Managing the Adoption of Asymmetric Bidirectional Firewalls: Seeding and Mandating

MHR. Khouzani
The Ohio State University, ECE
khouzani@ece.osu.edu

Soumya Sen
Princeton University, EE
soumyas@princeton.edu

Ness B. Shroff
The Ohio State University, ECE & CS
shroff@ece.osu.edu

***Abstract*—The security of the Internet can be significantly improved if Internet Service Providers adopt firewalls to monitor traffic entering and leaving access networks. But this process suffers due to 'free-riding', and hence, regulatory requirements and 'seeding' strategies are required to influence the adoption process. In this paper, we analytically derive the equilibrium adoption levels and relate them to the initial seeding and mandating condition, and explore the issues of incentive alignment across users, firewall developers, and regulators. We define different notions of optimality and analytically develop optimum seeding and mandating policies.**

I. Introduction

It is increasingly recognized that ISPs can play a significant role in mitigating the security threats in the Internet [1]–[6]. However, ISPs often have the incentives not to adopt proper security measures (e.g., firewalls) due to 'free-riding' on other's investments. Hence, there may be regulatory needs for ensuring security adoption. Regulators may also influence the adoption of security measures by providing subsidies through initial seeding of the market. In this work, we study the role of *seeding* and *mandating* in the adoption of *asymmetric, bidirectional* firewalls by ISPs. The current work provides insights into the adoption process and both policy guidelines and marketing of future firewalls, specifically:

- Characterizes the equilibrium adoption levels as close-form expressions, given the initial seeding profile and the regulatory constraints.
- Studies how mandates or seeding affect the equilibrium outcome and the implications for policy making.
- Explores the incentive alignment problems among ISPs, firewall developers, and regulators, and investigating the efficacy of mandating on the network security.
- Defines different notions of optimality and accordingly computes the optimum seeding and mandating policies.

Related Literature: The issue of adoption of technologies that exhibit network externalities has been studied in a variety of contexts, such as vaccination games [7], security games [2], [8], competition of network services [9], etc. These scenarios all share the characteristic that the population-level information about the system drives the (myopic) decisions made by individual agents. For example, previous works have used game-theoretic and epidemiological models [7], [10] to study the feasibility of universal vaccination when parents myopically decide on vaccinating their children given the relative risks and benefits of doing so.

These results demonstrate the harmful effect of 'positive externality' that leads to the 'free-riding' phenomenon and lower the network-wide security. But in addition to quantifying the equilibrium outcomes, researchers need to focus on means of managing the adoption of security measure in the face of such positive externalities. This work takes on this task by studying the role of seeding the market with free sample and mandating regulations as two policies in order to influence the diffusion process and the equilibrium outcomes in the context of bidirectional asymmetric security measures.

In particular, we focus on (a) the issues of incentive alignment among firewall vendors, regulators, and adopting ISPs, (b) investigate the limits and advantages of mandating and seeding in improving network security, and (c) define notions of optimality for the regulator and derive optimal seeding and mandating policies.

II. System Model

An ISP provides a gateway that connects its subnet to the Internet. A firewall software, installed at this gateway, can monitor the incoming (ingress) and outgoing (egress) traffic for detecting malicious activities and blocking security breach attempts. Intrusion prevention is achieved through different techniques, mainly classified as the following [11]: a) signature-based, b) statistical anomaly-based, and c) stateful protocol analysis. Adoption of such security measures has a stand-alone benefit for an ISP. Moreover, it can slow down the rate of attacks and provides positive externality to the rest of the ISPs (adopters and non-adopters) by improving the overall security in the network. Namely, the nodes in the subnets of other ISPs will be less likely to be targeted by an attack originating from the subnet of the protected ISPs. However, adoption of a firewall is not without cost: there can be a one-time purchase and installation fee. Moreover, a firewall may incur recurrent usage costs: it needs to be routinely maintained and updated; it can slow down the connection through latencies introduced by traffic monitoring; a firewall has *false positives*, that is, it occasionally blocks legitimate traffic. In what follows we provide a practical model that captures key attributes of adoption dynamics of firewalls. Our aim in this paper is to develop a *qualitative analysis* of the adoption process and the policies that can influence it. Hence, we make some technical assumptions along the way to keep the model analytically tractable.

