


January 2021

## Deterrence in Cyberspace: A Game-Theoretic Approach

Abderrahman Sokri

*Center for Operational Research and Analysis*, [abderrahmane.Sokri@unomaha.edu](mailto:abderrahmane.Sokri@unomaha.edu)

Follow this and additional works at: <https://digitalcommons.unomaha.edu/spaceanddefense>

 Part of the [Asian Studies Commons](#), [Aviation and Space Education Commons](#), [Defense and Security Studies Commons](#), [Eastern European Studies Commons](#), [International Relations Commons](#), [Leadership Studies Commons](#), [Near and Middle Eastern Studies Commons](#), [Nuclear Engineering Commons](#), [Science and Technology Studies Commons](#), and the [Space Vehicles Commons](#)

Please take our feedback survey at: [https://unomaha.az1.qualtrics.com/jfe/form/SV\\_8cchtFmpDyGfBLE](https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE)

### Recommended Citation

Sokri, Abderrahman (2021) "Deterrence in Cyberspace: A Game-Theoretic Approach," *Space and Defense*: Vol. 12: No. 0, Article 6.

DOI: 10.32873/uno.dc.sd.12.01.1084

Available at: <https://digitalcommons.unomaha.edu/spaceanddefense/vol12/iss0/6>

This Article is brought to you for free and open access by DigitalCommons@UNO. It has been accepted for inclusion in Space and Defense by an authorized editor of DigitalCommons@UNO. For more information, please contact [unodigitalcommons@unomaha.edu](mailto:unodigitalcommons@unomaha.edu).

---

## Deterrence in Cyberspace: A Game-Theoretic Approach

### Cover Page Footnote

Abderrahmane Sokri is data scientist at Defence Research and Development Canada, Center for Operational Research and Analysis. Copyright for this article remains with the Crown, Canada.

## Deterrence in Cyberspace: A Game-Theoretic Approach

Abderrahmane Sokri

*This novel application of the Stackelberg leader-follower game from economic theory illuminates situational constraints that point to a sweet spot, an optimal level of investment in cyber defense, for deterrence by denial.*

Deterrence is a form of persuasion intended to manipulate the cost-benefit analysis of would-be attackers and convince them that the cost of taking an action against the defender outweighs its potential benefit (Brantly, 2018; Wilner, 2017).<sup>1</sup> It is the prevention (of a target) from committing unwanted behavior by fear of the consequences (United States (US) Department of Defense (DoD), 2008; Taipale, 2010). Deterrence differs from compellence by focusing on prevention using *ex ante* actions. Compellence uses power to force an adversary, *post hoc*, to take a desired action under threat of possible escalation in the future (Brantly, 2018).

Two types of deterrence are generally used: deterrence by punishment and deterrence by denial. Deterrence by punishment hinges on the threat of retaliation against a potential attacker. This tit-for-tat or equivalent retaliation strategy adds to the attacker's perceived cost. Deterrence by denial sends a signal to potential challengers that they will be unsuccessful. This impenetrability strategy subtracts from the attacker's perceived benefits.

In the physical world, deterrence aims to dissuade specific actions against physical assets. In this space, the most common form of deterrence by punishment is the use of nuclear weapons. These weapons are inherently an existential threat against

potential challengers (Brodie et al., 1946; Brantly, 2018). An all-out nuclear war could be threatened but never fought to achieve reasonable political objectives (Freedman, 2004; Brantly, 2018). Deterrence by denial may include tightening defense around a critical infrastructure to deny attacker access. The target can be tightly defended by installing, for example, more security mechanisms and higher walls.

In the cyber domain, deterrence is more complex than in the physical domain. Digital attacks go beyond geographic and political boundaries. They are generally highly dynamic and imperceptible to the human senses (Moisan and Gonzalez, 2017; Sokri, 2019b). A cyber-attack may result in interception, degradation, modification, interruption, fabrication, or unauthorized use of an information asset. The information asset can be physically (e.g., hardware) or logically (e.g., software) based (Sokri, 2019a).

