TOWARD A GAME-THEORETIC METRIC FOR NUCLEAR POWER PLANT SECURITY

L.T. MACCARONE
University of Pittsburgh
Pittsburgh, United States of America

J.R. JAMES
Sandia National Laboratories
Albuquerque, United States of America

T.R. ORTIZ
Sandia National Laboratories
Albuquerque, United States of America

D.R. SANDOVAL
Sandia National Laboratories
Albuquerque, United States of America

R.J. BRUNEAU
Sandia National Laboratories
Albuquerque, United States of America

D.G. COLE
University of Pittsburgh
Pittsburgh, United States of America

C.C. LAMB
Sandia National Laboratories
Albuquerque, United States of America
cclamb@sandia.gov

Abstract

Modern commercial nuclear power plants depend on computer resources to ensure a profitable and safe operating environment. Because these computer resources are integrated with the plant, a strong cybersecurity program is essential to protect the plant's physical assets. The paper proposes a game-theoretic approach to select appropriate cybersecurity controls for plant systems and to quantify the security of the plant. The residual heat removal system of a boiling water reactor plant is analyzed as a case study. The cybersecurity of the system is cast as a stochastic game played by a defender and an attacker. The stochastic elements of the game model uncertainty in the outcomes of the players' actions. Methods are proposed to define the states of the stochastic game and to estimate the attacker's success rate. Challenges of solving the game are identified.

1. INTRODUCTION

Modern commercial nuclear power plants rely on the seamless integration of cyber components with underlying physical processes to achieve a profitable and safe operating environment. This integration of the cyber and physical worlds offers many improvements over traditional plant systems, like advanced instrumentation and control techniques and improved system monitoring and diagnostics [1, 2, 3]. While these technologies enable significant new capabilities, they can also introduce new vulnerabilities to plant systems.

Because the cyber and physical worlds are integrated, a successful cyber attack on a nuclear power plant could result in physical damage to the system, including severe consequences such as core damage. To protect a nuclear power plant, security teams must determine which cyber and physical defenses are required, and the degree to which those defenses must be implemented. A method is needed that can accomplish two things: (1) to quantify the security of nuclear power plants, and (2) to determine the defenses required to reasonably ensure the plant's safe operation. To achieve these objectives, in this paper we develop a game-theoretic approach to the cybersecurity problem.

Game theory is a mathematical approach used to analyze the interactions of multiple decision-makers, or players. Each player has a set of strategies from which to choose. The outcome of the game is dependent on the strategies selected by all the players. Each player receives a utility that is dependent on the outcome of the game.

The utility is a quantification of the game's outcomes that indicates the preferences of the players. Using game theory, we can determine which defense strategies to implement to optimize our utility.

Stochastic game theory will be implemented to study nuclear power plant cybersecurity. A stochastic game is a dynamic system that evolves as the players take action. As the game progresses through time, it traverses a finite set of states that describe the environment of the player's interaction. The transition from one state to the next is determined at random as a function of the actions selected by the players. This enables us to study the repeated interactions of decision-makers where the outcomes of their decisions are undetermined.

The residual heat removal (RHR) system of a boiling water reactor will be examined to demonstrate the construction of a stochastic game for nuclear power plant security. First, the players of the cybersecurity game will be identified. Second, the structure of the stochastic game will be defined. Third, the player's action sets and utility functions will be defined. Finally, the challenges of solving the game will be discussed and areas of future work will be identified.

2. STOCHASTIC GAME THEORY

A stochastic game is a dynamic system that evolves as the players take action [4]. The stochastic element of the game arises from uncertainty about the outcome of the players' actions. This section provides the theoretical framework for a stochastic game, and later sections describe the application of stochastic game theory to the cybersecurity of a residual heat removal system.

Let the set of players consist of a defender and an attacker denoted as $\mathbb{P} = \{D, A\}$. The defender's parameters will be designated with a "*D*" subscript, and the attacker's parameters will be designated with an "*A*" subscript.

Let the set of stochastic states be denoted $\mathbb{S} = \{s_1, s_2, \ldots, s_M\}$. Note that this set is finite with M stochastic states. These states describe the environment for the players' interactions. For example, one stochastic state may be a normal state where the plant is operating as intended and the attacker has not breached plant defenses. Another stochastic state may be a penetrated state where the attacker has breached plant defenses but has not damaged the system. From the penetrated state, the attacker may be able to escalate the attack and drive the game to a stochastic state where the plant is damaged.

Let the set of actions available to player A/D at state s be denoted $a_{A/D}(s)$. When both players select their actions, the resulting action profile $(a_D(s), a_A(s))$ determines the probability of transitioning from state s to every other state in $\mathbb{S}$. After an action profile has been selected, the transition to the next state is determined stochastically by a transition vector given by a function

$$p(s, a_D, a_A) = \big(p(s_1|s, a_D, a_A), p(s_2|s, a_D, a_A), \ldots, p(s_M|s, a_D, a_A)\big) \tag{1}$$

In this game, we assume the Markov property. The Markov property is the property that the probability distribution of future states is not dependent on the state history, only the current state [4].