We consider a continuous-time model with N interconnected ISP networks. Once an ISP purchases the firewall, it can un-adopt it by simply disabling the firewall. Subsequent adoptions are performed by enabling the firewall, and in

TABLE I
MAIN NOTATIONS IN THE MODEL

parameter	definition
N	number of ISPs
$x(t)$	fraction of the ISPs at time t that have obtained and enabled the firewall
$y(t)$	fraction of the ISPs at time t that are yet to obtain
γ	rate at which each ISP updates its adoption decision
$G_0(x)$	expected utility of a non-adopter ISP
$G_1(x)$	expected utility of an ISP that purchases and enables the firewall (hence includes the purchase fee)
$G_2(x)$	expected utility of an ISP that just enables its firewall
ζ	the value of x at the intersection of $G_0(x)$ and $G_1(x)$
ζ'	the value of x at the intersection of $G_0(x)$ and $G_2(x)$
Λ	rate of intrusion attempts on an ISP in the absence of any firewalls
μ	rate at which a successful intrusion to a subnet is detected and blocked
C_0	one-time purchase fee of the firewall
c	per unit time usage cost of the firewall
K_0	instantaneous cost upon a successful intrusion
k	cost (loss/damage) per unit time of intrusion
r	discount factor of the ISPs

particular, do not entail paying the firewall's one-time purchase fee. That is, for an ISP that has obtained the firewall, cost of a subsequent adoption only includes the firewall's recurrent usage cost. Hence, we need a model that distinguishes between the first adoption and subsequent re-adoptions. To do this, we introduce three different types of ISPs: (1) ISPs that have *obtained* and *enabled* the firewall; (2) ISPs that have *not obtained* it; and (3) ISPs that have *obtained* the firewall but have *disabled* it. Note that *obtaining* can be through purchasing it, or as we discuss in this paper, by being *seeded* for free. We will denote the *fraction* of ISPs of each type at time t by $x(t)$, $y(t)$ and $1-x(t)-y(t)$, respectively. The adoption state of the network at time t is represented by pair $(y(t),x(t))$. Table I contains all of the important notations that are used throughout the paper.

We assume that each ISP independently re-evaluates the rate of intrusion attempts on its subnet and accordingly updates its decision regarding the adoption of the firewall. These re-evaluations occur at independent random epochs that are according to i.i.d. Poisson processes with rate γ. We assume that the decisions of each ISP is its best response to the *current* measure of the intrusion rates, that is, assuming the current measure is not going to change.

Adoption decisions of an ISP are determined by comparing the expected utilities given each decision. Accordingly, we define three utilities: $G_0(x)$, $G_1(x)$ and $G_2(x)$: Given the current level of adoption x, $G_0(x)$ is the expected utility of an ISP that does not have the firewall and decides to stay unadopted; $G_1(x)$ is the expected utility of an ISP that does not have the firewall and decides to *purchase* and enable it; and finally, $G_2(x)$ is the expected utility of an ISP that already has the firewall and decides to enable it. Note that $G_1(x)$ and $G_2(x)$ differ only in the purchase fee of the firewall. Specifically, $G_2(x) = G_1(x) + C_0$, where C_0 denotes the (one-time) purchase fee of the firewall (also included in Table I).

Let c be the cost per unit time of using the firewall incurred by an adopter ISP due to maintenance, communication latencies, false positives, etc. For simplicity of exposition, we consider security breaches that do not propagate in the network. For instance, we will *not* consider attacks involving self-replicating malicious codes (known as *worms*) in this article. Hacking is a typical example of a non-replicating type of attack. We will refer to such attacks by the umbrella term of *intrusion* attempts. When a host in a subnet of an ISP is compromised, the ISP incurs an instantaneous cost of K_0 and a per unit cost of k that persist as long as the host is infiltrated by that specific hacker. We assume that the costs of different intrusions add up, that is, two concurrent intrusions from two independent intruders incur the ISP $2k$ costs per unit time, etc. The instantaneous cost may reflect the losses due to exposure of private information such as credentials (fingerprints, voice recognition, passwords, etc.), credit card information, or manipulation of data. On the other hand, the per unit time cost can represent the accumulation of eavesdropped data such as keystroke logs, accessing the network at the cost of the victim, slowdown of the victim's machine or the ISP's service, etc. The time it takes to remove an infection is according to an exponential random variable with rate μ. We assume that the machines are again susceptible to future attacks, since future attacks are likely to exploit new techniques.

Without prior knowledge, new security breaches can originate from the subnet of any of the ISPs. We assume that ISPs are homogeneous, that is, they assign the same parameters for costs and have similar subnet sizes, furthermore, that a target of an intrusion is chosen uniformly randomly from the space of IP addresses. These along with the assumption of homogeneous sizes of the subnets, imply that the target is equally likely to belong to the subnet of any of the ISPs.

The success probability of an intrusion attempt depends in part on the status of the ISPs of the attacker as well as the ISP of the target with regard to the adoption of the firewall.[1] Specifically, the highest chance of intrusion success is when *neither* of the ISPs have an enabled firewall, while the lowest likelihood is when *both* ISPs have (obtained and) enabled it. Based on the four different conditions for the adoption status of the ISPs of an attacker and its target, we define intrusion *success probabilities* π_0, π_1, Π_0 and Π_1 according to Table II. Namely, π_1 is the success probability of intrusion if both ISPs have enabled firewalls in place, Π_1 is the success probability of an intrusion if only the target's ISP has adopted the firewall, and so forth.

