A Game Theoretic Model of Strategic Conflict in Cyberspace

Harrison C. Schramm

David L. Alderson

W. Matthew Carlyle

Nedialko B. Dimitrov

December 11, 2013

Abstract

We study cyber conflict as a strictly competitive, two-person game in discrete time, where each player discovers new exploits according to an independent random process. Upon discovery, the player must decide if and when to exercise a munition based on that exploit. The payoff from using the munition is a function of time that is (generally) increasing. These factors create a basic tension: the longer a player waits to exercise a munition, the greater his payoff because the munition is more mature, but also the greater the chance that the opponent will also discover the exploit and nullify the munition. Assuming perfect knowledge and under mild restrictions on the time-dependent payoff function for a munition, we derive optimal exercise strategies and quantify the value of engaging in cyber conflict. Our analysis also leads to high level insights on cyber conflict strategy.

1 Introduction

Conflict in Cyberspace, or *cyber conflict*, is important at both strategic and tactical levels. In this paper we consider the strategic decisions made by states or other groups about when and how to engage in cyber conflict. The increasing dependency on interconnected networks both in military and civilian life means that little is beyond the reach of cyberspace. Cyberspace plays a central role in our social, economic, and civic welfare. It is, therefore, not surprising that the United States "has identified cyber security as one of the most serious economic and national security challenges we face as a nation" (United States Executive Office of the President 2010). Consequently, security and defense in cyberspace has become an increasingly large part of the defense budget (Stervstein 2011).

A defining characteristic of cyber conflict is the way in which weapons in cyberspace are discovered, developed and employed. Players search for mechanisms that can cause cyber systems to perform in ways not intended in their original design, called *exploits*, and, once found, develop them into one or more *cyber munitions*. These munitions can then be used as part of a cyber attack. In searching for exploits to use against an adversary, a player may also discover flaws in their own system and decide to fix or *patch* them so an adversary cannot use them. Moreover, a player could develop munitions based on an exploit that the adversary independently fixes, thereby making the munitions obsolete. Thus, collections of cyber munitions, or *arsenals*, are dynamic and their effectiveness depends on the relative state of knowledge of the opponents.

In this context, apparently simple questions, such as "how long should we hold a munition in development before using it in an attack?" and "how should we allocate limited resources to offense versus defense?" require novel analytical models. Moreover, the dynamic nature of cyber weapons

development and obsolescence makes it difficult to assess the potency of an arsenal; this is true for assessing our own arsenal as well as an arsenal belonging to an adversary. Clear, useful analysis at the national level is important both for making sound future investment decisions and for creating informed strategic and policy guidance.

To analyze the strategic decisions involved in cyber conflict, we present a highly stylized model using a game theoretic framework—we view cyber warfare as a game consisting of attacks that opposing players exercise at a time of their choosing. Each player discovers, develops, and chooses to exercise attacks to maximize the value of their cyber operations. Our analysis is independent of specific technologies, and does not assume an explicit cyber system or exploit.

More specifically, we model cyber warfare as a two-player Markov game (Thie 1983, Fudenberg and Tirole 1991) where the choices available to each player depend on the number of exploits known by each player and the strength of each player's munitions. In general, there may be multiple exploits that each player discovers, develops into munitions, and uses to attack, but we choose to focus this analysis on a scenario where there is only a single exploit to be discovered. At the beginning of this scenario, neither player knows the exploit. Each player discovers the exploit probabilistically, and upon discovery has a choice about whether to *attack* or *wait*. As soon as a player chooses to attack, then the game terminates and payoffs are determined. In general, the mechanics we develop are *strictly competitive* in the sense of Birmingham (1970). Our specific case studies are zero-sum. In all cases that we consider here, the decisions are chosen simultaneously from the action set of each player.

Using this highly stylized framework, we develop a preliminary theory of cyber games and present a few representative cases only. Because the data analysis required for this model to support real problem instances is highly

dependent on the context and/or possibly sensitive, we present only generic cases.

Under minimal assumptions, our analysis leads to a fundamental insight: *Success requires rapid action.* Our model shows that delays in taking action reduce the chance of a player's success in cyber conflict. Such delays can come from a variety of sources, including bureaucratic or command restrictions. A byproduct of our model is the calculation of how proficient a player must be in other areas to make up for delays in taking action; in most cases the required capability is unattainable. The immediate consequence of this is that command structures in cyberspace should be agile with the correct level of delegation of authority.

The organization of the paper is as follows: In Section 2, we review previous work. In Section 3 we present our model and analysis. In Section 4 we present numerical examples. In Section 5 we present some extensions to our model, and in Section 6 we conclude and discuss areas for future study.

2 Related work

The 2010 report, The Science of Cyber-Security JASON (2010), advocates a multi-disciplinary approach to the study of cyber security, and it specifically recommends borrowing ideas from other sciences, such as physics, cryptography, and biological sciences, including epidemiology. The JASONS introduce a two-player, stationary, discrete time model called the Forwarder's Dilemma as an example of what a game-theoretic analysis might look like. This game considers whether an administrator should forward another system's messages on their network and is similar both in format and solution to the well-known Prisoner's Dilemma found in Fudenberg and Tirole (1991), which, along with Thie (1983) form the basis of the general analysis presented here. Lye and Wing (2005) and Shen et al. (2007b) also consider

cyber attacks in the context of a game. A thorough survey of game theory and cyberspace is by Roy et al. (2010), which develops a taxonomy of cyber game theoretic models with two broad categories:

- Static vs. Dynamic. A 'one shot' cyber conflict game, where players choose plans of action and then execute them simultaneously, is a *static game*. A cyber conflict game with multiple stages and sequential decisions is a *dynamic game*.
- Available Information. Players may have exact, imperfect or no knowledge about their opponent's intentions or capabilities. If the players know the actions of other players once taken, the game is called a game with *perfect information*. If the players know the structure of the game and payoffs, but not the actions, this is called a game with *complete information*. Finally, a game in which the payoffs evolve in time in a random process is a *stochastic game*.

While game theory considers both cooperative and non-cooperative games, work to date on cyber conflict deals only with non-cooperative games. In the taxonomy of Roy et al. (2010), our proposed model is a non-cooperative, dynamic, stochastic game with perfect information. There are several other studies that consider game cyber conflict from a game theoretic point of view, including Shen et al. (2007a), Otrok et al. (2008), Jolyon (2006), and Liu et al. (2006).

The previous study which has the most in common with our approach is that of Lye and Wing (2005), which considers a two-player, stochastic game between an attacker and administrator. Their model considers cyber conflict at the machine level; it focuses on an attacker attempting to find the best policy among a portfolio of several attacks to damage a university computer network. This game theoretic model of Lye and Wing maps to the tactical level of conflict as opposed to our model that is focused at the

strategic level between two players engaged in cyber conflict. Other work has addressed aspects of our model. Cavusoglu et al. (2008) considers the optimum time for a system administrator to release a patch. Kannan and Telang (2005) considers the market value of new exploits, which is a factor in our analysis. Our analysis is different because we consider the two players as seeking to inflict the maximum damage on each other by use of a single exploit. Nguyen et al. (2009) considers a cyber game between an attacker and a defender with incomplete information; our work is different than prior approaches because we consider two players who may either attack or defend.