Cyber-attacks can be segregated into two main categories: targeted attack and opportunistic attack. A targeted attack requires a large effort and has the potential to cause significant damage to the defender. Denial of service and theft of information are typical targeted attacks. In contrast, an opportunistic attack has a number of intermediate targets, requires a small effort, and tends to cause less damage. A virus and spam e-mail are typical opportunistic attacks.

---

<sup>1</sup> Abderrahmane Sokri is data scientist at Defence Research and Development Canada, Center for

Operational Research and Analysis. Copyright for this article remains with the Crown, Canada.

The most challenging problem in cyber deterrence is the attribution dilemma (Wilner, 2017). Determining who to blame for an attack may be very difficult and time-consuming to do. Consequently, the credibility of any deterrence by punishment in digital space will depend on the blame attribution. (Glaser, 2011; Brantly, 2018). Since deterrence by denial does not require identification of potential attackers, it can be used to mitigate this dependency (Bordelon, 2016).

Cyber risk is present when a given threat meets a vulnerability in an information system allowing it to manifest. In this context, a threat is a potential cause of an unwanted occurrence while a vulnerability is a weakness in the information system (Sokri, 2019a; Zhang, 2012; Bowen et al., 2006). To minimize digital risk against an information asset, the defender should know at least two elements: (1) the probability of a successful attack and (2) the corresponding potential loss (Brantly, 2018; Glaser, 2011; Schneidewind, 2011; Branagan, 2012).

To protect their information assets against offensive cyber-attacks, policy makers are increasingly gravitating towards deterrence by denial (Taipale, 2010). A key decision-variable in digital deterrence by denial is the defender investment level in security. To protect a potential target, the defender can reduce the probability of a successful attack by investing in information security. The investment may, for example, reduce the vulnerability of the target.

The aim of this paper is to show how deterrence by denial as a defense strategy can be formulated in cyberspace using a sequential game with a disclosure mechanism. It shows the suitability of game theory to cyber deterrence. The paper extends existing

models by providing a new game formulation of deterrence using a more intuitive probability of a successful attack. It also combines stochastic simulation and game-theoretic approaches to handle uncertainty in the input data. A simulation could, for example, incorporate uncertainty on the model variables and parameters by changing their static values to statistical distributions.

Consider a sequential security game played between two adversarial agents: a defender  $D$  (the leader) and a strategic attacker  $A$  (the follower). The defender anticipates the attacker's reaction, determines, and credibly communicates the security investment to protect an information system. The defender can, for example, publicly release his level of investment in (1) detection and prevention techniques such as Antivirus software, Firewalls, and Intrusion Detection Systems (IDS) and (2) physical monitoring and inspection procedures (Sokri, 2019b). Revenue agencies usually use this tactic by revealing their auditing strategies to deter tax evasion (Cavusoglu et al., 2008).

The attacker observes the defender's decision and reacts with a certain level of willingness-to-attack. The true willingness-to-attack is latent and, therefore, not directly observable. It is modeled as the expected effort to be exerted by the attacker to compromise the system. The attacker's effort corresponds to the first activities of the cyber kill chain (Mihai et al., 2014). These activities particularly include (but are not limited to):

1. **Reconnaissance** – the process of collecting information about the system,
2. **Weaponization** – the process of analyzing the collected data to select the appropriate attack technique, and
3. **Delivery** – the process of transmitting the weapon to the targeted system.

Following this introduction, section 2, below, provides a comprehensive review of literature on security investment as a deterrence factor. Section 3, sets up a new game theoretic model of deterrence in cyberspace. Section 4, computes the Stackelberg equilibrium. Section 5 offers a formal discussion about the main results. Some concluding remarks are indicated in section 6.

## LITERATURE REVIEW

Identifying and understanding the factors influencing the decision to invest in information security is a key requirement for any effective deterrence and risk management in cyberspace. These factors form the pillars of the appropriate level of security investment. Security investment as a deterrence factor has been an active research area in the last decade. This literature can be divided into two main categories: decision theory and game theory approaches (Cavusoglu et al., 2008).

The decision-theoretic approach uses traditional risk analysis and cost–benefit perspectives for security investment decisions. This approach assesses the risk associated with security breaches and conducts a cost-benefit analysis to determine a certain level of security investment to mitigate the risk. While this approach can assess the economic value of intangible costs and benefits, it has two main limitations: (1) It does not determine the optimal security investment level. (2) It does not allow a defender’s security investment to influence the attacker’s behaviour.

