A Stochastic Game Model with Imperfect Information in Cyber Security

Harkeerat Bedi
Department of Computer Science, University of Memphis
hsbedi@memphis.edu
Dr. Sajan Shiva and Dr. Dipankar Dasgupta

Introduction

Game theory provides huge potential to address the cyber security problem. The interaction between the attacker and the defender (system administrator) can be considered as a game. One of the techniques proposed by prior researchers used stochastic game models to emulate network security games.

Background:

Stochastic game: Each player performs an action at a given state and receives a payoff. Based on the previous state & actions performed, the game moves to a new state.

Perfect Information: The state of the game is always known to each player. imperfect Information: The present state of the game is not always known.

Problem:

Prior researchers determined the Nash Equilibrium (NE) strategy for the defender considering the possible attack actions. However, they assumed that the players have perfect information about the current state of the game, which generally does not hold in reality.

Our contribution:

We compute Nash Equilibrium (NE) strategy for a zero-sum stochastic game with imperfect information. In particular:

> We present a static analysis and compute the best strategy of the system administrator in realistic scenarios.
> Our analysis and simulation experiments illustrate that the system administrator will be better off if he/she executes the strategy prescribed by the perfect information models.

Related Work and Motivation

<table>
<thead>
<tr>
<th>Related work</th>
<th>Stochastic game?</th>
<th>Perfect Information?</th>
<th>Zero-sum</th>
<th>General-sum</th>
<th>Type of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Lye 2005)</td>
<td>Yes</td>
<td>Perfect</td>
<td>General-sum</td>
<td>Static</td>
<td></td>
</tr>
<tr>
<td>(Alpcan 2003)</td>
<td>No</td>
<td>Imperfect</td>
<td>General-sum</td>
<td>Static</td>
<td></td>
</tr>
<tr>
<td>(Alpcan 2004)</td>
<td>No</td>
<td>Imperfect</td>
<td>General-sum</td>
<td>Dynamic</td>
<td></td>
</tr>
<tr>
<td>(Alpcan 2006)</td>
<td>Yes</td>
<td>Imperfect</td>
<td>Zero-Sum</td>
<td>One Numerical Example</td>
<td></td>
</tr>
<tr>
<td>(Paganon 2009)</td>
<td>Yes</td>
<td>Imperfect</td>
<td>General-sum</td>
<td>Dynamic</td>
<td></td>
</tr>
<tr>
<td>Our work</td>
<td>Yes</td>
<td>Imperfect</td>
<td>Zero-sum</td>
<td>Static</td>
<td></td>
</tr>
</tbody>
</table>

Prior stochastic game models for network security (Lye 2005) assume that the players have perfect information about the current state of the game, which generally does not hold in reality.

In real systems, a user sees a sensor (e.g., Intrusion Detection System (IDS)) to observe the current state of the system to decide the strategy.

It is widely believed that no real sensor can perfectly react the environment, i.e., usually there is a non-zero error probability. Therefore, in most cases, the above assumption about perfect information does not hold in real life.

To make the analysis simpler, we consider a game with only two states as illustrated in Figure 1.

The Game Model

Our model is an extension of the prior model (Lye 2005) and considers a player $\gamma$, $\gamma = 1, 2$, observes the game’s true state at a particular moment by an imperfect sensor device. That means, player $\gamma$ can view the true state $s_1$ or $s_2$ at any state in the information set $\bar{I}_s$ with some probability $\alpha = \frac{|s_1|}{|s|}$ or $\gamma = \frac{|s_2|}{|s|}$ being an element of $\bar{I}_s$.

Compared to the perfect information model, player $\gamma$’s action space may become wider, i.e., player $\gamma$ may take an action which is allowed at a state $s_k$, when $s_k \in \bar{I}_s$ belonging to the information set $\bar{I}_s$. Let $\bar{I}_s$ denote the set of possible actions of player $\gamma$ when his/her information set is $\bar{I}_s$. Then $\bar{I}_s = \bigcup_{s \in \bar{I}_s} \bar{I}_s$, where $\bar{I}_s$ denotes the action set of player $\gamma$ when he/she is sure that the true current state is $s_k$. Below we formally define the outcome of player $\gamma$ ’s extended action set $\bar{I}_s$ compared to $I_s$.

