Attackers and Defenders in Cybersecurity and their Optimal Investment Decisions

Austin Ebel[1] and Debasis Mitra[2]

[1] New York University, New York NY 10012, USA
abe5240@nyu.edu

[2] Columbia University, New York NY 10027, USA
debasismitra@columbia.edu

Abstract. Our thesis is that economics and investment policies are vital factors in determining the outcome of cybersecurity conflicts. For our economic framework we borrow from the pioneering work of Gordon and Loeb in which the Defender optimally trades-off investments for lower likelihood of its system breach. Our two-sided model has in addition an Attacker, assumed to be rational and also guided by economic considerations in its decision-making, to which the Defender responds. The model is a simplified adaptation of a model proposed during the Cold War for weapons deployment in the US. Our model is a Stackelberg game and, from an analytic perspective, a Max-Min problem. The complexity of the analysis is due to the non-convexity of the objective function in the optimization. The Attacker's possible actions add substantially to the risk to the Defender, and the Defender's rational, risk-neutral optimal investments in general substantially exceed the optimal investments predicted by the one-sided Gordon-Loeb model. We obtain a succinct set of three decision types that categorize all of the Defender's optimal investment decisions. Also, the Defender's optimal decisions exhibit discontinuous behavior as the initial vulnerability of its system is varied. The analysis is supplemented by extensive numerical illustrations. The results from our model open several major avenues for future work.

Keywords: Optimal security investment · Economics of cybersecurity · Max-Min optimization · Vulnerability mitigation

1 Introduction

In recent times the role of cybersecurity has grown at a pace comparable to that of information technology in society. This is to be expected since its failure has the potential to inflict damage wherever IT is extensively deployed, from national security to transportation, health care, banking, and energy. The number and variety of attackers, which range from state actors and terrorists to criminals and hackers, has grown similarly, as have their resources [1]. Just at the level of commercial enterprises, the projected current annual loss from cybercrime is estimated to be $945 billion, almost double the corresponding amount of $500

billion in 2018 [2]. IBM has estimated the cost of a (commercial) data breach in 2021 to be $4.24 million, up 10% from the previous year [3].

Cybersecurity defense and investments have kept pace as well. An example, at the national level in the USA, is the linked combination of the National Vulnerability Database [4], the Common Vulnerabilities and Exposures list, and the NIST Common Vulnerability Scoring System [5].

Our work in cybersecurity takes the economic perspective, and it examines the many facets of investment decision-making in the context of a model, which is simple to describe, and yet is of broad interest and rich in analytic and computational complexity. The model is an adaptation of a simplification of a model proposed during the Cold War for weapons deployment in the US [6]. Our model is two-sided, i.e., the Attacker and the Defender actively interact through the instrument of investments to further their respective goals of reaping gains from breaching the system, and minimizing expected losses by mitigating vulnerabilities. Both the Attacker and the Defender are assumed to be rational, risk-neutral and guided by economic consideration in their decision-making.

The Defender, the first mover, on inheriting a system with known initial vulnerability, decides on the amount that it will invest in improving the system defense by optimizing its objective function. Then, with knowledge of the upgraded system breach probability, the Attacker follows in attempting to breach the system by expending the amount of effort that optimizes its objective function.

We build our economic framework on the important, pioneering work of Gordon and Loeb [7] based on a one-sided model in which the Defender optimally invests in vulnerability mitigation. The Defender's investment decision is based on a known nonlinear function, $S(z, v)$, which gives the probability of a breach in its system from attacks, and which depends on the system's initial vulnerability (v), and the subsequent investment (z). For two specific classes of the function $S(z, v)$, they also prove the striking result that the optimum Defender's investment does not exceed $1/e$, i.e., about 37%, of the expected loss from a breach. Subsequent work [8,9] proved that the $1/e$ rule holds whenever $S(z, v)$ is log-convex in z. We assume log-convexity of the function $S(z, v)$ with respect to z.

Our two-sided model is a Stackelberg leader-follower game with the Defender as leader and Attacker as follower [10]. Our analysis parallels that of a Max-Min problem. The closest related work is Danskin [6]; however, our results here are quite different. Our goal is to extract a qualitative understanding of the quantitative processes in the optimization of investments. Our main obstacle is the provable nonconvexity of the Defender's objective function[3], which puts the problem nominally in the intractable category. However, there is enough structure that we are able to exploit to show that there are only three distinct types of optimum investment. We next look at the point of transitions in the Defender's

[3] It is worth recalling the view of an eminent researcher, R.T. Rockafeller: "The great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity." [11]

optimum decision types as the initial vulnerability, v, is varied. We show that the points at which transitions occur are solutions of Fixed Point equations. While the total number of transitions are small, decisions across transitions are typically sharply different. We also find that, in general, the investments predicted by the one-sided model of Gordon-Loeb [7] have qualitatively different characteristics, and also underestimate the investments necessary to prudently defend the system against attackers.

Related Work Anderson [12] gives an economic perspective of the many, often conflicting, factors that make realizing cybersecurity hard. Fedel and Roner [13] is a recent, comprehensive review of the literature on cybersecurity investments. It delineates four streams, of which the earliest originates with Gordon and Loeb [7]; this stream is also the simplest since it is specialized to single firms, i.e., one-sided settings. Hausken [14], working in the one-sided setting, generalizes the Defender's function $S(z, v)$ in Gordon and Loeb to have marginal returns on investment to be increasing initially and then decreasing, and shows that the $1/e$ rule is violated and also obtains four different types of optimal investment policies. Huang et al. [15] consider risk-averse Defenders and obtain results that apply when risk is extended to apply to securities.

Brown et al. [16] focus on models in the two-sided, Attacker-Defender setting for two and three sequential stages in which the objective functions represent costs and are linear in the efforts, and the solution methods are based on LP, LP relaxations and Mixed Integer Linear Programs.

For results from a detailed study of the costs of cybercrime see Anderson et al. [17]. The relationships between the average cost of an attack and the expected gain to the Attacker, and similar estimates for the Defender, are examined in Tiwari and Karlapalem [18]. Data from Japan's municipalities on vulnerabilities and investments are in Tanaka et al. [19].

Our paper does not undertake the analysis of networked systems, but it may provide a springboard for future work, and we take note of works with intersecting interests. Gueye and Marbukh [20] study a game in which the Defender selects a spanning tree from a network, and the Attacker selects a link to delete. Dziubinski and Goyal [21] combine graph theory with economics to obtain important insights on how to design optimally defensible networks in the Attacker-Defender setting. Strategic attacks, as in our paper, and random attacks yield different network structures; the paper does not address computations. Goyal and Vigier [22] and Acemoglu et al. [23] address contagion generated by a successful attack. In the former work a critical role belongs to the "network value function", e.g., its asymptotic behavior as the network becomes large determines the optimal network topology. The latter work, which considers a substantially broader class of networks, asks when are investments below or above the socially optimal, and, in an important result, it is shown that the nature of the externality is a key determinant.

Game-theoretic analyses are pervasive in network security, indeed all the aforementioned papers fall in this category. Less so are sequential Stackelberg

games, see Breton et al. [24], which model multiple stages; our model, for instance, is a special case.

A broad overview from the systems perspective of cybersecurity, including reconnaissance techniques which an Attacker may deploy to form estimates, is in Mazurczyk and Caviglione [25]. In the computer science literature on cybersecurity, attack graphs provide a natural framework and are frequently used, see Wang et al. [26] for an example. Moving Target Defense is an effective but costly technique wherein the Defender reconfigures its network periodically, see Lei et al. [27] for a survey of the literature.

The focus of Danskin [6] is on the technical difficulty that arises in Min-Max problems when the solution to the inner maximization is not differentiable. Danskin overcomes the problem by introducing directional derivatives, and goes on to obtain a general Lagrange multiplier principle for Min-Max problems. We note that the technical difficulty does not arise in our work, and, beyond the model, our works diverge.

To the best of our knowledge the results in this paper on a class of Attacker-Defender, two-sided models, the non-convex properties of the resulting objective function in the Min-Max optimization, and the implications of these properties on the structure of optimal investment policies have not been considered in prior work.

2 Preliminaries

2.1 Basic Variables, Functions, and Properties

We define the following variables related to the Defender's system:

1. $s =$ system breach probability, also referred to as (system) vulnerability.
2. $v =$ initial breach probability, also referred to as initial (system) vulnerability.
3. $z =$ incremental "effort" by system Defender to mitigate system vulnerability. We distinguish effort from its financial cost to the Defender, with the latter given by $d\,z$, where d is the Defender's unit cost of effort.
4. $s = S(z, v)$, $S(.,.)$ is the system breach probability function.

We make the following assumptions on the properties of the system breach probability function. Gordon-Loeb [7] also assume A1–A5, whereas A6 is new.