Without loss of generality, we let $\Pi_0 = 1$ and only consider the attempts that are successful in the absence of any firewall. However, we continue to use the *notation* Π_0 in our formulation for presentation purposes. In general, the following

[1]Note that intermediate routers do not monitor for threats and the only traffic monitoring for threats are at border (edge) ISPs.

TABLE II
SUCCESS PROBABILITIES OF AN INTRUSION ATTEMPT

		Host's ISP	
		Protected	Not Protected
Attacker's ISP	Protected	π_1	Π_1
	Not Protected	π_0	Π_0

ordering holds for the intrusion success probabilities:

$$0 \leq \pi_1 \leq \pi_0 \leq \Pi_1 \leq \Pi_0 \leq 1.$$

That Π_0 is the largest of the group is obvious, as it is the probability of success of an intrusion if no firewall is set up on both ISPs of the attacker and target (hence the most exposed scenario). $\pi_0 \leq \Pi_1$, since the primary goal of the firewall is to protect the subnet against the incoming threats and hence, a marketable firewall provides no less protection against the incoming threats than against the outgoing threats. $\pi_1 \leq \pi_0$, as π_1 is the success probability of an intrusion that has to bypass *both* firewalls of its own subnet's ISP and that of the victim's, while π_0 is the success probability of an intrusion that only has to bypass the firewall of the victim's ISP.

For a firewall whose mechanism of intrusion prevention is only signature-based, if both firewalls have access to the same signature database then $\pi_1 = \pi_0$, that is, if an intrusion can successfully bypass one of the firewalls, it will be able to bypass the other one as well. We will refer to this case as the *mutually inclusive* scenario. However, it could be that one of the firewalls is more up-to-date than the other, hence it is likely that $\pi_1 < \pi_0$. Also, anomaly detection mechanisms are in essence probabilistic and they have a *false negative* chance. The past traffic history of the two ISPs differ, hence the blocking events of the two firewalls may not be exactly mutually inclusive. In case the intrusion prevention outcomes of the firewalls are mutually independent, for $\Pi_0 = 1$, we have $\pi_1 = \pi_0\Pi_1$. Hence, we also have the following structural inequality:

$$0 \leq \pi_0\Pi_1 \leq \pi_1. \tag{1}$$

New mechanisms are proposed in which the firewalls in different ISPs "co-operate" to improve their detection and blocking chances (e.g. [12]). In such cases, it is (theoretically) possible for π_1 to be less than $\pi_0\Pi_1$. We, however, do not consider such cases in the current article.

Let Λ represent the rate of intrusion attempts on an ISP in the absence of any firewall in the network. Following the definition of the intrusion success probabilities, the rate of *successful* intrusion attempts on an ISP that *does not have* an enabled firewall is therefore $\Lambda(x\Pi_1 + (1-x)\Pi_0)$. This is because x fraction of the intrusion attempts have to successfully bypass the firewall of their own ISP to reach the decision-taking ISP, hence their success probability is Π_1 (recall that x is the fraction of the ISPs in the network that have adopted and enabled the firewall). The rest of the intrusion attempts, i.e. $(1-x)$ fraction of them, are confronted with no firewall and hence, their success probability is Π_0. Similarly, the rate of *successful* intrusion attempts on an ISP that *has* an enabled firewall is $\Lambda(x\pi_1 + (1-x)\pi_0)$. These are the two rates that each ISP can readily measure, then calculate its conditional utilities and accordingly make an adoption decision.

The utility of an ISP is a decreasing function of costs and losses due to potential future intrusions to its subnet. For ease of calculations, we assume *risk-neutral* ISPs. Hence, we can directly assume the negative of the costs to be the ISPs' utility. Let σ represent the state of the decision-taking ISP with respect to the intrusion, specifically, $\sigma \in \{0,1,2,\ldots\}$ indicates the number of ongoing intrusions in an ISP's subnet at the time of the ISP's decision-taking. Without loss of generality (by shifting the time coordinate) we can take a decision taking epoch to be at $t = 0$. We also consider discount factor r in calculation of the utilities, that is, costs incurred at time t in future are discounted at e^{-rt} when evaluated at present time. A larger r designate more shortsighted ISPs. A successful intrusion that occurs at time $t = 0$ incurs the following expected cost on the network:

$$\chi = K_0 + \int_0^\infty e^{-\mu t}\mu\, dt \times \left(\int_0^t e^{-r\tau} k\, d\tau\right) = K_0 + \frac{k}{\mu + r}$$

G_0 is composed of two parts: the cost due to ongoing intrusions, which we will call G_N^o, and the cost associated with future intrusions, which we represent by G_N^f. Hence, $G_0 = G_N^o + G_N^f$. Following Wald's equation, G_N^o is $-\mathbb{E}\sigma \times \chi$, since $\mathbb{E}\sigma$ is the expected number of ongoing intrusions, and χ is the expected cost of each of them. Recall the negative sign is to convert the cost to reward.