Al-Humaigani and Dunn (2003), for example, proposed a model to quantify the return on security investment (ROSI). The authors enumerated the fundamental components of ROSI for every organization and security

threat. They included what it costs to invest in information security spending (e.g., the cost of procuring the security tool or software, the losses in reputation and goodwill). They incorporated both the pre- and post- system implementation security measures.

In order to come through the first limitation of the decision-theoretic approach, Gordon and Loeb (2002) presented an economic model that determines the optimal amount to invest in information security. Their results indicate that defenders may be better off concentrating their efforts on information assets with midrange vulnerabilities. Extremely vulnerable information assets may be very expensive to protect. For some broad classes of security breach probability functions, results also indicate that optimal investment never exceeds 37% of the expected loss. Hausken (2006) examined the effect of different returns assumptions on the optimal level of investment. The author showed that optimal investment level may no longer be capped at 37% of expected loss. For an alternative class of security breach probability functions, the optimal investment can increase convexly in vulnerability and exceed 37%.

More recently, Mayadunne and Park (2016) used the expected utility approach to analyze information security investment decisions. They provided a comparison between the decisions made by a risk taking and a risk neutral decision maker. They found, for example, that for a group of information assets with equal value and varying vulnerabilities, the risk neutral decision maker will diversify security investment to a greater extent and the risk taker will invest a larger amount when protecting the high risk assets in the group.

The game-theoretical approach uses game oriented models to capture the strategic interactions between rational attackers and

defenders. Optimal investment in security is one of the defenders' resulting strategies. This approach has two main challenges: (1) Validity of the game-theoretic assumptions in cyberspace (e.g., rationality of players). (2) Complexity of the cyber domain scenarios (e.g., dynamic attacks and complex networks).

Cavusoglu et al. (2008), for example, argued that the old decision-theoretic approach is incomplete because it does not take into account the strategic nature of the interaction between attackers and defenders. The authors used a game-theoretic model to determine the optimal security investment level. Results indicate that the defender generally enjoys a higher payoff than that in the decision theory approach. The gap between the two results decreases over time and the rate of convergence depends on the defender learning model.

Wu et al. (2015) used game theory to model the relationship between the optimal information security investment and the characteristics of defenders' security environment. Results indicate that defenders are better off not investing in security (outside best practices) until the potential loss reaches a certain value. They should focus on the midrange of intrinsic vulnerabilities. When the potential loss is catastrophic, they should adopt other measures and stop investing in security.

More recently, Pan et al. (2017) suggested an optimal investment strategy using a game-theoretic framework. The authors concluded that the defender is better off using a single security level to protect all the information assets instead of using different security levels to protect different assets. The interested reader is referred to Sokri (2019a) and Sokri (2019b) for further information on game theory in cyber defense.

## A "STACKELBERG" DETERRENCE MODEL

The system is characterized by an inherent vulnerability  $v_0$ . Each successful attack can result in a potential loss  $l$  to the defender and a possible benefit  $b$  to the attacker. The loss/benefit occurring can be tangible (e.g., monetary loss/benefit) or intangible (e.g., loss/gain in reputation).

### *Probability of a successful attack*

Let  $i$  be the defender's security investment and  $t$  the attacker's level of effort to expend in hacking the defender. The compound probability  $p$  of a successful attack can be expressed as the product of the probability that the vulnerability may be exploited,  $v(i)$ , and the threat probability (i.e., the probability to receive an attack) (Wu et al, 2015):

$$(1) \quad p = v(i) \left( 1 - \exp\left(-\frac{t}{\mu}\right) \right),$$

where the expected effort  $t$  can be expressed in terms of time. The threat probability, also known as the probability of attack (prior to information about target vulnerability), is written in Equation 1 as the cumulative distribution function (CDF) of an exponentially distributed random variable evaluated at  $t$ . This CDF estimates the probability that the attacker's level of effort will be less than  $t$ . The parameter  $\mu$  represents the mean effort to attack (e.g., investigation, identification, weaponization done prior to knowledge of target defenses). It also represents the standard deviation of the distribution.