A1. $S(z, 0) = 0,\ \forall z \geq 0$

A2. $S(0, v) = v$

A3. $S_z(z, v) = \frac{\partial S(z,v)}{\partial z} < 0,\ \forall z$ and $\forall v \in (0, 1)$

A4. $S_{zz}(z, v) > 0,\ \forall v \in (0, 1)$

A5. $S(z, v) \to 0$ as $z \to \infty,\ \forall v \in (0, 1)$

A6. $S_v(z, v) > 0,\ \forall z$ and $\forall v \in (0, 1)$, i.e., system vulnerability increases with initial vulnerability.

We define the effort function $Z(.,.)$, where $z = Z(s,v)$. The function Z is the inverse of the function S, i.e., $z \equiv Z(S(z,v),v),\ \forall z \geq 0$. Our analysis is focused on the effort function and z, from which the behavior of the breach probability s can be inferred.

Examples Gordon-Loeb [7] define two classes of system breach probability functions, I and II. We have found these functions to be very useful in illustrating our analysis. We refer to these function classes as GL Class I and GL Class II.

1. GL Class I. $S(z,v) = \frac{v}{(\alpha z+1)^{\beta}}$, $\alpha > 0,\ \beta \geq 1$; $Z(s,v) = \frac{1}{\alpha}\left(\frac{v}{s}\right)^{1/\beta} - \frac{1}{\alpha}$
2. GL Class II. $S(z,v) = v^{\alpha z+1}$, $\alpha > 0$; $Z(s,v) = \frac{1}{\alpha}\frac{\log s}{\log v} - \frac{1}{\alpha}$ (natural log)

Since $\frac{\partial Z(s,v)}{\partial s} = Z_s(s,v) = \frac{1}{S_z(z,v)}$, hence from A3,

$$Z_s(s,v) < 0 \tag{2.1}$$

We will interpret $-Z_s(s,v) = \frac{\partial Z}{\partial(-s)}$ as the (positive) *marginal effort for invulnerability* (for fixed initial vulnerability). Analogously,

$$Z_v(s,v) = -\frac{S_v(z,v)}{S_z(z,v)} > 0 \tag{2.2}$$

is the marginal effort for maintaining invulnerability with increased initial vulnerability, with positivity following from A3 and A6. Similarly,

$$Z_{ss}(s,v) = -\frac{S_{zz}(z,v)}{\{S_z(s,v)\}^3} > 0 \tag{2.3}$$

Hence, $Z(s,v)$ is a convex, decreasing function of s. We also assume,

$$\text{A7.}\quad Z_{sv}(s,v) < 0\,, \tag{2.4}$$

i.e., the marginal effort for invulnerability increases with increased initial vulnerability. We call this property "effort complementarity", which is distinct from convexity. It is analogous to "cost complementarity" in the economics of multiproduct firms, which is the property that the marginal cost of producing product j decreases with increased production of product i, $i \neq j$ [28,29].

Furthermore, we assume the following factored form of $Z_s(s,v)$:

$$\text{A8.}\quad Z_s(s,v) = -\frac{f(v)}{g(s)} \tag{2.5}$$

where, following (2.1), $f(v)$ and $g(s)$ are positive $\forall v,s \in [0,1]$, and are unique to within constants of proportionality. It follows from (2.5) and (2.4),

$$f_v(v) > 0 \quad \text{and} \quad g_s(s) > 0 \tag{2.6}$$

The GL functions satisfy the above assumptions and the factored form.

(i) GL Class I.

$$f(v) = v^{1/\beta}; \quad g(s) = \alpha\beta s^{(\beta+1)/\beta} \tag{2.7}$$

$$Z_s(s,v) = -\frac{v^{1/\beta}}{\alpha\beta}\frac{1}{s^{(\beta+1)/\beta}}; \quad Z_{sv}(s,v) = -\frac{1}{\alpha\beta^2}\frac{1}{v^{(\beta-1)/\beta}}\frac{1}{s^{(\beta+1)/\beta}} \tag{2.8}$$

(ii) GL Class II.

$$f(v) = -\frac{1}{\log v}; \quad g(s) = \alpha s \tag{2.9}$$

$$Z_s(s,v) = \frac{1}{\alpha(\log v)s}; \quad Z_{sv}(s,v) = -\frac{1}{\alpha\left(\log^2 v\right)vs} \tag{2.10}$$

2.2 Log Convexity

We assume that $S(z,v)$ is log-convex in z for all fixed v, i.e., $\log S(z,v)$ is convex, which translates to, $\gamma(z,v) \geq 0$ where,

$$\gamma(z,v) = \frac{S_{zz}(z,v)\,S(z,v)}{\{S_z(z,v)\}^2} - 1 \tag{2.11}$$

We have introduced the function $\gamma(z,v)$ to parameterize log-convexity of $S(z,v)$. Making use of the factored form in (2.5), it follows that,

$$\gamma(s,v) = \frac{s\,g_s(s)}{g(s)} - 1 \tag{2.12}$$

Henceforth we abbreviate $\gamma(s,v)$ to $\gamma(s)$; it has an important role in our analysis.

It was shown independently by Baryshnikov [8] and Lelarge [9] that Gordon and Loeb's celebrated $1/e$ rule [7] holds if the system breach probability function, $S(z,v)$, is log-convex, and also that both Class I and II functions used by Gordon and Loeb to demonstrate the rule are log-convex. Note the following:

(i) GL Class I. $\gamma(s) \equiv 1/\beta$, a constant. Gordon-Loeb [7] require $\beta \geq 1$, so that $\gamma \leq 1$. In general, we will not place this restriction. However, there are some significant qualitative differences that appear in the analysis depending on whether $\gamma \leq 1$ or $\gamma > 1$. Hence, when the need arises to assume $\gamma \leq 1$, we will specify "GL Class I with $\gamma \leq 1$", and similarly for $\gamma > 1$, it being understood that otherwise γ is any nonnegative constant in GL Class I.

(ii) GL Class II. $\gamma(s) \equiv 0$.

For $g(s) = O(s^\delta)$ as $s \to 0$, it may be verified from (2.12) that

$$\gamma(s) \sim \delta - 1. \tag{2.13}$$

Hence,

$$\delta > (=)\ (<)\ 2 \ \text{ if } \ \gamma(0) > (=)\ (<)\ 1. \tag{2.13}$$

The above dependence of the asymptotic behavior of $g(s)$ as $s \to 0$ on the value of $\gamma(0)$ is important in the subsequent analysis.

In Sec. 5.2, we define the class of functions $\gamma(s)$ for which our results hold, and it is broad and extends beyond constants and the constraint $\gamma(s) \leq 1$.

3 Attacker and Defender: Model, Actions, Reactions

3.1 Model of Attacker

Let y represent a generalized measure of the aggregate effort that the Attacker deploys. In the simplest representation, y is the number of attempts, equal to the effort, that the Attacker makes to breach the Defender's system. Departures from this simple representation include breaching attempts of varying intensities. In the base representation where y is the number of independent attacks, the total cost to the Attacker will be assumed to be cy, where c is the Attacker's cost for unit effort. For a given system vulnerability s, we assume,

$$Pr[\text{system is breached}] = T(s, y) = 1 - (1 - s)^y \tag{3.1}$$

The key assumption that we are making is that the Attacker's attempts are independent Bernoulli trials with probability of system breach in each trial given by s, and that the Attacker is successful if at least one of the attacks succeed in breaching the system.

We let G denote the financial gain that the Attacker realizes if it succeeds. The Attacker's net expected gain for system vulnerability s and deployed effort y is therefore,

$$GT(s, y) - cy \tag{3.2}$$

We are adopting features of the model in Danskin [6], who states (in Chapter 4, Sec. 4), that the model was first posed during the Cold War at the RAND Corporation around 1951. In Danskin's model there are several types of attack units, whereas here there is only one, and the probability that an individual attack unit gets through is a function of the number of attack units, whereas here it is independent.

3.2 Attacker's Optimization Problem

The rational, risk-neutral Attacker will act to maximize its net expected gain, i.e.,

$$\max_{y \geq 0} \Big[GT(s, y) - cy \Big] \tag{3.3}$$

The Attacker is assumed to know s, which, for instance, it may estimate by deploying techniques described in Mazurczyk and Cavaglione [25]. Let $y^*(s)$ denote the solution to the Attacker's optimization problem, and let $T^*(s) = T(s, y^*(s))$

3.3 Model of Defender

Our model assumes that the Defender has knowledge of the Attacker's decision process and that the Defender is oblivious to risk, and hence its decisions are based on expected values only.

