A Game Theoretic Model of Strategic Conflict in Cyberspace

Harrison C. Schramm, David L. Alderson, W. Matthew Carlyle, Nedialko B. Dimitrov
Naval Postgraduate School, Monterey California, USA
hcschram@nps.edu
dlalders@nps.edu
mcarlyle@nps.edu
ndimitro@nps.edu

Abstract: We study cyber conflict as a two-person zero-sum 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.

Keywords: Cyber Conflict, Markov Game, Deterrence, Game Theory

1. Introduction

Conflict in Cyberspace, or *cyber conflict*, is important at both the 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” (National Security Council, 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 *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 use 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.

Using minimal assumptions, our model leads to two fundamental insights:

- ***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.

- ***Prospects for deterrence in cyber conflict may be limited.*** The ability of players to deter their opponents from attacking depends on an assured second strike. In cyber conflict, opposing players may have munitions based on the same exploit, and the first player to use the exploit effectively removes second strike munitions from the opponent's arsenal. Complicating factors to the cyber conflict game, such as an inability to identify the player who performed a cyber attack, or a player's ability to respond with kinetic munitions, also have an effect on deterrence in cyber conflict.

2. Related work

The JASONS 2010 report, The Science of Cyber-Security (JASON, 2010), recommends a variety of analytic approaches and 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 (Fudenberg and Tirole, 1991). Lye and Wing (2002) and Shen (2007) also consider cyber attacks in the context of a game. The most comprehensive survey of game theory and cyberspace is by Shiva et al. (2010). They develop a taxonomy of 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, this 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 Shiva et al. our proposed model is a non-cooperative, dynamic, stochastic game with perfect information.

The previous study which has the most commonality with our approach is that of Lye and Wing (2002). They consider a two-player, stochastic game between an attacker and administrator. Their model is 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.

Our work differs from previous work by abstracting cyber conflict away from individual machines and instruction sets in the same manner that Lanchester equations (Washburn and Kress, 2009) abstract physical conflict away from soldiers and weapons. 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

3.1 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 a munition based on that exploit is employed, then that munition becomes obsolete.

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 the time until the unaware player discovers the exploit. This uncertainty in discovery times is the fundamental tension that our model seeks to explore.

We model cyber warfare as a 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 our 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 probabilistically discovers the exploit, and when either player chooses to attack, then payoffs are determined and the game terminates.

3.2 Formulation

Our model focuses on a strategic cyber conflict between two players, where there is a single exploit 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 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 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 everyone's 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 $0 \le a_i(\tau) \le B_i$, where B_i is an arbitrary 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 is either waiting to discover the exploit $(\tau_i = 0)$ or he may know about the exploit $(\tau_i > 0)$ and be working to make his munitions more effective.
- 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. For each player *i*, the 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.

	$\tau_2 = 0$	$\tau_2 > 0$
$\tau_1 = 0$	$\theta_1 = \{W\};\ \theta_2 = \{W\}$ $\langle T,0,0 \rangle \begin{cases} \xrightarrow{(1-p_1(T))(1-p_2(T))} \langle T+1,0,0 \rangle \\ \xrightarrow{p_1(T)(1-p_2(T))} \langle T+1,1,0 \rangle \\ \xrightarrow{(1-p_1(T))p_2(T)} \langle T+1,0,1 \rangle \\ \xrightarrow{p_1(T)p_2(T)} \langle T+1,1,1 \rangle \end{cases}$	$\theta_1 = \{W\};\ \theta_2 = \{A,W\}$ $\langle T,0,\tau_2 \rangle \begin{cases} \xrightarrow{(1-p_1(T))} \langle T+1,0,\tau_2+1 \rangle \\ \xrightarrow{p_1(T)} \langle T+1,1,\tau_2+1 \rangle \end{cases}$
$\tau_1 > 0$	$\theta_1 = \{A,W\};\ \theta_2 = \{W\}$ $\langle T,\tau_1,0 \rangle \begin{cases} \xrightarrow{(1-p_2(T))} \langle T+1,\tau_1+1,0 \rangle \\ \xrightarrow{p_2(T)} \langle T+1,\tau_1+1,1 \rangle \end{cases}$	$\theta_1 = \{A,W\};\ \theta_2 = \{A,W\}$ $\langle T,\tau_1,\tau_2 \rangle \xrightarrow{1} \langle T+1,\tau_1+1,\tau_2+1 \rangle$

Table 1: Markov game action sets and state transitions as a function of $\langle T,\tau_1,\tau_2 \rangle$, the state of the game. The game always starts in $\langle T,0,0 \rangle$. As player i discovers the exploit, τ_i, becomes greater than zero and player i's action set includes attack.

