Advance Persistent Threat Defense:

History and Future in Four Solutions to One Game

Daniel Arce
University of Texas at Dallas, GR 31
800 W. Campbell, Rd
Richardson, TX 75080
darrce@utdallas.edu

Abstract

Advance Persistent Threats (APTs) represent a cybersecurity concern for governments and corporations alike. Over the past two decades, modes of APT defense have progressed from firewalls to defense-in-depth to the cyber kill chain and zero trust. We examine the history and future of cyber defense through an analysis of four solutions to one game: the cybersecurity enforcement game. The solutions (Stackelberg leader-follower, Nash, maximin, and correlated equilibrium) involve different assumptions about the behavior of APTs and Targets and the information they possess about each other. Given APTs may adjust their behavior in response to a change in the behavior of Targets, and vice-versa, we show increasing the sophistication of cybersecurity defense need not increase the Target's welfare.

1. Introduction

Cybersecurity plays an increasing role in national security. Howard (2023) defines cybersecurity as reducing the probability of material impact to an individual or organization due to a cyber event. Of particular concern for the Cybersecurity and Infrastructure Security Agency and the MITRE ATT&CK framework is the behavior of Advanced Persistent Threats (APTs).[1] APT is a term originally applied by security analyst Sean Carpenter at Sandia National Laboratories to China's 'Titan Rain' cyber breach campaign, which affected target nations and private enterprises during 2003 – 2006.

APTs are nation-state threat actors possessing resources and specialized skills to tailor methods for patiently probing specific targets for weaknesses; gain and maintain unauthorized access to systems to return at a later date; and compromise cryptographic protocols (Arce 2023). They are highly skilled and well-funded (possibly nation-sponsored) groups of cyber-attackers keen on long-term operational efforts and endowed with technological sophistication and R&D capabilities (Gilad, Pecht, and Tishler 2021). Stuxnet is an example of an APT-deployed virtual worm with an unprecedented four zero-day Windows exploits specifically designed to have physical consequences (Chen 2010). In addition, Stuxnet limited its rate of spread to maintain stealth. More recently, during 2024 into 2025 the startling sophistication of the TTPs involved in the Salt Typhoon and Flax Typhoon (HAFNIUM) cyber campaigns facilitated access to the U.S. Treasury, and telecoms and critical infrastructure in Australia, Japan, Vietnam, and the U.S, among myriad other targets.

Owing to the Internet being an "almost unique" computer system due to "varying relationships of collaboration and competition with each other," Papadimitriou (2001) endorses a game-theoretic approach to understanding the Internet. In the same year, Anderson (2001) makes an argument for using game theory in the context of information security, with Anderson and Moore (2006) claiming, "the tools of game theory and microeconomic theory are becoming just as important as the mathematics of cryptography to the security engineer." These claims have proved to be prescient; presently upwards of 40 reviews exist on the application of game theory to cybersecurity (Arce 2023). Recent

[1] ATT&CK is an acronym for Adversarial Tactics, Techniques, & Common Knowledge.

examples are Fedele and Roner (2022) and Collins, Xu, and Brown (2025).

Cybersecurity defense has evolved from firewalls to defense-in-depth to kill chains and zero trust. What is the effect of the change in defensive stance on the welfare of its practitioners? While each step in the evolution involves more sophistication on the part of the Target, malicious actors as well exhibit increasingly sophisticated tactics, techniques, and procedures (TTPs). It is especially true for APTs.

At their core, cybersecurity breaches by APTs are a form of illegal behavior. The economic theory of deterrence of illegal behavior begins with Becker (1965). Yet a valid critique of Becker is it is a nonstrategic analysis where the probability of being caught is exogenous. A game-theoretic approach instead endogenizes this probability. Within the context of APT deterrence the Target's monitoring TTPs and frequency of monitoring determine the probability of catching the APT. This in turn determines an APT's TTPs, including when to further penetrate the Target versus playing a waiting game. Similarly, the APT's actions influence the Target, and so on. Such a situation is known as an *inspection game* (Tsebelis, 1990; Fudenberg and Tirole 1992), thereby recognizing the interdependence between Target and APT. [2] As it is the building block for intrusion detection systems, Liu, Comanciu, and Man (2006), Chen and Leneutre (2009), Otrok et al. (2009), Gao et al. (2012), Gianini et al. (2013), and Rass and Zhu (2016), among others, introduce and examine cybersecurity versions of the inspection game.

Modeling cybersecurity attack and defense as an inspection game facilitates our analysis of the evolution of Target TTPs from firewalls to defense-in-depth to cyber kill chains and beyond. As originally conceived, firewalls are a form of perimeter cyber defense employed as the control point on the boundary between an organization's digital assets and the Internet. Firewalls inspect incoming and outgoing traffic based on IP protocols, ports, and packet filters. Firewalls lose their effectiveness if not continuously maintained and if more than one choke point needs monitoring. Over time, reliance on firewalls progressed to defense-in-depth and then cyber kill chains and zero trust.

A key insight stemming from this paper is each of the Target TTPs can be modeled as a

[2] This game is also known as an *enforcement game*. It is not to be confused with Dresher's (1962) sampling inspection game.

change in the information structure of the inspection game. Notably, a change in information structure manifests itself in terms of the solution concept applied to the same underlying cybersecurity inspection game. For example, firewalls correspond to a leader-follower structure, whereby the APT observes the Target's TTPs and acts accordingly. The associated equilibrium concept is Stackelberg equilibrium. In contradistinction, defense-in-depth corresponds to Nash play.

Liu, Comanciu, and Man (2006), Rass and Zhu (2016), and Huang and Zhu (2020) examine multistage games representing the Target's network, where each stage is a local inspection game between Target and APT. We employ a non-zero-sum version of Rass and Zhu (2016), where informational differences lead to different outcomes depending on the class of defense employed by the Target (firewall, defense-in-depth, kill chain, zero trust). The solutions (leader-follower, Nash, maximin, correlated) facilitate a comparison of how the Target fares across defensive classes. In this way, we examine the history of cybersecurity defense and potential future developments using four solutions to the cybersecurity inspection game. Hence, part of the contribution of the paper is the conceptualization of firewall defense, the cyber kill chain, and zero trust as solutions to the cybersecurity inspection game, thereby facilitating comparisons of how the Target fares under these three defensive classes. We also compare these three to Rass and Zhu (2016)'s conceptualization of Nash equilibrium as defense-in-depth.

How the Target fares across different defensive classes is an open question because as the Target changes its defense, APTs' TTPs change, and vice-versa. From a strategic perspective, defense-in-depth may result in a fallacy of composition against an ATP rather than generating synergy. Multiple layers of defense may increase the attack surface. Moreover, Wolff (2016) observes that encountering a strong overlapping defense may signal to an ATP that the Target is worthwhile. That is, the sophistication of defense may attract APTs. Hence the need for deception as part of defense-in-depth.

We characterize the similarities and differences of the four defensive stances using two criteria. The first is strategy equivalence, pertaining to the probability of monitoring at each level of the network. The second is payoff equivalence, measuring the Target's payoff at each level of the network. Our findings are as follows. The (probabilistic) level of Target

monitoring in the cybersecurity inspection game is the same when the Target acts as a Stackelberg leader versus Nash play. Yet the Target is better off under Nash play than as a Stackelberg leader. The two are Target strategy equivalent, but payoff non-equivalent. The difference has to do with the APT's information and its ability to react. An APT can act as a Stackelberg follower when it knows the Target's defense strategy, as is the case with defensive modes such as firewalls (Cavusoglu, Srinivasan, and Yue 2008). Kerckhoffs' (1883) Principle in cryptography – a cryptosystem must remain secure even if everything but the secret key is public knowledge, including the algorithm – is another example of creating a leader-follower environment (Collins, Xu, and Brown 2025). By contrast, defense-in-depth eliminates this advantage and induces Nash play. Consequently, the APT's second-mover advantage is eliminated, and this benefits the Target. As such, it is not so much the amount of monitoring that matters but the information the APT has about the monitoring.

