussion above, the defender's actions can be modeled by a vector $\boldsymbol{f} \in [0,1]^N$, where f_i is the percentage of node i's outgoing traffic that is being dropped. We assume this filtering is performed uniformly, either by outbound filtering across all egress points, or inbound filtering done by all other nodes which have agreed upon the same filtering level. In reality, this corresponds to the observation that filtering decisions are often source-based rather that destination-based. The attacker's actions are modeled by a vector $\boldsymbol{r} \in \mathbb{R}_+^N$, where r_i is the amount of effort spent by the attacker to breach node i. Increased effort exerted over a node leads to increased losses (e.g., through increased probability of compromising a node). The total amount of loss inflicted constitutes the attacker's profit. We further assume that the attacker has a limited amount of resources r, so that $\sum_{i=1}^{N} r_i \leq r$.

We will adopt the simplification of only considering indirect losses. The justification is that in large networks, the amount of direct loss sustained by a node is negligible compared to the total indirect losses it can inflict on the network. In a sense, the attacker's main objective is to contaminate a large set of nodes through network effects, rather than drawing utility from compromising selected nodes. Let L_{ij} denote the maximum loss per unit of effort that can be inflicted on node j through a breached node i, when node i's traffic is unfiltered. Note that filtering the traffic leaving a breached node does not protect the node itself against losses; it protects to some degree the rest of the network from indirect losses from that node. Thus, the attacker's utility is given by:

$$u_a^F(\boldsymbol{r}, \boldsymbol{f}) = \sum_{i=1}^{N} r_i \sum_{\substack{j=1 \\ j \neq i}}^{N} L_{ij} g^F(f_i) \ , \ \text{s.t.} \ \sum_{i=1}^{N} r_i \leq r \ . \tag{1}$$

Here, $g^F : [0,1] \to [0,1]$ is a risk function with respect to the filtering policy, which we will take to be linear ($g^F(f_i) = 1 - f_i$). To further illustrate, it is more natural to view a node as a network (a collection of individual machines or IP addresses); in this case the single defender becomes a convenient way to model consistent actions taken by different networks against other networks of known malicious activities. For instance, benign networks may adopt similar inbound filtering policies against a network known to send out large quantities of malicious traffic (e.g., given by the reputation blacklists). More specifically, a network may decide that all traffic from another network with a certain presence on the RBLs (percentage of its IPs listed) shall be filtered at a certain level (with some probability). In this case, the filtering level leads to linear reduction in risk and loss in value for the node. Alternatively, a network may decide that all traffic from listed IPs shall be blocked, in which case the amount of filtering is equal to the fraction of blacklisted IP addresses. However, with this interpretation, the reduction in risk and loss in value are no longer linear with respect to filtering levels. This is because targeted filtering is presumably more accurate, leading to higher risk reduction. We will revisit this case after deriving the equilibrium of our game, and explain how the results might also hold for the nonlinear case.

Define $L_i := \sum_{j \neq i} L_{ij}$ as the total indirect loss incurred by node i. We assume without loss of generality that users are indexed such that L_i is a decreasing sequence. Equation (1) can then be re-written as:

$$u_a^F(\boldsymbol{r}, \boldsymbol{f}) = \sum_{i=1}^{N} r_i L_i (1 - f_i) \ , \ \text{s.t.} \ \sum_{i=1}^{N} r_i \leq r \ . \tag{2}$$

From the defender's viewpoint, let V_i be the value associated with node i's traffic. Similar to the definition of L_i, $V_i = \sum_{j \neq i} V_{ij}$ is the value of traffic from node i to the rest of the nodes. Note that by filtering inbound traffic, the defender is inevitably losing a portion of a node's value, as she is filtering parts of the legitimate traffic as well. The defender's utility is thus given by:

$$u_d^F(\boldsymbol{r}, \boldsymbol{f}) = -u_a^F(\boldsymbol{r}, \boldsymbol{f}) + \sum_{i=1}^{N} V_i(1 - f_i) \ . \tag{3}$$

Together, $\mathcal{G}^F := \langle(\text{attacker, defender}), (\boldsymbol{r}, \boldsymbol{x}), (u_a^F, u_d^F)\rangle$ defines a one-shot simultaneous move filtering game with perfect information between an attacker and a defender. Theorem 1 characterizes the Nash equilibrium of the game $\mathcal{G}^F$.