Let L denote the financial loss to the Defender in the event of a system breach. Since d is the unit cost of the effort, the net financial cost of mitigation

is dz. Following the Attacker's actions and the Defender's mitigation efforts, the Defender's net expected financial loss is,

$$LT^*(s) + dz \tag{3.4}$$

which, upon substitution of the system breach probability function yields,

$$LT^*(S(z, v)) + dz \tag{3.5}$$

3.4 Defender's Optimization Problem

The rational, risk-neutral Defender will implement the solution to the following problem,

$$\min_{z \geq 0} \left[LT^*(S(z, v)) + dz \right] \tag{3.6}$$

We find it convenient to conduct the analysis with the decision variable s instead of z, which requires the replacement of the system breach probability function $S(z, v)$ by the effort function $Z(s, v)$, which has been introduced in Sec. 2.1. Since z and s are related by invertible functions, this transformation should not pose any fundamental problem. After the switch, the Defender's problem becomes,

$$\min_{s \leq v} \Phi(s, v) \tag{3.7}$$

where,

$$\Phi(s, v) = LT^*(s) + dZ(s, v) \tag{3.8}$$

We denote the solution to (3.7) by $s^*(v)$, also $z^*(v) = Z(s^*(v), v)$, which gives the Defender's optimum mitigation effort, and $\Phi^*(v) = \Phi(s^*(v), v)$.

4 Attacker's Problem: Solution and Discussion

4.1 Solution

The first order condition for optimality in the problem in (3.3), which is obtained by setting to zero the derivative of the objective function, yields,

$$\{-\log(1 - s)\}(1 - s)^y = \frac{c}{G} \tag{4.1}$$

The left-hand side is monotonic, decreasing in y and approaches 0 as $y \to \infty$. Hence, if its value for $y = 0$ is greater than c/G, then a unique positive solution $y^*(s)$ exists, and otherwise the solution to the Attacker's problem in (3.3) is $y^*(s) = 0$.

Define s_P to be the value of s such that the left-hand side of (4.1) at $y = 0$ is equal to c/G, i.e,

$$s_P = 1 - e^{-c/G} \tag{4.2}$$

We summarize here:

Proposition 4.1. *If $s > s_P$, then a unique positive solution to $y^*(s)$ to the Attacker's Optimization Problem in (3.3) exists, and satisfies,*

$$(1-s)^{y^*(s)} = \frac{1}{(G/c)\{-\log(1-s)\}} \tag{4.3, i}$$

i.e.,

$$y^*(s) = -\frac{\log\left[\log\left\{(1-s)^{-G/c}\right\}\right]}{\log(1-s)} \tag{4.3, ii}$$

If $s \leq s_P$, then $y^*(s) = 0$.

Also, the probability of a system breach that is a consequence of the Attacker's optimal effort,

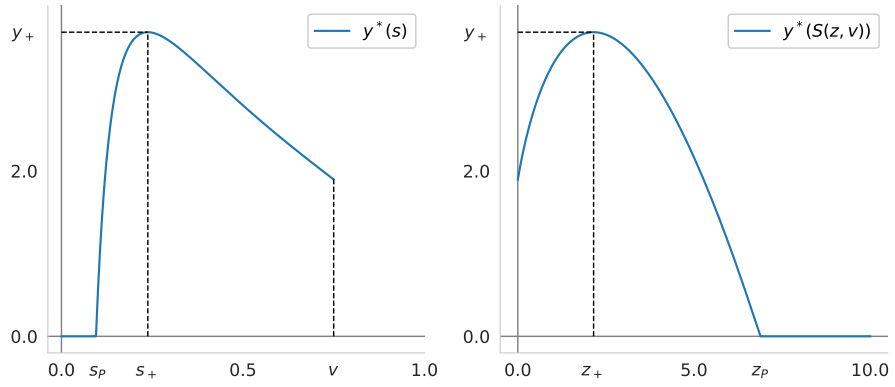
$$T^*(s) = 1 + \frac{1}{(G/c)\{\log(1-s)\}} \qquad \text{for } s > s_P \tag{4.4, i}$$

$$= 0 \qquad \text{for } s \leq s_P \tag{4.4, ii}$$

In the Attacker's problem, the system breach probability s is given and assumed to satisfy the constraint $s = S(z, v) \leq v$, the initial vulnerability. Note that if $v \leq s_P$ then $s \leq s_P$. To avoid trivialities, we assume that $v > s_P$.

4.2 Discussion, the Price of Deterrence

We examine the behavior of $y^*(s)$, the Attacker's optimal effort as a function of s, the system vulnerability, for $s > s_P$.



(a) $y^*(s)$ vs. s (b) $y^*(S(z, v))$ vs. z

Fig. 4.1: Visualizing rational Attacker investments and the Price of Deterrence; GL Class I, $v = 0.75$, $L = G = 10$, $\alpha = 1$, $\beta = 1$, $c = d = 1$; $\longrightarrow$ $y_+ \approx 3.68$, $s_+ \approx 0.24$, $z_+ \approx 2.15$, $s_P \approx 0.10$, $z_P \approx 6.88$

Differentiating the expression in (4.3, ii) with respect to s yields an expression for $\frac{dy^*(s)}{ds}$ from which it may be inferred (the detailed proof is omitted), for

$s > s_P$, that as s increases, $y^*(s)$ is initially monotonically increasing, reaches a stationary point, which we denote by s_+, and is thereafter monotonic, decreasing, as depicted in Fig. 4.1. We let the maximum value of $y^*(s)$ be y_+, i.e., $y_+ = y(s_+)$. The following result explicitly identifies s_+ and y_+.

Proposition 4.2. *The Attacker's maximum effort over all system vulnerabilities, y_+, is $1/e$, i.e., about 37% of the effort justified by the potential gain from system breach. This follows from,*

$$s_+ = 1 - e^{-ec/G}, \qquad y_+ = \frac{G/c}{e}, \qquad \text{so that} \quad (1 - s_+)^{y_+} = \frac{1}{e} \tag{4.6}$$

This result resembles Gordon-Loeb's $1/e$ result, and is due to the log-concavity of $T(s, y)$ in (3.1) with respect to y.

The corresponding behavior of $y^*(Z(s, v), v)$ as a function of the Defender's effort is readily inferred. We let z_P be the effort corresponding to s_P,

$$z_P = Z(s_P, v), \quad \text{i.e.,} \quad S(z_P, v) = s_P, \tag{4.7}$$

When G/c is large, as may be considered typical, then from (4.2), $(s_P)(G/c) \approx 1$. Now, s_P is the probability of system breach from the Attacker's unit effort, and G/c is the maximum effort that a rational Attacker will deploy. Hence, in this approximation regime, when $s \leq s_P$, the expected number of successful system breaches falls below one even when the Attacker mounts the maximum effort that is justified by the potential gain from a system breach. Consequently, in this case the rational, risk-neutral Attacker will not mount *any* attack.

We call z_P the "Price of Deterrence" (strictly it is dz_P). That is, if $z > z_P$, then the system breach probability falls below the threshold that justifies the Attacker to mount any attack.

It is important to keep in mind that the Price of Deterrence also depends on the initial vulnerability v, as is evident in (4.7).

Arms Race Analogy We draw attention to the Price of Deterrence, z_P in Fig. 4.1b: as the Defender's investment z increases, the Attacker's optimum effort, y^*, initially increases, i.e., the protagonists are locked in a classic arms race. However, if the Defender's investment exceeds z_+, the Attacker's economic interests no longer justify matching the escalation, and its investments decline. If the Defender's investment reaches the Price of Deterrence, then the Attacker throws in the towel, and cuts off all further investments. Note that the entire cycle consisting of the Attacker's expanding investments, followed by the partial withdrawal and finally total withdrawal while facing a Defender with superior economic resources is symptomatic of the point of view that economics is a substantial strategic factor in cybersecurity.

It is tempting to conjecture that the behavior in Fig. 4.1b mirrors the US – Soviet Union arms race during the Cold War until the economic collapse and dissolution of the Soviet Union ("Part of the logic proceeding with SDI was that, eventually, the arms race would cripple the Soviet economy. This is in fact what was happening." [30]). See also [31].

5 Defender's Problem: Analysis and Solution

5.1 Objective and Gradient Functions

Combining (3.8) and (4.4, i) yields the following expression for the Defender's objective function for $s > s_P$,

$$\Phi(s,v) = L\left[1 + \frac{1}{(G/c)\{\log(1-s)\}}\right] + dZ(s,v) \tag{5.1}$$

In deriving the derivative with respect to s, the system breach probability, we take note of the factored form of $Z_s(s,v) = -\frac{f(v)}{g(s)}$ in (2.5), to obtain,

$$\left[\frac{(1-s)\log^2(1-s)}{df(v)}\right]\frac{\partial}{\partial s}\Phi(s,v) = \frac{L/d}{G/c}\frac{1}{f(v)} - D(s) \tag{5.2, i}$$

where,

$$D(s) = \frac{(1-s)\log^2(1-s)}{g(s)} \tag{5.2, ii}$$

Note that the bracketed term on the left-hand side of (5.2, i) is positive for $0 < s < 1$, and the separation of v and s in the terms on the right-hand side. Let,

$$R = \frac{L/d}{G/c} \tag{5.3}$$

and refer to R as the Effective Loss to Gain Ratio. Now,

$$\Phi_s(s,v) = (<)\ (>)\ 0 \qquad \text{if} \quad D(s) = (>)\ (<)\ R\frac{1}{f(v)} \tag{5.4}$$

We refer to D(s) as the "gradient function".