The concept of an *zero day* attack is important in our development. Two references which discuss these in detail are Bilge and Dumitras (2012) and Patcha and Park (2007). An overall discussion of terminology and concepts in the cyber domain may be found in Szor (2005). A detailed view of attacks, by type is given in Hansman and Hunt (2005).

The goal of this paper is to provide a foundation from which to build more complex models towards the ultimate goal of integrating the cyber domain into the spectrum of conflict analysis, to support strategic models for decision makers at the national level.

3 Analysis

Our analysis starts with two simplifying assumptions. First, we assume that the two opponents are operating computer systems that are sufficiently similar such that the exploits are *symmetric* on each. In practice, we do not expect that two real opponents would be operating identical systems, however, modern software systems are often constructed from common components (e.g., operating systems, open source servers, or standard communication protocols) that it is reasonable to assume that the same vulnersbilities could be shared by both opponents.

Our second simplifying assumption is that there is only one exploit to be discovered, used, or patched. Extending our analysis to consider multiple exploits simultaneously would require a significant expansion of the state-space as well as additional decision variables and constraints, and we feel that this extra machinery, although more realistic, would detract from some of the basic insights we obtain regarding the timing of attack and patch decisions. We defer such analysis to future work.

3.1 Model Foundation

As defined previously, a computer system may contain exploits; these are unknown until discovered, after which they can be fixed in the form of a patch or weaponized into a munition. We model the life-cycle of a single cyber exploit as a four-stage process.

3.1.1 Discovery of the exploit

We model the discovery of a single exploit by each player as a random process, occurring independently for each player, which may depend on factors such as training, investment, experience and luck.

3.1.2 Development of munition

Once an exploit is discovered, a player can develop a munition based on the exploit. We assume that there is a relationship between the length of time that a player knows about an exploit and the effectiveness of the munition he develops based on that exploit. Munitions may only be developed for known exploits.

3.1.3 Employment

Once a munition is developed, it can be employed at will against an adversary in an attack.

3.1.4 Obsolescence

Consider a game between two players, Player 1 and Player 2. If Player 1 discovers an exploit in his system and patches it before Player 2 can develop and employ a munition based on that exploit, then that munition becomes obsolete. A patch can be thought of as an attack with value zero. While not explicitly examined by the perfect information model presented in this paper, when there is a lack of information, patches may play an important role in the player's strategy. For example, we may imagine a cyber conflict with multiple attacks and limited resources where a player would only choose to develop certain discovered exploits into munitions, while creating patches for the remainder in his portfolio. In a game of perfect information, as the one we consider in this paper, the ability to patch forces immediate action on the part of the adversary. The ability to patch is what leads to analytic results of Section 3.3.2.

Uncertainties about the obsolescence of a player's own arsenal are a key dimension in the analysis of cyber conflict. For the purposes of this analysis, we assume that a player who is aware of an exploit also knows whether the other player(s) are aware of the same exploit; this removes one type of uncertainty. For a player who is unaware of an exploit, we assume neither player knows how long it will be until the unaware player discovers the exploit. This uncertainty in discovery times is the fundamental tension that our model seeks to explore.

3.2 Model Formulation

Our model focuses on a strategic cyber conflict between two players, where there is a single exploit that is resident in the systems of each player at the beginning of the game, to be discovered. Let i index the players $i \in \{1, 2\}$. The mathematical notation used to describe the game falls into three broad categories: Discovery, Development, and Employment.

3.2.1 Discovery

Let T be the duration of time that an exploit has existed, which we also call the *clock time*. Without loss of generality, we assume that the game starts when the exploit is created. We create a discrete time model, with T increasing over the set of positive integers. If the exploit was part of the original system, then T is the age of the system. If the exploit was introduced as part of a software upgrade, then T is the age of the upgrade. Let d_i be player i's *discovery time*—it is the moment in clock time that player i discovers the exploit. We define $\tau_i = \max(0, T - d_i)$ to be the relative time that player i has known about the exploit; we call this player i's *holding time*. By definition, if player i is not aware of the exploit, then $\tau_i = 0$. We define a *state of the cyber game*, S, as:

$$S = \langle T, \tau_1, \tau_2 \rangle,$$

where the elements of this three-tuple represent how long the exploit has existed, how long Player 1 has known the exploit, and how long Player 2 has known the exploit, respectively.

3.2.2 Development

A player's success in cyber conflict depends both on his ability to discover exploits and his ability to develop effective munitions. We assume that at any moment following the discovery time d_i, player i has the ability to create

and deploy a perfectly effective patch. However, we assume that the act of deploying the patch effectively announces it to the adversary; so patching nullifies all munitions based on that exploit, and this ends the game for both sides. Let $p_i(T)$ denote the probability that player i discovers an exploit as clock time progresses from period T to period $T+1$. For convenience, let $q_i(T) = 1 - p_i(T)$. Let $a_i(\tau_i)$ be the value of an attack by player i using a munition developed using a holding time of τ_i. The value of an attack is a function of τ instead of T because we assume that once the exploit is known the effectiveness of the munition depends on holding time and not clock time. We impose two constraints on $a_i(\tau_i)$. First, we assume

$$a_i(0) = 0,$$

namely that if an exploit is not known, then an attack based on it has no value. Additionally, we assume

$$0 \leq a_i(\tau) \leq B_i,$$

where B_i is an arbitrary finite upper bound, thus disallowing cyber attacks with either a negative value or an infinite value.

3.2.3 Employment

Once a player has a cyber munition, he may choose to use it. Let $\theta_i(T)$ denote the action set of player i at time T. We define $\theta_i(T) \subseteq \{W, A\}$ where

- W: Wait. While a player is waiting, he has either not yet discovered the exploit ($\tau_i = 0$) or he knows about the exploit ($\tau_i > 0$) and may be developing his munition.
- A: Attack. When a player attacks he receives the value of his attack at that time. Attacking also broadcasts the attack's underlying exploit to all players.

A player who does not know the exploit has a singleton action set, $\{W\}$, and a player that does know the exploit has the full action set, $\{W, A\}$.

3.3 Zero sum game with perfect information

To fully specify the game, we must define action sets for each player, and the utilities for player's actions. We assume a zero sum strategic conflict, i.e., that any utility gain by one player results in an equal utility loss by the opponent. We use the convention that Player 1 is a maximizing player and Player 2 is a minimizing player. We assume that each player knows the state of the Markov game, S. But this perfect information assumption does not mean that a player knows the exploit. A player is still limited by his action set. For example, if the state of the game is $\langle T, 1, 0\rangle$, it means that: Player 1 knows the exploit, has a holding time of 1, and has an action set of $\{W, A\}$; while, Player 2 does not know the exploit, has a holding time of 0, and therefore has an action set of solely $\{W\}$.

