NEUP Final Report

Submitted for NEUP project

Support for Reactor Operators in Case of Cyber-Security Threats

Project Start Data: 10/01/2016

Project End Date: 09/30/2019

PI:

Carol Smidts (The Ohio State University)

Co-PIs:

Indrajit Ray (Colorado State University)

Quanyan Zhu (New York University)

Timothy Mabry (Originally with Framatome Inc.)

Timothy McJunkin (Idaho National Laboratory)

December 20, 2019

Summary

Increasing use of digital technology in nuclear power plants (NPPs) poses cyber-security as a crucial threat to public safety and to continuous energy production. Cyber-security risks are comprised of complex known and unknown interactions between various entities, system vulnerabilities, network protocols, human users and malicious attacks. There is little understanding or research geared towards plant operators' response under cyber-security threats and operational procedures to cope with such threats. This is particularly critical when a cyber-security event masquerades as a safety incident or an evolving accident. Instead of leading operators to the remediation of the accident, the masked cyber event may lead them to circumvent the safety mechanisms of the plant. Given these challenges, the objective of this project is to develop a tool to characterize abnormal NPP events as "cyber" or "safety" incidents, and to further develop this tool as a real-time operator aid to assist in response to the incident. This NEUP Final Report summarizes the research efforts towards the above goals and the results in the following three aspects.

First, a methodology for classifying abnormal NPP events as cyber-attacks or safety incidents was developed. This provides the basis for responding to potential cyber-attacks. In this methodology, the causal relationships between various variables (i.e., physical variables, network variables, and control variables) for a nuclear system under consideration are modeled using dynamic Bayesian networks. Based on observations of certain variables (i.e., the evidence), inference is performed to identify the state of hidden variables, which helps indicate the occurrence of cyber-attacks or safety incidents. The proposed methodology was applied to the full scope NPP simulator system at The Ohio State University as a case study. The result shows that the proposed methodology is capable of distinguishing between cyber-attacks and safety incidents.

Second, a game-theoretic approach for responding to cyber-attacks on nuclear power plants was developed. This is the second major goal of the project, and the method is used to recommend optimal responses to reactor operators in the case of cyber-attacks. We consider two cases in this research. In the first case, the interaction between the defender (i.e., the reactor operator) and the attacker is modeled as a two-player, nonzero-sum, Markov game with an infinite time horizon. In the second case, we consider a semi-Markov game between the defender and the attacker within a finite time horizon. The model in the second case enables us to consider the time sensitive response of the defender. In each case, the method for deriving the Nash equilibrium is developed. The Nash Equilibrium of the game provides the valid prediction of both players' actions because no single player can benefit from unilaterally deviating from the equilibrium policy if the other player adheres to his/hers, hence it provides the best response of the defender to cyber-attacks. The proposed method in each case is demonstrated through a simplified benchmark digital feedwater control system.

Third, a prototype display is developed to integrate the above information. The prototype display depicts the communication information to the operator and incorporates the cyber-security tool developed by the team into the interface. It consists of the regular display in current control rooms, as well as a panel used to indicate the occurrence of cyber-attacks or safety events and a panel used to guide the operator's response to detected cyber-attacks.

Table of Contents

Table of Figures

Table of Tables

1 Background

Increasing use of digital technology in nuclear power plants (NPPs) poses cyber-security as a crucial threat to public safety and to continuous energy production. Cyber-security risks are comprised of complex known and unknown interactions between various entities, system vulnerabilities, network protocols, human users and malicious attacks. There is little understanding or research geared towards plant operators' response under cyber-security threats and operation procedures to cope with such threats. This is particularly critical when a cyber-security event masquerades as a safety incident or an evolving accident. Instead of leading operators to the remediation of the accident, the masked cyber event may lead them to circumvent the safety mechanisms of the plant. Given these challenges, the objective of this project is to develop a tool to characterize abnormal NPP events as "cyber" or "safety" incidents, and to further develop this tool as a real-time operator aid to assist in response to the incident.

In a nuclear power plant, it is of vital importance to detect anomalies and classify them as safety events caused by system or component failures and accidental human errors or malicious cyber attacks. Incorrect diagnosis of the anomalies will prompt the operators to take inappropriate measures in responding to the anomalies, which may result in significant losses. Development of such an event classifier is the first major goal in this project and the research towards this goal is summarized in Section 2.

When an anomaly in a nuclear power plant is identified as a cyber attack, the operators need to take appropriate measures to mitigate the consequence caused by the attack. In responding to cyber attacks, the operators need not only to consider the effects of their own actions, but also to take into account the attacker's possible actions. In this project, a game theoretic approach for responding to cyber attacks on nuclear power plants is developed. The approach consists of two main elements, game theory which provides the optimal response policy of the operators, and probabilistic risk assessment which serves as a credible reward function used in the modeling of the game. In this project, we consider two models for describing the defender-attacker game: one is a Markov game with an infinite time horizon, and the other one is a semi-Markov game with a finite time horizon. The research related to these two models are summarized in Section 3 and Section 4, respectively.

As the third main goal of the project, a prototype display as the operator response support system is developed. The prototype display depicts the communication information to the operator and incorporates the cyber-security tool developed by the team into the interface. It consists of the regular display in current control rooms, as well as a panel used to indicate the occurrence of cyber-attacks or safety events and a panel used to guide the operator's response to detected cyber-attacks. The work towards this goal is summarized in Section 5.