Let $\eta := \Lambda(x\Pi_1 + (1-x)\Pi_0)$, i.e., the rate of successful intrusion attempts on a non-adopter ISP. The expected cost of the future intrusions, G_N^f, can be computed by conditioning on the first epoch at which a new successful intrusion attempt is made:

$$\begin{aligned} G_N^f(x) &= -\int_0^\infty e^{-\eta t}\eta\, dt \times \left(e^{-rt}\left(\chi + G^f\right)\right) \\ &= -G_N^f\frac{\eta}{\eta + r} - \chi\frac{\eta}{\eta + r} \Rightarrow G_N^f = -\frac{\chi}{r} \end{aligned}$$

The number of ongoing intrusions, σ, can be seen as the number of "requests" in an $M/M/\infty$ system with arrival (birth) rate η and service (death) rate μ. Hence $\mathbb{E}\sigma$ is simply η/μ. Combining all of the above computations leads to the following expression for G_0:

$$\begin{aligned} G_0(x) = G_N^o(x) + G_N^f(x) &= -\chi\frac{\eta}{\mu} - \chi\frac{\eta}{r} = -\chi\eta(\mu^{-1} + r^{-1}) \\ &= -\frac{\Lambda}{\mu r}\left(K_0(\mu + r) + k\right)\left(\Pi_0 - x(\Pi_0 - \Pi_1)\right) \end{aligned} \tag{2}$$

Calculation of the $G_1(x)$ is now straightforward: there are two components, first one is related to the cost of adoption, that is simply $C_0 + c/r$, and the other component comes from the cost of intrusions. The calculation of the latter component follows the same steps as in $G_0(x)$ except for accordingly changing the success probabilities of intrusion attempts. Hence, we can directly deduce:

$$G_1(x) = -C_0 - \frac{c}{r} - \frac{\Lambda}{\mu r}\left(K_0(\mu + r) + k\right)\left(\pi_0 - x(\pi_0 - \pi_1)\right) \tag{3}$$

Note that G_0 and G_1 turn out to be linear in x. Also, recall that $G_2(x) = G_1(x) + C_0$. A straightforward yet important property of the expected utilities is that all are increasing in the level of adoption:

Lemma II.1. *For any $x \in [0,1]$ we have: $\partial G_0(x)/\partial x$, $\partial G_1(x)/\partial x$, $\partial G_2(x)/\partial x \geq 0$. The equality holds only if $\Pi_1 = \Pi_0$.*

Hence, *positive externalities exist for adopters and non-adopters alike*. Moreover, the positive externalities vanish *only* when there is no protection against the outgoing threats.

We continue with a lemma that we will use later. In simple words, the lemma shows that *even though both adopter and non-adopters experience positive externalities of the firewall adopted by others, the non-adopters benefit more.* We will see the implication of this lemma on the seeding and mandating policies.

Lemma II.2. *For non-cooperating firewalls (hence inequality (1)) and when $\Pi_1 < \Pi_0$, $D(x) := \dfrac{d}{dx}(G_1(x) - G_0(x)) < 0$ at any point $x \in [0,1]$.*

Note that this lemma also implies $\dfrac{d}{dx}(G_2(x) - G_0(x)) < 0$ for any $x \in [0,1]$, since $G_2(x) = G_1(x) + C_0$. Also, for the case of no protection against outgoing threats, we have $\Pi_1 = \Pi_0$, which in turn implies $\pi_1 = \pi_0$. This leads to $D(x) \equiv 0$. The proof of this lemma follows.

Proof: Referring to (2) and (3), we expand $D(x)$:

$$D(x) = \frac{d}{dx}(G_1(x) - G_0(x)) = L\left[(\pi_0 - \pi_1) - (\Pi_1 - \Pi_0)\right]$$

Recall that without loss of generality, we can let $\Pi_0 = 1$ and hence, we have inequality (1), that is, $0 \leq \pi_0 \Pi_1 \leq \pi_1$. Hence:

$$\left[(\pi_0 - \pi_1) - (\Pi_0 - \Pi_1)\right] \leq -(1-\pi_0)(1-\Pi_1) \leq 0$$

■

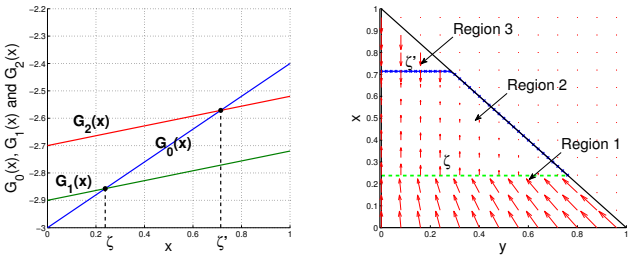
The above proof also reveals that $D(x) = 0$ holds if and only if $\Pi_1 = \Pi_0$, i.e., no protection against outgoing threats.