3.3.1 Markov game transitions

The discovery and development of attacks is modeled as transitions in the state of the Markov game. The game begins in the state $\langle 0, 0, 0\rangle$ and proceeds in discrete rounds. In each round, the clock time T increases deterministically. Each player i has holding time $\tau_i = 0$ until the player discovers the exploit. Exploit discovery happens with probability $p_i(T)$ for player i in round T. Once an exploit is discovered by a player, the player's holding time increases deterministically. The resulting transitions of the Markov game state are summarized in Table 1. A visual depiction of the states of the game is presented in Figure 1.

Let $V\langle T, \tau_1, \tau_2\rangle$ define the value of the game in state $\langle T, \tau_1, \tau_2\rangle$; this value represents the expected value to the players if they play the game starting

	$\tau_1 = 0$	$\tau_1 > 0$
$\tau_2 = 0$	$\langle T,0,0\rangle \xrightarrow{(1-p_1(t))(1-p_2(t))} \langle T+1,0,0\rangle$ $\xrightarrow{p_1(t)(1-p_2(t))} \langle T+1,1,0\rangle$ $\xrightarrow{p_2(t)(1-p_1(t)} \langle T+1,0,1\rangle$ $\xrightarrow{p_1(t)p_2(t)} \langle T+1,1,1\rangle$ $\theta_1 = \{W\}$ $\theta_2 = \{W\}$	$\langle T,\tau_1,0\rangle \xrightarrow{1-p_2(t)} \langle T+1,\tau_1+1,0\rangle$ $\xrightarrow{p_2(t)} \langle T+1,\tau_1+1,1\rangle$ $\theta_1 = \{A,W\}$ $\theta_2 = \{W\}$
$\tau_2 > 0$	$\langle T,0,\tau_2\rangle \xrightarrow{1-p_1(t)} \langle T+1,0,\tau_2+1\rangle$ $\xrightarrow{p_1(t)} \langle T+1,\tau_1+1,1\rangle$ $\theta_1 = \{W\}$ $\theta_2 = \{A,W\}$	$\langle T,\tau_1,\tau_2\rangle \xrightarrow{w.p.\,1} \langle T+1,\tau_1+1,\tau_2+1\rangle$ $\theta_1 = \{A,W\}$ $\theta_2 = \{A,W\}$

Table 1: Markov game state transitions and action sets as a function of the $\langle T,\tau_1,\tau_2\rangle$ state of the game. The game always starts in $\langle 0,0,0\rangle$. As player i discovers the exploit τ_i becomes greater than zero and i's action set includes attack. The payoffs of the players are described using two non-negative functions, $a_1(\tau_1)$ and $a_2(\tau_2)$. The function $a_i(\tau_i)$ describes the effectivness of a muntion developed by player i for τ_i time periods. We assume the game is zero sum, meaning that the payoffs can be described with a single number as opposed to two numbers – let a positive value be in faveor of player 1 and a negative value be in favor of player 2. If player 1 (2) plays "A", it contributes $a_1(\tau_1)$ ($-a_2(\tau_2)$) to the payoff. The payoffs are further described in Tables 2 and 3.

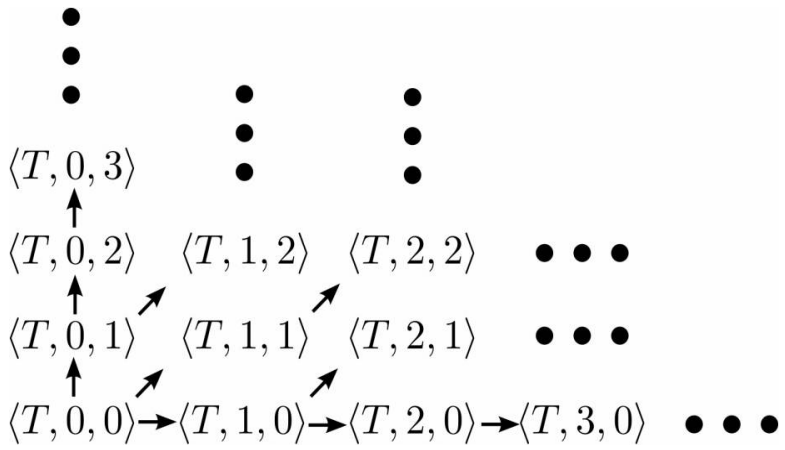

Figure 1: Diagram of states in the Markov Game. The arrows in the diagram show the possible transitions from one state to another, as described in Table 1. The horizontal axis describes increases in holding time for Player 1, τ_1,and the vertical axis describes increases in holding time for Player 2, τ_2.

at that state. Because the game is zero-sum, payoffs for both players can be described by a single value. To analyze the game, we seek to characterize this value function. In particular, $V\langle 0,0,0\rangle$ is the value of engaging in cyber conflict. We seek to characterize $V\langle T,\tau_1,\tau_2\rangle$ for every state of the Markov game. We proceed in our analysis by considering three cases on τ_1, τ_2.

3.3.2 Both players know the exploit

The case where both players know the exploit is characterized by $\tau_1 > 0, \tau_2 > 0$; here, both players have full action sets, meaning each may attack or wait. Table 2 represents the payoffs of the Markov game in such a state in matrix form. Each entry in the matrix contains a single real number, since the game is zero sum. If both players wait, the value is determined by future play. If one player attacks and the other waits, the attacking player receives the full value of his munition. If both players attack simultaneously, the difference of the attack values gives the result of the game.

	Player 2 plays: W	Player 2 plays: A
Player 1 plays: W	$V\langle T+1, \tau_1+1, \tau_2+1\rangle$	$-a_2(\tau_2)$
Player 1 plays: A	$a_1(\tau_1)$	$a_1(\tau_1) - a_2(\tau_2)$

Table 2: Payoff matrix for the Markov game when both players know the exploit. The payoff associated with "Wait, Wait" depends on the future play evolution of the game.