2 Dynamic Bayesian Networks based Abnormal Event Classifier for Nuclear Power Plants in case of Cyber Security Threats.

This work is presented in a detailed manner in the paper titled "Dynamic Bayesian Networks based Abnormal Event Classifier for Nuclear Power Plants in case of Cyber Security Threats", which will be submitted to Progress in Nuclear Energy Journal as a copyrighted document.

The implementation of digital systems that enable improved control, ease of diagnostics and a reduction in cost of maintenance, in Nuclear Power Plants has marked the beginning of a new era in the nuclear industry. However, in addition to the problems caused by safety events that occur due to component failures and accidental human errors, the deployment of digital systems gave rise to the possibility of malicious attacks such as the SQL slammer worm attack on the David-Besse nuclear power plant in 2003, and StuxNet in 2010. Such cyber-attacks can result in damages ranging from economic loss due to disturbances in operation to a worst-case scenario of severe environmental and life threats due to core damage. Significant research has been conducted on the cyber security aspects of Information Technology (IT) systems with a primary focus on confidentiality and privacy. However, in the case of Industrial Control Systems (ICSs) that support critical infrastructure, the potential impacts of cyber security events are on safety and not necessarily as much on confidentiality and privacy. As the number of reported occurrences of cyber-attacks has increased over the past several years, a number of research projects to protect Industrial Control Systems have been undertaken by academic, industrial and government communities. Several works in the past have presented Supervisory Control and Data Acquisition (SCADA) system specific Intrusion Detection Systems and Intrusion Prevention Systems [1]. Also, a set of cyber-attack detection systems by both the control and networks communities, which consider the physical behavior along with the network behavior of SCADA systems were presented in [2]. Signature based cyber-attack detection was presented in [3] and recently researchers have also explored machine learning based detection schemes [4]. Nuclear systems have well-defined response strategies for safety events, i.e., for abnormal events caused by component failures. A clever attacker can take advantage of such implementations, by masquerading his cyber-attack as a component failure to elicit an incorrect response from the operator. An incorrect diagnosis of a cyber-attack as a safety event that results in a wrong response can further aggravate the situation and cause significant damage to the power plant. Hence correctly classifying an abnormal event as either a safety event or a cyber-attack is of utmost importance.

As the first of three steps in the NEUP project "Support for Reactor Operators in Case of Cyber-Security threats," we developed an event classifier to differentiate between pure safety events that occur due to component failures, and cyber-attacks in real-time. Figure 1 depicts the architecture of the developed monitoring system for abnormal event classification. Three sets of variables, physical variables, control variables and network variables are collected from the operating environment using the data collection module. Physical variables represent measurements such as flowrate, pressure, temperature whereas control variables constitute input control signals to the actuators and network variables represent the features of digital signals such as frequency of packet transmission, information contained in the packets. These three sets of variables are collected using equipment that is not part of and hidden from the operating environment. Thereby we ensure that the data used by the monitoring system is always accurate and cannot be corrupted by any cyber-attacks on the operating environment. It is important to observe that these measured variables are inputs and outputs to the components in the operating environment. The system models module contains mathematical models for the physical components and the controllers in the operating environment. In addition, for software components such as controllers in the operating environment, identical controllers with identical software are installed on the monitoring system as part of the system models module to identify any abnormal behavior resulting from software faults. Using the data collected by the data collection module, the system models module generates a set of expected outputs for every component in

the operating environment. The expected values computed by the system models module and the measured outputs are compared in the anomaly detection module to generate diagnostic variables. These diagnostic variables are assigned discrete states Normal (N), Failed (F) or Compromised (C), based on the anomaly detection rules specified in the module. These diagnostic variables are provided as evidence to the recursively implemented 2 Time-Slice Dynamic Bayesian Network (2T-DBN) [5] to infer the marginal probabilities of different health states of the components. A physical component has two health states Normal (N) and Failed (F), whereas a digital component has three states, the third one being Compromised (C).

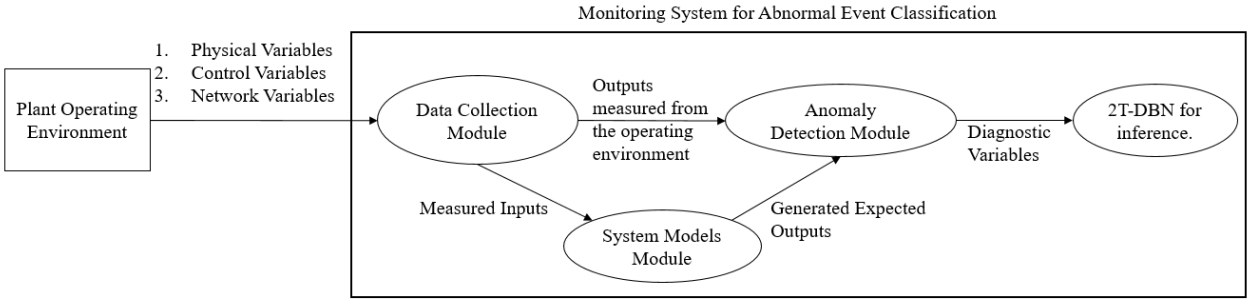


Figure 1. The architecture of the monitoring system for abnormal event classification.

