A Stochastic Game Model with Imperfect Information in Cyber Security

Sajjan Shiva, Sankardas Roy, Harkeerat Bedi, Dipankar Dasgupta, Qishi Wu

Department of Computer Science, University of Memphis, Memphis, TN, USA

sshiva, sroy5, hsbedi, ddasgupt, qishiwu@memphis.edu

Abstract

While there are significant advances in information technology and infrastructure which offer new opportunities, cyberspace is still far from completely secured. Recently, researchers have started exploring the applicability of game theory to address the cyber security problem. The interaction between the attacks and the defense mechanisms can be considered as a game played between the attacker and the defender (system administrator). One of the techniques that has been proposed in the literature used stochastic game models to emulate network security games and showed how to determine the best strategy for the defender considering the possible attack strategy used by the attacker. However, the prior research assumes that the players have perfect information about the current state of the game, which generally does not hold in reality. Our model relaxes this assumption and enriches the prior game models by enabling them to capture more realistic scenarios. In particular, this paper presents a theoretical analysis by which the defender can compute his/her best strategy to reach the Nash equilibrium of a stochastic game assuming imperfect sensory information. In addition, this paper shows that if the defender follows the strategy prescribed by the perfect information model, the Nash equilibrium is not achieved and the attacker's payoff can be higher. Our theoretical analysis is tested in simulation experiments and the results validate our approach.

Keywords

Network Security, Game Theory, Stochastic Games, Nash Equilibrium, Imperfect Information, Simulation

1. Introduction

With the explosive growth of the Internet and its extensive use in all sectors, security has become a challenge as hackers are finding new ways to launch multistage attacks to cause damage to information assets. Despite considerable effort from the research community this problem is far from completely solved. Recently, researchers have started exploring the applicability of game theory to address this problem (Roy 2010). Since game theory deals with scenarios in which multiple players with contradictory objectives compete with each other, it can provide a mathematical framework for analysis and modeling information system security challenges. The interaction between the attacks and the defense mechanisms can be considered as a game played between the attacker and the system administrator.

To model attacks and defense mechanisms, a stochastic game model has been proposed in the literature (Lye 2002; Lye 2005; Alpcan 2006). The state of the game probabilistically changes depending on actions taken by the players (i.e., type of attacks and defender's response) and the system configurations. During each state transition, each player gets a payoff or incurs some cost (negative payoff). Techniques exist by which a player can determine the best strategy to get the highest overall payoff considering all of the possible strategies of the adversary. Game theoreticians formulate the solution concept of a stochastic game by the notion of Nash equilibrium, and have already provided the analysis indicating the existence of the equilibrium (Filar 1997).

As stated, the prior stochastic game models for network security (Lye 2002; Lye 2005) assume that the players have perfect information about the current state of the game, which implies that the defender is always able to detect an attack and the attacker is always aware of the employed defense mechanism. In real systems, a player uses a sensor (e.g., the defender's sensor can be a part of the Intrusion Detection System (IDS)) to observe the current status of the system to decide the strategy. It is widely believed that no real sensor can perfectly read the environment, i.e., usually there is a non-zero error probability. So, in most cases, the above assumption about perfect information does not hold in real life.

Our paper relaxes this assumption and designs a stochastic game model which is able to capture more realistic scenarios. It considers that a player knows the system's true state at a particular moment with some error probability, i.e., at any given point of time the true state and a player's perception can be potentially different. With this additional constraint of imperfect information, this paper computes the best strategy for a player considering other players' choice of possible strategies.

In particular, this paper presents a theoretical analysis by which the defender can compute his/her best strategy to reach the Nash equilibrium of a stochastic game assuming the defender's sensor is imperfect. It is implicit that the defender knows the error probability of his/her sensor and the players' objectives are directly opposite, i.e., it is a zero-sum game. Moreover, our paper shows that if the defender follows the strategy prescribed by the perfect information model, then the Nash equilibrium is not achieved and the attacker's payoff can be more. Our algorithm for computing the best strategy runs offline well before the game is being played, i.e., our game analysis is static. Furthermore, our theoretical results are validated via simulation experiments in MATLAB.

The major contributions of this paper are summarized below:

a. We present a static analysis of an imperfect information zero-sum stochastic game and compute the best strategy of the system administrator in realistic scenarios.

b. Our analysis and simulation experiments illustrate that the system administrator will be betteroff if he/she takes our strategy compared to the scenario when he/she executes the strategy prescribed by the perfect information models.

The rest of the paper is organized as follows: Section 2 briefly presents the perfect information stochastic game model. Section 3 introduces our imperfect information stochastic game model and also provides analysis and simulation results. Section 4 discusses the related work, and Section 5 concludes the paper.

2. Background: A Stochastic Game Model

This section provides a brief overview of a stochastic game model as discussed elsewhere (Lye 2002; Lye 2005). For further details of the stochastic game model refer to (Filar 1997).

Lye et al. model the interaction between the attacks and the defense actions as a two players' ($k = 1,2$) game where player 1 is the attacker and player 2 is the system administrator (Lye 2002; Lye 2005). This infinite-horizon stochastic game model considers N states.