Theorem 1. *Assume $r \leq \sum {}^{V_i}/_{L_i}$. Define k to be the smallest integer such that $r \leq \sum_{i=1}^{k} {}^{V_i}/_{L_i}$. Define vectors $\boldsymbol{r}^*$, $\boldsymbol{f}^*$ as follows:*

$$(r_i^*, f_i^*) = \begin{cases} \left(\frac{V_i}{L_i}, 1 - \frac{L_k}{L_i}\right) & i < k \ , \\ \left(r - \sum_{j<k} r_j^*, 0\right) & i = k \ , \\ (0, 0) & i > k \ . \end{cases} \tag{4}$$

Then $(\boldsymbol{r}^*, \ \boldsymbol{f}^*)$ forms a Nash equilibrium for $\mathcal{G}^F$. Also if $L_i \neq L_j$ for $i \neq j$, and $\sum_{i=1}^{k} {}^{V_i}/_{L_i} \neq r$, then this Nash equilibrium is unique.

For $r > \sum {}^{V_i}/_{L_i}$, any $\boldsymbol{r}$ such that $r_i \geq {}^{V_i}/_{L_i}$ can constitute an NE. The defender's response in such equilibria is $f_i = 1$ for all i.

Note that at the Nash equilibrium, $u_a^F(\boldsymbol{r}^*, \boldsymbol{f}^*) = rL_k$. Therefore, the efficiency of the attacker is equal to L_k, where k is the strongest node under attack. It is also worth noting that the attacker will only dedicate a maximum of ${}^{V_i}/_{L_i}$ of resources to a node i; since beyond this point, the defender would filter that node completely. Consequently, ${}^{V_i}/_{L_i}$ can be viewed as the capacity, or saturation point, of each node, while $\sum {}^{V_i}/_{L_i}$ is the capacity of the network. When direct losses are not negligible, but still less that the total indirect losses, L_i can be redefined to include direct losses, and the results of Theorem 1 would still hold.

The game presented in this section can be viewed as a probabilistic filtering game. In other words, f_i represents the probability of blocking each unit of node i's outgoing traffic. It is also possible to consider the non-probabilistic, or binary, version of this game, where the defender's action space is $\{0, 1\}^N$. However, such games do not generally have a pure strategy Nash equilibrium. Another interesting observation is that at the NE, no nodes are being completely blocked. In fact, the maximum filtering level is $f_1 = 1 - {}^{L_k}/_{L_1}$. If this maximum is sufficiently small, then our assumption on the linearity of g^F is justified.

While our model does not restrict the type of malicious activities the attacker engages in, it helps to interpret the model in a more specific application context. We will use spam as an example. In this case the "single" attacker more aptly models a single spam campaign orchestrated by certain entity or entities. The attacker's effort translates into attempts toward acquiring bandwidth or processing power from a machine, either by purchasing or hijacking it. The indirect loss inflicted on other machines by an infected machine includes resources spent in processing or acting on spam traffic (e.g., from running the spam filter, storage, reading spams, to possible economic losses when taken in by spams).

3 Investment

In this section, we consider a similar strategic game between the defender and the attacker. However, the defender's action here is to choose a level of protection, effort, or investment in security, for each node, in order to mitigate the attacks. More precisely, the defender can choose to invest an amount $x_i \in [0, 1]$ on node i's security. This investment in turn decreases the effectiveness of the attacker's effort. The defender incurs a cost of $c_i > 0$ per unit of investment. Investing at level $x_i = 1$ is assumed to provide node i with perfect protection. The attacker's utility when the defender invests in security measures is given by:

$$u_a^I(\boldsymbol{r}, \boldsymbol{x}) = \sum_{i=1}^{N} r_i (1 - x_i) \sum_{\substack{j=1 \\ j \neq i}}^{N} L_{ij} (1 - x_j), \text{ s.t. } \sum_{i=1}^{N} r_i \leq r \ . \tag{5}$$