These equivalences or lack thereof are reversed when applying the cyber kill chain to the cybersecurity inspection game. As the cyber kill chain informs the Target about APT TTPs, both the Target and APT are better informed than in Nash play. The result is a maximin level of monitoring by the Target. This is Target strategy non-equivalence. Yet the increase in monitoring does not commensurately raise the Target's payoff. Instead, the Target does no better than under defense-in-depth, i.e., payoff equivalence. Effectively, the APT response to the cyber kill chain offsets the increased monitoring. APTs become more aggressive when facing the kill chain. APTs penetrate the next level of the network more often because, by the kill chain the Target knows the ATP's TTPs at the ATP's current level in the network, making waiting at a level riskier than under firewalls or defense-in-depth.

Payoff equivalence begs the question of whether this is as good as it gets for the Target? In investigating this question, we consider correlated strategies and correlated equilibrium. One rationale for this is correlated strategies inform players on a need-to-know basis, just as the principle of least privilege in zero trust is a need-to-know approach to information access. We show correlation can lead to a higher likelihood of the Target's most-preferred outcome in the cybersecurity inspection game as compared to Nash or Stackelberg play. At the same time, correlation cannot eliminate the persistent property that the Target's least-preferred outcome is more likely to occur than its most preferred outcome,

consistent with cybersecurity folk wisdom of attacker advantage. It is also consistent with the emerging viewpoint given in Arce (2020) and Shapiro (2023) that technical solutions such as firewalls, defense-in-depth, the cyber kill chain, and zero trust, may never be enough. Indeed, we characterize the limit of technical solutions for the cyber enforcement game in terms of the probability of the Target's most preferred outcome versus the probability of the APT's most preferred outcome.

2. Related Literature

In considering the history and future of cybersecurity defense, we begin with and then move beyond the common assumption that the Target acts as a Stackelberg leader. That is, the APT observes the Target's strategy and optimizes by finding its best reply to the Target strategy. The Target in turn optimizes given the APT's best replies to the Target's strategy. An example of a Stackelberg information structure is when Targets reveal their investments in firewalls, authentication systems, and monitoring and inspection procedures (Cavusoglu, Raghunathan, and Yue 2008).

By contrast, defense-in-depth involves redundant TTPs, such as access controls, integrity shells, online backups, virus monitors, virus traps, deception, etc. (Cohen 1992). Defense-in-depth provides sufficient redundancy to withstand new attack mechanisms, and, when properly integrated, results in synergy in the form of increased protection (Cohen 1992). In particular, when overlapping controls are orthogonal to each other the weakness of some controls can be mitigated by the strengths of others (Tirenin and Faatz 1999).

Recognizing the added coverage against known and unknown APT TTP's, Rass and Zhu (2016) characterize the outcome of defense-in-depth in terms of the Nash equilibrium of a zero-sum cybersecurity inspection game in strategic form, implying Target and APT have imperfect information about each other's TTPs. Liu, Comanciu, and Man (2006) and Huang and Zhu (2020) similarly use a variation on Nash equilibrium – Bayes-Nash equilibrium – to analyze defense-in-depth under conditions of incomplete information. Their informational comparative static is different than ours in they introduce Target uncertainty about whether a user is legitimate or an APT; and APT uncertainty about whether the Target is sophisticated (practices defensive deception) or not (called naïve).

In the kill chain model of cybersecurity, intrusions are not considered singular events, but rather phased progressions required to accomplish the APT's goal (Hutchins, Cloppert, and Amit 2011). Achieving that goal requires successfully completing seven distinct phases of APT activity: (i) reconnaissance, (ii) weaponization, (iii) delivery, (iv) exploitation, (v) installation, (vi) command and control (C2), and (vii) actions on objectives. These phases in turn define the cyber kill chain. Furthermore, any repetition of TTPs within a level of the network is a liability for the APT.

On a practical basis, MITRE's ATT&CK framework documents the known TTPs for each state of the kill chain for a given APT. In documenting an APT's specific implementation procedures used to deploy tactics, such intelligence, "is no longer ephemeral, is tied to known adversarial group behavior, and is conducive to designing impactful countermeasures" (Howard 2023). Indeed, the HAFNIUM campaigns exemplify the need for comprehensive threat intelligence to understand and counteract APT activities. From a theoretical perspective, the upshot is the cyber kill chain allows the Target to neutralize the ATP's TTPs. That is, the Target must do at least as well as its lower bound (maximin) payoff, thereby making the Target's payoff independent of the ATP's TTPs because the TTPs are known to the Target. This criterion is particularly appropriate for high-risk environments.

Although derived independently from Hutchins, Cloppert, and Amit (2011), Forrester Research's zero trust model of cybersecurity begins with the same premise – assume the adversary is already within the network. Consequently, the network itself cannot be trusted. Zero trust rests on the following three principles: (i) ensure all resources are accessed securely regardless of location. Security principles must apply equally within and outside the network; (ii) adopt a least-privilege strategy and strictly enforce access control. Treat all information on a need-to-know basis; (iii) inspect and log all external *and* internal traffic (Kindervag 2010). We therefore explore the correlated equilibrium of the cybersecurity inspection game because correlated equilibrium as well limits information on a need-to-know basis. Finally, while the cyber kill chain and zero trust were conceptualized over a decade ago, implementation of these concepts at scale is currently a work in progress with neither as ubiquitous as firewalls nor defense-in- depth. They are part of cybersecurity's future, as is machine learning, which we address in the conclusion.

3. The Cybersecurity Inspection Game

Rass and Zhu (2016) consider defense against an APT as a sequence of zero-sum inspection games, $G_K, G_{K-1}, \ldots, G_1$, where the subscript refers to the APT's current level within the Target's network, and at level $k+1$ the APT has the strategy of penetrating the next level, P_k, or waiting at the current level, W_{k+1}. The topology of the network is acyclic, implying state $k+1$ must be penetrated prior to k. The value of state k is S_k and S_0 is the APT's ultimate objective, with node/state k separated from S_0 by k edges. The value of a state strictly increases as the APT approaches S_0 within the network: $S_0 > S_1 > \cdots > S_K \geq 0$. Hence, in game G_{k+1}, S_k is the value of the next state/level of the network to be penetrated.

At G_{k+1}, the strategy of penetrating the next level with value S_k, is specified as P_k. Strategy P_k costs the APT $\pi_k > 0$. Alternatively, the APT can wait at G_{k+1}, which is strategy W_{k+1}. Strategy W_{k+1} costs $\omega_{k+1} > 0$. The potentiality that the APT may wait at level $k+1$, rather than attempt to penetrate level k, is consistent with the definitions of APT given in the introduction. Cost ω_{k+1} is consistent with the property that waiting at a level is a liability for the APT. The Target has two strategies at each G_{k+1}: monitor for penetration at level k (M_k) or not (N_k). Monitoring leads to the probability an APT is detected penetrating state k. Rass and Zhu (2016) instead assume the probabilities are parameters whose values differ with the state (k).

The game we analyze deviates from Rass and Zhu (2016) in two ways. First, the probability an APT penetrating state k is detected equals the probability the Target monitors state k, consistent with the standard approach in inspection games (Tsebelis, 1990; Fudenberg and Tirole, 1992). Hence, the probability of detection is now determined endogenously as an equilibrium value that is a function of the primitives of the game, whereas it is an exogenous parameter in Becker (1965), Chen and Leneutre (2009), and Rass and Zhu (2016). Second, Rass and Zhu (2016) consider the APT's payoff only and assume the game is zero-sum. We as well assume the payoffs are zero-sum in state values S_k and S_{k+1}, reflecting whether the APT successfully penetrates the state in question or not. The difference is the cost of a strategy for the Target need not be a benefit to the APT and the cost of a strategy for the ATP need not be a benefit to the Target. In this way, in addition to the APT's costs of penetrating and waiting, π_k and ω_{k+1}, the Target also has a cost of monitoring, $c_k > 0$. At each stage the game is no longer zero-sum, consistent with Chen and Leneutre

(2009) and Gianini et al. (2013). We do, however, maintain Rass and Zhu's (2016) network structure in that the counterpart strategy to P_k is not the strategy of not attacking – as is the case in Chen and Leneutre (2009) and Gianini et al. (2013) – with an associated payoff of zero. Instead, waiting at previously penetrated level $k+1, W_{k+1}$, nets a gain for successfully penetrating $k+1$ minus the liability for waiting at $k+1$.

Each stage of the cybersecurity inspection game is represented in strategic form by the game box in Figure 1.