The proposed abnormal event classifier was tested using a hardware in the loop experimental system connected to a full-scope nuclear power plant simulator. The experimental system is modeled on the digital feedwater control system (DFWCS) as presented in NUREG-CR 6942 [6]. A total of 19 cyber-attacks comprising 2 Denial of Service attacks, 3 Man in the Middle attacks and 14 malware-based attacks, and 7 safety events consisting 4 valve failures and 3 pipe leaks, were systematically designed. A Knowledge based approach was used in the construction of the 2T-DBN for abnormal event classification in the experimental system. The 2T-DBN had a total of 59 nodes and 82 arcs. The conditional probability tables were manually fine-tuned using 9 cyber-attacks and 2 safety events as reference. The abnormal event-classifier successfully detected the nature of the abnormal event in all the remaining 15 cases and successfully localized the cyber-attack or fault in 14 cases.

3 Markov Game with Infinite Horizon for Attack Response

This section is mainly based on a paper published in NPIC&HMIT 2019: Zhao, Y., Huang, L., Smidts, C. and Zhu, Q., 2019, January. A game theoretic approach for responding to cyber-attacks on nuclear power plants. In *11th Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies, NPIC and HMIT 2019* (pp. 399-410). American Nuclear Society.

3.1 Background

Digital systems are being more widely used in nuclear power plants. These systems are able to improve plant safety and operational performance, but also contain vulnerabilities that may expose critical digital assets to malicious cyber-attackers. There have been several cyber events in nuclear facilities[7], even though some of them are inadvertent events. The potential for serious consequences caused by cyber-attacks on nuclear power plants has drawn increasing attention to cyber-security issues in the nuclear industry. The US Nuclear Regulatory Commission (NRC) and relevant agencies have developed regulations or guidelines to strengthen the cyber-security posture of nuclear facilities. However, these efforts focus on how to prevent cyber-attacks from happening through measures, for instance, reducing the attack vector by elimination of known vulnerabilities, access control, and identification and authentication. To the authors' knowledge, little research has been done on how to respond to cyber-attacks on nuclear power plants. Successful responses should be able to minimize the incurred loss, or even to deter and terminate the attacks.

However, this is a challenging task for human operators. In the decision-making, the operators need to realize that the evolution of the plant depends not only on their own actions, but also on the attacker's actions. Besides, the operators need to take into account the long-term consequence due to the actions rather than myopically choose the optimal action for the current state. To address this problem, we propose a game theoretic approach for responding to cyber-attacks on nuclear power plants. The interaction between the defender (i.e., the operators) of the system and the attacker is modeled as a two-player, nonzero-sum, stochastic game which is represented as a competitive Markov decision process. The Nash Equilibrium derived for the game provides the valid prediction of both players' actions because no single player can benefit from unilaterally deviating from the equilibrium policy if the other player adheres to his/hers. Hence, the equilibrium policy provides the best response of the defender to cyber-attacks. The resulting tool can be used as a decision support system for the operators, and will help improve cyber-security capabilities of nuclear power plants. This research corresponds to ***Task 2.2*** and ***Task 2.3*** of the original proposal.

To identify credible system states and state transitions in the MDP model, we propose an approach based on analyzing all the possible interactions between the defender and the attacker starting from the initial normal system state. To obtain credible transition probabilities, we propose an approach based on probabilistic risk assessment to calculate the separated elements in a transition probability.

As a case study, the proposed approach is applied to a benchmark digital feedwater control system. The system was originally used for dynamic reliability analysis for digital systems and was described in a US NRC report [6]. Simplifications and modifications have been made for illustration of the proposed approach. The Markov decision process model for the system is developed and results (i.e., the Nash Equilibrium policy and state values) are obtained and discussed.

In the rest of the section, the game theoretic approach is described in Section 3.2. The approaches for identifying system states and state transitions and for obtaining transition probabilities are introduced in Section 3.3. In Section 3.4, we introduce the case study based on the feedwater control system, and provide discussions on the results. A summary of the research introduced in this section is provided in Section 3.5.

3.2 Game theoretic approach

We model the interaction between the defender and the attacker as a discounted stochastic game with an infinite time horizon $t = 1,2,\ldots,\infty$ and a finite state set $\mathbb{S}$. This game can be represented concisely as a

competitive MDP. A MDP is a tuple $(\mathbb{S}, \mathbb{A}_1, \mathbb{A}_2, \mathrm{T}, r_1, r_2, \beta)$ and the elements are explained in further detail in the following description.

The state set $\mathbb{S}$ contains all possible states of the system, e.g., normal operation, and the system states $s, s' \in \mathbb{S}$ represent the system states at time t and $t+1$, respectively. Player P_i can take different actions $a_i \in \mathbb{A}_i(s)$ at different states $s \in \mathbb{S}$, e.g., the attacker compromising different components in the system or the defender taking different protection measures. Thus, the action set $\mathbb{A}_i(s)$ is state dependent. Here i is used to indicate the two players, $i = 1$ for the defender, and $i = 2$ for the attacker.

At each state $s \in \mathbb{S}$, players taking action pair (a_1, a_2) results in the next state s' with the transition probability $p(s'|s, a_1, a_2)$, and in the meantime, a stage reward $r_i(s, a_1, a_2, s')$ for player i. $p(s'|s, a_1, a_2)$ is defined in the state transition probability function T. The stage rewards for the two players in general do not sum to zero, hence a nonzero-sum game. The action pair and the state transition may incur costs to the defender and the attacker. Each player i tries to decide a state-feedback mixed-strategy $\phi_i: \mathbb{S} \mapsto \Delta\mathbb{A}_i(s)$ where $\Delta\mathbb{A}_i(s) \coloneqq \left\{f: \mathbb{A}_i(s) \mapsto \mathbb{R}_+ \middle| \sum_{a_i \in \mathbb{A}_i(s)} f(a_i) = 1\right\}, \forall s \in \mathbb{S}$ represents the set of the probability distribution over the state-dependent action set. With a little abuse of notation, $\phi_i(s, a_i)$ is the probability of player i taking action a_i at state s, i.e., $\phi_i(s, a_i) \geq 0, \forall a_i \in \mathbb{A}_i(s), \forall s \in \mathbb{S}$ and $\sum_{a_i \in \mathbb{A}_i(s)} \phi_i(s, a_i) = 1, \forall s \in \mathbb{S}$. All these feasible mixed-strategies ϕ_i constitute the feasible policy set Φ_i.