Here, L_{ij} is the loss inflicted on node j per unit of attack on node i, when both are unprotected. The utility of the defender is given by:

$$u_d^I(\boldsymbol{r}, \boldsymbol{x}) = -u_a^I(\boldsymbol{r}, \boldsymbol{x}) - \sum_{i=1}^{N} c_i x_i \ . \tag{6}$$

We refer to the game $\mathcal{G}^I := \langle(\text{attacker, defender}), (\boldsymbol{r}, \boldsymbol{x}), (u_a^I, u_d^I)\rangle$ as the one-shot investment game with perfect information between an attacker and a defender.

In order to choose an optimal action, each player solves the KKT conditions for their respective optimization problem, assuming the other player's action is given. Therefore, at an NE, the following sets of conditions have to be satisfied:

$$\left\{ \begin{aligned} & (1 - x_i) \sum_{j \neq i} L_{ij} (1 - x_j) + \lambda_i - \eta = 0 \ , && \text{(7a)} \\ & \sum_{j \neq i} (r_i L_{ij} + r_j L_{ji})(1 - x_j) - c_i + \mu_i - \nu_i = 0 \ , && \text{(7b)} \\ & \lambda_i r_i = 0, \ \mu_i x_i = 0, \ \nu_i (1 - x_i) = 0 \ , && \text{(7c)} \\ & \eta \left(\sum_{i=1}^{N} r_i - r \right) = 0, \ \sum_{i=1}^{N} r_i \leq r \ , && \text{(7d)} \\ & r_i, \lambda_i, \mu_i, \nu_i, \eta \geq 0, \ 0 \leq x_i \leq 1 \ . && \text{(7e)} \end{aligned} \right.$$

A solution to the above system of equations indicates a Nash equilibrium for a given problem instance. Note that this problem has at least one NE, as the utilities are linear, and the action spaces are convex and compact [11]. To provide intuition on the properties of the equilibria of the game $\mathcal{G}^I$, we next propose a set of conditions on the problem parameters to simplify the KKT conditions in (7a)-(7e). We will then find the Nash equilibrium of the simplified game, and study its properties and dependence on the problem parameters.

Assumption 1. *For all i, and $j \neq i$, the loss L_{ij} can be written as $L_{ij} = \alpha_i \varLambda_j$, where $\varLambda_j$ is a parameter that quantifies the size of the target j, while α_i models the importance of node i.*

Assumption 2. *For all i, the unit cost of security investment c_i is proportional to the size of the node Λ_i, i.e., ${c_i}/{\Lambda_i}$ is a constant, $\forall i$.*

Without loss of generality, assume users are indexed such that α_i is a decreasing sequence, and that $\sum \Lambda_j = 1$. Then from Assumption 2, $\Lambda_i = {c_i}/{\sum c_i}$. We assume the number of nodes is large, so that we can approximate $\sum_{j \neq i} L_{ij} \approx \alpha_i \sum_j \Lambda_j = \alpha_i$, hence $\alpha_i \approx L_i$ from Section 2. Define $A := \sum \Lambda_j (1 - x_j)$, $B := \sum r_j \alpha_j (1 - x_j)$. Similarly for large networks, we can approximate $\sum_{j \neq i} \Lambda_j (1 - x_j) \approx A$ and $\sum_{j \neq i} r_j \alpha_j (1 - x_j) \approx B$, for all i. Using this approximation and Assumptions 1 and 2, we can characterize the Nash equilibrium of $\mathcal{G}^I$.