Because of the linearity of G_0 and G_1 in x, $G_1(x)$ can cross $G_0(x)$ at most once in the interval of $(0,1)$.[2] Similar comment applies to $G_2(x)$ and $G_0(x)$. Lemmas II.1 and II.2 further elucidate the nature of these potential two intersection points. For $C_0 > 0$, we have $G_1(x) < G_2(x)$ for all $x \in [0,1]$. Hence, in the most general case, as is illustrated by Fig. 1(a), we can introduce ζ and ζ' as the following:

$$\exists \zeta \ \& \ \zeta', \quad 0 < \zeta < \zeta' < 1, \ \text{s.t.}:$$
$$\begin{cases} G_0(x) < G_1(x) < G_2(x) \ \text{for } x \in [0,\zeta) & : \textit{Region } 3 \\ G_1(x) < G_0(x) < G_2(x) \ \text{for } x \in (\zeta,\zeta') & : \textit{Region } 2 \\ G_1(x) < G_2(x) < G_0(x) \ \text{for } x \in (\zeta',1] & : \textit{Region } 1 \end{cases}$$

We now have the necessary tools and terminologies to investigate our main objective of the paper, effects of seeding and mandating on the adoption process.

[2]Note that for nontrivial problems, that is if the firewall has some cost or if there is any protection, $G_0(x)$ and $G_1(x)$ cannot coincide everywhere on $[0,1]$.



(a) (b)

Fig. 1. A sample phase portrait for the adoption of firewalls. Region 1 is designated by $x < \zeta$, $x + y \leq 1$, where the ODE turns to $\dot{y}(t) = -\gamma y(t)$, $\dot{x}(t) = \gamma y(t) + \gamma(1 - x(t) - y(t))$; Region 2 is where $\zeta < x < \zeta'$, $x + y \leq 1$ and the ODE is $\dot{y}(t) = 0$, $\dot{x}(t) = \gamma(1 - x(t) - y(t))$; Region 3 is the region $\zeta' < x$, $x + y \leq 1$, where the ODE is transformed to $\dot{y}(t) = 0$, $\dot{x}(t) = -\gamma x(t)$. All the points in the set of (ζ', y) for $0 \leq y \leq 1 - \zeta'$ and $(x, 1-x)$ for $\zeta < x \leq \zeta'$ compose the equilibria of the system.

III. Seeding of the Firewalls

The regulator of the network (e.g. the government) or the developer of the firewall may choose to distribute free samples of the firewalls to a fraction of the ISPs in order to influence the equilibrium level of the adoption. This practice is known as *seeding*. In our context, seeding may be done for a variety of reasons: regulator might be interested in the social welfare or just in the security of the network, while the developer of the software seeks to increase its overall sale. The first step to investigate the effects of seeding is to relate the equilibrium of the adoption process to the value of initial seeding. The following theorem serves this purpose.

Theorem III.1. *Let the initial fraction of the ISPs that receive a free copy of the firewall be S. In the equilibrium we have (Fig. 1(b)):*

$$y^* = (1-S)(1-\zeta), \ x^* = \begin{cases} \zeta' & \frac{\zeta'-\zeta}{1-\zeta} \leq S \leq 1 \\ S + \zeta - \zeta S & S < \frac{\zeta'-\zeta}{1-\zeta} \end{cases}$$

Proof: With initial seeding S, the initial condition of the adoption process is $(y_0, x_0) = (1-S, 0)$. As long as $x(t) < \zeta$, the differential equation is in Region 1, i.e., $G_0(x) < G_1(x) \leq G_2(x)$. That is, ISPs that do not have the the firewall progressively purchase and activate it, and the ISPs that have the firewall (i.e., belong to the initial seeding) enable it. The corresponding differential equation is:

$$\begin{aligned} \dot{y} &= -\gamma y & y_0 &= 1 - S \\ \dot{x} &= \gamma y + \gamma(1 - x - y) = \gamma(1 - x) & x_0 &= 0 \end{aligned}$$

These are two first order linear ordinary differential equations, whose solutions are:

$$y(t) = (1-S)e^{-\gamma t}, \qquad x(t) = 1 - e^{-\gamma t}$$

The above solution is valid until time $t = T > 0$ at which $x(t) = \zeta$, i.e., the time at which $G_0(x(t)) = G_1(x(t))$. Since $\dot{x} > 0$ in both regions 1 and 2, the adoption process enters Region 2, where $G_0(x) < G_1(x)$. Hence, the ISPs that have not obtained the firewall will have no incentive to do so.