As in Wu et al. (2015), the defender's security investment does not directly affect the inherent threat probability. The defender can only reduce the first term, probability that the vulnerability may be exploited, using security investment  $i$ . That is,

$$(2) \quad v(i) = v_0 \exp(-\alpha i),$$

where the parameter  $\alpha > 0$ . Straightforward derivation leads to

$$(3) \quad \frac{v'(i)}{v(i)} = -\alpha,$$

which means that the parameter  $\alpha$  is the decay rate of the probability that the vulnerability may be exploited. It represents the rate at which vulnerability decreases with investment in cybersecurity. It can also be seen as a measure of investment productivity. It measures how efficiently security investment is used to reduce the asset vulnerability.

One can also readily see that  $v(i)$  satisfies the following three assumptions.

- Assumption 1.  $v(0) = v_0$ .
- Assumption 2.  $\lim_{i \rightarrow \infty} v(i) = 0$ .
- Assumption 3.  $v'(i) = \frac{dv(i)}{di} < 0$ ,  
 $v''(i) = \frac{d^2v(i)}{di^2} > 0, \forall i$ .

Assumption 1 states that if there is no investment in security, the vulnerability of the system will be the inherent vulnerability. Assumption 2 states that no finite investment can eradicate the vulnerability from information systems. Because of their complexity, perfect security is impossible (Wu et al., 2015). Assumption 3 states that the investment in security reduces the probability that the vulnerability may be exploited, but at a decreasing rate. Investment makes the system more secure, but with declining marginal return.

The probability of vulnerability exploitation is formulated in Equation 2 as an exponentially decreasing function of the security investment. Consequently, the probability of a *successful* attack can now be written as

$$(4) \quad p(i, t) = v_0 \exp(-\alpha i) \left( 1 - \exp\left(-\frac{t}{\mu}\right) \right).$$

This probability depends on the defender investment level and the attacker's effort level, in addition to the system's inherent vulnerability.

#### *Defender's loss and attacker's payoff*

In this game the defender seeks to find the optimal security investment that minimizes the following total cost

$$(5) \quad W_D = p(i, t)l + i,$$

where the first term of its right-hand side is the defender's expected loss due to a successful attack. The attacker seeks to maximize the following payoff

$$(6) \quad W_A = p(i, t)b - t,$$

where the first term of the right-hand side is the attacker's expected benefit and the second term represents the expected effort to compromise the system.

#### *Deterrence game's equilibrium*

This section characterizes the optimal solution to the deterrence game. As in the standard Stackelberg competition, the game is sequential: the defender moves first, committing to a strategy before the attacker reacts. The defender's strategic choice is to select the optimal security investment (deterrence by denial). The attacker's choice is to determine his appropriate level of effort. The outcome of this leader-follower interaction is called Stackelberg equilibrium. This equilibrium has been recognized as a sound theoretical framework for modeling the strategic interactions between attackers and defenders (Jain et al., 2010; Korzhyk et al., 2011; Kiekintveld et al., 2015; Acquaviva, 2017).

*Proposition 1.* The following condition is satisfied at equilibrium

$$(7) \quad \frac{\partial p(i,t)}{\partial t} b = -\frac{\partial p(i,t)}{\partial i} l.$$

**Proof.** Assuming an interior solution, the first-order condition (maximizing attacker payoff with respect to effort,  $t$ ) for the attacker optimization problem is

$$(8) \quad \frac{dW_A}{dt} = \frac{\partial p(i,t)}{\partial t} b - 1 = 0.$$

The optimality condition for the defender problem is

$$(9) \quad \frac{dW_D}{di} = \frac{\partial p(i,t)}{\partial i} l + 1 = 0.$$

Equations 8 and 9 lead to the equilibrium condition in the Proposition. ■

Fixing the defender's security investment to some strategy  $i$ , the first problem to be solved is to find the attacker's best response to  $i$ . In this optimization problem, the follower maximizes his expected benefit given  $i$ .