The stochastic game is represented by a tuple $(S, A^1, A^2, Q, R^1, R^2, \beta)$ whose elements are defined below.

- $S = \{\xi_1, \xi_2, \dots, \xi_N\}$ is the set of states. A state represents the current status of the whole system under consideration.

- $A^k = \{A^k{}_{\xi_1}, A^k{}_{\xi_2}, \dots, A^k{}_{\xi_N}\}, k = 1, 2$ where $A^k{}_{\xi_j} = \{\alpha^k_{j1}, \alpha^k_{j2}, \dots, \alpha^k_{jM^k}\}$ is the action set of player k at state ξ_j. It is assumed that $M^k = |A^k{}_{\xi_j}|$ for all $1 \le j \le N$.

- The state transition probabilities are represented by the function $Q: S \times A^1 \times A^2 \times S \to [0\ 1]$ which maps a pair of states and a pair of actions to a real number between 0 and 1. As an example, $Q(\xi_1, \alpha^1{}_1, \alpha^2{}_1, \xi_2) = 0.3$ is interpreted as the probability of state transition from state ξ_1 to ξ_2 given that player 1 takes action $\alpha^1{}_1$ and player 2 takes action $\alpha^2{}_1$.

- The reward of player k is determined by the function $R^k: S \times A^1 \times A^2 \to \mathbb{R}$ which maps a state and a pair of actions to a real number. As an example, $R^1(\xi_1, \alpha^1{}_1, \alpha^2{}_1) = 42$ is interpreted as the reward gained by the attacker at state ξ_1 given that attacker takes action $\alpha^1{}_1$ and player 2 takes action $\alpha^2{}_1$. Negative reward represents the cost incurred by a player.

- β, $0 < \beta < 1$ is the discount factor for discounting future rewards to calculate the overall payoff of a player in this infinite horizon game.

We now define the stationary strategy of a player. Stationary strategy is one that remains constant over time. Let $\Omega^{n} = \{ p \in R^{n} \mid \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1\}$ be the set of probability vectors of length n. Let the function $\pi^k: S \rightarrow \Omega^{M_k}$ denote the strategy for player k where $\pi^k(s) = [\pi^k(s, \alpha_1), \pi^k(s, \alpha 2, \ldots, \pi ks, \alpha Mk]$, while $\pi ks, \alpha i$ is the probability with which player k selects the action αi in state s. If π^k is such that $\forall s, i, \pi^k(s, \alpha_i)$ is 0 or 1, then π^k is called a pure strategy. Otherwise, π^k is called a mixed strategy.

During each state transition, player k gets a reward (defined by the function R^k) or incurs some cost (negative reward). To compute the overall payoff of player k, we consider the future moves which will change the present state to next states giving future payoff to player k. The overall payoff is computed by discounting the future payoff using the discount factor β. Let ${v^k}_{\pi^1, \pi^2}(s)$ denote the expected overall payoff of player k when the game starts at state s while the strategy of player 1 is π^1 and the strategy of player 2 is π^2. Let the vector ${v^k}_{\pi^1, \pi^2}$ denote the aggregate payoff of player k, where ${v^k}_{\pi^1, \pi^2} = [{v^k}_{\pi^1, \pi^2}(\xi_1), {v^k}_{\pi^1, \pi^2}(\xi_2), \ldots, {v^k}_{\pi^1, \pi^2}(\xi_N)]$.

Each player has the goal to maximize his expected payoff. The Nash equilibrium of this game is defined to be a pair of strategies (π_*^1, π_*^2) which simultaneously satisfy the following equations component-wise:

$$v^1{}_{\pi_*^1, \pi_*^2} \geq v^1{}_{\pi^1, \pi_*^2} \; \forall \pi^1 \in \Omega^{M_1}$$

$$v^2{}_{\pi_*^1, \pi_*^2} \geq v^2{}_{\pi_*^1, \pi^2} \; \forall \pi^2 \in \Omega^{M_2}$$

A stochastic game is called zero-sum if one player's reward at each state transition is equal and opposite of the other player's reward, i.e., for all i, j, m we have $R^1(\xi_i, \alpha^1_j, \alpha^2_m) = - R^2(\xi_i, \alpha^1_j, \alpha^2_m)$. It implies that for every pair of strategies the overall payoff of the players are same and opposite, i.e.,

$$\forall \pi^1 \in \Omega^{M_1}, \pi^2 \in \Omega^{M_2} \quad v^1{}_{\pi^1, \pi^2} = -v^2{}_{\pi^1, \pi^2}.$$

For a zero-sum stochastic game which has a Nash equilibrium,(π_*^1, π_*^2), the *value* of the game is considered as $v^1{}_{\pi_*^1, \pi_*^2}(s_1)$ where s_1 is the start state. Let V denote the value of the game.

We can compute the Nash equilibrium strategy of the players for a zero-sum stochastic game through a static analysis (offline analysis) of the game using the algorithm discussed in (Filar 1997). The algorithm used is basically an iterative non-linear optimization technique.