Figure 1: Representative Stage G_{k+1} of the Cybersecurity Inspection Game

↓Target/APT→	P_k	W_{k+1}
M_k	$S_k - c_k, -S_k - \pi_k$	$-S_{k+1} - c_k, S_{k+1} - \omega_{k+1}$
N_k	$-S_k, S_k - \pi_k$	$-S_{k+1}, S_{k+1} - \omega_{k+1}$

The payoffs at the (N_k, W_{k+1}) outcome represent the status quo. The APT has successfully penetrated level $k+1$ of the network, a necessary condition to penetrate level k. Hence, the payoffs are zero-sum in S_{k+1} with the APT also incurring its waiting cost, ω_{k+1}. At (N_k, P_k) the APT successfully penetrates level k. The payoffs are zero-sum in S_k with the APT incurring its penetration cost, π_k. At (M_k, P_k) the Target successfully deters the APT's attempt at penetrating level k. The payoffs are zero-sum in S_k, with the Target incurring its monitoring cost and the APT its penetration cost. Outcome (P_k, W_{k+1}) is a variation on the status quo at (N_k, W_{k+1}), where the Target is additionally incurring its cost for monitoring at level k.

The relative costs and benefits in the payoffs follow the standard assumptions for inspection games: (i) $S_k - c_k > 0$; (ii) $S_k - \pi_k > 0$; (iii) $S_{k+1} - \omega_{k+1} > 0$; (iv) $\pi_k > \omega_{k+1}$; and (v) $S_k - \pi_k > S_{k+1} - \omega_{k+1}$. Assumption (iv) corresponds to the intuition that it is more costly to penetrate the next level of the network rather than wait at the current level. Assumption (iv) makes it clear that, if the Target is not monitoring the next level, it is rational for the APT to penetrate the next level.

Given these benefits and costs, the game has a clockwise sequence of best replies. Specifically, if the Target monitors, the APT's best reply is to wait; if the APT waits, the Target's best reply is not monitor (thereby avoiding the cost of monitoring); if the Target does not monitor, the APT's best reply is to penetrate; and, finally, if the APT is penetrating, the Target's

best reply is monitoring.

This implies each stage has no pure strategy Nash equilibrium. At each stage the Target's local strategy (probability) of monitoring is $\lambda_M \in [0,1]$ and that for not monitoring is $\lambda_N \in [0,1]$, where $\lambda_M + \lambda_N = 1$.[3] [For simplicity, subscript k is suppressed.] Similarly, the APT's probability of penetrating is $\lambda_P \in [0,1]$ and that for waiting is $\lambda_W \in [0,1]$, where $\lambda_P + \lambda_W = 1$. [Again, we suppress subscripts k and $k+1$.] The expected payoffs for the target, $E_T[\lambda_M, \lambda_P]$, and APT, $E_A[\lambda_M, \lambda_P]$, at each stage are

$$E_T[\lambda_M, \lambda_P] = [S_k - c_k]\lambda_M\lambda_P + [-S_{k+1} - c_k]\lambda_M\lambda_W + [-S_k]\lambda_N\lambda_P + [-S_{k+1}]\lambda_N\lambda_W$$

which simplifies to

$$E_T[\lambda_M, \lambda_P] = 2S_k\lambda_M\lambda_P - c_k\lambda_M - S_k\lambda_P - S_{k+1}(1 - \lambda_P) \quad (1)$$

and,

$$E_A[\lambda_M, \lambda_P] = S_{k+1} - \omega_{k+1} + [(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]\lambda_P - 2S_k\lambda_M\lambda_P \quad (2)$$

Taken together, the sequence of games, $G_K, G_{K-1}, \dots, G_1$, constitutes an extensive form game. Consistent with the definition of strategies in extensive form games, ex ante each player specifies a (local) strategy for each G_k. The difference between our specification of strategies and payoffs and Chen and Leneutre (2009) is their strategies are K-tuples of probability distributions where the add-up condition is across all stages, e.g., $\lambda_{P_0} + \lambda_{P_1} + \cdots + \lambda_{P_{K-1}} \leq \overline{\lambda} \leq 1$ for the APT, and each player maximizes the sum of their payoffs over all stages. Our use of local strategies instead captures equilibrium behavior at each stage.

Result 1: the Nash equilibrium of each stage of the cybersecurity inspection game is

$$\lambda_M^* = \frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k}; \lambda_P^* = \frac{c_k}{2S_k} \quad (3)$$

Proof: all proofs of results and characterizations are in the appendix.

At this point we present the Nash equilibrium for baseline purposes. Discussion of when to expect the Nash equilibrium, and its characterization, occurs in Section 5. We now turn to the evolution of cybersecurity defense.

---

[3] Using local strategies recognizes the multistage nature of the game. By contrast, a mixed strategy for this game is a probability distribution over every possible K-tuple of pure strategies, one for each stage.

4. When the APT has an Informational Advantage

A common assumption in the analysis of cybersecurity games is the attacker – here, the APT – has an informational advantage in it observes the Target's defensive strategy. Given the Target's strategy, the APT selects its best reply. This is a leader-follower framework where the Target is the leader, the APT is the follower, and the solution concept is Stackelberg equilibrium. To be clear, the definition of Stackelberg equilibrium employed is:

> A Stackelberg equilibrium is a strategy profile in which the players select strategies in a given order and each player's strategy is a best response to the fixed strategies of the players preceding him … Such an equilibrium would not generally be Nash or [subgame] perfect (Rasmusen 2007).

In a Stackelberg equilibrium the Target solves the following optimization problem at each stage:

$$\max_{\lambda_M} \mathrm{E}_T[\lambda_M, br_A(\lambda_M)]$$

where $br_A(\lambda_M) \equiv$ APT's best reply correspondence (set valued function), and λ_P^F denotes the APT-as-follower's best reply probability of penetrating, P_k, given the Target-as-leader's probability of monitoring, λ_M^L.

The APT's best reply correspondence may identify a nonsingular set of best replies to a strategy by the Target. Two conventions exist for breaking the 'tie.' One is Strong Stackelberg equilibrium where where the Target commits to the leader strategy and the APT selects λ_P from its best replies to maximize the Targets' payoff. As the APT and Target's payoffs are zero sum in S_k and S_{k+1} the Strong Stackelberg convention is not in the APT's interest. Another convention is Weak Stackelberg equilibrium, where where the Target commits to the leader strategy and the APT selects λ_P from its best replies to mimimize the Targets' payoff. Such an APT is known as 'Byzantine' (Aiyer et al. 2005; Moscribroda, Schmid, and Wattenhofer 2006), which is a common assumption in cybersecurity. That is, among its best replies the APT selects the λ_P^F minimizing the Target's payoff.

Result 2: the Weak Stackelberg equilibrium with Target as leader and APT as (Byzantine) follower is

$$\lambda_M^L = \frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k} (= \lambda_M^*); \lambda_P^F = 1. \quad (4)$$

This equilibrium has the following characterization:

- The probability the Target monitors is the same for Stackelberg leadership and Nash equilibrium: $\lambda_M^L = \lambda_M^*$.
- The Target's expected payoff is $E_T[\lambda_M^L, \lambda_P^F] = \frac{(2S_k - c_k)[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})] - 2S_k^2}{2S_k}$.

In the next section we discuss these properties relative to the Nash equilibrium.

5. Defense-in-Depth

Defense-in-Depth refers to a multilayered system of redundant TTPs. In this section we analyze defense-in-depth relative to the situation it is designed to improve upon – firewall defense as Stackelberg leadership on the part of the Target. Specifically, we examine whether defense-in-depth is a welfare improvement. The question is open because, from a game-theoretic perspective, when the Target's or APT's information changes, or when their strategy changes, it may very well imply their counterpart's strategy changes.

The resulting outcome is the Nash equilibrium given by equation (3) in Result 1. As such, defense-in-depth has the following characteristics:

- Again, the probability the Target monitors is the same for both Stackelberg leadership, where the APT knows the Target's strategy; and the Nash equilibrium, where the APT has imperfect information about the Target's strategy: $\lambda_M^L = \lambda_M^*$. We call this Target strategy equivalence.
- The probability of monitoring state/level k (and, hence, detection) increases with the value of that state, $\frac{\partial \lambda_P^*}{\partial S_k} = \frac{\partial \lambda_P^L}{\partial S_k} = \frac{S_{k+1} + (\pi_k - \omega_{k+1})}{S_k^2} > 0$.[4] Furthermore, this increase is an increasing function of the value of the prior state, S_{k+1}.