Since both players aim to maximize their cumulative rewards in the long run, we consider the following cumulative expected objectives with discount factor $\beta \in (0,1)$ and initial state s^0. The discount factor assigns more weight to the immediate reward and less to the potential reward in the future. Note that the future state S^t (we use capital letters to represent random variables) at time t is a random variable with the support $\mathbb{S}$ and the notation E is the expectation over the mixed-strategy and the state randomness.

$$u_i(s^0, \phi_1, \phi_2) \coloneqq \sum_{t=0}^{\infty} \beta^t E\big(r_i(S^t, A_1, A_2, S^{t+1})\big),\ S^0 = s^0 \qquad 3\text{-}1$$

Each player i determines his strategy ϕ_i to maximize his cumulative reward $u_i(s^0, \phi_1, \phi_2)$. However, since the transition probability p and the stage reward r_i are coupled with the action of both players, the solution concept is the Nash equilibrium (NE) defined as follows.

Definition: The strategy pair $(\phi_1^* \in \Phi_1, \phi_2^* \in \Phi_2)$ is called a ε-Nash equilibrium (NE) if there exist no profitable unilateral deviations, i.e., $\exists \varepsilon > 0$ such that the state value for the defender, $v_1(s^0)$, and the state value for the attacker, $v_2(s^0)$

$$v_1(s^0) \coloneqq u_1(s^0, \phi_1^*, \phi_2^*) \geq u_1(s^0, \phi_1, \phi_2^*) - \varepsilon, \forall \phi_1 \in \Phi_1, \forall s^0 \in \mathbb{S} \qquad 3\text{-}2$$

$$v_2(s^0) \coloneqq u_2(s^0, \phi_1^*, \phi_2^*) \geq u_2(s^0, \phi_1^*, \phi_2) - \varepsilon, \forall \phi_2 \in \Phi_2, \forall s^0 \in \mathbb{S} \qquad 3\text{-}3$$

If $\varepsilon = 0$, it degenerates to the Nash equilibrium.

Define $h_i(s, a_1, a_2)$ as an intermediate variable used for convenience, and $\psi(v_1, v_2, \phi_1, \phi_2)$ as the objective function to be minimized in Equation 3-7

$$h_i(s, a_1, a_2) \coloneqq \sum_{s' \in \mathbb{S}} p(s'|s, a_1, a_2)\big(r_i(s, a_1, a_2, s') + \beta v_i(s')\big) \qquad 3\text{-}4$$

$$\psi(v_1, v_2, \phi_1, \phi_2) \coloneqq \sum_{\mathrm{i}=1}^{2} \sum_{s \in \mathbb{S}} [v_i(s) - \sum_{a_1 \in \mathbb{A}_1(s)} \sum_{a_2 \in \mathbb{A}_2(s)} \phi_1(s, a_1)\phi_2(s, a_2)h_i(s, a_1, a_2)] \qquad 3\text{-}5$$

and we use dynamic programming to characterize and capture the NE of the nonzero-sum game.

$$v_i(s) = \sum_{a_1 \in \mathbb{A}_1(s)} \phi_1^*(s, a_1) \sum_{a_2 \in \mathbb{A}_2(s)} \phi_2^*(s, a_2)\, h_i(s, a_1, a_2), \forall s \in \mathbb{S}, i \in \{1,2\} \quad 3\text{-}6$$

Since the transition reward is related to both s and s', we extend the existing methods[8] to obtain the following Theorem.

There exists at least one NE for the general-sum discounted stochastic game aiming at maximizing the utility in (1).

A strategy pair $(\tilde{\phi}_1 \in \Phi_1, \tilde{\phi}_2 \in \Phi_2)$ forms a NE strategy if and only if the point $\tilde{z} = (\tilde{v}_1, \tilde{v}_2, \tilde{\phi}_1, \tilde{\phi}_2)$ is the global minimum of the following nonlinear program with the objective function $\psi(\tilde{z}) = 0$.

If $\hat{z} = (\hat{v}_1, \hat{v}_2, \hat{\phi}_1, \hat{\phi}_2)$ is feasible for the nonlinear program with the objective function $\psi(\hat{z}) = \gamma > 0$. Then the strategy pair $(\hat{\phi}_1, \hat{\phi}_2)$ forms a ε-Nash equilibrium with $\varepsilon = \frac{\gamma}{1-\beta}$.

$$\min_{v_1, v_2, \phi_1, \phi_2} \psi(v_1, v_2, \phi_1, \phi_2) \quad 3\text{-}7$$

such that

$$(a). \sum_{a_2 \in \mathbb{A}_2(s)} \phi_2(s, a_2)\, h_1(s, a_1, a_2) \le v_1(s), \forall s \in \mathbb{S}, \forall a_1 \in \mathbb{A}_1(s) \quad 3\text{-}8$$

$$(b). \sum_{a_1 \in \mathbb{A}_1(s)} \phi_1(s, a_1)\, h_2(s, a_1, a_2) \le v_2(s), \forall s \in \mathbb{S}, \forall a_2 \in \mathbb{A}_2(s) \quad 3\text{-}9$$

$$(c). \sum_{a_i \in \mathbb{A}_i(s)} \phi_i(s, a_i) = 1, \forall s \in \mathbb{S}, i \in \{1,2\} \quad 3\text{-}10$$

$$(d).\ \phi_i(s, a_i) \ge 0, \forall s \in \mathbb{S}, i \in \{1,2\}, \forall a_i \in \mathbb{A}_i(s) \quad 3\text{-}11$$