3. Stochastic Game with Imperfect Information

The above game model assumes that the players have perfect information about the current state of the game. Our model presented in this section relaxes this assumption. Section 3.1 presents our imperfect information stochastic game model. Section 3.2 presents a static analysis and Section 3.3 provides the simulation results.

3.1 The Model

Our model is an extension of the prior model (Section 2) and considers that a player k ($k = 1,2$) observes the game's true state at a particular moment by an imperfect sensor device. That means, player k can view ξ_j as any state in the information set $I^k_{\xi_j}$ with some probability where $I^k_{\xi_j} = \{\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_p}\}$ with ξ_j being an element of $I^k_{\xi_j}$. Compared to the perfect information model, player k's action space may become wider, i.e., player k may take an action which is allowed at a state $\xi_{j_i} \neq \xi_j$ belonging to the information set, $I^k_{\xi_j}$.

Let $B^k_{\xi_j}$ denote the set of possible actions of player k when his/her information set is $I^k_{\xi_j}$. Then $B^k_{\xi_j} = \cup_{\xi_i \in I^k_{\xi_j}} A^k_{\xi_i}$ where $A^k_{\xi_i}$ denotes the action set of player k when he/she is sure that the true current state is ξ_i. Below we formally define the outcome of player k's extended action set $B^k_{\xi_j}$, compared to $A^k_{\xi_j}$ in

the previous model, when the true state is ξ_j. If player k takes an action $\alpha^k \in B^k_{\xi_j}$ when the true state is ξ_j but α^k is not in $A^k_{\xi_j}$, then in terms of the influence on state transition probability, α^k is equivalent to player k taking no action at state ξ_j. However, regarding the influence on player k's payoff α^k may not be equivalent to player k taking no action at state ξ_j depending upon the cost of the execution of α^k.

Formally, our model is represented by a tuple, $(S, I^1, I^2, E^1, E^2, A^1, A^2, B^1, B^2, Q, R^1, R^2, \beta)$ whose elements are defined below.

- $S = \{\xi_1, \xi_2, \dots, \xi_N\}$ is the set of states.
- $I^k = \{I^k{}_{\xi_1}, I^k{}_{\xi_2}, \dots, I^k{}_{\xi_N}\}$, $k = 1, 2$ where $I^k{}_{\xi_j}$ represents the information set of player k when the true state is ξ_j, i.e., $I^k{}_{\xi_j} = \{\xi_{j_1}, \xi_{j_2}, \dots, \xi_{j_p}\}$ (where p is an arbitrary positive integer) with the condition that $\xi_j \in I^k{}_{\xi_j}$.
- $E^k = \{E^k{}_{\xi_1}, E^k{}_{\xi_2}, \dots, E^k{}_{\xi_N}\}$, $k = 1, 2$ where the j-th set $E^k{}_{\xi_j}$ represents the error probabilities of k-th player's sensor at the true state ξ_j over the corresponding information set, $I^k{}_{\xi_j}$.
- $A^k = \{A^k{}_{\xi_1}, A^k{}_{\xi_2}, \dots, A^k{}_{\xi_N}\}, k = 1, 2$ where $A^k{}_{\xi_j} = \{\alpha^k_{j1}, \alpha^k_{j2}, \dots, \alpha^k_{j_M{}^k}\}$ is the action set of player k at state ξ_j.
- $B^k = \{B^k{}_{\xi_1}, B^k{}_{\xi_2}, \dots, B^k{}_{\xi_N}\}, k = 1, 2$ where $B^k{}_{\xi_j}$ represents the extended action set of player k at $I^k{}_{\xi_j}$. That means, $B^k{}_{\xi_j} = \bigcup_{\xi_i \in I^k_{\xi_j}} A^k_{\xi_i}$. By introducing identical actions we can make $|B^k{}_{\xi_j}|$ same for all $1 \le j \le N$. Let $T^k = |B^k{}_{\xi_j}|$.
- The state transition probabilities are represented by the function $Q: S \times B^1 \times B^2 \times S \to [0\ 1]$ which maps a pair of states and a pair of actions to a real number between 0 and 1. Our model assumes that for any state ξ_j if player k takes an action $\alpha^k_i \in B^k{}_{\xi_j}$ while α^k_i does not belong to $A^k{}_{\xi_j}$, then $Q(\xi_{j_1}, \alpha^k_{i1}, \alpha^l_{i2}, \xi_{j_2}) = Q(\xi_{j_1}, nop, \alpha^l_{i2}, \xi_{j_2})$ where l represents the other player.
- The reward of player k is determined by the function $R^k: S \times B^1 \times B^2 \to \mathbb{R}$ which maps a state and a pair of actions to a real number.
- $\beta, 0 < \beta < 1$ is a discount factor for discounting future rewards in this infinite horizon game.

We redefine the strategy function π^k of the perfect information model for this imperfect information model as $\pi^k: S \to \Omega^{T_k}$ where $\pi^k(s) = [\pi^k(s, \alpha_1), \pi^k(s, \alpha_2), \dots, \pi^k(s, \alpha_{T_k})]$. The definition of the payoff vector of player k ($v^k{}_{\pi^1, \pi^2}$) and the Nash equilibrium, (π^1_*, π^2_*) are similarly extended.