However, the remaining ISPs in the group of seeded ones that have not enabled their firewall continue to enable it, since still $G_2(x) > G_1(x)$. As we said, the fraction of the ISPs that have not obtained the firewall will not change, hence $y(t)$ will continue to retain its value at $t = T$, that is:

$$x(T) = 1 - e^{-\gamma T} = \zeta \Rightarrow e^{-\gamma T} = 1 - \zeta \Rightarrow y(T) = (1-S)(1-\zeta)$$

Hence, the dynamics of the adoption for $t \geq T$ changes to:

$$\begin{aligned} \dot{y} &= 0 & y(T) &= (1-S)(1-\zeta) \\ \dot{x} &= -\gamma(1-x-y) & x(T) &= \zeta \end{aligned}$$

Adoption of the firewalls continues according to the above equations until time $t = T'$ at which either (a) $(1 - x(T') - y(T')) = 0$, or (b) $x(T') = \zeta'$, whichever happens first. Note that in either one of these cases, $(y(T'), x(T'))$ is the equilibrium point (y^*, x^*) of the adoption. In case (a), this directly follows from the fact that at $t = T'$, we have $\dot{y} = \dot{x} = 0$. In case (b), we have $G_1(x(T')) = G_2(x(T'))$, hence the ISPs that have the firewall are indifferent between enabling and disabling their firewall, and hence any randomization between enabling and disabling is a valid decision. Therefore, $\dot{x}$ changes to $-\gamma(1-x-y)\theta + -\gamma x\theta'$ where $\theta, \theta' \in [0,1]$. Note that in Region 3, we have $\dot{x} = -\gamma x < 0$. Hence, $x = \zeta'$ is an equilibrium point. Since $\dot{y} = 0$ for $t \in (T, T']$, we have $y(T') = y(T) = (1-S)(1-\zeta)$. Now, if $y(T') + \zeta' < 1$, then (b) occurs before (a) and $x(T') = \zeta'$ is the equilibrium. If $y(T') + \zeta' > 1$, then (a) occurs first and $x(T') = 1 - y(T') = 1 - (1-S)(1-\zeta) = S + \zeta - \zeta S$. Note that the condition $y(T') + \zeta' = 1$ is equivalent to $(1-S)(1-\zeta) + \zeta' = 1$, or $S = 1 - \frac{1-\zeta'}{1-\zeta} = \frac{\zeta'-\zeta}{1-\zeta}$. This completes the proof. ■

Next, we discuss the corollaries of the Theorem III.1.

A. Can seeding benefit the developer of the firewall

At any t, the fraction of ISPs that have the firewall is $1 - y(t)$. From Theorem III.1, for any initial seeding S, $y^* = (1-S)(1-\zeta)$. Hence, $1-(1-S)(1-\zeta) = S + \zeta - \zeta S$ of the ISPs will have the firewall. However, S fraction of the ISPs were seeded, i.e. had received the firewall for free. Hence, the fraction of ISPs that *purchase* the firewall is $\zeta - \zeta S$, which is strictly decreasing in S. Therefore, *the optimum seeding for the firewall developer is $S = 0$, i.e, no seeding.*

B. Can seeding benefit the regulator

To answer this, we need to identify the objective of the regulator. In the following, we define two different objectives and subsequently investigate the effect of seeding on each.

1) Social Welfare: is defined as the average expected utility of all of the ISPs. Let us denote the social welfare utility by U. To compute U, in the light of Theorem III.1, we differentiate between the following two cases:

(a) $S < \dfrac{\zeta' - \zeta}{1-\zeta}$: In this case, all of the seeded ISPs enable their firewalls. Hence, $x^* - S$ fraction of the ISPs purchase the firewall and enable it, and y^* fraction of the ISPs never buy the firewall. For a fair comparison, we should subtract the cost of providing the seeds. This cost can be assumed to be shared among all of the ISPs, e.g. through taxes imposed by the regulator. Therefore, the average expected utility of the network is:

$$\begin{aligned} U &= (x^* - S)G_1(x^*) + SG_2(x^*) + y^*G_0(x^*) - C_0S \\ &= (x^* - S)G_1(x^*) + S(G_1(x^*) + C_0) + y^*G_0(x^*) - C_0S \\ &= x^*G_1(x^*) + y^*G_0(x^*) = x^*G_1(x^*) + (1-x^*)G_0(x^*) \end{aligned} \quad (4)$$