Note that the summation of elements of the transition vector is constrained to equal unity. This constraint removes the possibility of the game concluding before reaching an absorbing state (i.e. a state from which the probability of transitioning to every other state is zero). In this case, the only way to conclude the game is by reaching an absorbing state.

With each transition, each player earns an immediate utility. Let the immediate utility earned by player A/D be denoted $r_{A/D}(s_i, a_D, a_A, s_j)$. This notation reflects that the immediate utilities of each player are a function of the originating state, s_i, the action profile selected in s_i, (a_D, a_A), and the resulting state, s_j. The utilities used in this game will be measured in monetary units.

To evaluate the game, the immediate utilities must be aggregated by a cumulative utility function. We assume a discount factor, $\beta \in (0, 1)$, that assigns a larger weighting to the utilities that occur earlier in the game than to the utilities that occur later in the game [4]. This discount factor is selected based on the value that the players place on the utility earned earlier in the game compared to utility earned later in the game. Each player seeks to maximize his cumulative utility.

The solution concept for a stochastic game is the Nash equilibrium. At a Nash equilibrium, all of the players are playing strategies that are best responses to the strategies of the other players. This means that there is no incentive for any player to unilaterally deviate from the equilibrium. Within the context of a stochastic game, the Nash equilibrium can be interpreted as a prescription for play that identifies the optimal strategy for each player at each stochastic state.

The following sections will describe the practical design of a stochastic game to analyze the cybersecurity of an RHR system.

3. RESIDUAL HEAT REMOVAL SYSTEM

Stochastic game theory will be used to study the residual heat removal system (RHR) of a BWR/4 plant. The RHR system is used for both cooling and reactor vessel coolant inventory control. An overview of the RHR system is provided in Fig. 1. Several motor-operated valves (MOVs) are implemented throughout each of the systems. MOVs that are normally closed are shaded and MOVs that are normally open are not shaded. The RHR system consists of two interconnected systems that are nearly identical. Each RHR system has two RHR pumps that feed a recirculation pump [5, 6].