One major difference of this model from the perfect information game is as follows: As player k's sensor is not perfect, when his/her strategy π^k is executed in the true sense, his/her observed strategy (referred to as apparent strategy in the rest of this paper), $\pi^{k'}$ is different from π^k. We will illustrate this further in Section 3.2.

3.2 A Static Analysis for a Game with Two States

We now present a static analysis of our game model, by which a player can compute his/her best strategy offline. Only a zero-sum game is considered. This analysis considers the worst-case scenario from the defender's point of view. It is assumed that only the defender's sensor is erroneous while the attacker can perfectly observe the current state of the game. It is to be noted that our analysis can be easily extended to the case where the attacker's sensor is also imperfect. Furthermore, this analysis is restricted to a game of two states for the sake of simplicity. In the future work, this analysis will be extended for games with more than two states. We focus on the following game as illustrated in Figure 1.

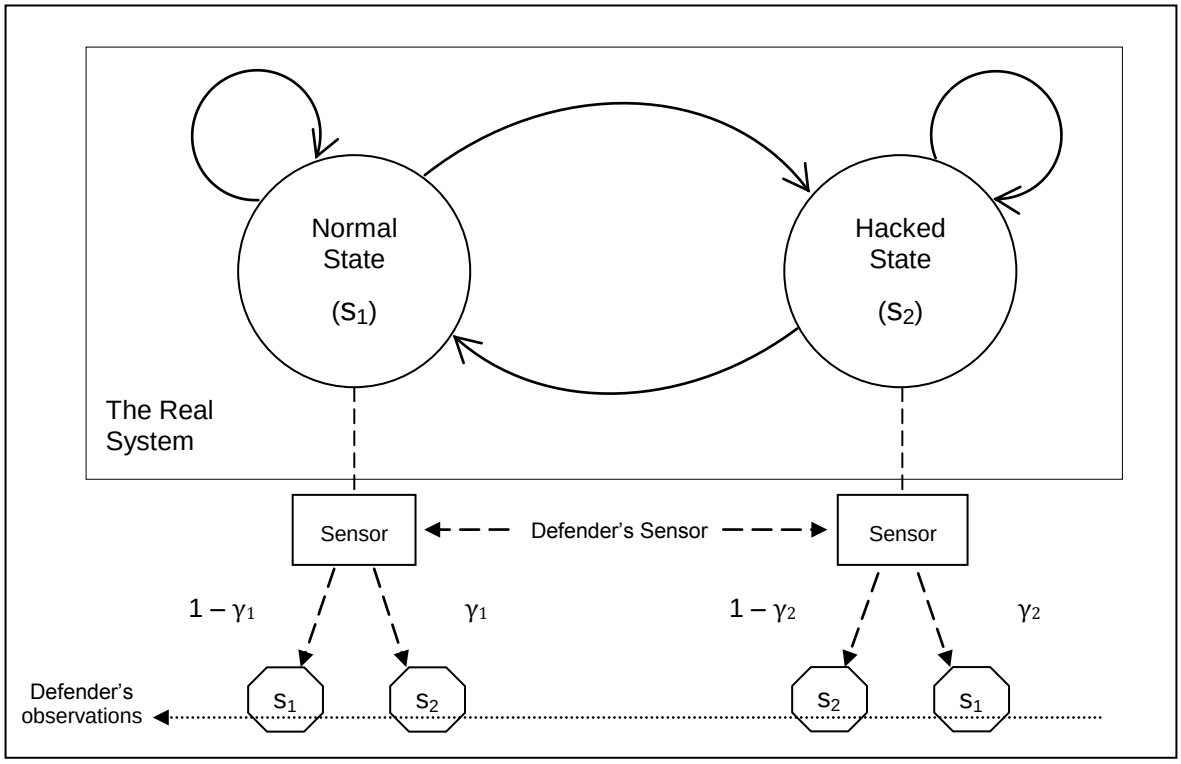

Figure 1: The state transition diagram and defender's observations — the same sensor is shown twice to indicate observations at different states

There are two states in this game. The system is either in *NormalState* (s_1) or in *HackedState* (s_2).The defender's sensor is imperfect and the error probability at state s_1 and s_2 are γ_1 and γ_2, respectively. That means, when the true state is s_1, with probability γ_1 the defender observes that as s_2, and when the true state is s_2, the defender observes the state as s_1 with probability γ_2. However, it is assumed that the sensor's error probabilities (γ_1 and γ_2) are known to the defender. On the other hand, the attacker's sensor observes the current state with no error.