[4] Note this has nothing to with the assumption $S_0 > S_1 > \cdots > S_K \geq 0$. The characterization that holds by this assumption is the probability of monitoring in state k is larger than for state $k + 1$.

Hence, rather than assuming a parametrized value of detecting increases with the value of the state, we establish the endogenous probability of monitoring increases with the value of the state.

- From equations (1) and (3), the Target's expected payoff is $E_T[\lambda_M^*, \lambda_P^*] = \frac{c_k S_{k+1} - c_k S_k - 2S_{k+1} S_k}{2S_k}$.
- Defense-in-depth is an improvement over Stackelberg leadership, $E_T[\lambda_M^*, \lambda_P^*] > E_T[\lambda_M^L, \lambda_P^F]$. That is, when the APT has the information to act as a Stackelberg follower, it has a second-mover advantage relative to the Target-as-leader. This advantage decreases the Target's payoff relative to the Nash equilibrium. We call this Target payoff non-equivalence.

This final characterization captures rammifications of the informational structure implied by a cybersecurity defense. Even with an identical level of monitoring, defense-in-depth does better because it changes the informational structure of the cybersecurity inspection game from Stackelberg leadership to Nash play. It is not just monitoring that matters but the informational context in which monitoring takes place.

6. Kill Chain Defense

The intent of Hutchins, Cloppert, and Amit's (2011) concept of the cyber kill chain is to turn the table on APTs by using strategies tailored to neutralize the ATP's TTPs. By the minimax theorem, $\max_{\lambda_M} \min_{\lambda_P} E_T[\lambda_M, \lambda_P] \leq \min_{\lambda_P} \max_{\lambda_M} E_T[\lambda_M, \lambda_P]$.[5] Hence, the lowest payoff a Byzantine APT can hold the Target's payoff to is maximin value, $\max_{\lambda_M^k} \min_{\lambda_P} E[\lambda_M, \lambda_P]$, where λ_M^k is the Target's kill chain strategy. Kill chain strategies begin with the premise the APT is already in the Target's network. By understanding an APT's TTPs the maximin strategy makes the Target's payoff independent of the ATP's TTPs, thereby neutralizing them. The maximin approach is particularly useful in high-risk environments where the Target needs to ensure a baseline level of security. Targets can use this strategy to mitigate the impact of APTs and reduce the likelihood of successful attacks.

[5] Solving $\min_{\lambda_P} \max_{\lambda_M} E_T[\lambda_M, \lambda_P]$ results in a strategy for the APT but not the Target. Furthermore, the Breton, Alj and Haurie (1988) procedure for building a strategy applies to maximin and not minimax.

Result 3: given a Byzantine APT, the Target's kill chain (maximin) strategy is

$$\lambda_M^k = \frac{S_k - S_{k-1}}{2S_k} \quad (5)$$

The Target's maximin payoff is a lower bound achievable by neutralizing the APT's TTPs. The Target can get a higher payoff depending upon what the APT does. As is always the case, the maximim solution only identifies the strategy for the maximizer; here, the Target. Yet the cyber kill chain situation requires the identification of the APT's strategy as well. Once again, we build the APT's strategy via the three-step procedure in Breton, Alj and Haurie (1988) used to prove Result 2 in the appendix.

The kill chain solution has the following characteristics:

- The associated strategy combination is $\left(\lambda_M^k, \lambda_P^k\right) = \left(\frac{S_k - S_{k-1}}{2S_k}, \frac{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k}\right)$.
- Under the kill chain defense, the probability of monitoring increases with the value of the state, $\frac{\partial \lambda_P^k}{\partial S_k} = \frac{S_{k-1}}{2S_k^2} > 0$. Once again, this rate of increase is increasing in the value of the prior state, S_{k-1}.
- The Target monitors more than under defense-in-depth: $\lambda_M^k > \lambda_M^*$; i.e., Target strategy non-equivalence. Consequently, as the Target is already monitoring more, the rate of increase of monitoring as a function of the value of the state is less than under defense-in-depth: $\frac{\partial \lambda_M^*}{\partial S_k} > \frac{\partial \lambda_M^k}{\partial S_k} > 0$.
- Even though the Target monitors more under the kill chain defense the Target does no better than under defense-in-depth: $E_T\left[\lambda_M^k, \lambda_P^k\right] = E_T[\lambda_M^*, \lambda_P^*]$; i.e., payoff equivalence.
- Given $E_T[\lambda_M^*, \lambda_P^*] > E_T\left[\lambda_M^L, \lambda_P^\ell\right]$, it also holds that $E_T\left[\lambda_M^k, \lambda_P^k\right] > E_T[\lambda_M^L, \lambda_P^F]$. The Target equally prefers both defense-in-depth and the cyber kill chain over firewalls.

What explains the payoff equivalence of kill chain versus defense-in-depth? The Target monitors more under the kill chain strategy than under defense-in-depth. More monitoring implies the Target incurs monitoring cost c_k more often. Furthermore, the APT's strategy differs as well.[6] This is the nature and value of game-theoretic analysis. When one

[6] The APT penetrates more if $[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})] > c_k$ and less if the inequality is reversed.

player changes its strategy, the other player(s) may change their strategy as well. Indeed, Wolff (2016) identifies cases where an increase in the sophistication of a Target's defense attracts APT interest because a sophisticated defense belies something worth protecting. Payoff equivalence is also a one-size-does-not-fit-all result in that Targets in high-risk environments will likely go the kill chain route, owing to its baseline level of maximin security, with others equally served by defense-in-depth.[7]

8. Improving the Situation

Thus far we consider three types of solutions corresponding to three models of cybersecurity defense: (Stackelberg) leader-follower, corresponding to firewalls; Nash equilibrium, corresponding to defense-in-depth; and maximin, corresponding to the cyber kill chain. In this section we consider a different noncooperative approach in the form of correlated strategies and correlated equilibrium.

Under local strategies the probability of each player's action is independent of the other players. Hence, the probability the (M_k, P_k) outcome occurs is $\lambda_M \cdot \lambda_P$, where $\lambda_M, \lambda_N, \lambda_P, \lambda_W \in [0,1],$ and $\lambda_M + \lambda_N = 1,\ \lambda_P + \lambda_W = 1$. By contrast, when strategies are correlated the probability distribution is on joint strategies, one for each player, $\rho_{MP}, \rho_{MW}, \rho_{NP}, \rho_{NW} \in [0,1]$, and $\rho_{MP} + \rho_{MW} + \rho_{NP} + \rho_{NW} = 1$.

The associated equilibrium concept – correlated equilibrium – is based on giving players information on a need-to-know basis. In particular, the players know the distribution, $\rho = (\rho_{MP}, \rho_{MW}, \rho_{NP}, \rho_{NW})$. In addition, for any joint strategy outcome realized under ρ, each player only knows their component of the realization. For example, if ρ results in (M_k, P_k), the Target only knows M_k resulted from ρ and the APT only knows P_k resulted from ρ. Dodis and Rabin (2007) call the need-to-know property of the realization of ρ *privacy preservation*.

In a correlated equilibrium, ρ provides the incentive for each player to follow ρ's realization, assuming the other player does as well. In terms of each player's (conditional) expected payoffs, ρ satisfies the following set of incentive compatibility constraints:

$$E_T[M_k|\rho_{MP}, \rho_{MW}] \geq E_T[N_k|\rho_{MP}, \rho_{MW}] \text{ and } E_T[N_k|\rho_{NP}, \rho_{NW}] \geq E_T[M_k|\rho_{NP}, \rho_{NW}] \qquad (6)$$

[7] Gilad and Asher Tishler (2024) derive a one-size-does-not-fit-all result for a different cybersecurity game.

for the Target, and, for the APT,

$$E_A[P_k|\rho_{MP},\rho_{NP}] \geq E_A[W_{k+1}|\rho_{MP},\rho_{NP}] \text{ and } E_A[W_{k+1}|\rho_{MW},\rho_{NW}] \geq E_A[P_k|\rho_{MW},\rho_{NW}] \quad (7)$$

As an example, if M_k is the realization for the Target, $E_T[M_k|\rho_{MP},\rho_{MW}] \geq E_T[N_k|\rho_{MP},\rho_{MW}]$ implies

$$\underbrace{[S_k - c_k]\rho_{MP} + [-S_{k+1} - c_k]\rho_{MW}}_{\text{Target's expected payoff for following realization } M \text{ given APT follows its realization } (P \text{ or } W).} \geq \underbrace{[-S_k]\rho_{MP} + [-S_{k+1}]\rho_{MW}}_{\text{Target's expected payoff for deviating to } N \text{ instead of } M.}$$

From the Target's perspective, its payoff is positive at (M_k, P_k) and is negative everywhere else. Hence, the Target is interested in the correlated equilibrium where $\rho_{MW}, \rho_{NP}, \text{and } \rho_{NW}$ are at their lower bounds, given by equilibrium constraints (6) and (7).