This leads to the following observation:

Theorem 1 *For any game state $\langle T, \tau_1, \tau_2\rangle$ such that $\tau_1 > 0$ and $\tau_2 > 0$, "Attack, Attack" is an iterated elimination of dominated strategies equilibrium with a value of $a_1(\tau_1) - a_2(\tau_2)$.*

Proof: Suppose $V\langle T+1, \tau_1+1, \tau_2+1\rangle \geq 0$. Then $V\langle T+1, \tau_1+1, \tau_2+1\rangle \geq -a_2(\tau_2)$ and $a_1(\tau_1) \geq a_1(\tau_1) - a_2(\tau_2)$ and "Attack" is a dominating strategy for Player 2. Given Player 2 plays "Attack", Player 1 must also play "Attack" and "Attack, Attack" is an iterated elimination of dominated strategies equilibrium. A symmetric argument holds if $V\langle T+1, \tau_1+1, \tau_2+1\rangle \leq 0$. ■

Theorem 1 results in the following corollary:

Corollary 1 *If the game starts in $\langle T, \tau_1, \tau_2\rangle$ with $\tau_1 > 0$ and $\tau_2 > 0$, the game terminates immediately and*

$$V\langle T, \tau_1, \tau_2\rangle = a_1(\tau_1) - a_2(\tau_2).$$

Interpreting the results of Theorem 1 and the above corollary, a game starting in $\langle T, 0, 0\rangle$, $T \geq 0$ ends optimally no later than one of the following states is reached: $\langle T, 1, \tau_2\rangle$ or $\langle T, \tau_1, 1\rangle$. However, the game may also end earlier, if a player who discovers the exploit chooses to attack before the second player has discovered the exploit. Because for each i, $a_i(\cdot)$ has a

	Player 2 Plays: Wait
Player 1 Plays: Wait	Y
Player 1 Plays: Attack	$a_1(\tau)$

Table 3: Payoffs for the case where player 1 knows the exploit and player 2 does not. Because player 2 does not know that the exploit exists, he may only wait, and the matrix reduces to a single column.

unique associated τ_i, for ease of exposition we drop the index i from future uses of τ. For the remainder of this paper, statements like $a_2(\tau)$ should be understood to mean $a_2(\tau_2)$.

3.3.3 Only one player knows the exploit

For simplicity, we develop the theory from a state where Player 1 has the exploit and Player 2 does not. The analysis follows identical lines in the opposing situation. In this case, Player 1 has a full action set and Player 2 may only wait to discover the exploit,

$$\theta_1 = \{A, W\}, \theta_2 = \{W\}.$$

Suppose the state of the game is $\langle T, \tau, 0\rangle$. We define

$$Y = (1 - p_2(T))V\langle T+1, \tau+1, 0\rangle + p_2(T)V\langle T+1, \tau+1, 1\rangle,$$

to be the expected utility if both players choose to wait at time T. Table 3 displays the payoffs in matrix form.

Player 1 prefers to attack if $Y \leq a_1(\tau)$. The fundamental analytic question is 'from which states does Player 1 prefer to attack?' If Player 2 discovers the exploit, the game transitions to the scenario described in section 3.3.2, and immediately concludes as specified in Theorem 1. We characterize states $\langle T, \tau, 0\rangle$ from which Player 1 prefers to attack as follows. We define

$v_\tau(h)$as the expected utility to Player 1 if he waits h time periods before attacking, starting in state $\langle T, \tau, 0\rangle$. In particular, we have:

$$\begin{aligned}
v_\tau(0) &= a_1(\tau) \\
v_\tau(1) &= q_2(T) \cdot a_1(\tau+1) + p_2(T) \cdot (a_1(\tau+1) - a_2(1)) \\
v_\tau(2) &= q_2(T+1)q_2(T) \cdot a_1(\tau+2) + \\
&\qquad p_2(T+1)q_2(T) \cdot (a_1(\tau+2) - a_2(1)) + \\
&\qquad p_2(T) \cdot (a_1(\tau+1) - a_2(1)),
\end{aligned}$$

leading to

$$v_\tau(h) = a_1(\tau+h) \cdot \prod_{k=0}^{h-1} q_2(T+k) + \sum_{k=0}^{h-1} (a_1(\tau+k) - a_2(1)) \cdot p_2(T+k) \prod_{j=0}^{k-1} q_2(T+j). \tag{1}$$

The definition of $v_\tau(h)$ allows us to evaluate the states from which Player 1 prefers to attack. Player 1 prefers to attack rather than wait in state $\langle T, \tau, 0\rangle$ if and only if the following holds:

$$a_1(\tau) = v_\tau(0) \geq v_\tau(h) \text{ for all } h \geq 1. \tag{2}$$

This statement mirrors our intuition that a player should attack if only if an immediate attack results in a higher utility than waiting for any number of turns before attacking.

Theorem 2 *If $a_1(\tau)$is concave and nondecreasing, and $p_2(T)$is nondecreasing, then $v_\tau(0) \geq v_\tau(1)$ implies that Player 1 should attack in state $\langle T, \tau, 0\rangle$ (i.e., Player 1 can never do better by waiting).*

Proof: We proceed by showing that the theorem assumptions imply that

$$v_\tau(0) \geq v_\tau(h) \text{ for all } h \geq 2.$$

Consider the quantity

$$
\begin{aligned}
v_\tau(h+1) - v_\tau(h) = a_1(\tau+h+1) \prod_{k=0}^{h} q_2(T+k) - \\
a_1(\tau+h) \prod_{k=0}^{h-1} q_2(T+k) + (a_1(\tau+h+1) - a_2(1)) p_2(T+h) \prod_{j=0}^{h-1} q_2(T+j) \\
= \prod_{k=0}^{h-1} q_2(T+k)[a_1(\tau+h+1) - a_1(\tau+h) - a_2(1)p_2(T+h)] \ .
\end{aligned}
$$

We know that $v_\tau(0) \geq v_\tau(1)$, which implies that

$$
\begin{aligned}
0 &\geq v_\tau(1) - v_\tau(0) \\
&= a_1(\tau+1) - a_1(\tau) - p_2(T)a_2(1) \\
&\geq a_1(\tau+h+1) - a_1(\tau+h) - p_2(T)a_2(1),
\end{aligned}
$$

where the last inequality came from the fact that $a_1(\cdot)$ is concave and nondecreasing. Continuing with the last expression above, we have

$$
\begin{aligned}
0 &\geq a_1(\tau+h+1) - a_1(\tau+h) - p_2(T)a_2(1) \\
&\geq a_1(\tau+h+1) - a_1(\tau+h) - p_2(T+h)a_2(1),
\end{aligned}
$$

where the last inequality came from the fact that $p_2(\cdot)$is nondecreasing and $a_2(1)$is nonnegative. Finally, multiplying both sides of the inequality by the positive number $\prod_{k=0}^{h-1} q_2(T+k)$, gives

$$
\begin{aligned}
0 &\geq \prod_{k=0}^{h-1} q_2(T+k)[a_1(\tau+h+1) - a_1(\tau+h) - p_2(T+h)a_2(1)] \\
&= v_\tau(h+1) - v_\tau(h)
\end{aligned} \tag{3}
$$

We can complete the proof as follows:

$$
\begin{aligned}
v_\tau(h) - v_\tau(0) = v_\tau(h) - v_\tau(h-1) + \\
v_\tau(h-1) - v_\tau(h-2) + \\
v_\tau(h-2) \ldots \\
v_\tau(1) - v_\tau(0).
\end{aligned}
$$