The last equality follows since from the proof of Theorem III.1 for case (a), we have $y^* = 1 - x^*$. Now, In order to find the best seeding from that maximizes the social welfare, we compute $\dfrac{dU}{dS}$. From (4), $\dfrac{dU}{dS}$ is equal to:

$$\frac{dx^*}{dS}\left(G_1(x^*) + x^*\frac{\partial G_1(x^*)}{\partial x^*} - G_0(x^*) + (1-x^*)\frac{\partial G_0(x^*)}{\partial x^*}\right)$$

In case (a), $\dfrac{dx^*}{dS} = (1-\zeta) > 0$. Hence, the roots and sign of the derivative is the same as the roots and sign of the expression inside parentheses. Replacing from (2) and (3) and some simplifications turns the expression inside parentheses to: $\frac{\Lambda}{\mu}(K_0(\mu + r) + k)(\Pi_0 - \pi_1) - C_0x^*$. The value of x^* for case (a) is $S + \zeta + (1-\zeta)S$. Hence, the optimum S in this region is:[3]

$$S_{\text{opt}} = \left[\frac{\frac{\Lambda}{\mu}(K_0(\mu + r) + k)(\Pi_0 - \pi_1) - C_0\zeta}{C_0(1-\zeta)}\right]_{[0, \frac{\zeta'-\zeta}{1-\zeta}]} \quad (5)$$

(b) $S > \frac{\zeta'-\zeta}{1-\zeta}$: In this case, $1 - (S + y^*)$ fraction of the ISPs purchase and enable the firewall, and the rest, never pay for it. Note also that in case (b), since $x^* = \zeta'$, we have $G_0(x^*) = G_2(x^*)$, i.e., the utility of non-adopters and enablers are the same. Hence, the social welfare utility (U) is:

$$(1 - (S + y^*))G_1(x^*) + (S + y^*)G_2(x^*) - C_0S = G_1(x^*) + C_0y^*$$

Note that in case (b), $x^* = \zeta'$ and does not vary with S. Therefore, $\dfrac{dU}{dS} = C_0\dfrac{dy^*}{dS} = -C_0(1-\zeta) < 0$. This shows that the optimum S does not belong to case (b). Hence, $S = S_{\text{opt}}$ as stated in (5) is the optimum seeding for social welfare.

2) Network Security: As a measure of how resistant the network is against intrusions, we define the *security utility* $V(x)$ to be the negative of the expected aggregate damage incurred on the network as a result of the intrusions. Indeed, higher values of V indicate a more secure network. Note specifically that $V(x)$ does not include the costs of the adoption of the firewall and/or seeding. Each of the ISPs that have enabled their firewall contributes $-G_1(x) - C_0 - c/r$ to the total cost of intrusions, and each ISPs that has not obtained the firewall or has not enabled it contributes $-G_0(x)$ to the total intrusion cost. Hence:

$$V(x^*) = x^*\left(G_1(x^*) + C_0 + c/r\right) + (1-x^*)G_0(x^*)$$

Similar arguments as for the social welfare utility show that $V(x^*)$ is maximized with respect to S for $S = \dfrac{\zeta' - \zeta}{1-\zeta}$ in case

[3]The notation $x_{[a,b]}$ represents *projection* of x onto interval of $[a,b]$, i.e., $x_{[a,b]} := \max\{a, \min\{x, b\}\}$.

(a). Note that for case (b), $x^* = \zeta'$ does not depend on S and hence, the value of $V(x^*)$ does not vary with S. This shows that the lowest S that achieves the highest security in the network is $\frac{\zeta' - \zeta}{1-\zeta}$. Note that as long as the regulator is paying for seeding, the firewall vendor also favors it.

IV. Mandating

A regulator with sufficient authority can enforce the ISPs to obtain and enable a firewall. However, no entity has full authority over the whole Internet. We consider the general case in which a fraction M of ISPs abide by the mandate. The initial state of the adoption process is $(y_0, x_0) = (1-M, M)$.

Finding the equilibrium in this case is easier than in seeding. First, note that for $x > \zeta$, $G_0(x) > G_1(x)$ and hence, no non-adopter ISP has incentive to purchase the firewall. Therefore, it immediately follows that for $M > \zeta$, the adoption process continues to be $(1-M, M)$. Here, we investigate the case of $M < \zeta$. At $t = 0$, the adoption state is in Region 1. Following the steps for seeding, $(y(t), x(t))$ is obtained as: $(y(t), x(t)) = ((1-M)e^{-\gamma t}, 1-(1-M)e^{-\gamma t})$. Therefore, $x(t)$ increases until time T at which $x(T) = \zeta$. Note that there is no node that has the firewall and has not enabled it, hence the adoption process cannot enter Region 2. Therefore, for $M < \zeta$, the equilibrium of adoption is $(y^*, x^*) = (1-\zeta, \zeta)$, which is achieved irrespective of the value of $M \in [0, \zeta]$, including specifically $M = 0$, i.e., no mandate. As a consequence, *any mandate authority on less than ζ fraction of the ISPs is futile.*

As for the seeding, we next consider two different objectives for the regulator and find the optimum fraction of ISPs to be mandated.