Result 4: the correlated equilibrium setting $\rho_{MW}, \rho_{NP}, \text{and } \rho_{NW}$ at their lower bounds generates the following probability for the (M_k, P_k) outcome

$$\rho_{MP} = \frac{c_k[(S_k-\pi_k)-(S_k-\omega_{k+1})]}{2S_k\{[(S_k-\pi_k)-(S_k-\omega)]+[(S_k+\pi_k)-(S_k-\omega_{k+1})]\}} \quad (8)$$

This equilibrium is of interest to the Target because

- Correlation can improve the likelihood of the Target's preferred outcome: $\rho_{MP} > \lambda_M^* \cdot \lambda_P^*$.

On the other hand, particularly disconcerting is the following property of this equilibrium, and, indeed, of any correlated or Nash equilibrium for the game:

- The (N_k, P_k) outcome is more likely than (M_k, P_k): $\rho_{NP} > \rho_{MP}$.

This inequality arises from the constraint $E_A[P_k|\rho_{MP},\rho_{NP}] \geq E_A[W_{k+1}|\rho_{MP},\rho_{NP}]$. As all Nash equilibria can be replicated by correlated equilibria, but not vice-versa, it is the property of any Nash equilibrium for the game as well.[8] In other words, *the APT's preferred outcome occurs more frequently than the Target's preferred outcome*. Indeed, accepting the idea the Target is fighting a losing battle is part of the cybersecurity folk wisdom of attacker advantage and is the rationale for zero-trust cybersecurity policies (Kindervag 2010). The argument for zero trust occurs within the context of assessing prior technical (coding) solutions to the cybersecurity problem, such as firewalls and defense-in-depth. The idea there may be "no

[8] Trivially, given the independence of the Nash local strategies, $\rho_{MP} = \lambda_M^* \cdot \lambda_P^*, \rho_{MW} = \lambda_M^* \cdot \lambda_W^*, \rho_{NP} = \lambda_N^* \cdot \lambda_P^*, \rho_{NW} = \lambda_N^* \cdot \lambda_W^*$ is a correlated equilibrium.

technical solution" to a social problem – dating to Hardin (1968) – also applies to cybersecurity (Arce 2020). It is similar to Shapiro's (2023) concept of "solutionism," which is the flawed argument that cybersecurity is primarily a technical problem requiring an engineering solution. Instead, Shapiro (2023) asserts cybersecurity is a human problem requiring understanding of human behavior and Arce (2020) proves it for the case of Internet platforms. As such, this section finishes with a discussion of how correlation has the flavor of a nontechnical solution.

Correlated equilibrium is an example of an 'as if' equilibrium. Other examples of 'as if' equilibria include the biological refinement of Nash equilibrium for symmetric two-player games known as evolutionary stable strategies (ESS), which Maynard Smith and Price (1973) formulate as a static characterization of a dynamic evolutionary process.[9] ESS was later shown to characterize the outcomes of the frequency-dependent replicator dynamic in biology, thereby validating the 'as if' reasoning of Maynard Smith and Price. Closer to the current analysis is the use of ε-Nash equilibrium to characterize outcomes 'as if' players are computationally limited (Kol and Naor 2008).[10] Here, we relate correlated equilibrium to zero trust via the following 'as if' reasoning. The need-to-know (privacy-preserving) realization of a correlated equilibrium strategy is 'as if' this need-to-know property translates into the principle of least privilege under zero trust, where all information and access is treated on a need-to-know basis.

Another 'as if' characterization also lends itself to considering correlated equilibrium as a potential nontechnical solution for the cybersecurity inspection game. Specifically, if the game is both (i) extended by an arbitrary but finite number of rounds of cheap talk (unmediated payoff-irrelevant preplay communication), and (ii) players have limited computational capacity, then Dodis et al. (2000) and Urbano and Vila (2002) show such preplay communication can implement any correlated equilibrium distribution among the players themselves without need for third party mediation.

---

[9] Using the symbol σ to denote a mixed strategy, a symmetric equilibrium (σ^*, σ^*) is an ESS if (i) $E[\sigma^*, \sigma^*] \geq E[\sigma, \sigma^*] \forall \sigma \neq \sigma^*$; and (ii) if $\mathrm{E}[\sigma^*, \sigma^*] = E[\sigma, \sigma^*]$, then $E[\sigma^*, \sigma] \geq E[\sigma, \sigma]$.

[10] Using σ again to denote a mixed strategy, joint strategy $(\sigma_i^*, \sigma_{-i}^*)$ is an ε-Nash equilibrium if no player can unilaterally deviate from $(\sigma_i^*, \sigma_{-i}^*)$ and increase its payoff by an amount greater than $\varepsilon > 0$: $E_i[\sigma_i^*, \sigma_{-i}^*] \geq E_i[\sigma_i, \sigma_{-i}^*] - \varepsilon \; \forall \sigma_i \neq \sigma_i^*, \forall i$. Typically, ε is a function of the polynomial time in which σ_i^* and σ_i can be calculated.

An example of costless preplay communication is the recommendation that zero-trust adopters announce the principle of least privilege as a means to change behaviors (Kindervag 2010). Hence, zero trust has a nontechnical component. Another example is the naming and shaming form of APT deterrence. Naming and shaming places significant importance on attribution, a problem that is not intractable (U.S. Department of Defense 2011; Baker 2012; Gilad, Pecht, and Tishler 2021). Indeed, naming and shaming is rarely a one round process as incentives exist for academics and commercial security providers to establish reputations by independently verifying attribution (Arce 2023). For example, the U.S. government often passes technical evidence to the nongovernment Cyber Threat Alliance (CTA), who confirm attribution using their own commercial product lines.

Consider the naming and shaming of the HAFNIUM group behind the Salt Typhoon and Flax Typhoon campaigns mentioned in the introduction. It is an example of extended cheap talk because such naming and shaming is unlikely to result in prison terms for members of HAFNIUM. Yet the logic for naming and shaming is clear through the 'as if' interpretation of correlated equilibrium and Result 4.

8. Conclusion

The evolution of cybersecurity defense is towards more sophistication. Yet whether more necessarily leads to better outcomes for the Target is subject to debate (Wolff 2015). Using one game – the cybersecurity enforcement (intrusion detection) game of a Target versus an APT – we consider four approaches to cybersecurity: firewalls, defense-in-depth, cyber kill chain, and zero trust. We evaluate these alternatives via four solutions to the game: Stackelberg leader-follower, Nash, maximin, and correlation. We compare the outcomes using two criteria: the Target's probability of monitoring and detection, and the Target's payoff. Our findings reveal more monitoring need not be better because the APT reacts to the Target defense and, together, the Target and APT strategies determine the Target's payoff.

One indicator of the value of theoretical modeling is whether some results initially appear to be counterintuitive, but the theory lends itself to explaining why this is not the case (Sandler 2001). That is, theory sheds new light on the situation. For example, the Target's Stackelberg and Nash solutions are strategically equivalent but payoff non-equivalent. The

Target monitors with the same probability as a Stackelberg leader or as a Nash player, but earns a higher payoff under Nash play, which we equate to defense-in-depth. The same monitoring probability for the Target leads to a payoff improvement for the Target under defense-in-depth because the APT has less idea of what the Target is doing and reacts accordingly. By contrast, defense-in-depth was initially promoted to create synergy among orthogonal TTPs with individual strengths and weaknesses. We do not dispute the potential synergistic properties of defense-in-depth but do suggest the associated change in information structure has a previously unrecognized role to play.