Since the program is nonlinear in both objective function ψ and constrains (a) and (b), the first challenge is to find the global optimal rather than local optimal. The equation $\varepsilon = \frac{\gamma}{1-\beta}$ provides a metric for the equilibrium approximation. Another challenge comes from the large dimensionality of the state space and action space, which can be solved by exploiting the sparsity and approximation[9,10]. The third challenge comes from the fact that the Theorem only guarantees the existence yet not the uniqueness of the NE. Principle and agent model with the Stackelberg equilibrium concept may be adopted in the future where the system operator announces his policy first and the attacker aware of the announced policy best-response to it.

3.3 MDP modeling

In this section, we discuss the approach for determining the set of system states, state transitions, and transition probabilities.

3.3.1 System states and state transitions

The system state $s \in \mathbb{S} \coloneqq (E_1, E_2, \dots, E_N)$ can be defined as a tuple of N elements constituting the system. Each element E_i has a finite number of potential values and e_i is its realization value. So each system state s corresponds to a combination of the states of the elements $(e_1, e_2, \dots, e_N)$. If each element has 2 possible values, there will be a total number of 2^N of combinations. It is unrealistic to analyze system states by considering all such combinations because of the difficulty in modeling and the computation complexity. Besides, some combinations do not actually exist in reality. For example, for the control system used in the case study in Section 3.4, which has one main computer and one backup computer, we should not consider the state where both computers are being used for control at the same time.

To identify credible system states, we propose an approach based on the analysis of the interactions between the defender and the attacker starting from the initial normal system state. The basic idea of the approach is to first identify all possible states resulting from the initial state by considering all the possible interactions between the defender and the attacker at the initial state. For the newly identified system states, we can conduct the same analysis and identify all the states resulting from these newly identified states. The analysis can be continued until no new state is found, which indicates that all possible system states have been identified. Besides the states, the transitions from each state can also be identified as a function of defender and attacker actions. The number of system states can be reduced further by grouping the identified states. For example, if a component is not used, whether the component is in a normal or compromised state no longer matters, so the two states can be grouped into one state. The procedure for the analysis is shown in Figure 3-1.

$\mathbb{S} = \emptyset$ \\ $\mathbb{S}$ is the set of system states we want to identify, and is set empty at start
$T = \emptyset$ \\ T here is only used to store the information about state transitions, not transition probabilities
$\mathbb{Q} = \{q_1\}$ \\ $\mathbb{Q}$ is a temporary set storing newly found states, and q_1 is the initial state
While $\mathbb{Q} \neq \phi$
  For $i = 1: |\mathbb{Q}|$ \\ for all the states in $\mathbb{Q}$
    If $q_i \in \mathbb{Q}$ \\ q_i is the ith state in $\mathbb{Q}$
      $\mathbb{Q} = \mathbb{Q} - q_i$ \\ remove q_i from $\mathbb{Q}$
      Else
        $\mathbb{S} = \mathbb{S} + q_i$ \\ add q_i to $\mathbb{S}$
        For $j = 1: |\mathbb{A}_1(q_i)|$ \\ for all the actions by the defender at state q_i
          For $k = 1: |\mathbb{A}_2(q_i)|$ \\ for all the actions by the attacker at state q_i
            \\ identify the state s' resulting from the state q_i, defender action $a_{1j}(q_i)$ and attacker action $a_{2k}(q_i)$. $e_{q_i,p}$ is the state of the pth element for system state q_i
            $s' = (e'_1, e'_2, \dots, e'_N) \leftrightarrow \left(q_i = (e_{q_i,1}, e_{q_i,2}, \dots, e_{q_i,N}), a_{1j}(q_i), a_{2k}(q_i)\right)$
            $\mathbb{Q} = \mathbb{Q} + s'$ \\ add newly found states to $\mathbb{Q}$
            $T = T + (q_i, a_{1j}(q_i), a_{2k}(q_i), s')$ \\ add the newly found state transition to T

Figure 3-1 The procedure for identifying system states and state transitions.

3.3.2 State transition probabilities

To formalize the approach, we subdivide the set $\mathbb{E}$ used for characterizing system states into three sub-sets. The first sub-set $\mathbb{E}_b = \{E_{b_1}, \dots, E_{b_m}\}$ contains the basic components that can be actually attacked by the attacker or controlled by the defender. Such components include a computer, and a sensor, etc. The second sub-set $\mathbb{E}_f = \{E_{f_1}, \dots, E_{f_n}\}$ contains elements representing system functions for which the states are affected by the states of the basic components. Such elements include the feedwater system, and the reactor shutdown system, etc. The defender or attacker can only affect system functions through manipulating components in the first sub-set. The third sub-set contains elements representing the consequences influenced by system functions. Here we focus on the reactor core, so there is only one variable E_r in the third sub-set. System states can be represented using the three sub-sets as, $S = (S_b, S_f, S_r)$, where $S_b = (E_{b_1}, \dots, E_{b_m})$, $S_f = (E_{f_1}, \dots, E_{f_n})$, and $S_r = E_r$. We still use S and S' to represent state variables at time t and $t + 1$, respectively.