*Proposition 2.* Assuming an interior solution, the optimal effort the attacker is willing to exert is given by

$$(10) \quad t = -\alpha\mu i + \mu \ln\left(\frac{bv_0}{\mu}\right)$$

**Proof.** After substitution for  $p(i, t)$ , Equation 6 becomes

$$(11) \quad W_A = v_0 \exp(-\alpha i) \left(1 - \exp\left(-\frac{t}{\mu}\right)\right) b - t.$$

Computing the derivative of  $W_A$  with respect to  $t$ , equating to zero, and solving leads to the expression of  $t$  as a function of  $i$ . ■

*Proposition 3.* The attacker's level of effort is a decreasing function in the defender's investment.

**Proof.** The derivative of  $t$  with respect to  $i$  is

$$(12) \quad \frac{dt}{di} = -\alpha\mu < 0.$$

■  
*Proposition 4.* Assuming an interior solution, the defender optimal security investment level is given by

$$(13) \quad i = \frac{1}{\alpha} \ln(\alpha l v_0).$$

**Proof.** Equations 4 and 5 imply that

$$(14) \quad W_D = v_0 \exp(-\alpha i) \left(1 - \exp\left(-\frac{t}{\mu}\right)\right) l + i.$$

The expression of  $t$  in Equation (10) is equivalent to

$$(15) \quad \exp\left(-\frac{t}{\mu}\right) = \frac{\mu}{bv_0} \exp(\alpha i).$$

Substituting for  $\exp\left(-\frac{t}{\mu}\right)$  from Equation 15 in Equation 14, computing the derivative of  $W_D$  with respect to  $i$ , equating to zero, and solving provides the equilibrium strategy in the Proposition. ■

*Proposition 5.* The attacker's optimal level of effort is given by

$$(16) \quad t = \mu \ln\left(\frac{b}{\alpha\mu l}\right).$$

**Proof.** Substituting for  $i$  from Equation 13 in Equation 10 leads to the result. ■

*Proposition 6.* The defender should not invest in security beyond best practices until the potential loss reaches

$$(17) \quad l^* = \frac{1}{\alpha v_0}.$$



**Proof.** To have a positive investment,  $\ln(\alpha l v_0) > 0$ . This is possible only if  $\alpha l v_0 > 1$ , which leads to the condition in the Proposition. ■

*Proposition 7.* The attacker should not exert any effort until the potential benefit reaches

$$(18) \quad b^* = \alpha \mu l.$$

**Proof.** To have a positive effort,  $\ln\left(\frac{b}{\alpha \mu l}\right) > 0$ . This is possible only if  $\frac{b}{\alpha \mu l} > 1$  which leads to the condition in the Proposition. ■

*Proposition 8.* The defender's optimal security investment level is an increasing concave function of the potential loss,  $l$ .

**Proof.** As shown in Equations 19 and 20, the first derivative of  $i$  with respect to  $l$  is positive and the second derivative is negative, respectively.

$$(19) \quad i'(l) = \frac{di(l)}{dl} = \frac{1}{\alpha l} > 0.$$

$$(20) \quad i''(l) = \frac{d^2i(l)}{dl^2} = -\frac{1}{\alpha l^2} < 0.$$

Consequently,  $i$  is a concave function in  $l$  that increases at decreasing rate. ■

*Proposition 9.* The defender's optimal investment spent on information security as a fraction of potential loss  $l$  is given by

$$(21) \quad r(l) = \frac{i}{l} = \frac{1}{\alpha l} \ln(\alpha l v_0).$$

**Proof.** Dividing the expression of  $i$  in Equation 13 by  $l$  leads to the result.

*Proposition 10.* The fraction  $r(l)$  is an increasing function in the potential loss  $l$  for

$l \leq l^{**} = \frac{e}{\alpha v_0} \approx \frac{2.718}{\alpha v_0}$ . It is decreasing for  $l \geq l^{**}$  with a horizontal asymptote at  $y = 0$ .