The action spaces of the players, A^1 and A^2 are as follows where a denotes 'attack', na denotes 'no attack', d denotes 'defense' and nd denotes 'no defense'. The first row in A^1 or A^2 represents the actions available in state s_1 and the second row is for s_2.

$$A^1 = \begin{bmatrix} a & na \\ a & na \end{bmatrix} \ and \ A^2 = \begin{bmatrix} d & nd \\ d & nd \end{bmatrix}.$$

In this game, each player's extended action space (Section 3.1) remains same as the original action set. The strategy of the player k is represented by the probability distribution with which player k selects the available actions. The strategies of the players are represented by the following matrices π^1 and π^2:

$$\pi^1 = \begin{bmatrix} {\pi^1}_{11} & {\pi^1}_{12} \\ {\pi^1}_{21} & {\pi^1}_{22} \end{bmatrix} \ and \ \pi^2 = \begin{bmatrix} {\pi^2}_{11} & {\pi^2}_{12} \\ {\pi^2}_{21} & {\pi^2}_{22} \end{bmatrix}.$$

As an example, ${\pi^1}_{11}$ represents the probability with which player 1 selects action a and ${\pi^1}_{12}$ represents the probability with which player 1 selects action na at s_1 and ${\pi^1}_{11} + {\pi^1}_{12} = 1$.

State Occurrence Ratio (r_1,r_2): A stochastic game involves state transitions. The proportion of times a state s_i will occur during the whole play is called its occurrence ratio and is denoted by r_i. The value of r_i depends on the state transition probability function Q and the true strategies π^1 and π^2.

Given true strategies π^1 and π^2, we can compute the effective state transition probability matrix P whose dimension is $|S| \times |S|$. The element $P(i,j)$ represents the probability with which state s_i will switch to state s_j . Here, P is a 2×2 matrix.

We can compute $r_1$and $r_2$as follows. From basic theory of stochastic game (Filar 1997) we know that $P^i(1,j)$ represents the probability that state s_j will occur at the ith transition.

$$r_1 = \lim_{n\to\infty} \frac{P(1,1) + P^2(1,1) + \ldots + P^n(1,1)}{n}$$

$$r_2 = \lim_{n\to\infty} \frac{P(1,2) + P^2(1,2) + \ldots + P^n(1,2)}{n}$$

As expected from the above two expressions we get $r_1 + r_2 = 1$. As the defender's sensor is not perfect, he/she can observe different occurrence ratios. As the attacker's sensor is perfect, now onwards the term 'apparent' only relates to the defender.

Apparent State Occurrence Ratio ($r_1{}', r_2{}'$): The apparent occurrence ratios of state s_1 and s_2 are as follows.

$$r_1^{'} = (1-\gamma_1)r_1 + \gamma_2 r_2$$

$$r_2^{'} = \gamma_1 r_1 + (1-\gamma_2)r_2$$

We stress the fact that the defender's true strategy, π^2 is different from his/her apparent strategy, $\pi^{2'}$, which he/she observes being executed. We represent $\pi^{2'}$ as follows.

$$\pi^{2'} = \begin{bmatrix} \pi^{2'}{}_{11} & \pi^{2'}{}_{12} \\ \pi^{2'}{}_{21} & \pi^{2'}{}_{22} \end{bmatrix}.$$

As an example, $\pi^{2'}{}_{11}$represents the apparent probability of action d and $\pi^{2'}{}_{12}$ represents the apparent probability of action nd at s_1. Note that $\pi^{2'}{}_{11} + \pi^{2'}{}_{12} = 1$.

The defender's apparent strategy, $\pi^{2'}$ is determined by his/her true strategy, π^2, sensor error probabilities (γ_1,γ_2) and the true state transition ratios, (r_1, r_2) as described in the following matrix equation. The matrix IIF is called the imperfect information factor and represents the influence of the sensor's errors.

$$\begin{bmatrix} \pi_{11}^{2'} & \pi_{12}^{2'} \\ \pi_{21}^{2'} & \pi_{22}^{2'} \end{bmatrix} = IIF \,.\, \begin{bmatrix} \pi_{11}^{2} & \pi_{12}^{2} \\ \pi_{21}^{2} & \pi_{22}^{2} \end{bmatrix} \quad \ldots (1)$$

$$where\ IIF = \begin{bmatrix} \dfrac{(1-\gamma_1)\,r_1}{(1-\gamma_1)\,r_1 + \gamma_2\,r_2} & \dfrac{\gamma_2\,r_2}{(1-\gamma_1)\,r_1 + \gamma_2\,r_2} \\ \dfrac{\gamma_1\,r_1}{\gamma_1\,r_1 + (1-\gamma_2)\,r_2} & \dfrac{(1-\gamma_2)\,r_2}{\gamma_1\,r_1 + (1-\gamma_2)\,r_2} \end{bmatrix}$$

We recall from Section 2 that Nash equilibrium strategies (π_*^1, π_*^2) of the players can be computed using the algorithm discussed in (Filar 1997). To reach this equilibrium the defender has to execute his/her apparent strategy $\pi_*^{2'}$ after computing it using equation (1). In equation (1), he/she has to replace $\pi^2$by π_*^2.