The causal relationships between the variables are shown in Figure 3-2. The relationships imply that basic components states and defender and attacker actions at the current time will influence basic components states at the next time step. The basic components states define system function states, so this causal relationship is placed at the same time. System functions at current time step influence the state of the reactor core at the next time step. The time step here actually means stage, so the time interval between two stages can be long. The state of the reactor core also depends on its previous state because if it's already damaged, it can no longer be returned to be OK. The causal relationships in Figure 3-2 can be extended,

for instance, the other consequences besides core damage can be considered, and more layers can be added to the framework. However, the analysis follows the same approach introduced below.

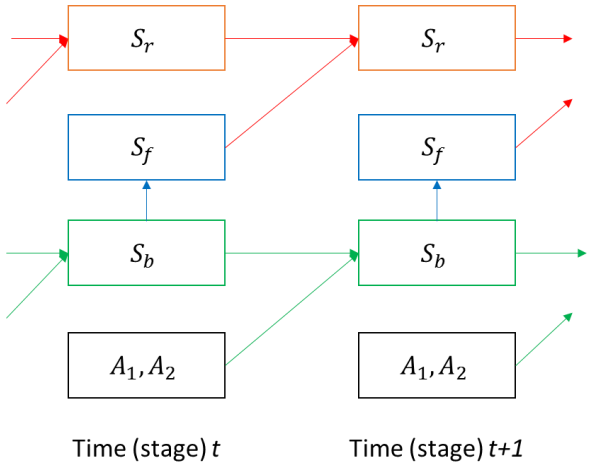


Figure 3-2 The causal relationships between different variables characterizing system states.

By separating S into S_b, S_f, and S_r, and S' into S'_b, S'_f, and S'_r, the transition probability can be expressed as:

$$P(S'|S, A_1, A_2) = P\left((S'_b, S'_f, S'_r)\middle|(S_b, S_f, S_r), A_1, A_2\right) \quad 3\text{-}12$$

Based on the causal relationships in Figure 3-2 and using the chain rule of probability, Equation 3-12 can be simplified to be Equation 3-13.

$$P(S'|S, A_1, A_2) = P(S'_b|S_b, A_1, A_2) \times P\left(S'_f\middle|S'_b\right) \times P\left(S'_r\middle|S_f, S_r\right) \quad 3\text{-}13$$

The first term on the right side of Equation 3-13 quantifies the interaction between the defender and the attacker before any concrete damage has been realized. This transition probability can be obtained by expert judgment, statistics, or simulations[11], or based on the Common Vulnerability Scoring System (CVSS) scores[12].

The second term on the right side of Equation 3-13 quantifies the relationship between basic components states and the states of system functions. This probability can be obtained through fault tree modeling and analysis. Fault tree analysis is one essential element of PRA. It quantifies the relationship between a top event and a set of basic events, and is able to obtain the probability of the top event based on the information about the basic events. One limitation of traditional fault tree analysis is that it is based on Boolean logic so it is not able to handle probabilistic or soft relations between basic events and the top event. For example, the basic event *sensors failed and used* in Figure 3-3 below may lead to the top event *feedwater system failure* with a probability that is smaller than 1. Traditional fault tree can be extended by integrating other techniques, e.g., Bayesian networks, to address this limitation.

The third term on the right side of Equation 3-13 quantifies the relationship between the state of the reactor core and its influencing variables. This transition probability can be obtained through event tree analysis which is another essential element of PRA. An event tree captures the relationship between the states of system functions (i.e., failed or success) during the evolution of an accident and the final consequence (e.g., core damage or not, containment failure or not). Based on the states of system functions, event tree analysis provides us with the probability of the consequence, which in this research is core damage. The third term in Equation 3-13 can be specified further as below, where $P\left(S'_r\middle|S_f\right)$ is obtained from event tree analysis.

$$P(S_r'|S_f, S_r) = \begin{cases} 1 & for\ S_r' = core\ damage\ if\ S_r = core\ damage \\ 0 & for\ S_r' = core\ OK\ if\ S_r = core\ damage \end{cases} \quad 3\text{-}14$$

$$(S_r'|S_f, S_r) = \begin{cases} P(S_r'|S_f) & for\ S_r' = core\ damage\ if\ S_r = core\ OK \\ 1 - P(S_r'|S_f) & for\ S_r' = core\ OK\ if\ S_r = core\ OK \end{cases} \quad 3\text{-}15$$

The approach for obtaining state transition probabilities is illustrated in Figure 3-3 through an example taken from the case study in Section 3.4. In this example, the actions are *use approximate model* for the defender and *no action* for the attacker. The first term equals 1 because state 16 covers all the basic components states. Mapping the basic components states for state 2 in the MDP to the basic events in the fault tree, we obtain the second term, which equals 1. Assuming the reactor is successfully shut down and the probability that the auxiliary feedwater system successfully responds to the transient is 0.9999, we obtain the probability of core damage from the event tree as 0.0001. Multiplying the three terms, we finally obtain the transition probability as 0.0001.