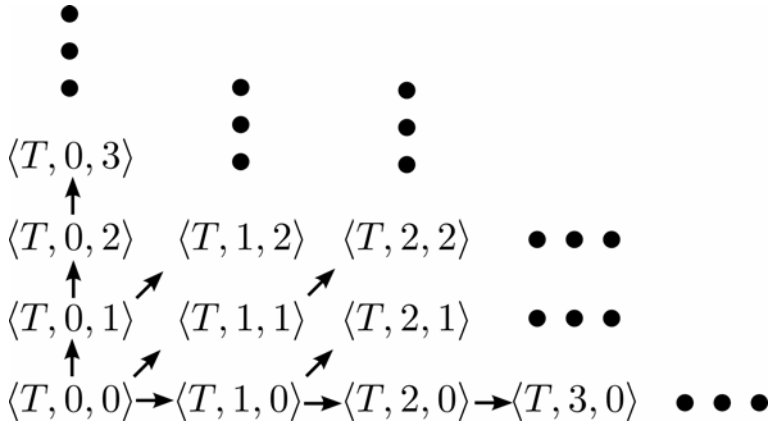

Figure 1. Diagram of states in the Markov Game. The arrows in the diagram show the possible transitions from one state to another, as described in Table 1. The horizontal axis describes increases in holding time for Player 1, τ_1,and the vertical axis describes increases in holding time for Player 2,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

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 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

In this case, we have $\tau_1>0,\tau_2>0$ and 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 sum of the munition values gives the result of the game.

Table 2: Payoff matrix for the Markov game when both players know the exploit. The payoff associated with "Wait, Wait" depends on future evolution 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*	$a_1(\tau_1)$	$a_1(\tau_1)-a_2(\tau_2)$

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 $a_1(\tau_1)\geq a_1(\tau_1)-a_2(\tau_2)$, therefore "Attack" is a dominating strategy for Player 2. Given Player 2 chooses "Attack", Player 1 must also play "Attack" and "Attack, Attack" is an 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 state $\langle T, \tau_1, \tau_2 \rangle$ 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 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

$$Y = (1 - p_2(T))V\langle T+1, \tau_1+1, 0 \rangle + p_2(T)V\langle T+1, \tau_1+1, 1 \rangle,$$

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

	Player 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. By definition, Player 2 has a singleton action set and the matrix reduces to a single column.

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:

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

$$
v_\tau(h) = a_1(\tau+h)\cdot\prod_{k=0}^{h-1} q_2(T+k) + \sum_{k=0}^{h-1}\left[\left(a_1(\tau+k+1) - a_2(1)\right)\cdot p_2(T+k)\prod_{j=0}^{k-1} q_2(T+j)\right] \qquad (1)
$$

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:

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

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

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

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 $k^*=\min_k\{a_1(k+1)-a_1(k)\leq p_2a_2(1)\}$. The set in the definition of k^* is never empty when $a_1(\cdot)$ is bounded, concave, and nondecreasing and $p_2a_2(1)$ 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:

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

where we have introduced the γ values 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.

4.1 Scenario 1: Constant a_i functions

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

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

Because $a_i(\tau)$ is concave and increasing 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:

$$p_1a_1(1) > p_2a_2(1)$$

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 a_1

Suppose Players 1 and 2 have attack functions such that:

$$a_i(0) = 0$$
$$a_1(\tau) = \tau \quad 1 \leq \tau \leq 5$$
$$a_1(\tau) = 5 \quad \forall \tau \geq 5$$
$$a_2(\tau_2) = c \quad \forall \tau_2 \geq 1$$

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:

$$k_1^* = \begin{cases} 1 & \text{if } p_2c \geq 1 \\ 5 & \text{otherwise} \end{cases}$$