Example We introduce the "universal" function $\xi(s)$ here to illustrate a gradient function and also because it will be useful later in the analysis:

$$\xi(s) = \frac{(1-s)\log^2(1-s)}{s} \tag{5.5}$$

For GL Class II functions, for which $g(s) = \alpha s$, see (2.9), the gradient function is $D(s) = \xi(s)/\alpha$. In the interval $[0,1]$, the corner points of $\xi(s)$ are $\xi(0) = \xi(1) = 0$. The function $\xi(s)$ has the following canonical shape: a unique maximum $\hat{\xi} = \max_{0<s<1}\xi(s)$; monotonic, increasing in $[0, \hat{s}_\xi)$, and monotonic, decreasing in $(\hat{s}_\xi, 1]$, where we have denoted the location of the maximum by $\hat{s}_\xi$, i.e., $\xi(\hat{s}_\xi) = \hat{\xi}$. We shall say that such functions have an "inverted-U shape"[4]. In the example of $\xi(s)$ in (5.5), $\hat{\xi} \approx 0.64$ and $\hat{s}_\xi \approx 0.8$. Note that such functions are not necessarily concave.

[4] Use of the term "inverted-U" function or relationship has precedents, for example, see [32].

5.2 Shape of the Gradient Function

Here we obtain the shape of the gradient function $D(s)$. A key determinant is the non-negative function $\gamma(s)$, which was introduced in (2.12). Taking the derivative of $D(s)$ in (5.2, ii) with respect to s, and making use of (2.12) gives,

$$\left[\frac{g(s)}{\log^2(1-s)}\right]\frac{d}{ds}D(s) = H(s;\gamma) \tag{5.6}$$

where the bracketed term on the left-hand side is positive for $s \in (0,1)$, and (H for *Hessian*),

$$H(s;\gamma) = -\left\{\frac{2}{\log(1-s)} + \frac{1}{s}\right\} - \gamma(s)\left(\frac{1}{s} - 1\right) \tag{5.7}$$

We proceed below with separate analyses for $0 \le \gamma(s) \le 1$ and $1 < \gamma(s),\ s \in [0,1]$. Let,

$$\Gamma_1 = \{\gamma(s) : 0 \le \gamma(s) \le 1,\ \text{and}\ \ 0 \le \gamma_s(s),\ \forall s \in [0,1]\}$$

Observe that $\gamma(s)$ corresponding to GL Class I functions with $\gamma \le 1$, and GL Class II functions are subsumed in the set Γ_1.

Proposition 5.1. *If $\gamma(s) \in \Gamma_1$ then the Hessian $H(s;\gamma)$ monotonically decreases, i.e., $H_s(s;\gamma) < 0,\ s \in (0,1)$, from non-negative to negative with increasing $s,\ s \in [0,1]$. Hence, $\exists$ unique $\hat{s} \in (0,1)$ such that $H(\hat{s};\gamma) = 0$, where $D_s(\hat{s}) = 0$ and $D_s(s) > (<)\ 0$ for $s < (>)\ \hat{s}$. In the special case of $\gamma(s) \equiv 1$, $\hat{s} = 0$.*

The proof is in Appendix A.1. From (5.6), $D(\hat{s}) = \max_{0\le s\le 1} D(s)$, and we let $\hat{D} = D(\hat{s})$. The proposition establishes that for $\gamma(s) \in \Gamma_1$, the gradient function $D(s)$ has the inverted-U shape with the peak value of $\hat{D}$ at $\hat{s}$. In the special case of $\gamma(s) \equiv 1$, the peak is located at $\hat{s} = 0$, and the shape of the entire gradient function $D(s), s \ge 0$, coincides with the segment of the general inverted-U shape to the right of the peak, i.e., $s \ge \hat{s}$.

We can prove that in general $D(s)$ is not concave. This fact gives weight to the gradient function's inverted-U shape since this property will prove to be adequate for deducing key properties of stationary points of the Defender's objective function, which is considered in Sec. 5.4.

Next, we investigate the behavior of the gradient function $D(s)$ when the characteristics of the function $\gamma(s)$ are complementary to that of Prop. 5.1. Let,

$$\Gamma_2 = \{\gamma(s) : 1 < \gamma(s),\ \forall s \in [0,1]\}$$

Observe that $\gamma(s)$ corresponding to GL Class I functions with $\gamma > 1$ are subsumed in Γ_2.

Proposition 5.2. *If $\gamma(s) \in \Gamma_2$, then $H(s;\gamma) < 0$ and $D_s(s) < 0,\ \forall s \in (0,1)$. Also, $D(s) \to \infty$ as $s \to 0$.*

The proof is in Appendix A.1. For $\gamma(s) \in \Gamma_2$, $D(s) \to \infty$ as $s \to 0$ which represents a significant qualitative difference from $\gamma(s) \in \Gamma_1$; $D(s)$ is monotonic, strictly decreasing for $s > 0$. We let $\hat{s} = 0$.

We now build on the results of Prop. 5.1 and 5.2 to obtain the shape of the gradient function for a more general class of functions $\gamma(s)$,

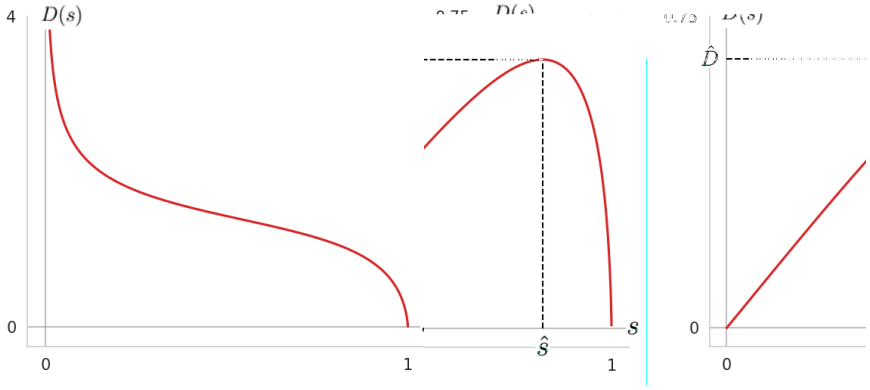
$$\Gamma = \{\gamma(s) : \gamma(s) \geq 0, \text{ and } \gamma(s) \leq 1 \Rightarrow \gamma_s(s) \geq 0, \ s \in [0,1]\} \tag{5.8}$$

In the analysis of Γ we allow for the possibility of a single cross-over of behavior from Γ_1 to Γ_2 to avoid getting immersed in details. Note that if for any s', $\gamma(s') > 1$, then no crossing below the $\gamma = 1$ level is possible for $s > s'$.

The following proposition is proved in A.1.

Proposition 5.3. *Assume that $\gamma(s) \in \Gamma$. If $\gamma(0) > 1$, then Prop. 5.2 applies. If $\gamma(0) \leq 1$, then $D(s)$ is inverted-U shaped with a unique maximum.*

An interpretation of the proposition is that if $\gamma(0) > 1$, then the shape of the gradient function is what is expected from Prop. 5.2 for $\gamma(s) \in \Gamma_2$, and if $\gamma(0) \leq 1$, then the gradient function is inverted-U shaped, as stated in Prop. 5.1 for $\gamma(s) \in \Gamma_1$. See examples of the gradient function in Fig. 5.1.



(a) GL Class I ($\gamma = 1.25$) (b) GL Class II ($\gamma = 0$)

Fig. 5.1: Visualizing $D(s)$ for GL Class I and II functions

5.3 Stationary Points of the Defender's Objective Function

The stationary points, i.e., local maxima and minima, of the Defender's objective function $\Phi(s, v)$ for the variable s and fixed v, the initial vulnerability, are solutions of $\Phi_s(s, v) = 0$. Therefore, following (5.2, i), the stationary points are solutions of,

$$D(s) = R\frac{1}{f(v)} \tag{5.9}$$

If $\gamma(s) \in \Gamma_2$, then from Prop. 5.2 it is known that $D(s) \to \infty$ as $s \to 0$ and also that $D(s)$ is monotone, strictly decreasing with increasing s. Hence, a unique solution $s_1(v)$, $0 < s_1(v) < 1$, exists. We next focus on $\gamma(s) \in \Gamma_1$ for which, from Prop. 5.1, $D(s)$ is inverted-U shaped. Since $f(v)$ increases with v, for any

solution of (5.9) to exist, the initial vulnerability v needs to be sufficiently high. In fact, any stationary point exists if and only if,

$$\hat{v} \leq v \tag{5.10}$$

where $\hat{v}$ is defined by the following relation in which $\hat{D}$ is the maximum value of $D(s)$.

$$\hat{D} = R\frac{1}{f(\hat{v})}, \quad \text{i.e.,} \quad \hat{v} = f^{-1}\left(\frac{R}{\hat{D}}\right) \tag{5.11}$$

For GL Class I functions, $f(v) = v^{1/\beta}$, see (2.7), so that, for $\gamma \leq 1$, i.e., $\beta \geq 1$, $\hat{v} = (R/\hat{D})^{\beta}$. For GL Class II functions, recall from Sec. 5.2 that $\hat{D} = \hat{\xi}/\alpha$ and $f(v) = -1/\log v$, hence, $\hat{v} = \exp(-\hat{\xi}/(\alpha R))$.