Each of the paired terms on the right hand side is smaller than zero, by Equation (2), thus we have

$$v_\tau(h) - v_\tau(0) \leq 0,$$

completing the proof. ■

For the remainder of this paper we assume stationary probabilities $p_i(T) = p_i \; \forall T$. Theorem 2 shows that $v_\tau(0) \geq v_\tau(1)$ is sufficient to prefer Attack at a holding time of τ while Equation (1) shows that $v_\tau(0) \geq v_\tau(1)$ is necessary to prefer Attack at τ. Therefore, from state $\langle T, 1, 0\rangle$ player 1 waits for $k^* = \min_k \{v_k(0) \geq v_k(1)\}$ turns before attacking. Substituting the definition of $v_\tau(\cdot)$ we can write this as $k^* = \min_k \{a_1(k+1) - a_1(k) \leq p_2 a_2(1)\}$. The set in the definition of k^* is never empty when $a_1(\cdot)$ is bounded, concave, and nondecreasing and $p_2 a_2(1)$ is not identically zero, meaning that Player 1 will eventually prefer to attack. We conclude that

$$V\langle T, 1, 0\rangle = v_0(k^*) \tag{4}$$

While we presume that most cases will have nondecreasing a_1, a_2, p_1, p_2 functions, there is no reason that it must be so. Nondecreasing functions model situations where the passage of time brings increased capability, both in development and detection. However, there may be interesting, and operationally relevant, cases where the functions are decreasing. Although we do not present detailed results here, the value functions in these alternate situations may evaluated directly by using Equations (1) and (2).

3.3.4 Neither player has the exploit

In this case, the game has been in play for an unknown amount of time and $\tau_1 = \tau_2 = 0$; therefore both players have singleton action sets,

$$\theta_1 = \{W\}$$
$$\theta_2 = \{W\}.$$

Using the theory previously developed, the value of the game given Player 1 discovers the exploit first is: $V\langle T,1,0\rangle$. Similarly, if Player 2 discovers the exploit first the value is:$V\langle T,0,1\rangle$. In the case where both players simultaneously discover the exploit $V\langle T,1,1\rangle = a_1(1) - a_2(1)$. Because the state $\langle T,0,0\rangle$transitions into previously analyzed states, we are only concerned with the first transition. For stationary discovery probabilities the next state transition probabilities out of $S = \langle T,0,0\rangle$ are:

$$\begin{aligned}
\Pr\{\text{next state is}\,\langle T,1,0\rangle\} &= \gamma_{1,0} = \frac{p_1(1-p_2)}{p_1(1-p_2)+p_2(1-p_1)+p_1p_2} \\
\Pr\{\text{next state is}\,\langle T,0,1\rangle\} &= \gamma_{0,1} = \frac{p_2(1-p_1)}{p_1(1-p_2)+p_2(1-p_1)+p_1p_2} \\
\Pr\{\text{next state is}\,\langle T,1,1\rangle\} &= \gamma_{1,1} = \frac{p_1p_2}{p_1(1-p_2)+p_2(1-p_1)+p_1p_2},
\end{aligned}$$

where we have introduced the γ notation for brevity. The value of the game starting from $\langle T,0,0\rangle$ is

$$\begin{aligned}
V\langle T,0,0\rangle &= \gamma_{1,0}V\langle T,1,0\rangle - \gamma_{0,1}V\langle T,0,1\rangle + \gamma_{1,1}V\langle T,1,1\rangle \\
&= \gamma_{1,0}v_0^1(k^{1*}) - \gamma_{0,1}v_0^2(k^{2*}) + \gamma_{1,1}\left(a_1(1) - a_2(1)\right),
\end{aligned} \tag{5}$$

where the negative sign comes from the fact that Player 1 is a maximizing player, and Player 2 is a minimizing player; and $v_0^1(\cdot)$, k_1^* denote results of Equations (3) and (4) if Player 1 is the first to discover the exploit while $v_0^2(\cdot)$, k_2^* denote the results of Equations (3) and (4) if Player 2 is the first to discover the exploit.

4 Numerical analysis

In this section, we consider some concrete examples of the theory developed in the previous section. Unless otherwise specified, we assume $p_i(T) = p_i, \forall T$ and $p_i \neq 0$. As a notational convenience we will denote the value of any particular example as V^n where n is the example number. The examples in the following sections differ in the functional form we assume for the $a_i(\cdot)$

functions. In each section, we give a possible interpretation on where that particular functional form may arise.

4.1 Scenario 1: constant effectiveness functions

Suppose that Players 1 and 2 both have attack value functions such that:

$$a_i(0) = 0$$
$$a_i(\tau) = c_i \ \forall \ \tau \geq 1.$$

The case of constant $a_i(\cdot)$ functions represents the case where the maximum value of the attack is realized in the first turn after the exploit is discovered. This would represent cases where either the munition development team is very quick in comparison to the other activities in the game, or the munition had been previously developed and awaiting a suitable vulnerability to make it viable.

Because $a_i(\tau)$ is concave and nondecreasing for both players, we can use Theorem 2 to compute the optimal attack time for each player, k_i^* for $i = 1, 2$, which is 1 for both players. We may directly compute the value of the game using Equation (5):

$$V^1 = \frac{p_1(1-p_2)a_1(1) - p_2(1-p_1)a_2(1) + p_1p_2(a_1(1) - a_2(1))}{p_1(1-p_2) + p_2(1-p_1) + p_1p_2}$$

In particular, Player 1 will have a positive expected payoff if and only if:

$$p_1 a_1\,(1) > p_2 a_2\,(1)$$

In this case, a player may make up for a deficiency in either discovery or development by being strong in the other area. Because $0 \leq p_i \leq 1$ these tradeoffs are implicitly limited.

4.2 Scenario 2: linearly increasing effectiveness

Suppose Players 1 and 2 have attack functions such that:

$$\begin{aligned} a_i(0) &= 0 \\ a_1(\tau) &= \tau \quad 1 \le \tau \le 5 \\ a_1(\tau) &= 5 \quad \forall \tau \ge 5 \\ a_2(\tau_2) &= c \quad \forall \tau_2 \ge 1. \end{aligned}$$

Linearly increasing $a_i(\cdot)$ considers linear improvement in the value of the attack function with time invested. This case generalizes the constant $a_i(\cdot)$ example in the previous section. While both examples demonstrate a maximum value of the attack–the maximum cap of 5 in the linear growth above–a linear increase requires some time before the maximum can be achieved.

This function is also concave, increasing and we may use Theorem 2 to determine the optimal attack time, k_i^*, for both players. Specifically, $k_2^* = 1$ and k_1^* is dependent on the values of p_2 and c as follows:

$$k_1^* = \begin{cases} 1 & \text{if } p_2 c \ge 1 \\ 5 & \text{otherwise} \end{cases}$$

As verification, we compute the values of $v_\tau(h)$ for $h = 1, 2...5$. We see in Figure 2 that the maximizing value is $h = 5$. For example, if $a_2(1) = 1, p_2 = .2$

Knowing k^* for both players, we may compute the value of the game, $V^2\langle T, 0, 0\rangle$ as a function of p_1; see Figure 3.