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

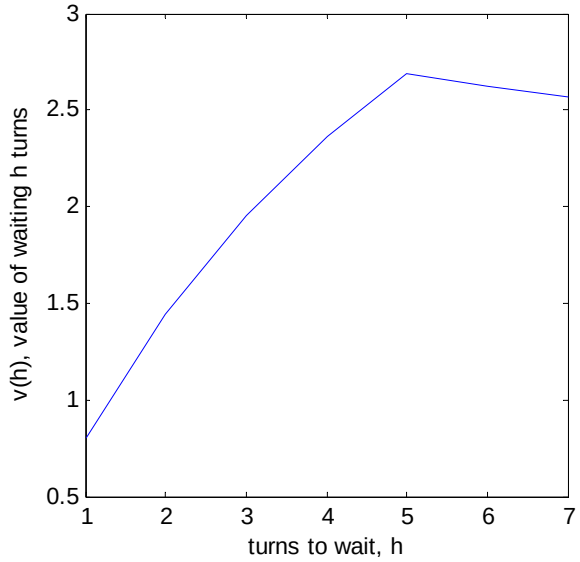

Figure 2: Value of Scenario 2 from Player 1's point of view. The vertical axis plots the value, $v_\tau(h)$, as a function of the number of time periods Player 1 waits before attacking, h. The value function increases to the point $h = 5$, and decreases afterward. By Theorem 2, this implies that Player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

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.

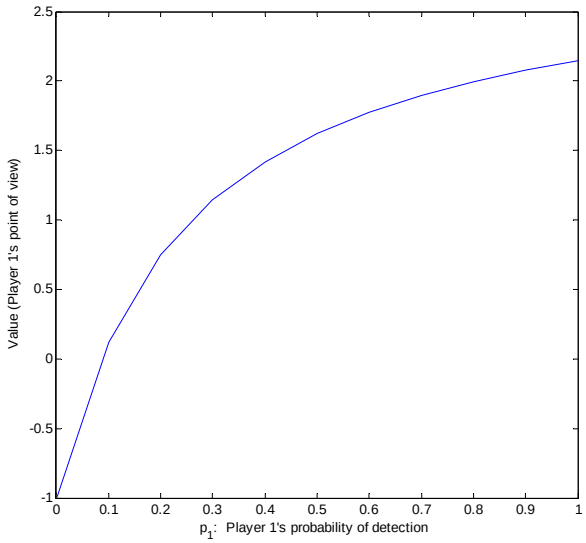

Figure 3: Value of Scenario 2 as a function of Player 1's probability of discovering the Exploit, p_1. Here we see that the value of the game is a concave function of Player 1's probability of detecting the exploit; increases in detection probability at low detection values provide a bigger increase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

4.3 Scenario 3: Non-Monotone a_1

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. In this case, we cannot use Theorem 2 to compute the optimal attack time. However, we may compute the optimal attack time directly, by computing the value of holding for each possible holding period, as depicted in Figure 5.

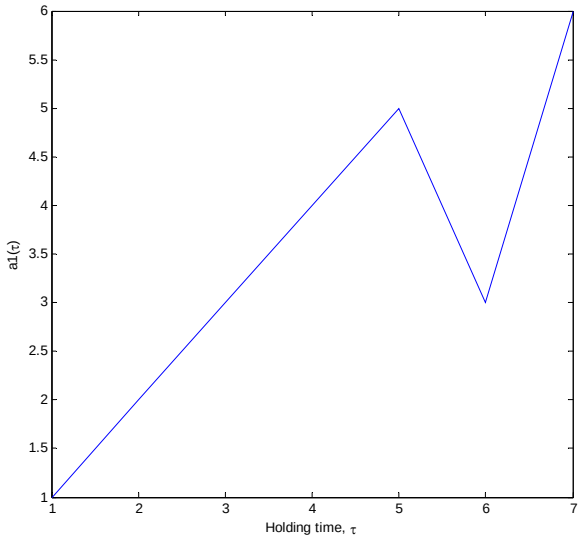

Figure 4: The function a_1 for Scenario 3. Unlike our previous examples, the value of Player 1's attack has a dip at $\tau_1 = 6$. In this scenario, Theorem 2 no longer applies in finding the optimal attack time k_1^*.

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.