Observe that for both Class I with $\gamma \leq 1$ and Class II functions, $\hat{v}$ is independent of initial vulnerability v, and importantly, increases with the effective loss to gain ratio, R.

If $\gamma(s) \in \Gamma$ and $\gamma(0) < 1$, then there exist two stationary points of $\Phi(s, v)$, denoted $s_2(v)$ and $s_1(v)$, where,

$$s_2(v) < \hat{s} < s_1(v) \quad \text{if } \hat{v} < v, \tag{5.12}$$

and, in the special case of $v = \hat{v}$, $s_2(v) = \hat{s} = s_1(v)$. It is easy to see that $s_1(v)$ ($s_2(v)$) increases (decreases) with v.

If $\gamma(s) \in \Gamma$ and $\gamma(0) > 1$, then there is a single stationary point of $\Phi(s, v)$, denoted by $s_1(v)$ which increases with v. Hereafter it is convenient to consider only $\gamma(s) \in \Gamma$ and $\gamma(0) \leq 1$, and accommodate the case of $\gamma(s) \in \Gamma$ and $\gamma(0) > 1$ by noting that in the latter case $\hat{s} = 0$, and consequently $s_2(v)$ does not exist in $[0, 1]$.

Using (5.2, i) to obtain the sign of $\Phi_s(s, v)$ in the intervals defined by the stationary points, we obtain,

Proposition 5.4. *Assume $\gamma(s) \in \Gamma$ and v is constant. Then,*

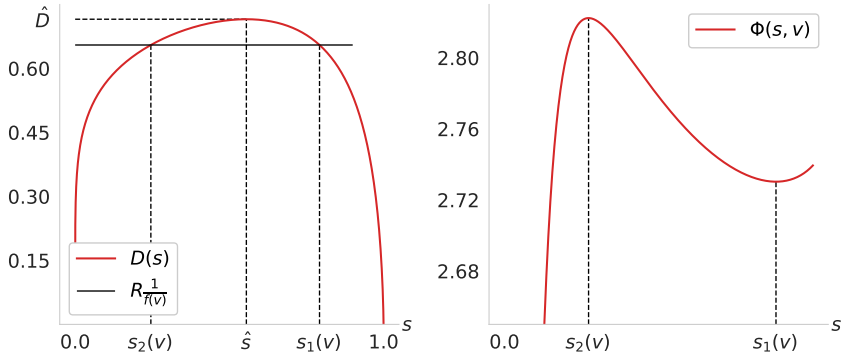
$$\begin{aligned} \Phi_s(s, v) &> 0 \quad \text{for } s \in (0, s_2(v)) \\ &< 0 \quad \text{for } s \in (s_2(v), s_1(v)) \\ &> 0 \quad \text{for } s \in (s_1(v), 1) \end{aligned} \tag{5.13}$$

That is, $s_2(v)$ is a local maximum and $s_1(v)$ is a local minimum of $\Phi(s, v)$ as s is varied in $(0, 1)$ with v fixed. See Fig. 5.2 for an example.

The above result illuminates a central feature of the objective function, i.e., the presence of intervals of local concavity and convexity in one variable (s), with initial vulnerability (v) adding a separate dimension of complexity.

5.4 Solution to the Defender's Problem

We compose the solution to the problem stated in (3.7). The main task is to deduce the implications of the constraint $s \leq v$ in the Defender's problem in (3.7).



(a) $D(s); R\frac{1}{f(v)}$ vs. s (b) $\Phi(s, v)$ vs. s

Fig. 5.2: Visualizing criteria for the existence of stationary points in the Defender's objective function; GL Class I, $v = 0.9,\ R = 0.6,\ \alpha = 1,\ \beta = 1.2$

If $v < \hat{v}$ then, from (5.2, i), $\Phi_s(s, v) > 0,\ \forall s$, and,

$$\Phi^*(v) = \Phi(s_P, v) = dz_P \tag{5.14}$$

Recall from the discussion in Sec. 4.2 that dz_P is the Price of Deterrence. If $v > \hat{v}$, then from Prop. 5.4, we need to consider two disjoint cases, $v < s_1(v)$, and $s_1(v) < v$. From Prop. 5.4,

$$\text{If } \ \hat{v} < v < s_1(v) \ \text{ then } \ \Phi^*(v) = \min\left[\Phi(s_P, v), \Phi(v, v)\right] \tag{5.15}$$

so that $s^* = s_P$ or v. Since $\Phi(v, v)$ is obtained from $s = v$, i.e., $z = 0$,

$$\Phi(v, v) = LT^*(v) = L\left[1 + \frac{1}{(G/c)\log(1 - v)}\right], \tag{5.16}$$

a positive quantity since we have assumed (see Sec. 4.1) that $v > s_P$. Note that $\Phi(v, v)$ is the Defender's net expected loss if it makes no effort to reduce vulnerability from its initial value, and $\Phi(s_P, v)$ is the corresponding quantity if it makes the effort necessary to reduce vulnerability to the level where the Attacker is deterred from making any effort. It is not obvious a priori which is the superior solution for the Defender. Hence we have (5.17).

In contrast,

$$\text{If } \ s_1(v) < v \ \text{ then } \ \Phi^*(v) = \min\left[\Phi(s_P, v), \Phi(s_1(v), v)\right] \tag{5.17}$$

so that $s^* = s_P$ or $s_1(v)$.

We define three "Decision Intervals" (DI) in the range of values of v:

DI 1. $v < \hat{v}$
DI 2. $\hat{v} < v < s_1(v)$ (5.18)
DI 3. $s_1(v) < v$

DI 2 and 3 may be empty, e.g., DI 3 is empty if $v \leq s_1(v)$, $\forall v$. Also, as we shall see, the DI may be interleaved. To describe the Defender's investment decision, we employ the following mnemonics:

1. "all in" to describe $\Phi^*(v) = \Phi(s_P, v)$
2. "none" if $\Phi^*(v) = \Phi(v, v)$
3. "some" if $\Phi^*(v) = \Phi(s_1(v), v)$

Now, by combining Prop. 5.4 and the constraint $s \leq v$, we obtain the following result for the Defender's problem in (3.7),

Proposition 5.5. *The Defender's optimal investment decision depends on the DI in which v exists, as follows: DI 1: "all in"; DI 2: "all in" or "none"; DI 3: "all in" or "some".*

Note the role of the parameter R, the effective loss to gain ratio, in the Defender's decision. For instance, since $\hat{v}$ is monotonic, increasing with R, if R (and therefore $\hat{v}$) is sufficiently large then $v < \hat{v}$, in which case only Interval 1 is of interest, and the Defender's decision is "all in".

The above observation on the role of R also illuminates the potential of spreading false estimates of loss (L) and gain (G) to nudge the adversary to making damaging decisions. This is a rich line of investigation that, however, is outside the scope of this paper.

6 Defender's Optimal Investment Policy

The preceding section has shown how to obtain the Defender's optimal investment for any given value of the initial vulnerability, v. Here, we investigate patterns in optimal investment decisions for the entire range of values of v. From the reader's perspective, an illuminating component is the illustrations of results and properties, which are based on Gordon-Loeb's Class I and II system breach probability functions. Lastly, this section compares and contrasts the optimum investment policies that are obtained from the model of Gordon-Loeb and our model.

6.1 Fixed Point Equations

Section 5.4, and specifically (5.18), has shown that the transitions between Decision Intervals (DI) 2 and 3 occur when $\{s_1(v) - v\}$ changes sign, where v is the initial vulnerability and $s_1(v)$ is the unique local minimum of the Defender's objective function $\Phi(s, v)$ when s, the system breach probability, is varied with v held constant. Hence, we investigate the condition in which equality holds, and this leads to the Fixed Point Equation (FPE) below. Since $s_1(v)$, when it exists, is obtained as the solution of the equation $D(s_1(v)) = \frac{R}{f(v)}$, $s_1(v) \geq \hat{s}$, the transition points in v, i.e., solutions of $s_1(v) = v$, are solutions of the following FPE,

$$D(x) = \frac{R}{f(x)}, \quad \hat{s} < x < 1 \tag{6.1}$$

Note that any solution of (6.1) will satisfy $x > \hat{v}$. Since we will not encounter more than two Fixed Points, we denote the solutions of (6.1) by v^L and v^H. If two solutions exist, we assume that $\hat{v} < v^L < v^H$, and if there is only one, then $\hat{v} < v^H$.