Theorem 2. *Assume $\alpha_i \neq \alpha_j$ for $i \neq j$. Let $\boldsymbol{r}^*$ and $\boldsymbol{x}^*$ be an equilibrium point for the game $\mathcal{G}^I$, and let $\boldsymbol{\lambda}^*$, $\boldsymbol{\mu}^*$, $\boldsymbol{\nu}^*$, η^* and A^*, B^* be the corresponding parameters. Then there exists some $1 \leq k \leq N$ such that $r_k^* \leq {(c_k - B^* \Lambda_k)}/{A^* \alpha_k}$, and,*

$$(r_i^*, x_i^*) = \begin{cases} \left({(c_i - B^* \Lambda_i)}/{A^* \alpha_i}, 1 - {\eta^*}/{A^* \alpha_i}\right) & i < k \ , \\ (0, 0) & i > k \ . \end{cases}$$

If $r_k^* < {(c_k - B^* \Lambda_k)}/{A^* \alpha_k}$ then $x_k^* = 0$, and if $r_k^* = {(c_k - B^* \Lambda_k)}/{A^* \alpha_k}$, then any $0 \leq x_k^* \leq 1 - {\alpha_{k+1}}/{\alpha_k}$ constitutes an NE.

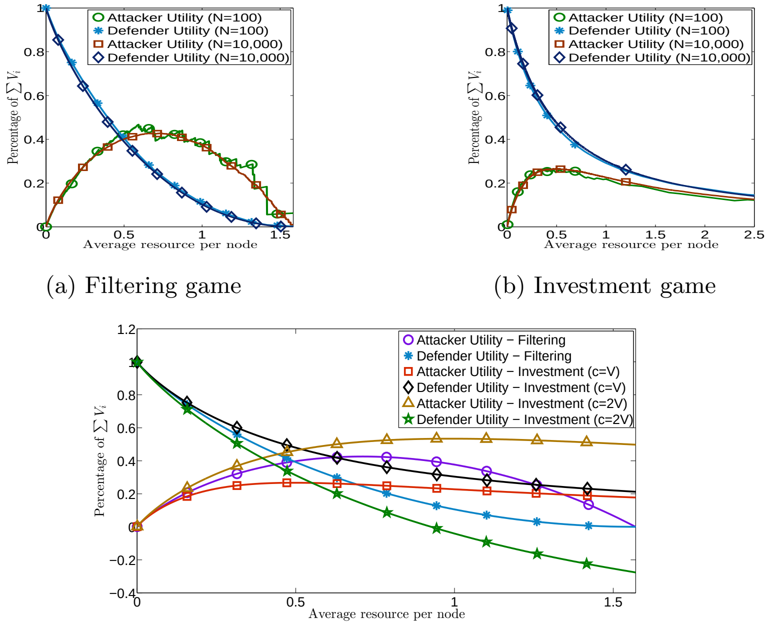
Consider an instance of $\mathcal{G}^I$ where $r_k^* = {(c_k - B^* \Lambda_k)}/{A^* \alpha_k}$ and $x_k^* = 0$. Using Theorem 2, we can find the equilibrium point for such a case, where k nodes have been completely saturated by the attacker, and the defender chooses not to secure the k^{th} node. We can represent this equilibrium point as a function of k. Let $\boldsymbol{r}(k)$ and $\boldsymbol{x}(k)$ denote the NE, and $r^I(k) = \sum r_j(k)$ be the corresponding parameter. Defining $D(k) := \sum_{j=1}^{k} c_j \frac{\alpha_k}{\alpha_j}$ and $E(k) := \sum_{j=k+1}^{N} c_j$, we have:

$$\begin{cases} r^I(k) = \displaystyle\sum_{j=1}^{k} r_j(k) = \frac{1}{\alpha_k} \frac{D(k)}{2D(k) + E(k)} \sum_{j=1}^{N} c_j \ , & \text{(8a)} \\ u_a^I(k) := u_a^I(\boldsymbol{r}(k), \boldsymbol{x}(k)) = D(k) \dfrac{D(k) + E(k)}{2D(k) + E(k)} \ , & \text{(8b)} \\ u_d^I(k) := u_d^I(\boldsymbol{r}(k), \boldsymbol{x}(k)) = \dfrac{D^2(k)}{2D(k) + E(k)} - \displaystyle\sum_{j=1}^{k} c_j \ . & \text{(8c)} \end{cases}$$

4 Numerical Results

To illustrate the results of Section 2, we generate a network of N nodes by drawing V_i and L_i independently from a Rayleigh distribution, and plot the utilities of both parties at the NE, as a function of the attacker's resources.