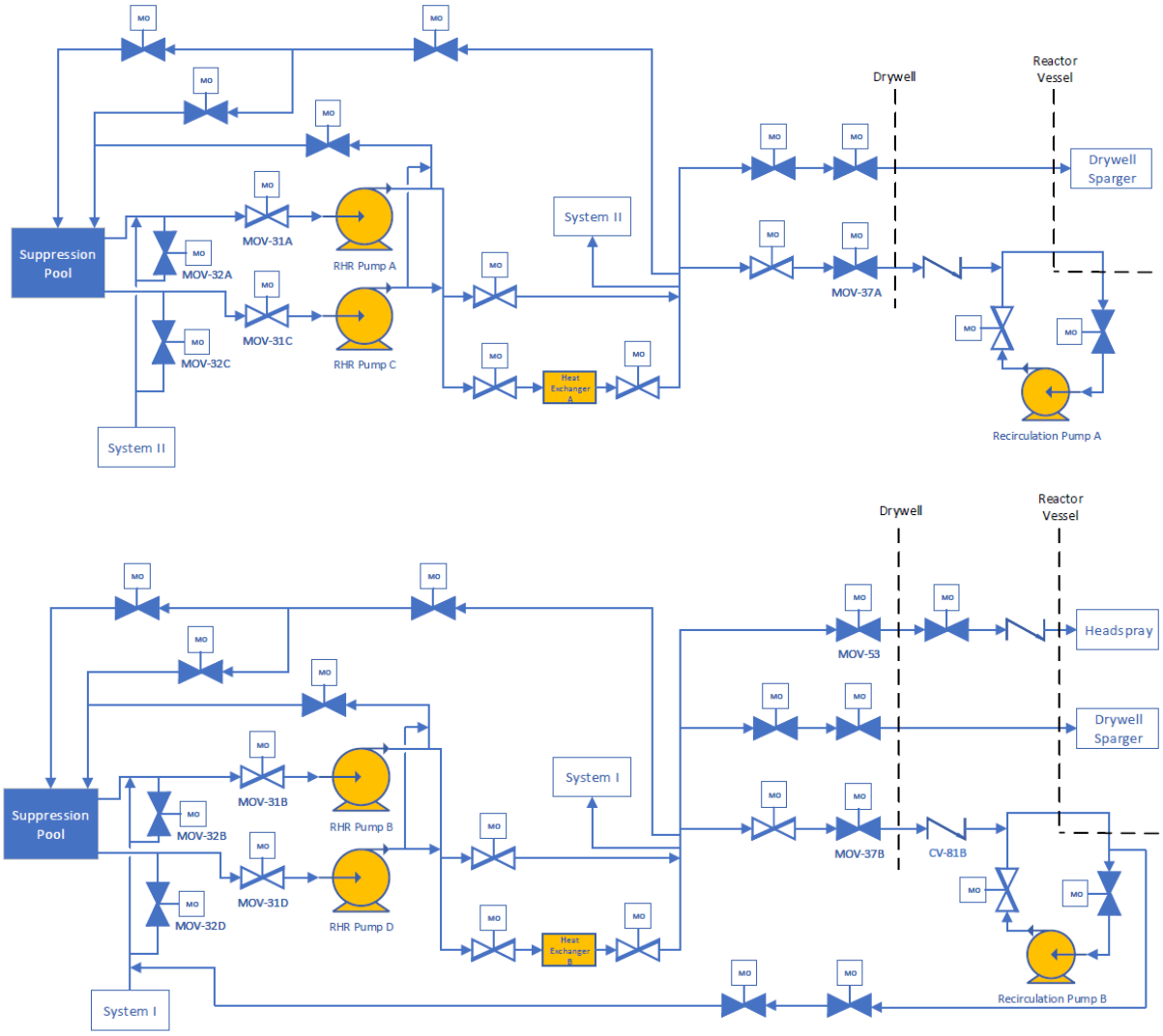

FIG. 1. RHR System 1 (top) and System 2 (bottom).

This work will consider the RHR operating in the low pressure coolant injection (LPCI) mode. The LPCI mode is automatically initiated during a loss of coolant (LOCA) accident to maintain reactor water level. The RHR pumps take water from the suppression pool and discharge to the reactor vessel. Two of the four RHR pumps must inject to meet the required LPCI flow for a design base LOCA [6].

The postulated network topology of the RHR system is shown in Fig. 2. This is an isolated network within the plant. We have assumed that the RHR system is controlled by two programmable logic controllers (PLCs). PLC-1 controls all of the components in RHR System 1, including all MOVs and pumps A and C. PLC-2 controls all of the components in RHR System 2, including all MOVs and pumps B and D. The two PLCs are connected by a network switch.

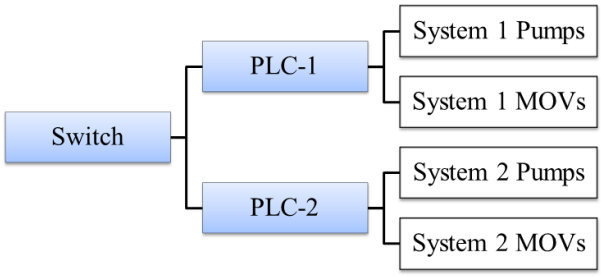

FIG. 2. RHR network topology.

4. DESCRIPTION OF PLAYERS

In this stochastic cybersecurity game, we assume that one attacker is playing against one plant defender. Both players will be defined by their capabilities, limitations, and goals. Various libraries exist that describe the motivations and capabilities of several classes of threat agents. A possible set of attack profiles for nuclear facilities is provided in [7]. Tiers of attackers for military applications are described in [8]. The threat agent library developed by Intel Corporation describes threats to information security in terms of a set of attributes [9]. Using a given threat library, the threat agent of concern can be identified and described.

Once the threat agent has been identified, the agent's goal(s) can be defined. The definition of the attacker's goals is important, because they will be used to define the stochastic state space. Examples of attackers and their goals (with respect to nuclear facilities) are provided in Table 1.

TABLE 1: EXAMPLE THREAT AGENTS AND THEIR GOALS [7].

Threat Agent	Goals
Recreational Hacker	Fun; status development
Nation State	Intelligence collection; access for later actions; technology theft
Terrorist	Access for later actions; impact public opinion; chaos

The attacker in this game will be a government cyberwarrior. A government cyberwarrior is defined as "a state-sponsored attacker with significant resources to affect major disruption on a national scale" [9]. The attacker's capabilities and limitations are summarized below.

- The attacker does not have internal access to the plant
- The attacker does not have any legal limitations
- The attacker has significant resources and persists long-term
- The attacker is adept and is capable of sophisticated attacks

The goal of the attacker in this game is to cause core damage. A real attacker may have multiple goals and may prioritize those goals. We have selected one goal for ease of demonstration. The methods presented in these work are readily applicable for attackers with multiple goals.

The defender will be a nuclear power plant cybersecurity team. The defender is capable of implementing industry-standard cybersecurity practices. The goals of the defender are summarized below.

- Maintain power generation
- Prevent release of radioactive materials
- Prevent damage to plant equipment
- Prevent employee injury or loss of life
- Prevent unauthorized disclosure of plant data

A discount factor must be assigned to capture the players' preferences for utility earned earlier in the game relative to utility earned later in the game. We assume a discount factor of $\beta = 0.95$ to characterize both the defender's need for sustained plant operation and the attacker's willingness to focus on long-term rewards.