6.2 Gordon-Loeb Class I and II System Breach Probability Functions

In the case of GL Class I functions, the FPE is

$$\frac{(1-x)\log^2(1-x)}{\alpha\beta x^{(\beta+1)/\beta}} = \frac{R}{x^{1/\beta}}, \quad \hat{s} < x < 1 \tag{6.2, i}$$

or equivalently,

$$\xi(x) = R\alpha\beta, \quad \hat{s} < x < 1 \tag{6.2, ii}$$

where $\xi(.)$ is the "universal" function defined in (5.5).

Proposition 6.1. *For GL Class I functions,*

(i) If $\hat{\xi} < R\alpha\beta$, then no Fixed Point exists, and $s_1(v) < v, \ \forall v \in (\hat{v}, 1)$.

(ii) If $\xi(\hat{s}) < R\alpha\beta < \hat{\xi}$, then two Fixed Points, v^H and v^L exist. Moreover $s_1(v) < v, \ \forall v \in (\hat{v}, v^L)$; $v < s_1(v), \ \forall v \in (v^L, v^H)$; $s_1(v) < v, \ \forall v \in (v^H, 1)$.

(iii) If $R\alpha\beta < \xi(\hat{s})$, then one Fixed Point, v^H, exists. Moreover, $v < s_1(v), \ \forall v \in (\hat{v}, v^H)$; $s_1(v) < v, \ \forall v \in (v^H, 1)$.

The proof is in A.2. Recall from Prop. 5.5 that DI 2 corresponds to $\hat{v} < v < s_1(v)$, and D3 to $s_1(v) < v$.

In the case of GL Class II functions, $D(s) = \frac{\xi(s)}{\alpha}$, so that $\hat{D} = \frac{1}{\alpha}\hat{\xi}$ and $\hat{s} = \hat{s}_\xi$. (Recall that $\hat{\xi} \approx 0.64$ and $\hat{s}_\xi \approx 0.8$.) Since $s_1(\hat{v}) = \hat{s}$, it follows that $s_1(\hat{v}) = \hat{s}_\xi$. Also, since $D(\hat{s}) = \frac{R}{f(\hat{v})} = -R\log\hat{v}$, it follows that $\hat{v} = \exp(\frac{-\hat{\xi}}{\alpha R})$. The FPE in (6.1) translates to,

$$\xi(x) = -\alpha R\log(x), \quad \hat{s} < x < 1 \tag{6.3}$$

Proposition 6.2. *For GL Class II functions,*

(i) If $\alpha R < \frac{\hat{\xi}}{-\log\hat{s}_\xi} \approx 2.87$, then no Fixed Point exists, and $v < s_1(v), \ \forall v \in (\hat{v}, 1)$.

(ii) If $\frac{\hat{\xi}}{-\log\hat{s}_\xi} < \alpha R$, then one Fixed Point, v^H, exists, and $s_1(v) < v, \ \forall v \in (\hat{v}, v^H)$; $v < s_1(v), \ \forall v \in (v^H, 1)$.

Proof: It is easy to see that there exists one Fixed Point if and only if,

$$D(\hat{s}) < \frac{R}{f(\hat{s})}, \quad \text{i.e.,} \quad \frac{\hat{\xi}}{-\log\hat{s}_\xi} < \alpha R$$

In the absence of any Fixed Point, the sign of $\{s_1(v) - v\}$ is invariant for $v \in (\hat{v}, 1)$, and may be obtained from $v = \hat{v}$. If a Fixed Point v^H exists, the sign of $\{s_1(v) - v\}$ changes at v^H.

6.3 General Results on the Existence of Fixed Points

We give here a result that holds for $\gamma(s) \in \Gamma_1$, the set defined in Sec. 5.2, which subsumes $\gamma(s)$ for GL Class I with $\gamma \leq 1$ and GL Class II. The result on the existence of solutions to (6.1) follows directly from the previously established properties of $D(x)$ and $f(x)$ for $x \in (\hat{s}, 1)$.

Proposition 6.3. *Assume* $\gamma(s) \in \Gamma_1$. *If,*

$$\frac{R}{f(\hat{s})} < \hat{D} \tag{6.4}$$

then there exists a unique Fixed Point, unless $\frac{R}{f(1)} = 0$, *in which case no Fixed Point exists.*

Proof: Recall from Prop 5.1 that $D(x)$ and $\frac{R}{f(x)}$ are strictly decreasing with increasing $x \in (\hat{s}, 1)$. Note that if (6.4) holds, then $D(\hat{s}) > \frac{R}{f(\hat{s})}$, and since $D(1) = 0 \leq \frac{R}{f(1)}$, it follows that a solution to (6.1) exists if the latter inequality is strict, and no solution exists if equality holds.

For GL Class II functions, $\frac{R}{f(1)} = 0$, and hence from the above proposition, no Fixed Point exists if (6.4) holds, as also stated in Prop. 6.2.

In general, no Fixed Point exists if R, the effective loss to gain ratio, is sufficiently large, specifically, $R > R_c$, where,

$$R_c = \max_{0 \leq x \leq 1} [D(x)\, f(x)] \tag{6.5}$$

When no Fixed Point exists, the Defender's optimal decision is "all in" and invariant for all v. This property is evident in the examples in Fig. 6.1a and 6.1b.

6.4 Visualizing Fixed Point Equations

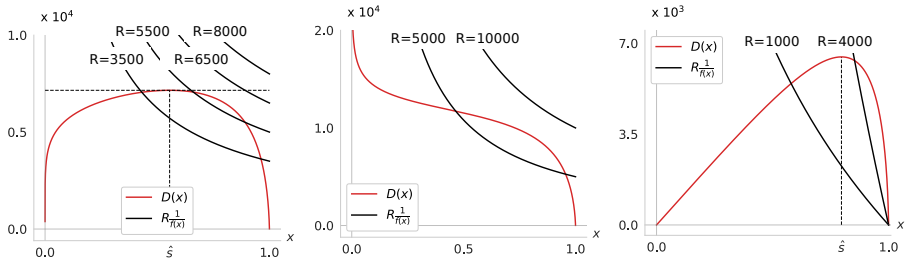
We now graphically illustrate various scenarios related to the existence of Fixed Points, and also provide corroboration of Prop. 6.1 and 6.2.

Fig. 6.1 shows plots of the left and right-hand sides of the FPE in (6.1). The role of R is highlighted. Fig. 6.1a and 6.1b are for GL Class I with $\gamma < 1$ and $\gamma > 1$, respectively, and Fig. 6.1c is for GL Class II. Observe that for $\gamma > 1$, $D(s) \to \infty$ as $s \to 0$, and $D_s(s) < 0,\ \forall s \in (0, 1)$ as stated in Prop. 5.2.

6.5 Attacker-Defender Model Examples

We proceed to obtain plots of the optimal Attacker and Defender investments for various initial vulnerabilities, v.

For Fig. 6.2a we consider a GL Class I function where $R = 5,000$, $\alpha = 0.0001$, and $\beta = 1.1$. Since the Effective Loss to Gain Ratio, $R = \frac{L/d}{G/c}$, we let $L = \$100,000$, $G = \$70,000$, $d = \$1$, and $c = \$3,500$. That is, we assume that the Defender has more to lose than the Attacker has to gain, and, in the

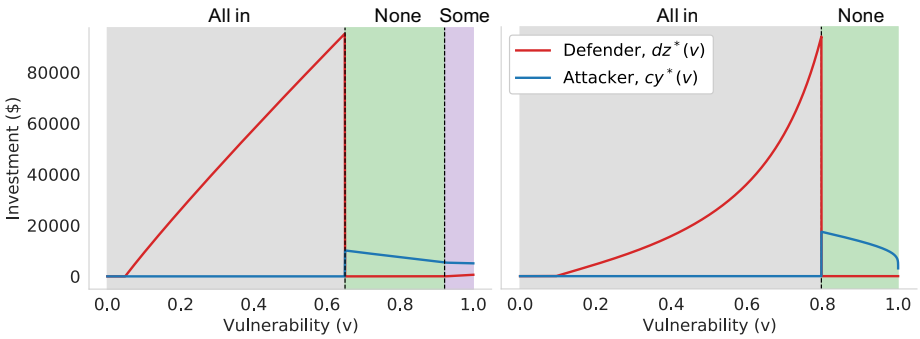


(a) $\beta = 1.2$, ($\gamma \approx 0.83$) (b) $\beta = 0.9$, ($\gamma \approx 1.1$) (c) $\gamma = 0$

Fig. 6.1: Visualizing the existence of Fixed Points of GL Class I and II functions, and the role of R; $\alpha = 0.0001$.

simplest representation, a single breach attempt by the Attacker costs $\$3,500$. In this case, the local minimum, $s_1(v)$, of the Defender's objective function plays an important role to induce an "all or some" policy in certain intervals of v. As v increases, the Defender's policies transition from "all in", to "all or none", and then to "all or some" at the $\hat{v}$ and v^H interval boundaries, respectively. This example highlights the complexity of behavior that arises from our simple two-sided model.