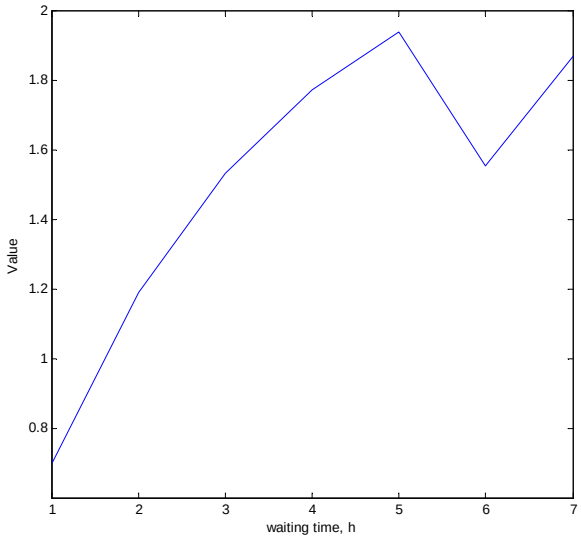

Figure 5: Player 1's value as a function of waiting time, h in Scenario 3. We see that the payoff for waiting to $h = 7$ is less than executing at $h = 5$.

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

5 Extensions and applications

In this section, we explore the operationally relevant implications of our model.

5.1 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.

5.1.1 One player delayed action

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 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 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:

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

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 $0<\tau\le 10$ 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 below:

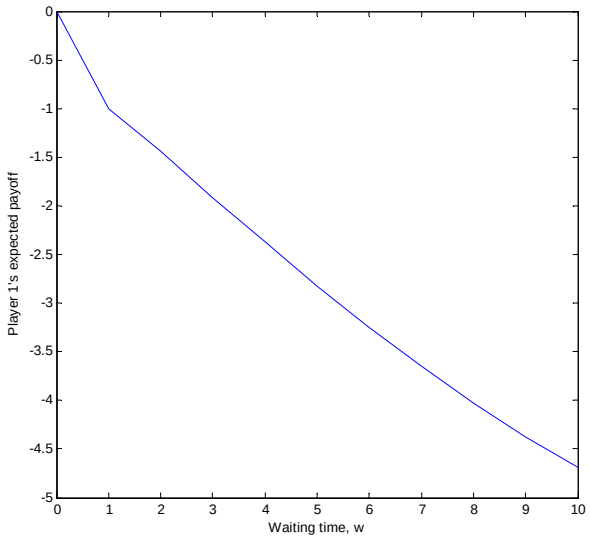

Figure 6: Player 1's utility curve as a function of waiting time w against an evenly matched opponent. We see that Player 1's utility drops off rapidly from an expected value of zero, with the implication that waiting is costly.

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 2 shows the adjustment required in this example; for waiting times longer than 5 periods, even perfect detection does not achieve parity.

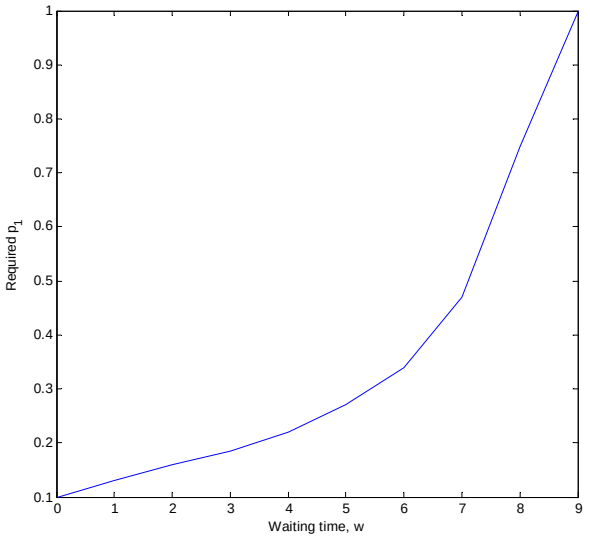

Figure 7: Player 1's required detection probability p_1 required to achieve $V^w\langle 0,0,0\rangle = 0$ as a function of waiting time, w. Player 1's required capability increases rapidly and, because p_1 may never be greater than 1, parity is unachievable after $w = 5$

The lesson of Figures 7 and 8 is that waiting times are costly and adversely affect one's prospects in cyber conflict.