We now discuss the benefit of our approach compared to the perfect information model. If the defender follows the perfect information model he/she executes π_*^2 as the apparent strategy. In that case, the defender ends up playing the true strategy π^2 given by the following matrix equation.

$$\pi^2 = IIF^{-1}.\pi_*^2$$

As a result, the true strategy π^2 deviates from the Nash equilibrium strategy when IIF is not an identity matrix. So, the equilibrium is not reached and the attacker can gain higher payoff as shown by

our simulation results. Moreover, there exists such a stochastic game for which no feasible π^2 exits corresponding to the Nash equilibrium strategy, π_*^2. Some of our simulation experiments illustrate such a game.

3.3 Simulation

We validate the above analysis using simulation experiments as discussed below.

3.3.1 Simulation Framework

We simulate a stochastic game being played between an attacker and a system administrator using MATLAB. We implement an application that is able to produce the pair of optimal strategies for a zero-sum game with imperfect information. This application is based on the modified Newton's method as described under article 3.3 in (Filar 1997). An iterative non-linear optimization algorithm is used. The input to this algorithm includes the state transition matrix and the reward matrix. As this is a zero-sum game, only the first player's reward matrix is given as input.

To compute the output, the modified Newton's method requires solving a *matrix* game in each iteration. This functionality is achieved by using an additional component that generates the optimal strategies and the value for a zero-sum matrix game as in (Williams 1966).

3.3.2 Simulation Results

We demonstrate the feasibility and effectiveness of our model by using games as discussed in Section 3.2. Figure 1 displays the two system states and the transitions possible among them. The actions possible by the attacker during either state are $a\ (attack)$ or $na\ (no\ attack)$. The $attack$ action indicates the execution of an attack with the motivation to bring the network to *HackedState* or to continue further attacking in *HackedState*. The actions possible by the defender during either state are $d\ (defense)$ or $nd\ (no\ defense)$. The $defense$ action indicates the execution of a restore process with the motivation to bring back the network to *NormalState* from the *HackedState* or to strengthen the *NormalState* by increasing the monitoring level. The na or nd action indicates an instance of no action. We set the discount factor β to 0.75 and defender's sensor's two error probabilities, γ_1 and γ_2 as 0.1 and 0.05, respectively.

Our first experiment shows that perfect information models (Lye 2002; Lye 2005) *can give higher payoff to the attacker compared to our model.* The state transition probabilities and the reward matrices are shown in Figure 2. This style of representation is based on that in (Filar 1997). The rows for each state represent the actions possible by attacker and columns represent the actions possible by defender. Each element is divided by a diagonal into two halves where the upper half represents the reward to the attacker from that state and the lower half represents the state transition probabilities when the corresponding actions are performed by both players. For example, in *NormalState*, when the attacker and defender both perform their first actions, the reward to the attacker is 10 and the probability of the network remaining in *NormalState* is 0.7 and changing to *HackedState* is 0.3 (First row in Figure 2).

	Defender's Action 1 (d)	Defender's Action 2 (nd)
Attacker's Action 1 (a)	10 (0.7, 0.3)	40 (0.7, 0.3)
Attacker's Action 2 (na)	200 (1, 0)	0 (1, 0)

The Rewards (to the attacker) and State Transition Probabilities at *NormalState*

	Defender's Action 1 (d)	Defender's Action 2 (nd)
Attacker's Action 1 (a)	200 (0.4, 0.6)	55 (0, 1)
Attacker's Action 2 (na)	45 (0.8, 0.2)	550 (0, 1)

The Rewards (to the attacker) and State Transition Probabilities at *HackedState*

Figure 2: Specifications of the game in the first experiment

We calculate the pair of true optimal strategies, which are optStrat_1 = [0.8696 0.1304; 0.8735 0.1265] (for the attacker) and optStrat_2 = [0.3557 0.6443; 0.7014 0.2986] (for the defender). The value of the game V (the attacker's payoff when the game starts from *NormalState*) is found to be 284.5637. However, since the defender's sensor is faulty, he/she cannot directly execute this true strategy. The apparent strategy for the defender is computed as appStrat_2 = [0.3708 0.6292; 0.6620 0.3380] using equation (1). Apparent strategy for only the defender is considered in our example as it is assumed that only the defender is uncertain about the present state of the system and not the attacker.

Our model suggests the defender to execute the apparent strategy (appStrat_2) and the value of the game thus obtained is V which is 284.5637. It is verified that in reality, the true strategy optStrat_2 gets executed every time this apparent strategy is played by the defender. Therefore the Nash equilibrium is attained and the value of the game (V) remains the same as previous. Since the Nash equilibrium is attained, if the defender adheres to appStrat_2, the attacker cannot gain a higher payoff than V if he alters his strategy.

If the defender were to follow a game model based on perfect information, optStrat_2 would be his/her apparent strategy. This scenario was also simulated and it was observed that the game is not in Nash equilibrium. This was observed by setting the attacker's strategy to Strat_1 = [0 1; 0 1] and the value of the game (V_A) obtained was 422.8347, which is higher than V. Note that the increment in the attacker's gain can be much higher depending on the specification of the particular game (e.g., reward matrices and transition probabilities). It was also verified that, if the attacker adheres to optStrat_1 (which corresponds to the Nash equilibrium), then the value of the game remains the same as expected.