4.3 Non-monotone effectiveness

Suppose that $a_2(1) = 1$, $p_2 = .3$, and Player 1's value function has a single dip, specifically $a_1(\tau) = (1,\ 2,\ 3,\ 4,\ 5,\ 3,\ 6)$ as shown in Figure 4. Non-monotone $a_i(\cdot)$, is an operationally important case. The non-monotonicity is caused not by "lost learning" on the attacker's part, but rather captures

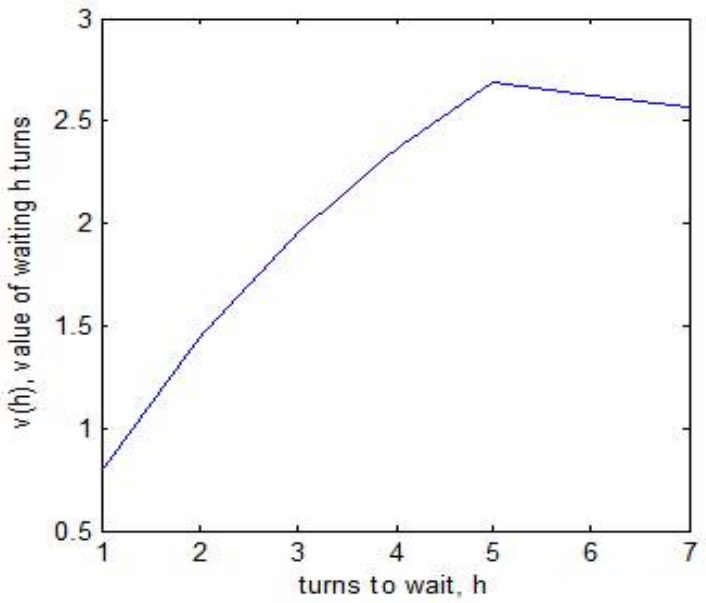

Figure 2: Value of Scenario 2 from Player 1's point of view. The vertical axis plots the value, $v_\tau(h)$, as a function of the number of time periods Player 1 waits before attacking, h. The value function increases to the point $h = 5$, and decreases afterward. By Theorem 2, this implies that Player 1's optimal attack time, k_1^*, is 5.

the idea that there are times–due to political, military, or environmental considerations–when an attack may have a payoff before and after some specific event or circumstance, but not be as effective during the circumstance.

Because $a_1(\tau)$ is not concave and increasing, we cannot apply to Theorem 2. Here we need to actually compute the numeric values of $v_\tau(h)$. Performing this calculation, we see that ${k_1}^* = 5$ and it is not advisable to wait through the non-increasing region.

A decision maker may want to know what value of $a_1(7)$ would change Player 1's decision? We answer this question by performing a line search on

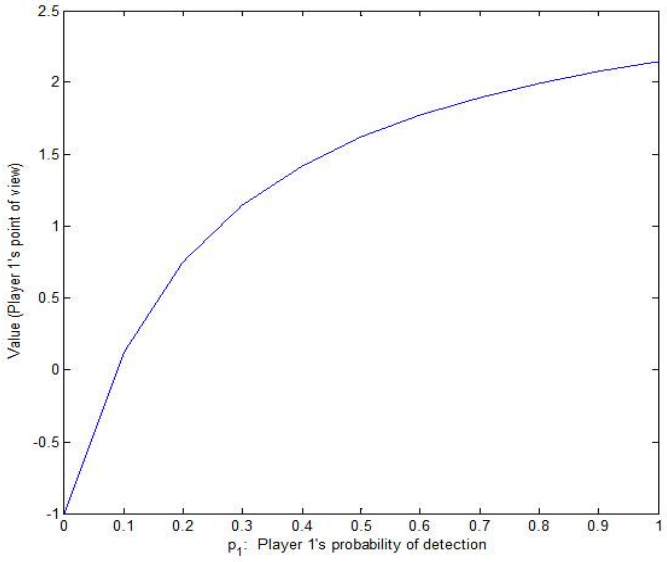

Figure 3: Value of Scenario 2 as a function of Player 1's probability of discovering the Exploit, p_1. Here we see that the value of the game is a concave function of Player 1's probability of detecting the exploit; increases in detection probability at low detection values provide a bigger increase in the game value than increases in detection probability at high detection values.

$a_1(7)$ and determine the threshold value is ≈ 6.6.

5 Extension: delayed action

It may be the case that a player discovers an exploit and cannot take action; specifically, he is unable (or not allowed) to attack, patch, or work towards development of a munition for some predetermined fixed time after discovery of an exploit. This may be due to legal, policy, or organizational limitations.

Suppose Player 1 has a rule where he must wait w time periods after discovery before any attack, patch or development of a munition. Consistent with our previous definition of perfect information, if Player 2 has the

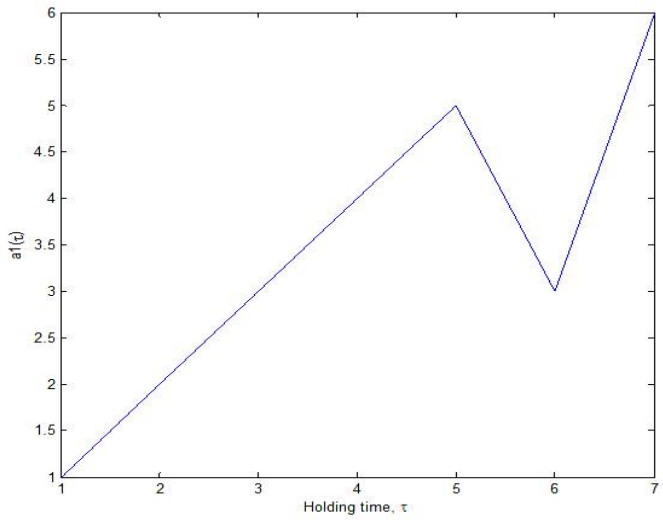

Figure 4: The effectiveness function for Scenario 3. Unlike our previous examples, the value of Player 1's attack has a dip at $\tau_1 = 6$. In this scenario, Theorem 2 no longer applies in finding the optimal attack time k_1^*.

exploit, he learns if Player 1 knows the exploit. Player 2 also knows the existence and duration of Player 1's delay rule. We wish to understand the value of this delayed version of our game, which we denote as $V^w\langle\cdot\rangle$. If both players have the exploit, Player 2 can wait and exercise his munition the turn before Player 1 is able to begin work; therefore,

$$V^w\langle T, 1, 1\rangle = -a_2\,(w-1)\,.$$

If Player 2 has the exploit and Player 1 does not, Player 2 may continue developing his munition until Player 1 discovers the exploit, and an additional $(w-1)$ time periods before attacking; therefore,

$$V^w\langle T, 0, 1\rangle = -\sum_{i=0}^{\infty} p_1(1-p_1)^i a_2\,(i+w).$$

Finally, if Player 1 has the exploit and Player 2 does not, there are two possibilities. First, Player 1 may retain sole knowledge of the exploit until