5.2 Deterrence

In the preceding subsection, we advise belligerents in cyber conflict to develop and execute their attacks quickly–a stance that is incompatible with the notion of "crisis stability" (Kent, 1989) of classical deterrence theory. Can deterrence in cyber conflict be achieved, and if so, how? Several scholars ask this question, notably (Sterner, 2011). In this paper we consider one aspect of cyber deterrence.

5.2.1 A short review of strike stability

The concept of strike stability was developed during the Cold War to understand which sets of circumstances would lead to nuclear conflict. The original papers describe the development and application of this theory to nuclear arms. Kent (1989) describes a game which has many similarities with the one described herein; two players are faced with the decision of 'attacking' or 'not attacking'; they make this decision by weighing the benefits of going 'first' or 'second' with the assumption that the other player will surely retaliate with whatever force he has left. The closer the ratio of costs of going second to going first is to one, the more stable the system is because the decision maker is indifferent to striking first or striking second and may be deterred. Low values of strike stability indicate a large disadvantage to attacking second and therefore lead to instability. Deterrence requires both sides to choose non-action (Wait in our model) at each decision epoch.

5.2.2 Strike stability for cyber conflict

The analysis of section 4 shows that if a player has the ability to attack, he eventually does with certainty. This means that cyber conflict with perfect information and a single exploit is deterrence unstable. Intuitively this is because there is no second strike. Theorem 1 is sufficient to demonstrate that the single-attack case is deterrence unstable; the first player to attack receives the reward of his development to date, and the non-attacking player is left with and empty arsenal.

Considerations outside the model we consider may provide some degree of deterrence in reality. For example, military, economic or diplomatic consequences or large cyber munition arsenals may provide some guarantee of a second strike. Such guarantees, while important to deterrence are outside the bounds of our current work. Nevertheless, without these external guarantees deterrence in cyber conflict does not exist.

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

- Demonstrate a framework for analyzing the problem;
- Demonstrate that in cyber conflict, idle wait times are damaging, and provide a means to calculate their disutility; and

- Show implications for deterrence in cyber conflict.

This paper considered a single attack in discrete time with perfect information—three idealizations that help us begin to tackle the problem of cyber conflict. Of these three, the perfect information assumption appears to be the richest area to explore in the future, and with this exploration come considerations of credibility, reputations, and risk taking.

References

Falliere, N., Murchu, L., and Chien, E. (2011) *W.32 Stuxnet Dossier,* Mountain View: Symantec Corporation.

Fudenberg, D. and Tirole, J. (1991) *Game Theory*, Cambridge: MIT Press.

JASON(2010) *The Science of Cyber Security.* McLean: MITRE Corporation, JSR-10-102.

Kent, G, and Thaler, D., (1989) *First Strike Stability: A Methodology for Evaluating Strategic Forces*, Santa Monica, RAND.

Lye, K., and Wing, J. (2002) 'Game Strategies in Network Security', *International Journal of Information Security,* vol. 4, pp. 71-86.

National Security Council (2010) *The Comprehensive National Cybersecurity Initiative.* [Online] http://www.whitehouse.gov/cybersecurity/comprehensive-national-cybersecurity-initiative

Shen, D., Chen, G., Blasch, E. and Tadda, G. (2007) *A Markov game theoretic approach for cyber situational awareness*. SPIE's Defense and Security Symposium, Orlando, FL 9-13 April.

Shiva, S., Dasgupta, D., and Wu, O. *Game theoretic Approaches to Protect Cyberspace,* Technical Report No. CS-10-001 Memphis: University of Memphis.

Sterner, E. (2011) 'Deterrence in Cyberspace'. *Strategic Studies Quarterly* Spring, pp. 68-80.

Sternstein, A. (2011) *The White House's 2012 budget devotes a greater percentage of IT funds to cybersecurity,* [Online], Available: http://www.nextgov.com/nextgov/ng_20110216_3295

Szor, P. (2005*) The Art of Computer Virus Research and Defense*, Upper Saddle Lake, NJ: Symantec Press.

Thie, P. (1983) *Markov Decision Processes,* Lexington: UMAP expository monograph series.

Washburn, A, and Kress, M (2009) *Combat Modeling*, New York, NY: Springer Press.