5. IDENTIFICATION OF STOCHASTIC STATES

The stochastic state space must be able to characterize the plant over the operating range of consideration and throughout the course of all postulated attacks. For complex systems with many digital devices, the unaltered

state space may be too large for the security game to be tractable. For example, consider a system with 25 components, each of which has two operational states: operational or nonoperational. For this system, there would be a minimum of $2^{25} = 33,554,432$ states corresponding to each possible combination of the individual components' operational states. A large state space presents challenges both in computation and in interpretation of results. Thus, a method is needed to constrain the size of the stochastic state space.

System-theoretic process analysis (STPA) is used to manage the size of the stochastic state space. STPA is a hazard analysis technique that considers both component failures and unsafe interactions of system components to model accident causation. Readers are encouraged to refer to [10] for greater detail and multiple example applications. The application of STPA to stochastic game theory is summarized in the following four steps.

1. Define the purpose of analysis. This step involves three sub-tasks: identifying losses, identifying system-level hazards, and identifying system-level constraints.

   Losses are consequences that are unacceptable to stakeholders. These losses should be aligned with the goals of the attacker. Examples of losses for a nuclear facility include core damage, radiation release, lost generation, personnel injury or death, and major equipment damage. In this work, the loss is core damage.

   Hazards are system states that can lead to a loss under a specific set of conditions. Hazards that may lead to core damage for an RHR system operating in LPCI mode are:

   (a) Loss of flow path alignment capability to RHR subsystems
   (b) Damage to RHR pumps
   (c) Inadequate flow for intended operation

   Constraints are conditions that must be met to prevent hazards. A constraint statement is often the inversion of the corresponding hazard statement. The constraints for the RHR system are:

   (a) Flow path alignment capability to RHR subsystems must be maintained
   (b) RHR pumps must not be damaged
   (c) Adequate flow must be maintained

2. Model the control structure. The control structure is a functional model of the interactions of the controllers with the controlled process through feedback and control actions. The control structure is identical to the network topology shown in Fig. 2.
3. Identify the unsafe control actions. Unsafe control actions (UCAs) are control actions that can lead to a hazard under certain conditions. It is important to note that the safety of a control action is dependent on the conditions under which the action takes place. A UCA describes the source of the control action, the type of control action, and the context of the control action.
4. Identify loss scenarios. A loss scenario summarizes the events that can lead to UCAs and hazards. There are many ways that an attacker could cause a UCA in the RHR system. For brevity, we summarize the combinations of UCAs that will lead to hazards in Table 2. In these examples, multiple UCAs must occur to cause a hazard.

   When nuclear power plants assess loss scenarios such as core damage, there are several systems and plant functions that must also fail for the loss to occur. In this example, "inadequate flow for intended operation" represents just one hazard that if combined with other system hazards could lead to core damage. Although analysis of core damage states is beyond the scope of the paper, failure of the RHR system to provide its intended function could contribute to a core damage state.

At this stage, we can generate a first iteration of the stochastic state space. The stochastic state space is shown in Fig. 3. For visual clarity, all transitions to the normal state are not shown. Every state except for the hazard states can transition to the normal state.

The game begins with an initial state where the plant is operating as expected. In this game, the normal state represents a LOCA scenario where the RHR would be required to operate in LPCI mode. From the normal state, the game may self-transition to the normal state or it may transition to one of several hacked states. For the game to transition from the normal state to a hacked state, the attacker's offensive action must be successful, and the defender's cybersecurity control action must be unsuccessful.

The set of hacked states is representative of all possible combinations of components that can be hacked by the attacker. In this game, there are three devices that can be hacked by the attacker: PLC-1, PLC-2, and the switch. Because there are three components that can be hacked, there are seven $(2^3 - 1)$ hacked states. From some hacked states, the game may transition to one of several hazard states. This occurs if the attacker is able to successfully use his privileges on the hacked device(s) to cause a hazard to the system. The game may also transition back to the normal state if the plant defender successfully expunges the attacker from the hacked components. It is also possible that the defender might only be able to expunge the attacker from a subset of the hacked components, and the game may transition to another hacked state. For ease of demonstration, we will assume that if the defender is

TABLE 2: HAZARDOUS CONDITIONS.

Case	Unsafe Control Actions	Hazards
1	(a) PLC-1 sends close signal to MOV-37A when LOCA conditions are indicated and reactor pressure is below threshold. (b) PLC-2 sends close signal to MOV-37B when LOCA conditions are indicated and reactor pressure is below threshold.	(a) Loss of flow path alignment capability to RHR subsystems
2	(a) PLC-1 does not send open signal to MOV-31A when LPC initiation signal is received. (b) PLC-1 does not send open signal to MOV-31B when LPCI initiation signal is received. (c) PLC-1 does not send open signal to MOV-31C when LPCI initiation signal is received.	(a) Damage to RHR pump (b) Loss of flow path alignment capability to RHR subsystems
3	(a) PLC-2 does not send start signal to Pump B during LPCI mode. (b) PLC-2 does not send start signal to Pump D during LPCI mode. (c) PLC-1 does not send open signal to MOV-37A when LOCA conditions are indicated and reactor pressure is below threshold.	(a) Loss of flow path alignment capability to RHR subsystems (b) Inadequate flow for intended operation
4	(a) PLC-1 sends stop signal to Pump A when reactor cooling has not been achieved and all valves in suction path are fully open. (b) PLC-1 sends stop signal to Pump C when reactor cooling has not been achieved and all valves in suction path are fully open. (c) PLC-2 sends stop signal to Pump B when reactor cooling has not been achieved and all valves in suction path are fully open.	(a) Inadequate flow for intended operation