	Defender's Action 1 (d)	Defender's Action 2 (nd)
Attacker's Action 1 (a)	80 (0.7, 0.3)	-20 (0.7, 0.3)
Attacker's Action 2 (na)	100 (1, 0)	0 (1, 0)

The Rewards (to the attacker) and State Transition Probabilities at *NormalState*

	Defender's Action 1 (d)	Defender's Action 2 (nd)
Attacker's Action 1 (a)	300 (0.4, 0.6)	100 (0, 1)
Attacker's Action 2 (na)	300 (0.8, 0.2)	100 (0, 1)

The Rewards (to the attacker) and State Transition Probabilities at *HackedState*

Figure 3: Specifications of the game in the second experiment

We now discuss our second experiment which shows the existence of such a game where strategies suggested by perfect information models could not be executed. For the game in Figure 3 the true optimal strategy obtained for defender was optStrat_2 = [0 1; 1 0]. Apparent strategies for all possibilities of true strategies were calculated and it was observed that none of them were equal to optStrat_2. This is illustrated by Figure 4 that shows the Euclidian distance between the calculated apparent strategy (1st column) and the true optimal strategy optStrat_2 (1st column). We observe that no point in the graph touches the XY plane, which signifies that no possibility of true strategies can lead to an apparent strategy equal to optStrat_2. This result shows that it is not always possible for the defender to execute the strategy (optStrat_2) prescribed by the perfect information model.

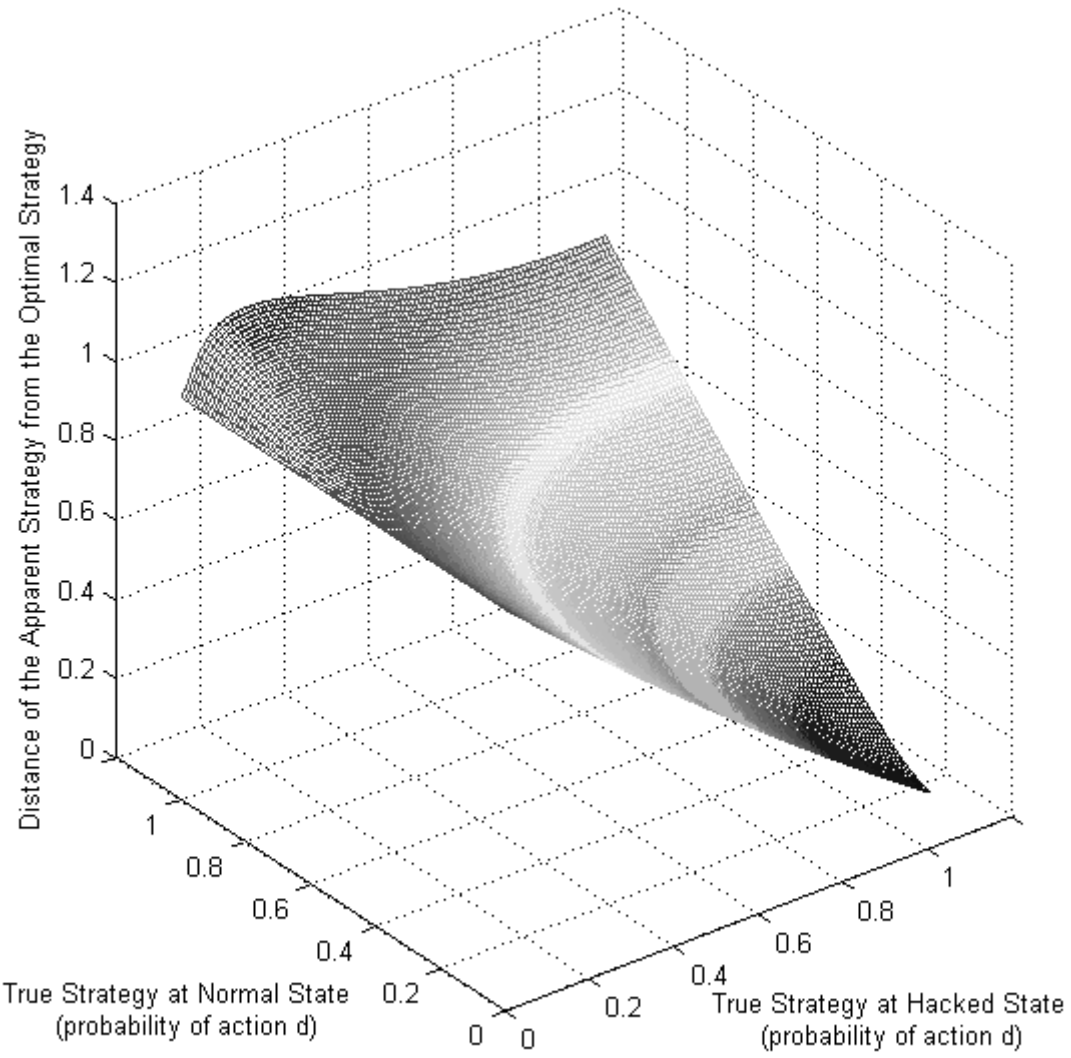

Figure 4: The second experiment result — this plot implies that the defender cannot execute an apparent strategy which is same as the *optimal strategy*

4. Related Work

(Hamilton 2002) outlined the areas of game theory which are relevant to information warfare. (Liu 2005) presented a methodology to model the interactions between a DDoS attacker and some defense mechanism such as 'pushback'. The following papers are most relevant to our work.