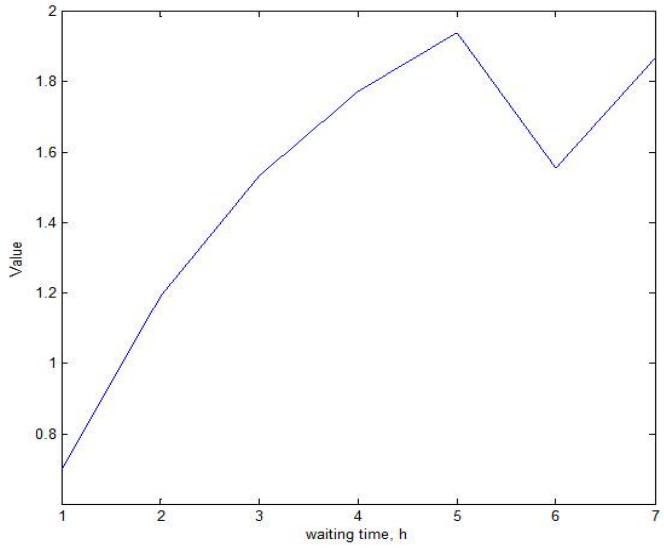

Figure 5: Player 1's value as a function of waiting time, h in Scenario 3. We see that the payoff for waiting to $h = 7$ is less than executing at $h = 5$.

the end of the waiting period, or, second, Player 2 may discover the exploit during Player 1's forced delay time; therefore,

$$V^w \langle T, 1, 0 \rangle = (1 - p_2)^w V \langle T, 1, 0 \rangle - \sum_{i=1}^{w-1} p_2 (1 - p_2) a_2 (w - i).$$

We may combine these expressions to write:

$$\begin{aligned} V^w \langle T, 0, 0 \rangle = \gamma_{1,0} \left[(1 - p_2)^w V \langle T, 1, 0 \rangle - \sum_{i=1}^{w-1} p_2 (1 - p_2) a_2 (w - i) \right] \\ - \gamma_{0,1} \left[\sum_{i=0}^{\infty} p_1 (1 - p_1)^i a_2 (i + w) \right] - \gamma_{1,1} a_2 (w - 1) . \end{aligned} \tag{6}$$

The implication of this is that unproductive waiting times are damaging to a player's prospects in cyber conflict.

Consider the specific example of two evenly matched players with bounded, linear development functions, thus: $p_1 = p_2 = .1$, $a_1 (\tau) = a_2 (\tau) = \tau$ for $0 < \tau \leq 10$ and $a_1 (\tau) = a_2 (\tau) = 10$ for $\tau > 10$. By symmetry, $V \langle T, 0, 0 \rangle = 0$ for this game when neither player is forced to wait. Now consider the case

where Player 1 has a waiting time w. We plot the player 1's expected payoff as a function of w in Figure 6.

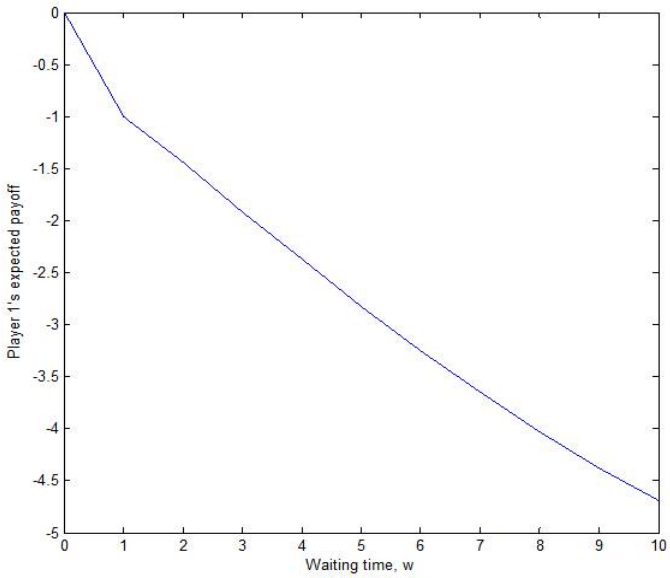

Figure 6: Player 1's utility curve as a function of waiting time w against an evenly matched opponent. We see that Player 1's utility drops off rapidly from an expected value of zero, with the implication that waiting is costly.

We can also ask 'How good does Player 1's detection probability $p_1$need to be in order to make up for a given waiting time w?' Figure 7 shows the adjustment required in this example; for waiting times longer than 5 periods, even perfect detection does not achieve parity.

The lesson of Figures 6 and 7 is that waiting times are costly and adversely affect one's prospects in cyber conflict.

6 Conclusion and future work

We have developed and exercised a limited, stylized model. Real situations, of course, have many differences from the idealized mathematics; the utility

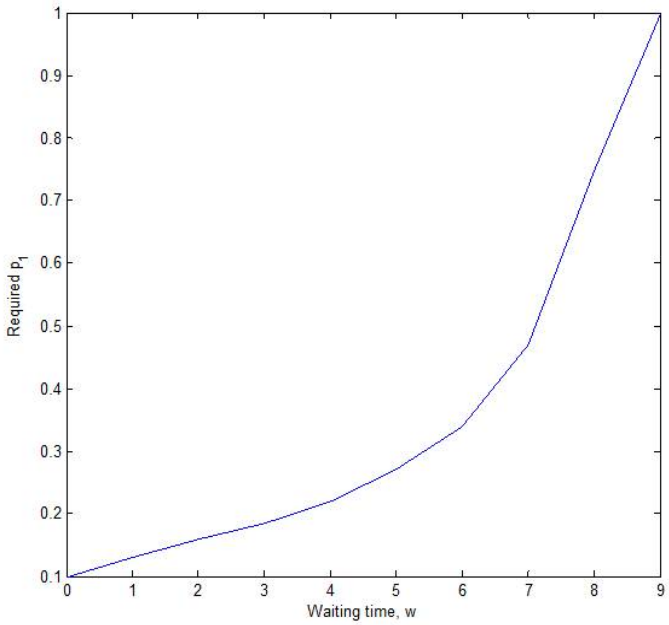

Figure 7: Player 1's required detection probability p_1 required to achieve $V^w\langle 0,0,0\rangle = 0$ as a function of waiting time, w. Player 1's required capability increases rapidly and, because p_1 may never be greater than 1, parity is unachievable after $w = 5$

of this work is to define the cyber conflict problem with perfect information. Additionally, we:

- Demonstrate a framework for analyzing the problem, which may be extended to a wider class of military problems within and beyond cyberspace, and
- Demonstrate that in cyber conflict, idle waiting times are damaging, and we provide a means to calculate their disutility.

This paper considered (a) a symmetric problem involving (b) a single attack that takes place in (c) discrete time and (d) with perfect information—four idealizations that help us begin to tackle the problem of cyber conflict.

Relaxing each of these assumptions provides avenues for additional research. Of these, we believe that relaxing assumption (d) perfect information appears to be the richest area to explore in the future, and with this exploration come considerations of credibility, reputations, and risk taking.