Before we formally define the outcome of player $\gamma$ ’s extended action set $\bar{I}_s$, compared to $I_s$ in the previous model, when the true state is $s_1 \in \bar{I}_s$ but $s_1 \notin I_s$ then in terms of the influence on state transition probability, $\gamma$ is equivalent to player $\gamma$ taking no action at state $s_1$.

However, regarding the influence on player $\gamma$ ’s payoff $\pi'$ may not be equivalent to player $\gamma$ taking no action at state $s_1$ depending upon the cost of the execution of $s_1$.

Prior researchers determined the Nash Equilibrium (NE) strategy for the defender considering the possible attack actions. However, they assumed that the players have perfect information about the current state of the game, which generally does not hold in reality.

We assumed the defender knows the sensor’s error probabilities (False Positive $\gamma_{FP}$ and False Negative $\gamma_{FN}$) and the true strategies $\gamma_1$ and $\gamma_2$ when the true state is $s_1$.

Our work is an extension of the prior model (Lye 2005) and considers that a player $\gamma$, $\gamma = 1, 2$, observes the game’s true state at a particular moment by an imperfect sensor device. That means, player $\gamma$ can view the true state $s_1$ or $s_2$ at any state in the information set $\bar{I}_s$ with some probability $\alpha = \frac{|s_1|}{|s|}$ or $\gamma = \frac{|s_2|}{|s|}$ being an element of $\bar{I}_s$.

Compared to the perfect information model, player $\gamma$’s action space may become wider, i.e., player $\gamma$ may take an action which is allowed at a state $s_k$, when $s_k \in \bar{I}_s$ belonging to the information set $\bar{I}_s$. Let $\bar{I}_s$ denote the set of possible actions of player $\gamma$ when his/her information set is $\bar{I}_s$. Then $\bar{I}_s = \bigcup_{s \in \bar{I}_s} \bar{I}_s$, where $\bar{I}_s$ denotes the action set of player $\gamma$ when he/she is sure that the true current state is $s_k$. Below we formally define the outcome of player $\gamma$’s extended action set $\bar{I}_s$ compared to $I_s$.

Before we formally define the outcome of player $\gamma$’s extended action set $\bar{I}_s$, compared to $I_s$ in the previous model, when the true state is $s_1 \in \bar{I}_s$ but $s_1 \notin I_s$ then in terms of the influence on state transition probability, $\gamma$ is equivalent to player $\gamma$ taking no action at state $s_1$.

However, regarding the influence on player $\gamma$’s payoff $\pi'$ may not be equivalent to player $\gamma$ taking no action at state $s_1$ depending upon the cost of the execution of $s_1$.

Optimization

To compute the payoff, the modified Newton’s method requires solving a system of linear equations. This functionality is achieved by using an additional component that generates the payoff to the attacker compared to our model.

To compute the output, the modified Newton’s method requires solving a system of linear equations. This functionality is achieved by using an additional component that generates the payoff to the attacker compared to our model.

Optimization

To compute the payoff, the modified Newton’s method requires solving a system of linear equations. This functionality is achieved by using an additional component that generates the payoff to the attacker compared to our model.

Simulation Results

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 (Farre 1997). An iterative non-linear optimization algorithm is used. The input to this algorithm includes the state transition matrix and the payoff matrix. As this is a zero-sum game, the only player’s current 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 game matrix as in (Wolfe 1968). The value of $\gamma$, $\gamma = 1, 2$, is given as input.

The value of $\gamma$, $\gamma = 1, 2$, is given as input.

Result Summary:

> Our first experiment shows that perfect information models (Lye 2005) can give higher payoff to the attacker compared to our model.

> Our second experiment shows the advantage of such a game where strategies suggested by perfect information models could not be executed.

Limitations and Scope of Future Work

> We assumed the defender knows the sensor’s error probabilities (False Positive and False Negative ratios).

> Our current analysis is for only two states and each players' action set has only two actions.

> We performed a static ( offline) analysis: Strategies are determined before the game is played.