(Lye 2002; Lye 2005) proposed a perfect-information stochastic general-sum game and computed the Nash equilibrium using simulation. Unfortunately, the used equilibrium computation algorithm for this general-sum game was not available in these papers.

(Alpcan 2003) proposed a general-sum, static, finite game with dynamic information. Moreover, (Alpcan 2004) presented an imperfect information repeated game with `finite steps' or `infinite steps'. They analyzed the Nash equilibrium in the general-sum setting.

(Alpcan 2006) captured the operation of the IDS using a finite-state Markov chain. With a few numerical examples, tools such as minimaxQ and naive Q-learning were used to find the best strategies of the players. (Nguyen 2009) viewed the security problem as a general-sum repeated game. This model considers that the players cannot make perfect observations of each other's previous actions.

The following table compares our paper with the prior body of work. The dimensions used for the comparison include the type of analysis (static, dynamic or none) present in the work.

Work	Stochastic game?	Perfect information?	Zero-sum / general-sum game	Type of analysis
1. (Lye 2002; Lye 2005)	Yes	Perfect	General-sum	Static
2. (Alpcan 2003)	No(static game)	Imperfect	General-sum	Static
3. (Alpcan 2004)	No(repeated game)	Imperfect	General-sum	Dynamic
4. (Alpcan 2006)	Yes	Imperfect	Zero-Sum	Only Numerical Examples
5. (Nguyen 2009)	No(repeated game)	Imperfect	General-Sum	Dynamic
6. Our work	Yes	Imperfect	Zero-sum	Static

5. Conclusion and Future Work

Techniques that were proposed in the literature used stochastic game models to emulate network security game, and showed how to determine the best strategy for the defender considering the possible attack strategy used by the attacker. However, the prior research work assumed that the players have perfect information about the current state of the game, which generally does not hold in reality. Our model relaxed this assumption and enriched the prior game models by enabling them to capture more realistic scenarios. This paper presented a theoretical analysis using which the system administrator can compute his/her best strategy to reach the Nash equilibrium of a stochastic game even if the IDS sensor is imperfect. Our theoretical results were validated via simulation experiments.

This paper presented a static analysis to compute the best stationary strategy of the players. It was not discussed how the equilibrium can be reached during the game being played. Our aim is to investigate an answer to this question in the future work.

6. Bibliography

(Alpcan 2003) Alpcan T. and Basar T. "A game theoretic approach to decision and analysis in network intrusion detection." *Proc. of the 42nd IEEE Conference on Decision and Control.* Maui, HI, 2003. 2595--2600 Vol.3.

(Alpcan 2004) Alpcan T. and Basar T. "A game theoretic analysis of intrusion detection in access control systems." *Proc. of the 43rd IEEE Conference on Decision and Control.* Bahamas, 2004. 1568--1573 Vol.2.

(Alpcan 2006) Alpcan T. and Baser T. "An intrusion detection game with limited observations." *12th Int. Symp. on Dynamic Games and Applications.* Sophia Antipolis, France, 2006.

(Filar 1997) Jerzy A. Filar and Koos Vrieze. *Competitive Markov decision processes.* New York: Springer, 1997.

(Nguyen 2009) Kien C. Nguyen, Tansu Alpcan and Tamer Başar. "Security Games with Incomplete Information." *Proc. of the 2009 IEEE International Conference on Communications (ICC 2009).* Dresden, Germany, 2009.

(Lye 2005) Lye, Kong-wei and Jeannette Wing. "Game strategies in network security." *International Journal of Information Security* 4 (02 2005): 71--86.

(Lye 2002) Lye, Kong-Wei and Jeannette Wing. "Game Strategies in Network Security." *Proceedings of the Workshop on Foundations of Computer Security.* 2002.

(Liu 2005) P. Liu, W. Zang, and M. Yu. "Incentive-based modeling and inference of attacker intent, objectives, and strategies." *ACM Transactions on Information and System Security (TISSEC).* 2005.

(Hamilton 2002) S. N. Hamilton, W. L. Miller, A. Ott, and O. S. Saydjari. "Challenges in applying game theory to the domain of information warfare." *Proceedings of the 4th Information survivability workshop (ISW-2001/2002).* 2002.

(Roy 2010) Sankardas Roy, Charles Ellis, Sajjan Shiva, Dipankar Dasgupta, Vivek Shandilya, and Qishi Wu. "A Survey of Game Theory as Applied to Network Security." *to appear in 43rd Hawaii International Conference on System Sciences.* Hawaii, 2010.

(Williams 1966) Williams, J. D. *The Compleat Strategyst.* New York: McGraw-Hill, 1966.