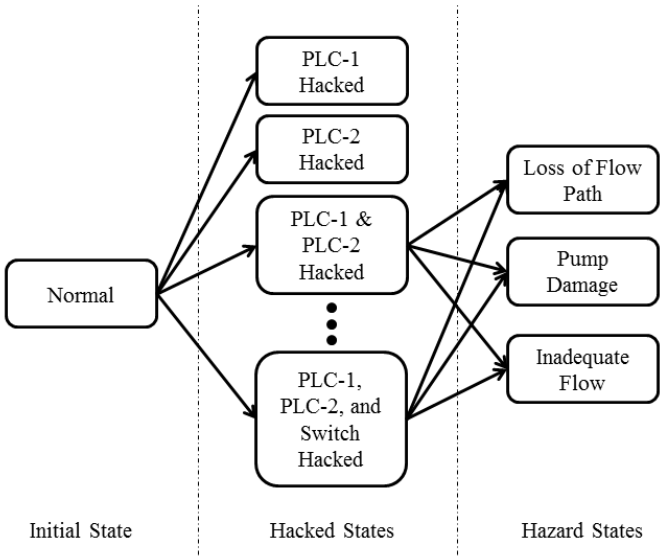

FIG. 3. The stochastic state space.

able to expunge the attacker from one hacked component, then he is able to expunge the attacker from all hacked components.

The set of hazard states includes all hazards identified in STPA that are aligned with the goals of the attacker. These hazards may cause a loss, but only under specific sets of conditions. Modeling the transition from a hazard to a loss is beyond the scope of this work. The hazard states are absorbing states, meaning that no transitions to other states are possible from the loss states. Reaching a hazard state signals the end of the game.

6. SELECTION OF ACTIONS

The action set available to each player must be defined for each stochastic state. To define accurate action sets, the game designer must have a thorough understanding of both the industrial control system devices and the players. An understanding of the control system devices is required to identify device vulnerabilities, cybersecurity control actions that can address those vulnerabilities, and malicious actions that can exploit those vulnerabilities.

The attacker and defender each have several choices to make regarding each component in the system. For

the defender, these choices address the configurations of the devices that can be hacked. For the attacker, these choices address the attack vector for bypassing the defender's cybersecurity control actions. The choices available to the players in the normal state are summarized in Table 3.

TABLE 3: DEFENDER'S AND ATTACKER'S CHOICES IN THE NORMAL STATE.

	PLC-1 & PLC-2	Switch	Communication Network
Defender's Choices	Authentication: on/off	Authentication: on/off	Encryption: on/off
	Wireless: on/off	Firewall: on/off	
Attacker's Choices	Join: yes/no	Join: yes/no	Decryption: yes/no
	Wireless Exploit: yes/no	Attack: yes/no	

Note that there are many levels of encryption and the encryption level determines the computational resources required for decryption. For the purpose of this work, we assume there is one industry-standard level of encryption available to the defender.

Note that the action sets available to the attacker and defender in the normal state must include every possible combination of choices pertaining to all of the devices in the system. For example, one complete action of the defender is to enable authentication on the PLCs and the switch, disable wireless on the PLCs, enable the firewall on the switch, and enable encryption on the communication network between the switch and PLCs.

We have chosen to augment the initial stochastic state space to allow the attacker to decide whether or not to attempt to password-crack a device if an attempt to join that device is unsuccessful. This augmentation of the state space reflect the attacker's tendency to first determine whether authentication is required before attempting a password crack. The creation of additional states is necessary to preserve the Markov property of the game.

In the hacked states, the attacker has the option to manipulate the hacked devices to attempt to cause a hazard. We assume that the attacker has stealthily hacked components to arrive at a hacked state, therefore the defender can not take action to return the game to the normal state. The hazards for this system and their corresponding unsafe control actions are summarized in Table 2. It is important to note that all hazards in this system require that the attacker has the ability to manipulate both PLCs. The hazard states are absorbing states, therefore no actions are available to either player.

7. DEFINING STATE TRANSITIONS

State transition probabilities must be defined at each state for each possible action profile at that state. Each attack and defense action can be described by a set of attributes. Using these attributes, the state transition probabilities will then be estimated. In this work, we use the exploitability metrics defined by the common vulnerability scoring system (CVSS), version 3.1 [11]. There are four metrics: (1) AV, The attack vector metric measures the context required for an attacker to exploit a vulnerability. A vulnerability that can be exploited remotely will have a larger AV score than a vulnerability that must be exploited locally. (2) AC, The attack complexity metric measures the conditions outside of the attacker's control that must be met for a vulnerability to be exploited. (3) PR, The privileges required metric measures the privileges the attacker must have to exploit the vulnerability. (4) UI, The user interaction metric measures whether another user besides the attacker to participate in the exploitation of the component. The possible values of AV, AC, PR, and UI are summarized in Table 4.