Hence, it is in the Target's interest to recognize scenarios where its TTPs result in the Target-as-leader and ATP-as follower, calling for actions to reduce ATPs' second-mover advantage. For example, in the future machine learning is expected to figure prominently in cybersecurity. Most machine learning methods do not recognize the existence of active adversarial opponents. Evasive ATP TTPs in the wild need not follow the same patterns as training data. Even adversarial machine learning methods are likely to commit the Target to a strategy before the adversary takes its actions (Zhou, Kantarcioglu, and Xi 2021). Hence, in either case the information structure allows the APT to act as a Stackelberg follower with the Target-as-leader committing to its machine learning model.[11]

Both defense-in-depth and the cyber kill chain increase the Target's payoff relative to firewalls. Yet in comparing the two we do not find more sophistication is better. Strategic non-equivalence (more monitoring in the kill chain) leads to payoff equivalence. An unanticipated effect of the kill chain is it spurs the APT to be more aggressive. Consequently, the kill chain is more appropriate for high-risk environments demanding a baseline level of security independent of the APT's actions. Targets should balance the implementation of advanced cybersecurity measures with practical considerations, ensuring they do not inadvertently increase the attack surface or attract more aggressive APTs.

Overall, the message is the TTPs of new cybersecurity modes have informational implications that need to be recognized and understood. Hence, when a Target introduces new TTPs, they should ask themselves whether these TTPs make the Target or APT a Nash

[11] The Target has a remaining problem in much of the output from machine learning is unexplainable owing to the impenetrability of the underlying logic for generating weights.

player, a Stackelberg leader or follower, a Maximiner, or none of the above? That is, new TTPs may generate a new information structure. For example, defense-in-depth requires implementing strategies accounting for imperfect information. Zero trust requires measures to ensure a need-to-know information structure. By focusing on the informational implications, Targets can enhance their cybersecurity posture and better defend against APTs and other sophisticated threats.

Another related question is, how does costless communication – rather than costly signals – manifest itself and matter for the success of new Target TTPs? Within this context, we provide a rationale for informing users of the principle of least privilege when implementing zero trust, and naming and shaming in cybersecurity defense. Other ways to change the information structure and forms of costless-communication-as-cyber-defense are topics for future research.

Appendix: Proofs.

Nash Equilibrium: proof of Result 1 in section 3 and its accompanying characterization in section 5.

Nash equilibrium in probabilistic strategies satisfies the indifference property of equating the expected payoff of the pure strategies the other player plays with positive probability. For the Target's local strategy, this implies $E_A[\lambda_M, P_k] = E_A[\lambda_M, W_{k+1}]$:

$$[-S_k - \pi_k]\lambda_M + [S_k - \pi_k](1 - \lambda_M) = S_{k+1} - \omega_{k+1}$$

$$\lambda_M^* = \frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k}$$

Similarly, for the APT's mixture $E_T[M_k, \lambda_P] = E_T[N_k, \lambda_P]$:

$$[S_k - c_k]\lambda_P + [-S_{k+1} - c_k](1 - \lambda_P) = [-S_k]\lambda_P + [-S_{k+1}](1 - \lambda_P)$$

$$\lambda_P^* = \frac{c_k}{2S_k}$$

Substituting λ_M^* and λ_P^* into the formula for $E_T[\lambda_M, \lambda_P]$ in equation (1):

$$E_T\left[\lambda_M^*, \lambda_P^*\right] = 2S_k\left(\frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k}\right)\left(\frac{c_k}{2S_k}\right) - c_k\left(\frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k}\right) - S_k\left(\frac{c_k}{2S_k}\right) - S_{k+1}\left(\frac{2S_k - c_k}{2S_k}\right)$$

$$E_T[\lambda_M^*, \lambda_P^*] = \frac{c_k S_{k+1} - c_k S_k - 2S_{k+1}S_k}{2S_k} \quad \blacksquare$$

Target as Stackelberg leader: proof of Result 2 and its accompanying characterization.

Case 1: $\lambda_M^L > \lambda_M^* > 0$, implying $\lambda_P^F = 0$ is the APT's best reply. Given $\lambda_P^F = 0$, the Target-as-leader's optimal strategy is N_k; i.e., $\lambda_M^L = 0$, which contradicts $\lambda_M^L > \lambda_M^* > 0$.

Case 2: $\lambda_M^L < \lambda_M^* < 1$, implying $\lambda_P^F = 1$ is the APT's best reply. Given $\lambda_P^F = 1$, the Target-as-leader's optimal strategy is N_k; i.e., $\lambda_M^L = 1$, which contradicts $\lambda_M^L < \lambda_M^* < 1$.

These two cases imply the Target commits to $\lambda_M^L = \lambda_M^*$ as the Stackelberg leader. By the indifference property on Nash strategies, $E_A\left[\lambda_M^*, \tilde{\lambda}_P\right] = E_A\left[\lambda_M^*, \hat{\lambda}_P\right] \forall \tilde{\lambda}_P, \hat{\lambda}_P \in [0,1]$. That is, the follower's best reply to the leader's Nash strategy is a correspondence, rather than a function. However, given the Target's commitment to λ_M^*, a Byzantine APT selects λ_P^F from its set of best replies to minimize $E_T[\lambda_M^*, \lambda_P^F]$. Moreover, as $E_T[\lambda_M^*, \lambda_P]$ is linear in λ_P^F, only the extreme values of λ_P^F need be considered to identify the minimum. Here, $E_T[\lambda_M^*, P_k] < E_T[\lambda_M^*, W_{k+1}]$:

$$[S_k - c_k]\lambda_M^* + [-S_k](1 - \lambda_M^*) < [-S_{k+1} - c_k]\lambda_M^* + [-S_{k+1}](1 - \lambda_M^*)$$

$$\lambda_M^* < \frac{S_k - S_{k+1}}{2S_k}$$

Substituting the value of λ_M^*:

$$\frac{(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})}{2S_k} = \frac{S_k - S_{k+1} - (\pi_k - \omega_{k+1})}{2S_k} < \frac{S_k - S_{k+1}}{2S_k}$$

Finally, we show $\lambda_P^F = 1 \Rightarrow P_k$ satisfies Breton, Alj and Haurie's (1988) three-step procedure for validating Weak Stackelberg follower best replies. First, identify the APT's best replies to the Target's-as-leader's strategy, $br_A(\lambda_M^L)$. Second, find the APT's maximum payoff for these best replies. Call this payoff $\mathrm{E}_A^*[\lambda_M^L] = max\{E_A[\lambda_M^L, \lambda_P] | \lambda_P \epsilon br_A(\lambda_M^L)\}$. Third, λ_P^F must satisfy $E_A[\lambda_M^L, \lambda_P] \geq \mathrm{E}_A^*[\lambda_M^L] - \varepsilon$ for some $\varepsilon \geq 0$. Operationalizing this procedure:

First, as $E_A[\lambda_M^L]$ is linear in λ_P^F, only extreme values of λ_P^F need be considered for $\mathrm{E}_A^*[\lambda_M^L]$. Second, as λ_M^L is the same as the Target's Nash strategy, $E_A[\lambda_M^L, P_k] = E_A[\lambda_M^L, W_{k+1}]$, implying $\mathrm{E}_A^*[\lambda_M^L] = E_A[\lambda_M^L, P_k] = E_A[\lambda_M^L, W_{k+1}]$. Third, trivially $E_A[\lambda_M^L, P_k] \geq \mathrm{E}_A^*[\lambda_M^L] - \varepsilon$ $\forall \varepsilon \geq 0$. ■

Kill chain: proof of Result 3 and its accompanying characterization.