A. Social Welfare

The social welfare utility, U in this case is $x(G_1(x)) + (1-x)G_0(x)$. Following the same steps that lead to (5), the optimum M is found to be:

$$M_{opt} = \left[\frac{\frac{\Lambda}{\mu}(K_0(\mu + r) + k)(\Pi_0 - \pi_1)}{C_0 + c/r} \right]_{[0,1]}$$

B. Network Security

The network security utility V is equal to $x(G_1(x) + C_0 + c/r) + (1-x)G_0(x)$. This expression is maximized for the largest possible x, which can be achieved for $M = 1$. Hence, from the viewpoint of the network security, mandating all of the ISPs to purchase and install the firewall is optimum.

V. Further Discussions

We introduce two features of any general security management policy and investigate them in seeding and mandating. If the utility of *none* of the ISPs decreases, and the utility of some of them increases, compared to when a specific policy is not applied, we say that ISPs are *better-off* with that policy. If after applying a policy, *none* of the ISPs have any strict preference to change their decision with regards to the adoption of the firewall, we say that ISPs are *happy* with the policy.

Now, it is straightforward to verify that ISPs are both better-off and happy with the seeding policy. They are better-off because without seeding, the value of x^* at equilibrium would be ζ, while with any seeding $S > 0$, we have $x^* > \zeta'$. Also recall that both adopters and non-adopters benefit from higher x. Moreover, ISPs are happy with seeding because they reach an equilibrium, and by definition, they are at their best individual choice.

On the other hand, with optimum mandating, ISPs can be unhappy. For instance, for $M > \zeta'$ the ISPs that are enforced to buy and keep their firewall enabled, strictly prefer to disable it. This is while each ISP is better off compared to the case if they are freely allowed to choose.

VI. Conclusion

We presented an analytical model for the incentive-compatible adoption of asymmetric, bidirectional firewalls among ISPs, using which, we investigated the efficacy and characteristics of two main adoption management policies: seeding and mandating. Our results show that the positive externalities of the firewalls for both adopters and non-adopters can lead to non-trivial implications with regards to the implementation of these policies. For instance, seeding beyond a certain threshold of ISPs is non-beneficial and mandating bellow a threshold of ISPs is futile. We also showed that ISPs, firewall developers and the regulator can have un-aligned incentives in regards to the preferred choice of the policy. Our future research will focus on combining the use of firewalls and cyber-insurance, as well as generalizing the model to non-homogeneous ISPs.

References

[1] J. François, G. Moura, and A. Pras, "Cleaning your house first: Shifting the paradigm on how to secure networks," *Managing the Dynamics of Networks and Services*, pp. 1–12, 2011.

[2] L. Chen, T. Longstaff, and K. Carley, "The economic incentives of providing network security services on the internet infrastructure," *Journal of Information Technology Management*, vol. 15, no. 3-4, pp. 1–13, 2004.

[3] J. Bauer and M. van Eeten, "Cybersecurity: Stakeholder incentives, externalities, and policy options," *Telecommunications Policy*, vol. 33, no. 10, pp. 706–719, 2009.

[4] R. Anderson, R. Böhme, R. Clayton, and T. Moore, "Security economics and the internal market," *Report to the European Network and Information Security Agency*, 2008.

[5] M. Van Eeten, J. Bauer, H. Asgharia, S. Tabatabaie, and D. Rand, "The role of internet service providers in botnet mitigation an empirical analysis based on spam data," 2012.

[6] S. Hofmeyr, T. Moore, S. Forrest, B. Edwards, and G. Stelle, "Modeling internet-scale policies for cleaning up malware," *Arxiv preprint arXiv:1202.4008*, 2012.

[7] C. Bauch and D. Earn, "Vaccination and the theory of games," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 36, p. 13391, 2004.

[8] M. Lelarge, "Coordination in Network Security Games," in *Proc. of INFOCOM*. IEEE, 2012, pp. 2856–2860.

[9] S. Sen, Y. Jin, R. Guerin, and K. Hosanagar, "Modeling the dynamics of network technology adoption and the role of converters," *IEEE/ACM Trans. on Networking*, vol. 18, no. 6, pp. 1793–1805, 2010.

[10] T. Reluga, C. Bauch, and A. Galvani, "Evolving public perceptions and stability in vaccine uptake," *Mathematical biosciences*, vol. 204, no. 2, pp. 185–198, 2006.

[11] M. Whitman and H. Mattord, *Principles of information security*. Course Technology Ptr, 2011.

[12] G. Zhang and M. Parashar, "Cooperative defense against DDoS attacks," *J. of Research and Practice in IT*, vol. 38, pp. 69–84, 2006.