For each action profile, we use these CVSS metrics to estimate the transition probabilities. Note that we assume that the defender's actions are always implemented correctly. That is, if the attacker is able to circumvent the defender's protections, it is due to the attacker's own ability, not because the defender has failed to execute his actions correctly. This assumption may not be valid for complex cybersecurity control actions requiring advanced knowledge or extensive procedures.

For demonstration purposes, we consider the following action profile occurring in the normal state. The actions selected by the defender and attacker are summarized in Table 5. We use the CVSS metrics to evaluate each of the vulnerabilities that the attacker has attempted to exploit in his action selection. The attacker's attempted exploitations target PLC-1's authentication, PLC-2's authentication, the switch's authentication, and the switch's firewall. The vulnerabilities and their parameters are summarized in Table 6.

Based on these CVSS parameters, we estimate the probability of the attacker's success. This estimate is dependent on expert opinion and is an area of ongoing research. These estimates may be validated using representative capture-the-flag games with cybersecurity professionals. The estimated attack success rates are also shown in Table 6.

We can now determine the transition vector from the normal state for the given action profile. We assume that each component of the attacker's action is independent, that is, the attacker's efforts to hack each device are independent from the efforts to hack the other devices. The numerical probability of transitioning to each of the

TABLE 4: DESCRIPTION OF CVSS METRICS QUOTED DIRECTLY FROM [11].

Attack Vector	Description
Network (*N*)	The vulnerable component is bound to the network stack and the set of possible attackers extends up to and including the entire Internet.
Adjacent (*A*)	The vulnerable component is bound to the network stack, but the attack is limited at the protocol level to a logically adjacent topology.
Local (*L*)	The vulnerable component is not bound to the network stack and the attacker's path is via read/write/execute capabilities.
Physical (*P*)	The attack requires the attacker to physically touch or manipulate the vulnerable component.

Attack Complexity	Description
Low (*L*)	Specialized access conditions or extenuating circumstances do not exist. An attacker can expect repeatable success when attacking the vulnerable component.
High (*H*)	A successful attack depends on conditions beyond the attacker's control. A successful attack cannot be accomplished at will, but requires the attacker to invest in some measurable amount of effort in preparation or execution.

Privileges Required	Description
None (*N*)	The attacker is unauthorized prior to attack, and therefore does not require any access to settings or files of the the vulnerable system to carry out an attack.
Low (*L*)	The attacker requires privileges that provide basic user capabilities that could normally affect only settings and files owned by a user.
High (*H*)	The attacker requires privileges that provide significant control over the vulnerable component allowing access to component-wide settings and files.

User Interaction	Description
None (*N*)	The vulnerable system can be exploited without interaction from any user.
Required (*R*)	Successful exploitation of this vulnerability requires a user to take some action before the vulnerability can be exploited.

TABLE 5: DESCRIPTION OF CVSS METRICS QUOTED DIRECTLY FROM [11].

	PLC-1	PLC-2	Switch	Communication Network
Defender Action	Wireless - on Authentication - off	Wireless - on Authentication - on	Authentication - on Firewall - off	Encryption - on
Attacker Action	Wireless Exploit - no Join - yes	Wireless Exploit - no Join - yes	Join - yes Attack - yes	Decryption - no

TABLE 6: CVSS EVALUATION OF ATTACKER'S ACTION.

CVSS Metric	PLC-1 Authentication	PLC-2 Authentication	Switch Authentication	Switch Firewall
AV	Adjacent	Adjacent	Adjacent	Adjacent
AC	Low	High	High	Low
PR	None	High	High	None
UI	None	None	None	None
Success Rate	0.98	0.00	0.00	0.90

hacked states is given below. Note that the "*N*" subscript of the state variable s indicates the normal state, the subscript "1" indicates that PLC-1 has been hacked, "2" indicates that PLC-2 has been hacked, and "*S*" indicates that the switch has been hacked.

$$p(s_N|s_N,a_D,a_A) = (1-0.98)(1-0.90) = 0.0020 \qquad p(s_S|s_N,a_D,a_A) = 0.90(1-0.98) = 0.018$$
$$p(s_1|s_N,a_D,a_A) = 0.98(1-0.90) = 0.098 \qquad p(s_{1S}|s_N,a_D,a_A) = 0.98(0.90) = 0.88$$
$$p(s_2|s_N,a_D,a_A) = 0.00 \qquad p(s_{2S}|s_N,a_D,a_A) = 0.00$$
$$p(s_{12}|s_N,a_D,a_A) = 0.00 \qquad p(s_{12S}|s_N,a_D,a_A) = 0.00$$

Note that transitioning from the normal state to a hazard state is not possible.