As a reference point, in all the following simulations, we set $\mathbb{E}[V_i] = 1$. Also the scaling of L_i does not have an effect on the overall shape of the curve; it only affects the maximum capacity of the network. Therefore, we will let $\mathbb{E}[L_i] = 1$ throughout. Moreover, we will present the utilities of both parties as



(a) Filtering game

(b) Investment game

(c) Comparison of NEs of $\mathcal{G}^F$ and $\mathcal{G}^I$

Fig. 1: Defender and attacker utilities at the NEs of $\mathcal{G}^F$ and $\mathcal{G}^I$

a percentage of the total value of the network, i.e., $\sum V_i$. The reason for this choice of scaling is that in the absence of any attack (i.e., $r = 0$), the defender obtains the entire value of the network at equilibrium. Therefore, the vertical axis depicts the fraction of network value lost as a result of the attacks. Finally, in order to obtain a better comparison among networks of different size, we scale the horizontal axis by the number of nodes N, resulting in plots that illustrate utilities as a function of the average attack resource per node.

Figure 1a plots the attacker and defender utilities under the filtering game $\mathcal{G}^F$, for two networks of size $N = 100$ and $N = 10,000$. An important aspect to this plot is that the utility of the attacker, $u_a(\boldsymbol{r}^*, \boldsymbol{f}^*)$ is not necessarily an increasing function of the total attack power r. In other words, the most successful attacker is not necessarily the one with the highest attack power. This observation can be intuitively explained as follows: assume an attacker with high r decides to spend only a smaller amount $r' < r$ of his attack resources. If the defender's response is such that the NE corresponding to r' is realized, a smaller number of nodes would be filtered, and both parties would receive a higher utility. Nevertheless, the attacker's action will no longer be a best response to the defender's strategy in this scenario, as the attacker has access to additional resources to further attack the unfiltered nodes. As the availability of these resources is common knowledge, and thus known to the defender, she will not under-filter the system against more powerful attackers. This increased filtering of nodes against a more powerful attacker in turn limits the attacker's ability to profit from the network, and ultimately, reduces his utility.

It is also interesting to note the sudden drops in the attacker's utility, which are more easily observable for $N = 100$. These drops correspond to points where the attacker's total power is such that exactly k nodes have been completely saturated ($r_i^* = {}^{V_i}/_{L_i},\ 1 \leq i \leq k$), following which an attacker with more attack power would start putting his resources into the $k + 1^{\text{th}}$ node. As a result, the defender's filtering becomes more aggressive by limiting nodes under attack to an effective loss of L_{k+1} (i.e., $L_i(1 - f_i^*) = L_{k+1},\ 1 \leq i \leq k$), hence the drop in the attacker's utility. The defender's utility, however, is always decreasing in r.

Figure 1b illustrates the utilities at the NE of $\mathcal{G}^I$ as a function of ${}^{r}/_{N}$, by plotting $r^I(k)$, $u_a^I(k)$ and $u_d^I(k)$ from (8a)-(8c). The parameters of the game are generated similar to the filtering game simulations, i.e., α_i (which is parallel to L_i in $\mathcal{G}^F$) and c_i are drawn from a Rayleigh distribution with unit mean.

One important aspect of the investment game is that the x-axis extends further than the filtering game. In other words, the capacity of the network is larger in comparison to $\mathcal{G}^F$. An intuitive explanation for this phenomenon is the presence of internalities when nodes protect themselves via investment. When a node is blacklisted, the rest of the network is protected against attacks targeting that node, but this action does not protect the node itself. Therefore, filtering is an action that has externality, but not internality. This is not the case for investment, since investing in security protects oneself, as well as the rest of the network. When the attacker is powerful, a large portion of the network is investing in security, and the defender is well-protected by internalities. Thus the capacity of each node is relatively large.