**Proof.** The first derivative of  $r(l)$  with respect to  $l$  is

$$(22) \quad r'(l) = \frac{dr(l)}{dl} = \frac{1}{\alpha l^2} (1 - \ln(\alpha l v_0)).$$

It is straightforward to show that

$$r'(l) = 0 \text{ for } l = \frac{e}{\alpha v_0}, \quad r'(l) \geq 0, \text{ for } l \leq \frac{e}{\alpha v_0},$$

$$\text{and } r'(l) \leq 0, \text{ for } l \geq \frac{e}{\alpha v_0}.$$

Hence, the potential loss  $l$  that maximizes the fraction  $r(l)$  is given by

$$(23) \quad l^{**} = \frac{e}{\alpha v_0} \approx \frac{2.718}{\alpha v_0}.$$

Using the l'Hopital rule,

$\lim_{l \rightarrow \infty} r(l) = 0$ , which shows that  $r(l)$  has a horizontal asymptote at  $y = 0$ . ■

To deal with uncertainty in the input data, a Monte Carlo simulation could represent each uncertain parameter as a probability distribution.

## INVESTMENT IN CYBERSECURITY AT EQUILIBRIUM

A parsimonious game-theoretical model is used in this paper to characterize deterrence in cyberspace. A Stackelberg game is played to capture the strategic nature of this interaction and provide clear insights about it. The suggested mechanism involves disclosing the defender's investment information to the potential attacker. The game's logic and results crucially depend on the timings of each move. The defender moves first, anticipates the strategic behavior of the attacker, and decides on the security investment. The attacker observes the defender's level of investment and determines a certain effort level. By revealing the security investment strategy, the defender becomes able to control the attacker's

incentive and deter (or reduce the effort behind) potential attacks.

Assuming an interior solution, Proposition 1 characterizes the first-order optimality conditions for the defender and attacker strategies. It compares, at equilibrium, magnitude decline in expected defender loss from extra security investment to magnitude increase in expected attacker benefit from extra effort. Stackelberg interaction joins their fates.

At equilibrium, marginal reduction in defender's expected loss due to additional investment precisely balances marginal increase in the attacker's expected benefit attributable to additional effort. In order to reach this decision point, the attacker as follower must be able to measure the magnitude of loss to the defender from a successful cyber attack. In the Stackelberg interaction, attacker does have a clue from observing optimal defender security investment, which is tied to defender assessment of cost in the event of disruption. Physical properties of the cyber system's vulnerability must also be common knowledge.

Propositions 4 and 5 define the attacker's optimal level of effort and the defender's optimal investment, respectively. Proposition 4 relates the defender's strategy to three parameters:

- the inherent vulnerability  $v_0$
- the decay rate in the vulnerability due to investment  $\alpha$ , and
- the defender's potential loss  $l$ .

Proposition 5 shows that the attacker's strategy depends on two other parameters in addition to  $\alpha$  and  $l$ , namely the mean level of effort  $\mu$  (independent of system vulnerability) and the attacker potential benefit  $b$  from system disruption.

The derivative of the attacker's expected effort  $t$  with respect to the defender's investment  $i$  in Equation 12 indicates that the parameters  $\alpha$  and  $\mu$  and their interaction effect are the key factors in cyber deterrence, that is, in sharply affecting adversaries' attack plans through denial. Equation 12 shows that the higher the two parameters the more likely the attacker is to be deterred through additional defender investment. The parameter  $\alpha$  measures the speed at which security investment translates into a reduction of the asset's vulnerability to attacks. An increase in the parameter  $\alpha$  for any given level of investment will decrease the probability that inherent vulnerability may be exploited, lessen the probability of a successful attack, and, therefore, result in a reduction in the attacker's level of effort. At the same time, the influence of additional investment on reducing attacker effort even further will rise. Equation 12 also shows that opportunistic attacks (with small  $\mu$ ) are harder to influence than targeted attacks (with high  $\mu$ ). Extensive initial interest in the targeted system leads potential attackers to be discouraged at a steeper rate once they learn of additional defender investment.

Propositions 6, 8 and 9 characterize the defender's optimal security investment level  $i$  as a function of the potential loss  $l$ . These propositions highlight the following key findings:

- The defender should not invest in security beyond best practices until the potential loss reaches a given value;
- The optimal security investment increases with the expected loss at a decreasing rate;
- The optimal investment in security as a fraction of potential loss  $l$  has a horizontal asymptote at  $y = 0$ . This means that, for very large potential losses, the optimal amount to spend on information security does not keep

pace; it is far smaller than the potential loss.