Transition probabilities must be estimated in a similar manner for action profiles occurring in the hacked states. To cause a hazard either both PLC-1 and PLC-2 must be hacked, or the switch must be hacked and the communication between the switch and PLCs must be unencrypted. We assume that if the attacker has achieved one of these two conditions, the probability that he will be able to cause a hazard state is 0.7. This estimate is dependent on threat intelligence data that is specific to the adversary of concern. If the attacker's attempt to cause a hazard is unsuccessful, we assume that the defender returns the game to the normal state.

No analysis is required for transitions from the hazard states. The hazard states are absorbing states that signal the end of the game.

8. DEFINING UTILITY FUNCTIONS

The utility function is a qualitative representation of the players' preferences. The general form of player A/D's immediate utility function resulting from a transition from s_i to s_j after action profile (a_D, a_A) is

$$r_{A/D}(s_i, a_D, a_A, s_j) = \phi_{A/D}(s_j) - E_{A/D}(a_{A/D}) \tag{2}$$

The immediate utility function has two terms. We will explain each term within the context of a successful cyber-attack resulting in a transition from the normal state to a hacked state. The first term depends only on the resulting state, s_j. In our example, the resulting state is a hacked state, therefore it is reasonable that the attacker will achieve a reward and the defender will incur a penalty, i.e. $\phi_A(s_j) \geq 0$ and $\phi_D(s_j) \leq 0$. The second term quantifies the cost of the player's action. In our example, the attacker incurs a cost to attack, thus $E_A(a_A) < 0$. Similarly, there is an implementation cost for the defender's action, therefore $E_D(a_D) < 0$.

Accurately quantifying these terms can be challenging. Research into the economics of cybersecurity is a developing field [12, 13]. The economic impact of cyber-attacks is discussed in [14]. A cost-benefit analysis of cyber-terrorism can be found in [15]. Applications of security economics to critical infrastructure are explored in [16]. The economics of information security investment are discussed in [17].

For the purposes of demonstration, we will assume the values summarized in Table 7 and Table 8. In Table 7, the first row indicates the state transition that results in each transition utility parameter. In Table 8, the actions corresponding to PLC-1, PLC-2, the switch, and the communication network occur in the normal state. The actions corresponding to the hazard occur in the hacked states where the attacker has hacked PLC-1 and PLC-2 or has hacked the switch and can manipulate the PLCs over the communication network. Note that the defender does not have an expense utility parameter for the hazard condition because this cost is captured in the transition utility parameter if the attacker's efforts fail.

TABLE 7: TRANSITION UTILITY PARAMETERS FOR BOTH PLAYERS.

	Normal to Hacked	Hacked to Normal	Hacked to Loss of flow path	Hacked to Pump damage	Hacked to Inadequate flow
ϕ_D	-100	50	-10^4	-10^6	-10^4
ϕ_A	500	-300	10^5	10^5	10^5

TABLE 8: EXPENSE UTILITY PARAMETERS FOR BOTH PLAYERS.

	PLC-1 & PLC-2	Switch	Communication Network	Hazard
E_D	Authentication On: -10 Wireless Off: -5	Authentication On: -10 Firewall On: -50	Encryption On: -50	N/A
E_A	Join: -5 Wireless Exploit: -300	Join: -5 Attack: -500	Decryption: -100	Hazard: -500

The cumulative utility function is defined to analyze the outcome of the game. Each player seeks to maximize his cumulative utility. The cumulative utility function of player A/D is defined as

$$u_{A/D}(s_N, \pi_A, \pi_D) = \sum_{t=0}^{\infty} \beta^t \mathbb{E}[r_{A/D}(s^t, a_D^t, a_A^t, s^{t+1})], \quad s^0 = s_N \tag{3}$$

The parameter $\pi_{A/D}$ denotes player A/D's strategy — the discrete probability distribution that is assigned to the action set of that player at each state in the stochastic game. The solution of the game provides optimal values of $\pi_{A/D}$. The cumulative utility function is also dependent on the initial state, s^0. In this game, the initial state is the normal state, s_N. The function $\mathbb{E}[\cdot]$ denotes the expectation over the states and strategies.

9. SOLUTION CHALLENGES

This paper has presented several tools to construct a stochastic game for nuclear power plant cybersecurity. There are several challenges that must be addressed to solve this stochastic game.

First, solving the game identifies the optimal strategies for the defender and attacker, π_D and π_A. Each strategy assigns a probability to the actions available to that player in each state of the game. This presents the challenge that the parameter space is large. In this game, we must optimize over 576 total probabilities (256 probabilities for the defender and 320 for the attacker).

Second, we want to ensure that the solution is the global optimum. The global optimum is the minimizing solution from the set of all possible solutions. Finding the global optimum is difficult for many optimization methods. This is particularly so for this problem with its large search space.