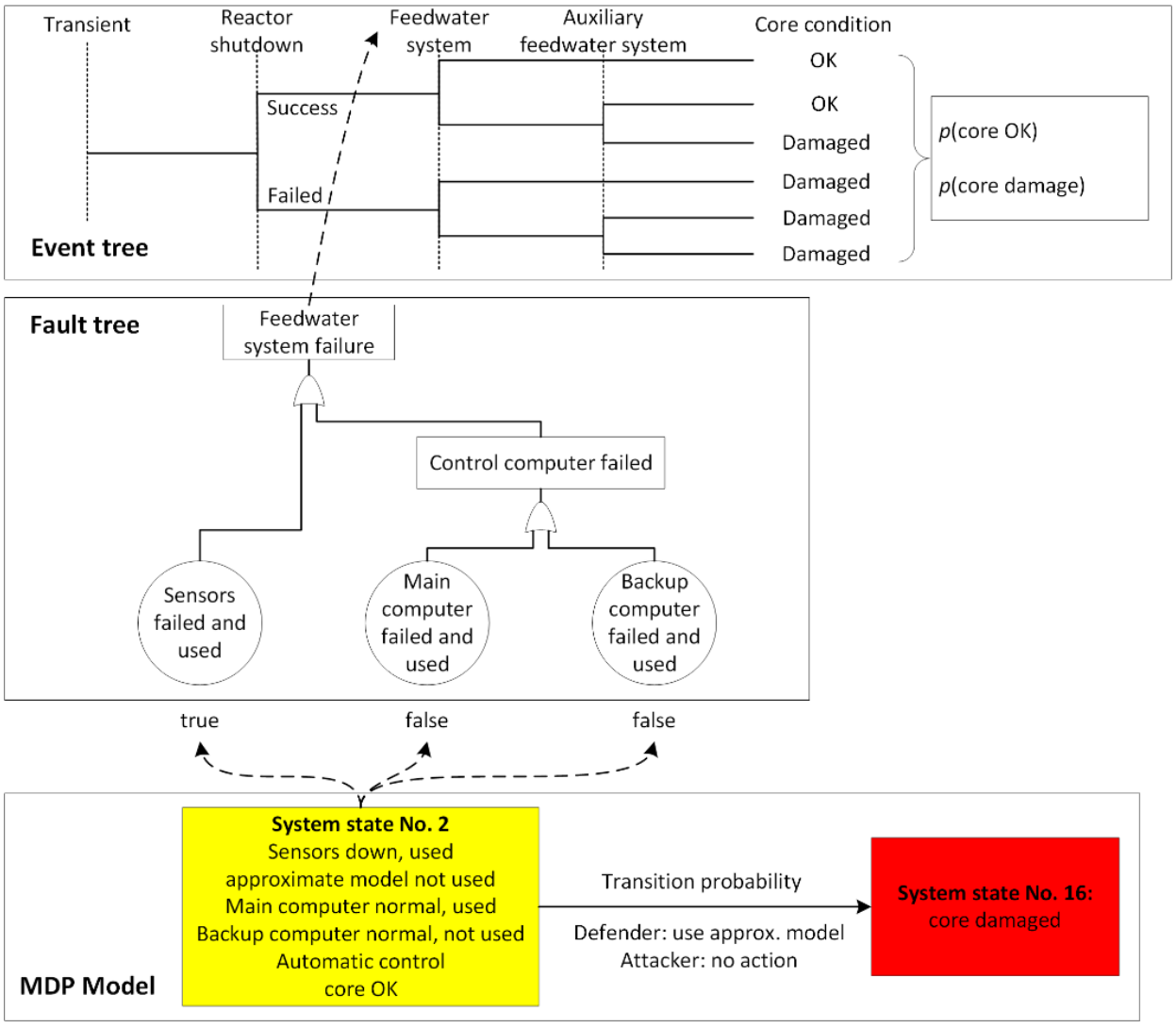


Figure 3-3 Illustration of the approach to obtaining state transitions probabilities.

3.4 Case study

3.4.1 Introduction to the system

The system under study is a digital feedwater control system. During the normal operation and transients caused by abnormal events, e.g., a cyber-attack, the system is used to maintain sufficient water flow to the steam generator, which cools the reactor core. Failure of this system may lead to dryout of the steam generator and the damage of the core. The system in this case study is taken from the US NRC report NUREG/CR-6942[6], where it was originally used as a benchmark system for dynamic reliability analysis of digital systems in nuclear power plants. Simplifications and modifications have been made to the original system to simplify the analysis.

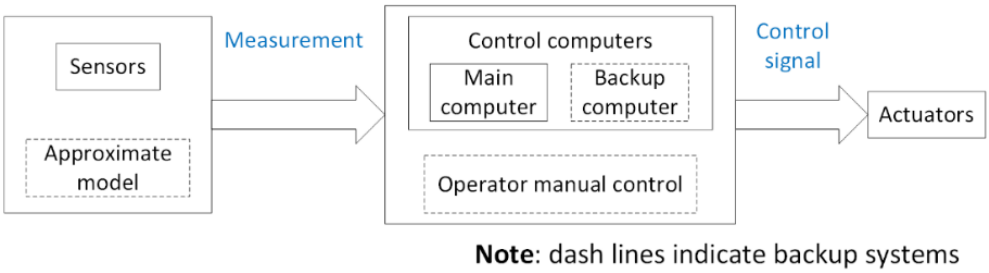


Figure 3-4 The digital feedwater control system.

The function of the system is similar to the one of a Programmable Logic Controller. It consists of three main components as shown in Figure 3-4. The sensors provide the information on the state of the plant, e.g., water level in the steam generator, feedwater flow, and steam flow. The information is sent to the computer that implements the control algorithm. There are two computers that can be used. Under normal operation of the plant, the main computer is used. In the case of main computer failure, the backup computer takes over the control. In certain cases, the operators will take over the control, so the control transitions from the mode of automatic control to manual control. The control signal calculated in either the main computer or the backup computer or obtained from the operators is sent to field actuators, e.g., feedwater pump, and feedwater flow regulating valves. In this research, we focus on cyber-attacks on sensors, the main computer, and the backup computer. A potential type of attack on sensors is false data injection attacks[13]. The two computers can be attacked by installing malware, as in the attack on Iranian nuclear facilities[14].