These findings are on par with the deterrence literature. They are particularly consistent with the study conducted by Gordon and Loeb (2002).

The formalism in Equations 4, 5, and 6 is grounded theoretically such that the model could be repeated or extended using different probability distributions. Its underlying mathematics is clear and conceptually based. Variations of the probability distribution will provide qualitatively the same findings. The numerical values of these findings will, of course, depend on the values of the deterrence model parameters.

A myopic approach such as a simultaneous game or a decision-theoretic technique would produce different results. Under a simultaneous game, players make single decisions before seeing the other player's moves (as in the famous Prisoner's Dilemma [PD]) and possibly under incomplete information about the other player's payoff from certain outcomes. Attackers, for example, are not able to observe the outcome of previous actions before responding. The main characteristic of myopic approaches is the non-cooperative, monotonic relationship between defender investment level and attacker effort. Both players rationally defect in PD-type games. When one cost variable increases, the other increases and vice versa; net payoffs in equilibrium for both decline. In this situation, attackers are never deterred, *per se*, because myopic approaches lack disclosure mechanisms. A deeper understanding of this interaction will be generated in future works.

## CONCLUSION

Deterrence is used to prevent unwanted actions by influencing the cost-

benefit analysis of potential attackers. The most common form of deterrence in cyberspace is deterrence by denial. Deterrence by denial sends a signal to would-be attackers that they will be unsuccessful. In this defense strategy, the defender reduces the probability of a successful attack by investing in information security. While the credibility of deterrence by punishment depends on blame attribution, deterrence by denial does not require this knowledge.

This paper used a sequential game theoretic approach with a disclosure mechanism (Stackelberg competition) to formulate a deterrence strategy in cyberspace. It derived the defender's optimal security investment level and the attacker's level of effort. The factors influencing the decision to invest in cybersecurity were identified and discussed. To deal with uncertainty in the input data, the model invites parametric analysis using Monte Carlo simulation.

Results for the equilibrium indicate that effectiveness of the security investment ( $\alpha$ ) and the category of attack ( $\mu$ ) and their interaction effect are the key factors in cyber deterrence. The more effective the security investment in reducing vulnerability and the higher attacker initial interest in the target, the more likely attacker is to be deterred by additional investment. Targeted attacks aiming at significant damage to the defender are more manageable by security investment than opportunistic attacks.

The defender's optimal security investment level ( $i$ ) as a function of potential loss ( $l$ ) indicates that investment in cybersecurity as a deterrence strategy will top out after the middle part of losses. At very high levels of loss, there is a numbing effect; optimal investment does not change much with additional increments of loss.

Deterrence in the cyber domain is more complex than in the physical field. Further efforts should be undertaken to understand it in order to influence potential attackers' behaviors. Examples of such studies include (but are not limited to)

- application of the model to a real-world cyber-security problem using real-life parameters,
- analyzing the interaction between defenders and attackers in dynamic scenarios,
- assessing the risk to the defender of a disclosure strategy,
- including deception mechanisms to enhance security,
- developing models to deal with bounded rationality of human adversaries,
- combining game theoretic models such as this Stackelberg version with other techniques and tools to make the formalism more realistic and tractable; techniques may include numerical simulation and genetic algorithms; tools may consist of firewalls and anti-virus software.