Third, we must ensure the solution's uniqueness. The existence of a Nash equilibrium is guaranteed, but its uniqueness is not [18]. If there are multiple Nash equilibria, additional factors must be considered to determine which equilibrium is most credible. Credibility considers the players' preferences to predict which equilibrium is most likely to occur; factors can include focal points, risk dominance, and Pareto dominance [19].

10. CONCLUSION

A strong cybersecurity program is essential to protect the assets of commercial nuclear power plants. Security teams at commercial nuclear power plants need a method to prioritize cybersecurity controls to secure the plant. A stochastic game-theoretic approach was presented to meet this need. A game-theoretic approach enables security teams to assess the costs and benefits of cybersecurity controls while also considering the goals and abilities of the adversary.

System-theoretic process analysis (STPA) was used to define the state space of a stochastic game. STPA enables security teams to understand how an attacker with access to digital components might cause a hazard to occur. By defining system losses, hazards, and unsafe control actions, security teams can identify scenarios whereby an attacker could damage the plant.

The Common Vulnerability Scoring System (CVSS) was used to estimate the probability of an attack's success. The CVSS exploitability metrics serve as a lens for security experts to estimate state transition probabilities for each action profile in the game. This stage is dependent on expert opinion and future work is needed.

The construction of a stochastic game for a residual heat removal system was presented. Solving the game is challenging given the large parameter space of the attacker's and defender's strategies. Other challenges include ensuring that the optimization identifies the global solution and ensuring that the solution is unique. A solution to the game will be presented in future work.

Game-theory provides a mathematical framework to analyze the security of complex systems such as commercial nuclear power plants. Game theory offers a holistic approach to security by considering the cyber, physical, and human elements of system security. By implementing these game-theoretic techniques, security teams can prioritize security measures to ensure the safe and profitable operation of the plant.

ACKNOWLEDGEMENTS

The authors would like to thank Andrew J. Clark of Sandia National Laboratories for his contributions.

The work conducted at the University of Pittsburgh was supported by Sandia National Laboratories Contract 2084579: Civilian Nuclear Power Plant Systems Game and Risk Modeling.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

[1] IAEA, Computer Security of Instrumentation and Control Systems at Nuclear Facilities, IAEA Nuclear Security Series No. 33-T (2018).

[2] VINOD, S., et al., Symptom based diagnostic system for nuclear power plant operations using artificial neural networks, Reliability Engineering & System Safety, **82** 1 (2003) 33-–40.

[3] FARBER, J., et al., Using kernel density estimation to detect loss-of-coolant accidents in a pressurized water reactor, Nuclear Technology, **205** 8 (2018) 1043-–1052.

[4] FILAR, J., VRIEZE, K., Competitive Markov Decision Processes, Springer-Verlag, New York, United States of America (1996).
[5] UNITED STATES NUCLEAR REGULATORY COMMISSION, BWR/4 Technology Manual (R-104B), report, U.S. NRC, Washington, D.C., United States of America (2018).
[6] GENERAL ELECTRIC, "Residual heat removal system", General Electric Systems Technology Manual, report, General Electric (1996).
[7] IAEA, Computer Security at Nuclear Facilities, IAEA Nuclear Security Series No. 17 (2011).
[8] UNITED STATES DEFENSE SCIENCE BOARD, Resilient Military Systems and the Advanced Cyber Threat, report, U.S. DSB, Washington, D.C., United States of America (2013).
[9] CASEY, T., Threat Agent Library Helps Identify Security Risks, report, Intel Corporation (2007).
[10] LEVESON, N., THOMAS, J., STPA Handbook, report (2018).
[11] FIRST, Common Vulnerability Scoring System v3.1, report, FIRST.Org, Inc. (2019).
[12] ARMIN, J., 2020 Cybercrime economic costs: no measure no solution, Proceedings - International Conference on Availability, Reliability and Security, (2015) 701-–710.
[13] ANDERSON, R., MOORE, T., "Information security economics - and beyond", Advances in Cryptology - CRYPTO 2007. Lecture Notes in Computer Science, vol 4622. Springer, Berlin, Heidelberg (2007) 68–91.
[14] CASHELL, B., et al., The Economic Impact of Cyber-Attacks, report, Congressional Research Service, Washington, D.C., United States (2004).
[15] GIACOMELLO, G., et al., Bangs for the buck: a cost-benefit analysis of cyberterrorism, Studies in Conflict and Terrorism, **27** 5 (2004) 387–408.
[16] ANDERSON, R., FULORIA, S., Security economics and critical national infrastructure, Economics of Information Security and Privacy, (2010) 55–66.
[17] GORDON, L., LOEB, M., The economics of information security investment, ACM Transactions on Information and System Security, **5** 4 (2002) 438-–457.
[18] ZHAO, Y., et al., A game theoretic approach for responding to cyber-attacks on nuclear power plants, Proceedings - NPIC & HMIT, (2019) 399-–410.
[19] FUDENBERG, D., TIROLE, J., Game Theory, MIT Press, Cambridge, United States of America (1991).