For Fig. 6.2b consider a GL Class II function where $R = 10,000$ and $\alpha = 0.0001$. Specifically, let $L = \$100,000$, $G = \$100,000$, $d = \$1$, and $c = \$10,000$. Since no solution to the FPE exists, the Defender's optimal decisions for all values of v lie entirely in Decision Interval 2, i.e., it should choose to either pay the Price of Deterrence or invest nothing.



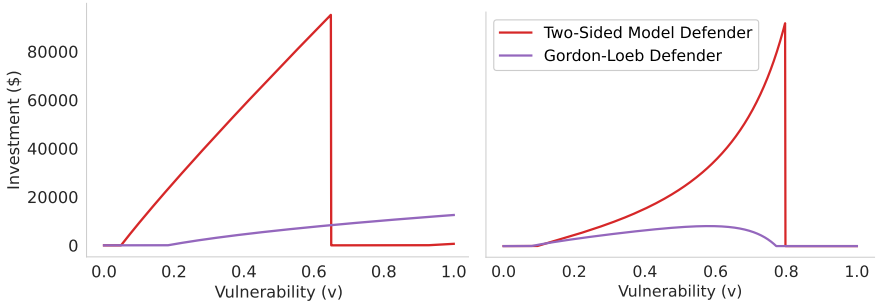
(a) GL Class I, $R = 5,000$, $\alpha = 0.0001$, $\beta = 1.1$ (b) GL Class II, $R = 10,000$, $\alpha = 0.0001$

Fig. 6.2: Attacker and Defender's optimal investments and Defender's decision types vs. initial system vulnerability, v.

6.6 Comparisons with the Gordon-Loeb Model

There are several interesting features worth highlighting in Fig. 6.2a and 6.2b. In contrast to the results from the model of Gordon-Loeb [7], optimal investment curves here are not guaranteed to be smooth. This reflects the stark transitions in decision type, e.g., from "all in" to "none", as the initial system vulnerability is varied.

Furthermore, we see a substantial departure from the celebrated $1/e$ result [7]. In fact, the results from our model indicate that a rational, risk-neutral Defender may find it in its best interest to invest up to the potential loss of L due to a system breach.[5] For an appreciation of these differences, compare the optimal Defender investment curve for our Attacker-Defender model with that from the Gordon-Loeb model shown in Fig. 6.3.



(a) GL Class I, $R = 5,000$, $\alpha = 0.0001$, $\beta = 1.1$

(b) GL Class II, $R = 10,000$, $\alpha = 0.0001$

Fig. 6.3: Comparison of Defender investments from the Gordon-Loeb and our two-sided models.

Observe that the inclusion of a rational, risk-neutral Attacker has a profound impact on the Defender's investment strategy. We find that, in general, for broad ranges of the system's initial vulnerability the optimal investments from the Gordon-Loeb model [7] underestimate what is necessary to prudently defend the system. This is because any attack, regardless of severity, serves to increase system breach likelihood (and therefore increase the Defender's expected loss) by some degree. Thus, we argue that it is crucial to consider the goals and resources of an Attacker so as to capture this far-reaching characteristic of cybersecurity and arrive at more realistic estimates of investments in system security.

[5] Of note is that the upper limit to rational investment is $LT(S(0, v))$, the Defender's expected loss taking into consideration the reaction of the Attacker, however $T(S(0, v))$ may tend arbitrarily close to 1.

7 Conclusions

We conclude with some observations on directions for future work. We have noted earlier that the estimates of potential gain (G), loss (L) and unit costs (c, d) play important roles, via the Effective Loss to Gain Ratio parameter R, in the determination of the optimal investment policy. It follows that a non-standard dimension of cyberwarfare is the promotion of false estimates by adversaries, and a better understanding of how to deal with this possibility will be useful. We have assumed a risk-neutral Defender, and a question of interest, albeit one that will introduce added complexity, is the nature of the optimal investment for a risk-averse Defender. In a similar vein, it will be insightful to obtain the Defender's decision based on worst-case analysis to allow for a possibly irrational Attacker.

The deployment costs of attacks and defense have been modeled as linear functions of the respective efforts. Realistically the costs may be expected to be nonlinear, and to also include large fixed costs, which are known to significantly affect market structures and pricing [33]. It should be worthwhile to delineate areas in cybersecurity where fixed costs matter and where they do not.

A natural progression of our static problem formulation would be to dynamic, sequential Stackelberg games. Much is known [24], but not where the focus is on fundamental understanding in the context of the economics of cybersecurity. Other avenues for generalizing our model is to networked systems, where attack graphs provide a possible framework [26]. Similarly, Moving Target Defense [27] is also a promising direction for generalization.

Appendix

A.1

Proof of Prop. 5.1: Taking the derivative with respect to s,

$$s^2 H_s(s;\gamma) = \frac{-2\left\{s^2 - (1-s)\log^2(1-s)\right\}}{(1-s)\log^2(1-s)} - \{1-\gamma(s)\} - \gamma_s(s)s(1-s) \tag{A1.1}$$

Now denoting the numerator of the first term on the right hand side (rhs) by $-2F(s)$, it can be shown that,

$$F(s) = s^2 - (1-s)\log^2(1-s) > 0, \quad s \in (0,1), \qquad = 0, \quad s = 0. \tag{A1.2}$$

To see this, observe that $F(s)$ is strictly convex since,

$$\frac{1}{2}(1-s)F_{ss}(s) = -\{s + \log(1-s)\} > 0, \quad s \in (0,1)$$

Also, $F_s(0) = 0$, hence $F_s(s) > 0,\ s \in (0,1)$. Since $F(0) = 0$, (A1.2) follows.

Using (A1.2) and (A1.1), it follows that if $0 \le \gamma(s) < 1$ and $0 \le \gamma_s(s),\ \forall s \in [0,1]$, then $H_s(s;\gamma) < 0,\ s \in [0,1]$. Also, $0 \le H(0;\gamma)$ and $H(1;\gamma) < 0$. Hence

there exists a unique $\hat{s} \in (0,1)$ such that $H(s;\gamma) = 0$, i.e., $D_s(\hat{s}) = 0$ and $D_s(s) > (<)\ 0$ for $s < (>)\ \hat{s}$.

Proof of Prop. 5.2: For $1 < \gamma(s),\ s \in [0,1]$, from (4.6),

$$\begin{aligned}[-s\log(1-s)]\, H(s;\gamma) &= 2s + (2-s)\log(1-s) \\ &\quad + \{\gamma(s)-1\}\,(1-s)\log(1-s) && \text{(A1.4)} \\ &\le 2s + (2-s)\log(1-s) && \text{(A1.5)} \\ &< 0, \quad s \in (0,1] && \text{(A1.6)}\end{aligned}$$

To see how (A1.6) follows from (A1.5), let $h(s)$ denote the expression on the rhs of (A1.5), and verify that $h_{ss} < 0$, i.e., $h(s)$ is strictly concave in $(0,1]$. Also, $h_s(0) = 0$ and $h(0) = 0$, which in combination with strict concavity yields (A1.6).

Since the bracketed term in the lhs in (A1.4) is positive for $s \in (0,1]$, we have $H(s;\gamma) < 0$ and $D_s(s) < 0$, $s \in (0,1]$. In the definition of $D(s)$ in (5.2, ii), note that as $s \to 0$, $\log^2(1-s) \sim s^2$, and, from (2.12-2.13), $g(s) = O(s^\delta)$. Hence, as $s \to 0$,

$$D(s) \to O(s^{2-\delta}) \tag{A1.7}$$

From (2.13), when $\gamma(0) > 1$, $\delta > 2$, and consequently $D(s) \to \infty$ as $s \to 0$.

Proof of Prop. 5.3: Now suppose $\gamma(0) \le 1$, $\forall s \in [0,1]$, then Prop. 5.2 will apply. Next, suppose the following alternative: $\gamma(0) < 1$ and $\exists s_c, s_c \in (0,1]$ such that $\gamma(s) < 1$ for $s < s_c$ and $\gamma(s) > 1$ for $s > s_c$, i.e., there is a cross-over at s_c. In this case, $\gamma(s) > 1$, $\forall s \in [s_c, 1]$. Continuing with this case,

(i) Since $\gamma(s) > 1$, $s \in (s_c, 1]$, from Prop. 5.2, $H(s;\gamma) < 0$. In particular, $H(s_c,\gamma) < 0$.
(ii) For $s < s_c$, since $\gamma(s) \le 1$, and $\gamma_s(s) \ge 0$, from Prop. 5.1, $H_s(s;\gamma) < 0$, $s \in [0, s_c]$.

Combining (i) and (ii), since $H(0;\gamma) > 0$ and $H(s_c;\gamma) < 0$, it follows that $\exists s_B, s_B < s_c$, such that $H(s_B;\gamma) = 0$. Hence from (5.6),

$$\frac{dD(s_B)}{ds} = 0 \tag{A1.8}$$

Since $D_s(s) > (<)\ 0$ for $s < (>)\ s_B$, $D(s)$ is inverted-U shaped with its maximum at s_B.