## REFERENCES

- Al-Humaigani, M., and Dunn, D. (2003). A model of return on investment for information systems security. *IEEE 46th Midwest Symposium on Circuits and Systems*, Vol. 1, pp. 483-485.
- Acquaviva JR (2017). Optimal Cyber-Defence Strategies for Advanced Persistent Threats: A Game Theoretical Analysis. Master Thesis, the Pennsylvania State University.
- Bordelon, E.B. (2017). Approaching Cyber Warfare: Geopolitics, Deterrence, and International Law. Senior Honors Thesis 639, Cyber War and Geopolitics, Liberty University, Virginia.
- Bowen, P., Hash, J., and Wilson, M. (2006). *Information Security Handbook: A Guide for Managers*. National Institute of Standards and Technology (NIST) Special Publication 800-100.
- Branagan, M. (2012) A risk simulation framework for information infrastructure protection. Ph.D. Dissertation, Queensland University of Technology, Australia.
- Brantly, A.F. (2018). The cyber deterrence problem. 10<sup>th</sup> International Conference on Cyber Conflict (CyCon), IEEE, pp. 31-54.
- Cavusoglu, H., Raghunathan, S., and Yue, W.T. (2008). Decision-Theoretic and Game-Theoretic Approaches to IT Security Investment. *Journal of Management Information Systems*, Vol. 25:2, pp. 281-304.
- Freedman, L. (2004). *Deterrence*. Polity Press, Cambridge.
- Glaser, C.L. (2011). Deterrence of Cyber-attacks and US National Security. Report GW-CSPRI-2011-5, The George Washington University, Washington, D.C.
- Gordon, L. A., and Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC)*, Vol. 5:4, pp. 438-457.
- Hausken, K. (2006). Returns to information security investment: The effect of alternative information security breach functions on optimal investment and sensitivity to vulnerability. *Information Systems Frontiers*, Vol. 8:5, pp. 338-349.
- Jain M, Tsai J, Pita J, Kiekintveld C, Rathi S, Ordone, F, and Tambe, M. (2010). Software assistants for randomized patrol planning for the LAX airport police and the Federal Air Marshals Service. *Interfaces*, Vol. 40:4.
- Kiekintveld C, Lisy V, and Pibil R. (2015). Game-theoretic foundations for the strategic use of honeypots in network security. In *Cyber Warfare*. Springer, p. 81–101.
- Korzhyk D, Yin Z, Kiekintveld C, Conitzer V, and Tambe M. (2011). Stackelberg vs. Nash in

- Security Games: An Extended Investigation of Interchangeability, Equivalence, and Uniqueness. *Journal of Artificial Intelligence Research*, Vol. 41.
- Mayadunne, S., and Park, S. (2016). An economic model to evaluate information security investment of risk-taking small and medium enterprises. *International Journal of Production Economics*, Vol. 182, pp. 519-530.
- Mihai, I.C., Pruna, S. and Barbu, I.D. (2014). Cyber Kill Chain Analysis. *Int'l J. Info. Sec. & Cybercrime*, Vol. 3.
- Moisan, F. and Gonzalpez, C. (2017). Security under Uncertainty: Adaptive Attackers Are More Challenging to Human Defenders than Random Attackers. *Frontiers in Psychology*, Vol. 8:982.
- Pan, C., Zhong, W., and Mei, S. E. (2017). Investment strategy analysis of information systems with different security levels. IEEE 2nd International Conference on Big Data Analysis, pp. 703-708.
- Pereira, J. P., & Ferreira, P. (2011). Next Generation Access Networks (NGANs) and the geographical segmentation of markets. ICN Tenth International Conference on Networks, pp. 69-74.
- Schneidewind, N.F. (2009). *Systems and Software Engineering with Applications*. Wiley-IEEE Press.
- Sokri, A. (2019a). Cyber Security Risk Modelling and Assessment: A Quantitative Approach. Proceedings of the 19th European Conference on Cyberwarfare and Security (ECCWS19), 4-5 July 2019, Coimbra University, Portugal.
- Sokri, A. (2019b). Game theory and cyber defence. In: *Games in Management Sciences*, Pineau, P.-O. and Taboubi, S. (eds) Springer International Series in Operations Research & Management Science.
- Taipale, K.A. (2010). Cyber-deterrence. Law, Policy and Technology: Cyberterrorism, Information, Warfare, Digital and Internet Immobilization, IGI Global.
- US DoD (2008). Joint Publication 1-02, Department of Defense Dictionary of Military and Associated Terms, Washington, D.C.
- Wilner, A. (2017). Cyber deterrence and critical-infrastructure protection: Expectation, application, and limitation. *Comparative Strategy*, Vol. 36:4, pp. 309-318.
- Wu, Y., Feng, G., Wang, N., and Liang, H. (2015). Game of information security investment: Impact of attack types and network vulnerability. *Expert Systems with Applications*, Vol. 42:15-16, pp. 6132-6146.
- Zhang, J. (2012). Information Security Risk Management Framework: China Aerospace Systems Engineering Corporation. Master Thesis, University of South Australia.