In this analysis, we have used "holding time" as a proxy for development effort, the idea being that the longer a player holds the exploit, the more effort that goes into developing the munition. In general, the development of a munition will depend on many kinds of resources, including labor, computational resources, equipment, and time. If multiple, independent exploits exist, then we would also want to include in our model the additional resource management decisions surrounding the simultaneous development of multiple munitions (or patches). If multiple, *synergistic* exploits exist (i.e., the combined effect of using two exploits simultaneously is significantly different from the sum of the effects of the two exploits if employed singly), then the model would need to account for this as well.

Extending the analysis here to consider the asymmetric (and admittedly more realistic) case where opponents do not share the same vulnerabilities will be an important topic of future work. In practice, we expect that real opponents may share some, but not all vulnerabilities. In a perfect information setting, extending the model we present to asymmetric vulnerabilities involves expanding the state space of the Markov game. The expanded state could include data on how many of the shared, own, or opponent's vulnerabilities are known. Further realism could be added by differentiating individual vulnerabilities, at the cost of complicating the analysis.

Perhaps the most important avenue of future work is exploring imperfect information settings. The Markov game framework we present provides a strong basis for extensions into imperfect information. Specifically, imperfect information could be modeled as a belief probability distribution over the state space of the Markov game. In other words, the players do not

know exactly which state of the Markov game they are in. Instead, each player has their own probability distribution over the state space of the game–based on the information the player has observed over the course of the game. In each round, each player would make decisions based their own belief distribution, and the results of both players' decisions would impact the belief distributions for the next round of play. Exploring simple versions of the cyber conflict game under imperfect information would be the most promising future avenue for yielding insights into the problem, though it is clearly analytically difficult.

Other important next steps in game theoretic modeling of cyber conflict will be to determine functional forms for the rate of detecting vulnerabilities and the rate of developing exploits because these create the two fundamental tensions explored in this paper. In considering the rate and longevity of vulnerabilities, we recommend following the approach taken in Bilge and Dumitras (2012). For the development of munitions, we recommend considering software development generally; a good starting place is Nikula et al. (2010).

Future work may also focus on finding, analyzing and using data to populate the model we propose. The sources of such information might be sensitive, or even classified; however, we note that absolute values the parameters (e.g., the probability of detecting an exploit) are not required in order for the model to be insightful. Rather, it might be sufficient to analyze sensitivity to a range of values for the models to yield insight about the competitive dynamics at work in cyber conflict.

Acknowledgments

Commander Schramm would like to thank the Office of Naval Intelligence for their support. Professors Alderson and Carlyle would like to thank the

Office of Naval Research for their support.

References

I. Adler, C. Daskalakis, and C.H. Papadimitriou. A note on strictly competitive games. *Lecture notes in computer science.*, 5929:471–474, 2009. ID: 501353377.

L. Bilge and T. Dumitras. Before we knew it: An empirical study of zero-day attacks in the real world. *Proc ACM Conf Computer Commun Secur Proceedings of the ACM Conference on Computer and Communications Security*, pages 833–844, 2012. ID: 820188668.

Robert L Birmingham. The generality of neutral principles: A game-theoretic perspective. *California Law Review*, pages 873–884, 1970.

H. Cavusoglu, H. Cavusoglu, and J. Zhang. Security patch management: Share the burden or share the damage? *Management Science*, 54(4): 657–670, 2008. ID: 4909178689.

D. Fudenberg and J. Tirole. *Game theory*. MIT Press, Cambridge, Mass., 1991. ISBN 0262061414 9780262061414. ID: 23180038.

S. Hansman and R. Hunt. A taxonomy of network and computer attacks. *Computers & Security Computers & Security*, 24(1):31–43, 2005. ID: 4647009749.

JASON. Science of cyber-security. Technical Report JSR-10-102, McLean: MITRE Corporation, 2010.

C. Jolyon. A game theoretic formulation for intrusion detection in mobile ad hoc networks. *International Journal of Network Security*, 2(2):131–137, 2006. ID: 5113581049.

K. Kannan and R. Telang. Market for software vulnerabilities? think again. *Management Science*, 51(5):726–740, 2005. ID: 4894310187.

Y. Liu, C. Comaniciu, and H. Man. Modelling misbehaviour in ad hoc networks: a game theoretic approach for intrusion detection. *INTERNATIONAL JOURNAL OF SECURITY AND NETWORKS*, 1(3/4): 243–254, 2006. ID: 206285869.

K.W. Lye and J.M. Wing. Game strategies in network security. *INTERNATIONAL JOURNAL OF INFORMATION SECURITY*, 4:71–86, 2005. ID: 210274579.

K.C. Nguyen, T. Basar, and T. Alpcan. Security games with incomplete information. *IEEE Int Conf Commun IEEE International Conference on Communications*, 2009. ID: 469205464.

Uolevi Nikula, Christian Jurvanen, Orlena Gotel, and Donald C Gause. Empirical validation of the classic change curve on a software technology change project. *Information and Software Technology*, 52(6):680–696, 2010.

H. Otrok, N. Mohammed, L. Wang, M. Debbabi, and P. Bhattacharya. A game-theoretic intrusion detection model for mobile ad hoc networks. *COMPUTER COMMUNICATIONS -GUILDFORD THEN AMSTERDAM- BUTTERWORTH SCIENTIFIC LIMITED THEN ELSEVIER-*, 31(4):708–721, 2008. ID: 225809316.

A. Patcha and J.M. Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, 51(12):3448–3470, 2007. ID: 442571055.

S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu. A survey of game theory as applied to network security. *HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*, 43:880–889, 2010. ID: 613807484.

D. Shen, G. Chen, J. B. Cruz, E. Blasch, M. Kruger, and INTELLIGENT AUTOMATION INC ROCKVILLE MD. Game theoretic solutions to

cyber attack and network defense problems., 2007a. ID: 318687048.

D. Shen, G. Chen, J. B. Cruz, L. Haynes, M. Kruger, and E. Blasch. A markov game theoretic data fusion approach for cyber situational awareness [6571-16]. *PROCEEDINGS- SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, 6571:65710F, 2007b. ID: 210282321.

E. Sterner. Retaliatory deterrence in cyberspace. *Strategic Studies*, 62, 2011.

Aliya Stervstein. Pentagon seeks $3.2 billion for revised cyber budget. http://www.nextgov.com/nextgov/ng_20110324_2474.php?oref=rss, 2011.

Peter Szor. *The Art of Compuer Virus Research and Defense*. Symantec Press, Upper Saddle Lake, NJ, 2005.

Paul R. Thie. *Markov decision processes*. COMAP Inc., Lexington, Mass., 1983. ISBN 0912843047 9780912843049. ID: 10485516.

United States Department of Defense. Department of defense strategy for operating in cyberspace, 2011. ID: 741354228.

United States Executive Office of the President. The comprehensive national cybersecurity initiative, 2010. ID: 537418651.

A. Washburn and M. Kress. *Combat Modeling*. Springer Press, 2009.