A.2

Proof of Prop. 6.1: First we show that,

$$s_1(\hat{v}) = \hat{s} < \hat{v} \tag{A2.1}$$

To see this, note from the defining relation $D(\hat{s}) = \frac{R}{f(\hat{v})}$ that,

$$\xi(\hat{s}) = R\alpha\beta \left(\frac{\hat{s}}{\hat{v}}\right)^{1/\beta} \tag{A2.2}$$

since $D(s) = \xi(s)/(\alpha\beta s^{1/\beta})$. Now suppose, contrary to (A2.1), $\hat{s} \geq \hat{v}$. Then from (A2.2), $\xi(s) \geq R\alpha\beta$, which contradicts $\hat{\xi} < R\alpha\beta$.

If $\hat{\xi} < R\alpha\beta$, then from (6.2, ii), clearly no Fixed Point exists. Now noting (A2.1), in the absence of any Fixed Point, it must be that $s_1(v) < v$, $\forall v \in (\hat{v}, 1)$. Hence (i) is proved.

Next we show that,

$$\hat{s} < \hat{s}_\xi \quad \text{and} \quad \xi(\hat{s}) < \hat{\xi} \tag{A2.3}$$

From the relation between $D(s)$ and $\xi(s)$ given above,

$$\alpha\beta D_s(s) = \frac{1}{s^{1/\beta}} \left[\xi_s(s) - \frac{1}{s}\xi(s) \right]. \tag{A2.4}$$

Since $D_s(\hat{s}) = 0$, it follows that,

$$\xi_s(\hat{s}) = \frac{1}{\hat{s}}\xi(\hat{s}) > 0 \tag{A2.5}$$

Hence, from the inverted-U shape of $\xi(s)$ and the fact that its maximum value of $\hat{\xi}$ is when $s = \hat{s}_\xi$, (A2.3) follows.

If $R\alpha\beta < \hat{\xi}$, then from the inverted-U shape of $\xi(s)$, it follows that there exist two solutions, ξ_1 and ξ_2, to the solution $\xi(s) = R\alpha\beta$, and we let $\xi_2 < \hat{s}_\xi < \xi_1$. Making use of (A2.3), the possibilities are,

$$\hat{s} < \xi_2 < \hat{s}_\xi < \xi_1, \quad \text{and} \quad \xi_2 < \hat{s} < \hat{s}_\xi < \xi_1. \tag{A2.6}$$

In the former case, two Fixed Points exist, $v^L = \xi_2$ and $v^H = \xi_1$, and in the latter case only one Fixed Point exists, $v^H = \xi_1$. (In the latter case, since $\xi_2 < \hat{s}$, ξ_2 violates the constraint that Fixed Points satisfy.) We can precisely characterize the separation of the two cases in (A2.6). The two cases correspond respectively to,

$$\xi(\hat{s}) < \xi(\xi_2) = R\alpha\beta < \hat{\xi}, \quad \text{and} \quad \xi(\xi_2) = R\alpha\beta < \xi(\hat{s}) < \hat{\xi}. \tag{A2.7}$$

The former case corresponds to (ii) of the Proposition, and the latter to (iii).

In both cases (ii) and (iii), the sign of $\{s_1(v) - v\}$ between the transition points follows from the number of Fixed Points in each case, in combination with the sign of $\{s_1(\hat{v}) - \hat{v}\}$, which has been established in (A2.1).

Acknowledgements

DM gratefully acknowledges the support of the National Institute of Standards and Technology, and, especially, NIST's Applied and Computational Mathematics Division. DM also gratefully acknowledges the benefit of discussions with Vladimir Marbukh.

References

1. D. Clark, T. Berson, H. Lin, At the nexus of cybersecurity and public policy: Some basic concepts and issues, National Academies Press (2014).

2. Z. M. Smith, E. Lostri, The hidden costs of cybercrime, https://www.mcafee.com/enterprise/en-us/assets/reports/rp-hidden-costs-of-cybercrime.pdf (2020).
3. IBM, Cost of a data breach report 2021, https://www.ibm.com/downloads/cas/OJDVQGRY (2021).
4. NVD, National vulnerability database, https://en.wikipedia.org/wiki/National_Vulnerability_Database (2021).
5. P. Mell, K. Scarfone, S. Romanosky, et al., A complete guide to the common vulnerability scoring system version 2.0, in: Published by FIRST-forum of incident response and security teams, Vol. 1, 2007, p. 23.
6. J. M. Danskin, The theory of max-min and its application to weapons allocation problems, Vol. 5, Springer Science & Business Media, 2012.
7. L. A. Gordon, M. P. Loeb, The economics of information security investment, ACM Transactions on Information and System Security (TISSEC) 5 (4) (2002) 438–457.
8. Y. Baryshnikov, It security investment and gordon-loeb's 1/e rule., in: WEIS, 2012.
9. M. Lelarge, Coordination in network security games: a monotone comparative statics approach, IEEE Journal on Selected Areas in Communications 30 (11) (2012) 2210–2219.
10. R. S. Gibbons, Game theory for applied economists, Princeton University Press, 1992.
11. R. T. Rockafellar, Lagrange multipliers and optimality, SIAM review 35 (2) (1993) 183–238.
12. R. Anderson, Why information security is hard-an economic perspective, in: Seventeenth Annual Computer Security Applications Conference, IEEE, 2001, pp. 358–365.
13. A. Fedele, C. Roner, Dangerous games: A literature review on cybersecurity investments, Journal of Economic Surveys 36 (1) (2022) 157–187.
14. K. Hausken, Returns to information security investment: The effect of alternative information security breach functions on optimal investment and sensitivity to vulnerability, Information Systems Frontiers 8 (5) (2006) 338–349.
15. D. C. Huang, Q. Hu, R. S. Behara, An economic analysis of the optimal information security investment in the case of a risk-averse firm, International Journal of Production Economics 114 (2008) 793–804.
16. G. Brown, M. Carlyle, J. Salmeron, K. Wood, Defending critical infrastructure, Interfaces 36 (6) (2006) 530–544.
17. R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. Van Eeten, M. Levi, T. Moore, S. Savage, Measuring the cost of cybercrime, The economics of information security and privacy (2013) 265–300.
18. R. K. Tiwari, K. Karlapalem, Cost tradeoffs for information security assurance., in: Workshop on the Economics of Information Security (WEIS), 2005.
19. H. Tanaka, K. Matsuura, O. Sudoh, Vulnerability and information security investment: An empirical analysis of e-local government in japan, Journal of Accounting and Public Policy 24 (1) (2005) 37–59.
20. A. Gueye, V. Marbukh, A game-theoretic framework for network security vulnerability assessment and mitigation, in: Decision and Game Theory for Security: Third International Conference, GameSec 2012, Budapest, Hungary, November 5-6, 2012. Proceedings 3, Springer, 2012, pp. 186–200.
21. M. Dziubinski, S. Goyal, Network design and defence, Games and Economic Behavior 79 (2013) 30–43.
22. S. Goyal, A. Vigier, Attack, defence and contagion in networks, The Review of Economic Studies 81 (4) (2014) 1518–1542.

23. D. Acemoglu, A. Malekian, A. Ozdaglar, Network security and contagion, Journal of Economic Theory 166 (2016) 536–585.
24. M. Breton, A. Alj, A. Haurie, Sequential stackelberg equilibria in two-person games, Journal of Optimization Theory and Applications 59 (1988) 71–97.
25. W. Mazurczyk, L. Caviglione, Cyber reconnaissance techniques, Communications of the ACM 64 (3) (2021) 86–95.
26. L. Wang, T. Islam, T. Long, A. Singhal, S. Jajodia, An attack graph-based probabilistic security metric, in: Data and Applications Security XXII: 22nd Annual IFIP WG 11.3 Working Conference on Data and Applications Security London, UK, July 13-16, 2008 Proceedings 22, Springer, 2008, pp. 283–296.
27. C. Lei, H.-Q. Zhang, J.-L. Tan, Y.-C. Zhang, X.-H. Liu, Moving target defense techniques: A survey, Security and Communication Networks 2018 (2018).
28. W. J. Baumol, J. C. Panzar, R. D. Willig, Contestable markets: An uprising in the theory of industry structure: Reply, The American Economic Review 73 (3) (1983) 491–496.
29. J. C. Panzar, Technological determinants of firm and industry structure, Handbook of industrial organization 1 (1989) 3–59.
30. J. Swift, Soviet-american arms race, `https://www.historytoday.com/archive/soviet-american-arms-race` (2009).
31. I. Goodwin, The price of victory in cold war is $5.8 trillion for nuclear arms and delivery systems, says panel., Physics Today 51 (8) (1998) 49–51.
32. P. Aghion, N. Bloom, R. Blundell, R. Griffith, P. Howitt, Competition and innovation: An inverted-u relationship, The quarterly journal of economics 120 (2) (2005) 701–728.
33. R. R. Braeutigam, Optimal policies for natural monopolies, Handbook of industrial organization 2 (1989) 1289–1346.