To conclude this section, we look at the utilities of both parties under the investment and filtering games in Figure 1c. To this end, we set $N = 10,000$, and compare the two games under two different security cost vectors $c = V$ or $c = 2V$, the latter indicating relatively costly security measures.

First, we note that as expected, investing in network security is preferred by the defender when the cost of it is sufficiently low. The more surprising result is however in the trend of the attacker's utility under the different protection models. We see that with low attack power, both filtering and security yield similar utility to the attacker, as no considerable filtering or protection has yet been introduced by the defender. As the attack power grows, the attacker who is facing filtering gains a higher utility. Intuitively, this is also a consequence of the internality of investment actions. To further illustrate, note that an unfiltered attack on node i yields a payoff of $\sum_{j \neq i} L_{ij}$ per unit of effort. In contrast, an attack on an unprotected node i yields a payoff of $\sum_{j \neq i} L_{ij}(1 - x_j)$ per unit of effort. Lastly, for very powerful attackers, the attacker facing investment is more successful. This is due to the fact when the defender chooses to filter nodes, the network gets increasingly close to being fully saturated under high attack power. However, under investment, it takes more resources for the the attacker to saturate all nodes, leaving him more room to gain profit.

5 Conclusion

In this paper, we compare the efficacy of two security options, namely inbound/outbound filtering based on RBLs and investing in self-protection methods, by a defender controlling a set of nodes facing a resource constrained strategic attacker. Specifically, our models take into account the indirect losses inflicted on neighboring nodes by a compromised node, loss of value due to the inevitable filtering of parts of the legitimate traffic, and the higher cost of self-protection as compared to filtering. Our analysis and simulation results show that the defender chooses to invest in security measures over filtering only when the cost of investing is sufficiently low. On the other hand, the attacker's potential to benefit in the face of each protection method is determined by his total attack power. Highly resourceful attackers are less successful when facing filtering actions rather than investment actions.

The current work can be continued in several directions. It would be interesting to study filtering and investment actions under less restrictive conditions, e.g. nonlinear risk functions with respect to the filtering policy, and taking diminishing returns into account when considering attacks originating from multiple sources (the latter can be modeled by setting the attacker's profit to a concave function of the sum in Equation (1)). The same game form can also be analyzed a dynamic framework, as both attacker and defender actions can be affected by the history of past events, including previous attack patterns, node takedowns, and the amount of time a node stays blacklisted. Modeling information asymmetries among the players, and strategic interactions among multiple non-cooperative attackers and/or defenders, are other possible extensions of the current model.

References

1. Spam data and trends, Q1 2009
2. Spam volumes drop by two-thirds after firm goes offline (2008), http://voices.washingtonpost.com/securityfix/2008/11/spam_volumes_drop_by_23_after.html
3. Fultz, N., Grossklags, J.: Blue versus red: Towards a model of distributed security attacks. In: Financial Cryptography and Data Security. Springer (2009)
4. Hausken, K.: Income, interdependence, and substitution effects affecting incentives for security investment. Journal of Accounting and Public Policy 25(6) (2006)
5. Hausken, K.: Strategic defense and attack of complex networks. International Journal of Performability Engineering 5(1) (2009)
6. Hausken, K.: Strategic defense and attack of series systems when agents move sequentially. IIE Transactions 43(7) (2011)
7. Inc., C.S.: SpamCop Blocking List - SCBL (May 2011), http://www.spamcop.net/
8. Laszka, A., Felegyhazi, M., Buttyán, L.: A survey of interdependent security games. CrySyS 2 (2012)
9. Networks, B.: Barracuda Reputation Blocklist (May 2011), http://www.barracudacentral.org/
10. Nochenson, A., Heimann, C.L.: Simulation and game-theoretic analysis of an attacker-defender game. In: Decision and Game Theory for Security. Springer (2012)
11. Osborne, M.J., Rubinstein, A.: A course in game theory. MIT press (1994)