Equation (1) can be rewritten as

$$E_T\left[\lambda_M^k, \lambda_P\right] = \left[2S_k\lambda_M^k - S_k + S_{k+1}\right]\lambda_P - c_k\lambda_M^k - S_{k+1}$$

As a Byzantine APT sets λ_P to minimize the Target's payoff, $E_T\left[\lambda_M^k, \lambda_P\right]$, the Target maximizes its payoff by minimizing the effect of λ_P on $E_T\left[\lambda_M^k, \lambda_P\right]$. That is, it chooses λ_M^k to eliminate the influence of λ_P on $E_T\left[\lambda_M^k, \lambda_P\right]$ by setting the coefficient on λ_P in brackets equal to zero. Setting $2S_k\lambda_M^k - S_k + S_{k+1} = 0$ yields $\lambda_M^k = \frac{S_k - S_{k+1}}{2S_k}$. ■

To build λ_P^k we again use the three-step procedure in Breton, Alj and Haurie (1988), as described in the proof of Result 2. First, given $\lambda_M^k = \frac{S_k - S_{k-1}}{2S_k}$, the APT's best reply will be an extreme value of λ_P because $E_A\left[\lambda_M^k, \lambda_P\right]$ is linear in λ_P:

$$E_A\left[\lambda_M^k, P_k\right] = [-S_k - \pi_k]\left(\frac{S_k - S_{k-1}}{2S_k}\right) + [S_k - \pi_k]\left(\frac{S_k + S_{k-1}}{2S_k}\right) = 0$$

$$E_A\left[\lambda_M^k, W_{k+1}\right] = S_{k+1} - \omega_{k+1} > 0$$

Clearly, $E_A\left[\lambda_M^k, W_{k+1}\right] > E_A\left[\lambda_M^k, P_k\right]$. Second, $\mathrm{E}_A^*\left[\lambda_M^k\right] = S_{k+1} - \omega_{k+1}$. Third, given the expression for $E_A[\lambda_M, \lambda_P]$ in equation (3), and setting $\varepsilon = 0$, $E_A\left[\lambda_M^k, \lambda_P\right]$ must satisfy

$$S_{k+1} - \omega_{k+1} + [(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]\lambda_P - 2S_k\lambda_M^k\lambda_P \geq S_{k+1} - \omega_{k+1}$$

Canceling the $S_{k+1} - \omega_{k+1}$ term on each side and substituting in the value for λ_M^k

$$[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]\left(\frac{S_k - S_{k-1}}{2S_k}\right) - 2S_k\left(\frac{S_k - S_{k-1}}{2S_k}\right)\lambda_P \geq 0$$

$$\frac{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k} \geq \lambda_P$$

Although $\lambda_P = 0$ satisfies this inequality, the ATP never penetrates. An ATP that never penetrates is not of interest nor is it Byzantine. Hence, λ_P is instead set to its upper bound:

$$\lambda_P^k = \frac{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k} \blacksquare$$

Proof that the kill chain outcome and the Nash outcome generate the same payoff for the Target: $E_T\left[\lambda_M^k, \lambda_P^k\right] = E_T[\lambda_M^*, \lambda_P^*]$. The value of $E_T[\lambda_M^*, \lambda_P^*]$ is given in the characterization of the Nash equilibrium. The value of $E_T\left[\lambda_M^k, \lambda_P^k\right]$ is derived from Equation (1) with the values for λ_M^k and λ_P^k.

$$\begin{aligned} E_T\left[\lambda_M^k, \lambda_P^k\right] = {} & 2S_k\left(\frac{S_k - S_{k-1}}{2S_k}\right)\left(\frac{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k}\right) - c_k\left(\frac{S_k - S_{k-1}}{2S_k}\right) \\ & - S_k\left(\frac{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k}\right) \\ & + S_{k+1}\left(\frac{2S_k - [(S_k - \pi_k) + (S_{k+1} - \omega_{k+1})]}{2S_k}\right) \\ = {} & \frac{c_k S_{k+1} - c_k S_k - 2S_{k+1}S_k}{2S_k} = E_T[\lambda_M^*, \lambda_P^*] \end{aligned}$$

Canceling the c_k terms on both sides, the $2S_k > 0$ denominators, and expanding terms:

$S_k[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})] - S_{k+1}[(S_k - \pi_k) + (S_{k+1} - \omega_{k+1})] - S_k[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})] - 2S_{k+1}S_k + S_{k+1}[(S_k - \pi_k) + (S_{k+1} - \omega_{k+1})] = -2S_{k+1}S_k \blacksquare$

Correlated equilibrium: proof of Result 4 and its accompanying characterization.

When M_k is the realization of ρ, $E_T[M_k|\rho_{MP},\rho_{MW}] \geq E_T[N_k|\rho_{MP},\rho_{MW}]$ requires

$$[S_k - c_k]\rho_{MP} + [-S_{k+1} - c_k]\rho_{MW} \geq [-S_k]\rho_{MP} + [-S_{k+1}]\rho_{MW}$$

$$\rho_{MP} \geq \frac{c_k}{2S_k - c_k} \cdot \rho_{MW} \quad \text{(A1)}$$

When P_k is the realization of ρ, $E_A[P_k|\rho_{MP},\rho_{NP}] \geq E_T[W_{k+1}|\rho_{MP},\rho_{NP}]$ requires

$$[-S_k - \pi_k]\rho_{MP} + [S_k - \pi_k]\rho_{NP} \geq [\,S_{k+1} - \omega_{k+1}]\rho_{MP} + [\,S_{k+1} - \omega_{k+1}]\rho_{NP}$$

$$[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]\rho_{NP} \geq [(S_k + \pi_k) + (S_{k+1} - \omega_{k+1})]\rho_{MP}$$

Which defines a lower bound on ρ_{NP} in terms of ρ_{MP}:

$$\rho_{NP} \geq \frac{[(S_k+\pi_k)+(S_{k+1}-\omega_{k+1})]}{[(S_k-\pi_k)-(S_{k+1}-\omega_{k+1})]} \cdot \rho_{MP} \quad \text{(A2)}$$

as the coefficient on ρ_{MP} is greater than one, this immediately provides the characterization $\rho_{NP} \geq \rho_{MP}$.

When N_k is the realization of ρ, $E_T[N_k|\rho_{NP},\rho_{NW}] \geq E_A[M_k|\rho_{NP},\rho_{NW}]$ requires

$$[-S_k]\rho_{NP} + [-S_{k+1}]\rho_{NW} \geq [S_k - c_k]\rho_{NP} + [-S_{k+1} - c_k]\rho_{NW}$$

which defines a lower bound on ρ_{NW} in terms of ρ_{NP}:

$$\rho_{NW} \geq \frac{2S_k - c_k}{c_k}\rho_{NP}$$

Substituting the lower bound for ρ_{NP} in equation (A2):

$$\rho_{NW} \geq \frac{2S_k - c_k}{c_k} \cdot \frac{[(S_k+\pi_k)+(S_{k+1}-\omega_{k+1})]}{[(S_k-\pi_k)-(S_{k+1}-\omega_{k+1})]} \cdot \rho_{MP} \quad \text{(A3)}$$

When W_{k+1} is the realization of ρ, $E_A[W_{k+1}|\rho_{MP},\rho_{NP}] \geq E_A[P_k|\rho_{MP},\rho_{NP}]$ requires

$$[\,S_{k+1} - \omega_{k+1}]\rho_{MW} + [\,S_{k+1} - \omega_{k+1}]\rho_{NW} \geq [-S_k - \pi_k]\rho_{MW} + [S_k - \pi_k]\rho_{NW}$$

$$[(S_k + \pi_k) + (S_{k+1} - \omega_{k+1})]\rho_{MW} \geq [(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]\rho_{NW}$$

which gives a lower bound on ρ_{MW} in terms of ρ_{NW}:

$$\rho_{MW} \geq \frac{[(S_k-\pi_k)-(S_{k+1}-\omega_{k+1})]}{[(S_k+\pi_k)+(S_{k+1}-\omega_{k+1})]}\rho_{NW}$$

Substituting the lower bound on ρ_{NW} in equation (A2)

$$\rho_{MW} \geq \frac{2S_k - c_k}{c_k} \cdot \rho_{MP} \qquad \text{(A4)}$$

By definition, $\rho_{MP} + \rho_{MW} + \rho_{NP} + \rho_{NW} = 1$. Substituting the lower bound for ρ_{MW} in (A4), for ρ_{NP} in (A2), and ρ_{NW} in (A3), the add up condition for ρ becomes:

$$\rho_{MP} \cdot \left[1 + \frac{2S_k - c_k}{c_k} + \frac{[(S_k + \pi_k) + (S_{k+1} - \omega_{k+1})]}{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]} + \frac{2S_k - c_k}{c_k} \cdot \frac{[(S_k + \pi_k) + (S_{k+1} - \omega_{k+1})]}{[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}\right] = 1$$

That is,

$$\rho_{MP} = \frac{c_k[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})]}{2S_k[(S_k - \pi_k) - (S_{k+1} - \omega_{k+1})] + 2S_k[(S_k + \pi_k) - (S_{k+1} - \omega_{k+1})]} \quad \blacksquare$$

References

Aiyer, A.S., L . Avisi, A. Clement, M. Dahlin, J-P Martin and C. Porth 2005. BAR Fault Tolerance for Cooperative Services. *Proceedings of the 20th ACM Symposium on Operating Principles*, New York: ACM, pp.45-58.

Anderson, Ross 2001. Why Information Security is Hard. An Economic Perspective. *Proceedings of the Seventeenth Annual Computer Security Applications Conference*, New Orleans: IEEE, pp.358-365.

Anderson, Ross and Tyler Moore 2006. The Economics of Information Security. *Science* 314(5799) 610-613.

Arce, Daniel 2020. Cybersecurity and Platform Competition in the Cloud. *Computers & Security* 93: 1-8.

Arce, Daniel 2023. Cybersecurity for Defense Economists. *Defence and Peace Economics* 34(6) 705-725.

Baker, Stewart 2012. Cybersecurity and the Attribution Problem: Good News at Last? https://www.skatingonstilts.com/ skating-on-stilts/2012/10/my-entry.html

Becker, Gary S. 1962. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76(2) 169-217.

Breton, M., A. Alj, and A. Haurie 1988. Sequential Stackelberg Equilibria in Two-Person Games. *Journal of Optimization Theory and Applications* 59(1): 71-97.

Cavusoglu, Huseyin, Srinivasan Raghunathan, and Wei T. Yue 2008. Decision-Theoretic and Game-Theoretic Approaches to IT Security Investment. *Journal of Management Information Systems* 25(2) 281-304.

Chen, Lin and Jeann Leneutre 2009. A Game Theoretical Framework on Intrusion Detection in Heterogenous Networks. *IEEE Transactions on Information Forensics and Security* 4(2) 165-178.

Chen, Thomas M. 2010. Stuxnet, the Real Start of Cyber Warfare? *IEEE Network* 24 (6): 2–3.

Cohen, Frederick B. 1992. Defense-in-Depth Against Computer Viruses. *Computers & Security* 11: 563-579.

Collins, Brandon, Shouhuai Xu, and Philip N. Brown 2025. Game-Theoretic Cybersecurity: the Good, the Bad, and the Ugly. *arXiv*.

Dodis, Yevgeniy, Shai Halevi, and Tal Rabin 2000. A Cryptographic Solution to a Game Theoretic Problem. In Bellare, M. (ed.), *Advances in Cryptology – CRYPTO 2000. Lecture Notes in Computer Science*, vol.1880, Berlin: Springer, 112-130.

Dodis, Yevgeniy and Tal Rabin 2007. Cryptography and Game Theory. Chapter 8 in Noam Nisan et al. (eds.) *Algorithmic Game Theory*, Cambridge: Cambridge University Press, pp.181-205.

Dresher, Melvin 1962. *A Sampling Inspection Problem in Arms Control Agreements: A Game-Theoretic Analysis*. Memorandum RM-2972 (ARPA Order No.189-61), Santa Monica: RAND Corporation.

Fedele, Alessandro and Cristian Roner 2022. Dangerous Games: A Literature Review on Cybersecurity Investment. *Journal of Economic Surveys* 36(1) 157-187.

Fudenberg, Drew and Jean Tirole 1992. *Game Theory*, Cambridge: MA: MIT Press.

Gao, Zhaoyu, Haojin Zhu, Suguo Du et al. 2012. PDMS: A Probabilistic Misbehavior Detection Scheme in DTN. *2012 IEEE International Conference on Communications – ICC*. Ottawa, CA: IEEE, pp.4970-4974.

Gianini, Gabriele, Ernest Damiani, Tobias R. Mayer, et al. 2013 Many-player Inspection Games in Networked Environments. *7th IEEE International Conference on Digital Ecosystems and Technologies – DEST*. Menlo Park, CA: IEEE, pp.1-6.

Gilad, Amitai, Eyal Pecht, and Asher Tishler 2021. Intelligence, Cyberspace, and National Security. *Defence and Peace Economics* 32(1) 18-45.

Gilad, Amitai and Asher Tishler 2024. Measuring and Mitigating the Risk of Advanced Cyberattackers. *Decision Analysis* 21(4) 215-234.

Hardin, Garrett 1968. The Tragedy of the Commons. *Science* 162(3559) 1243-1248.

Howard, Rick 2023. *Cybersecurity First Principles*, Hoboken, NJ: Wiley.

Huang, Linan and Quanyan Zhu 2020. A Dynamics Games Approach to Proactive Defense Strategies Against Advanced Persistent Threats in Cyber-Physical Systems. *Computers & Security* 98: 1-15.

Hutchins, Eric, Michael Cloppert, and Rohan Amin 2011. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. *Proceedings of the 6th International Conference on Information Warfare and*

Security, Leigh Armistead and Edith Cowan (eds.). Reading, UK: Academic Publishing, pp.113-125.

Kal, Gillat and Moni Naor 2008. Cryptography and Game Theory: Designing Protocols for Exchanging Information. In R. Canteti (ed), *Theory of Cryptography – TCC 2008*. *Lecture Notes in Computer Science*, vol.4948. Berlin: Springer, pp.320-339.

Kerckhoffs, A. 1883. La Cryptographie Militaire. *Journal Des Sciences Militaries* 9: 161–191.

Kindervag, John 2010. *No More Chewy Centers: Introducing the Zero Trust Model of Information Security*, Cambridge, MA: Forrester Research.

Liu, Yu, Cristina Comanciu, and Hong Man 2006. A Bayesian Game Approach for Intrusion in Wireless Ad Hoc Networks. *Proceedings from the 2006 Workshop on Game Theory for Communications Networks – GamesNets '06*, New York: ACM, pp.4 – es.

Maynard Smith, John, and George R. Price. 1973. The Logic of Animal Conflict. *Nature* 241: 15-18.

Moscibroda, Thomas, Stefan Schmid, and Roger Wattenhofer 2006. When Selfish Meets Evil: Byzantine Players in a Virus Inoculation Game. *Proceedings of the 25th ACM Symposium on Principles of Distributed Computing*, New York: ACM, 35-44.

Otrok, Hadi, Benwen Zhu, Hamdi Yahyaoui, and Prabir Bhattacharya 2009. An Intrusion Detection Game Theoretical Model. *Information Security Journal: A Global Perspective* 19(5) 199-212.

Papadimitriou, Christos H. 2001. Algorithms, Games, and the Internet. *Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing – STOC '01*, Hersonissos, Crete: ACM, pp. 749-753.

Rasmusen, Eric 2007. *Games and Information*, 4th Edition, Malden, MA: Blackwell.

Rass, Stefan and Quanyan Zhu 2016. GADAPT: A Sequential Game-Theoretic Framework for Designing Defense-in-Depth Strategies Against Advanced Persistent Threats. *GameSec* 2016, *LCNS* 9996, edited by Q. Zhu et al. Springer International, pp.314-326.

Sandler, Todd 2001. *Economic Concepts for the Social Sciences*. Cambridge, UK: Cambridge University Press.

Shapiro, Scott J. 2023. *Fancy Bear Goes Phishing*. *The Dark History of the Information Age in Five Extraordinary Hacks*. NY: Farrar, Straus, and Giroux.

Tirenin, W and D Faatz 1999. A Concept for Strategic Cyber Defense. *Proceedings of the IEEE Military Communications Conference* (*MILCOM '99*), New York: IEEE Press, 458-463.

Tsebelis, George 1990. Are Sanctions Effective? A Game Theoretic Analysis. *Journal of Conflict Resolution* 34(1) 3-28.

U.S. Department of Defense 2011. *Cyberspace Policy Report*.

Urbano, Amparo and Jose E. Vila 2002. Computational Complexity and Communication: Coordination in Two-Player Games. *Econometrica* 70(5) 1893-1927.

Wolff, Josephine 2015. Perverse Effects in Defense of Computer Systems: When More is Less. *Journal of Management Information Systems* 33(2) 597-620.

Zhou, Yan, Murat Kantarcioglu, and Bowei Xi, 2021. A Game Theoretic Perspective on Adversarial Machine Learning and Related Cybersecurity Applications. In Charles A. Kamhoua et al. (eds.) *Game Theory and Machine Learning for Cyber Security*, Hoboken, NJ: Wiley for IEEE.