The benchmark system was originally used for reliability analysis, rather than cyber-security analysis, so there was no particular consideration of countermeasures against cyber-attacks. In our research, an approximate linear model is considered as such a countermeasure[15]. It approximates the dynamics of the plant and serves as a backup to the sensors. However, the output of the model is only approximate so use of the model may cause disturbances to the plant. We assume the objective of the attacker is to damage the core of the reactor and the defender aims to minimize the damage. However, the proposed approach is not limited to these assumptions.

3.4.2 MDP modeling

The elements used for defining system states in the case study and their states are listed in Table 3-1. In this case study, we do not explicitly include the feedwater system because its states can be easily defined by its components.

Table 3-1 Components considered for the system and their states.

Component No.	Component name	Component states
1 (SENS)	Sensors	Normal, used; Normal, not used; Compromised, used; Compromised, not used.
2 (MC)	Main computer	Normal, used; Normal, not used; Compromised, used; Compromised, not used.
3 (BC)	Backup computer	Normal, used; Normal; not used; Compromised, used; Compromised, not used.
4 (CM)	Control mode	Automatic control; Manual control.
5 (AM)	Approximate model	Not used; Used.
6 (RC)	Reactor core	OK; Damaged.

The actions available to the defender include:

1. Use the output of the approximate model instead of the sensors in automatic or manual control.

The approximate model is able to provide information about the status of the plant, with a similar function as the sensors. However, the information is approximate, so use of the model may cause disturbances and even lead to core damage. This model is specifically designed to defend against cyber-attacks, so in this research we assume it cannot be compromised by the attacker. Again, the proposed approach is not limited to this assumption.

2. Use the backup computer to implement the control algorithm.

The backup computer will provide the same function as the main computer.

3. Control the system manually.

Manual control is immune to cyber-attacks, but may cause disturbances to the system.

4. No action.

This option is provided just to include the fact that the defender may not take any action.

The actions available to the attacker include:

1. Compromise sensors

False data can be injected into sensors, which results in false inputs to the computers or the defender. Control signals generated or decisions made based on these false inputs will lead to inadequate feedwater flow to the steam generator which may lead to core damage.

2. Compromise main computer

Compromised computer will lead to inappropriate control signals sent to field actuators, even if the inputs from the sensors or the approximate model are correct.

3. Compromise backup computer

The same reasoning holds for an attack on the main computer.

4. No action

The same reasoning holds as for the defender.

All the possible combinations of the states of the elements will lead to $4 \times 4 \times 4 \times 2 \times 2 \times 2 = 512$ system states. We followed the approach introduced in Section 3.3.1 to identify credible system states. For this system, the initial state is the normal system state, where all the elements listed in Table 3-1 are in the normal state, the sensors provide the measurements of the plant status and the main computer is used for automatic control. Finally, we identified 16 system states. The states are listed in Table 3-2, along with the states of the 6 elements.

Table 3-2 Sixteen system states identified for the case study.

System state No.	State of the 6 elements					
	1 (SENS)	2 (MC)	3 (BC)	4 (CM)	5 (AM)	6 (RC)
1	1	1	2	1	1	1
2	3	1	2	1	1	1
3	1	3	2	1	2	1
4	1	x*	1	1	1	1
5	3	3	2	1	1	1
6	3	x	1	1	1	1
7	x	1	2	1	2	1
8	1	x	3	1	1	1
9	x	3	2	1	2	1
10	3	x	3	1	1	1
11	1	x	x	2	1	1
12	x	x	1	1	2	1
13	3	x	x	2	1	1
14	x	x	3	1	2	1
15	x	x	x	2	2	1
16	x	x	x	x	x	2

* x means that the corresponding element can be in any state.

The transition probabilities are calculated following the approach introduced in Section 3.3.2. For simplicity, we assume that the probability of success for each of the attacker's actions is 0.5. The defender's action is assumed always successful with probability of 1. In any case where the sensors, the main computer, or the backup computer is compromised and still being used, the feedwater system is failed and the core will be damaged with a probability of 0.0001. This probability is obtained based on the event tree shown in Figure 3-3, which was used in WASH-1400[16]. In WASH-1400, the probability that the auxiliary feedwater system fails ranges from 10^{-5} to 10^{-3}, so we use 10^{-4} in this case study. When the approximate model is used or the feedwater system is under manual control, we assume the feedwater system fails with a probability of 0.2. So, the core will be damaged with a probability of 0.2×10^{-4} in this case.

Regarding the rewards, we assume the reward for the defender to take any action is 0, the reward in the case of core damage is on the order of the loss of one-year electricity production, which is about -10^8 dollars. For the attacker, we assume the cost for taking any action to compromise a component is the risk of exposing himself/herself, and we assume the corresponding reward as -10^{-3} dollars. The reward for the attacker in the case of core damage is 10^8 dollars. In the case study, a discount factor $\beta = 0.7$ is used. Methods for more credible quantification of these elements will be investigated in future research.

As an illustration of the resulting MDP model, the transitions from the 11th state are shown in Figure 3-5. The notation on the links are in the format of *<defender action, attacker action>, transition probability, <defender reward, attacker reward>*.

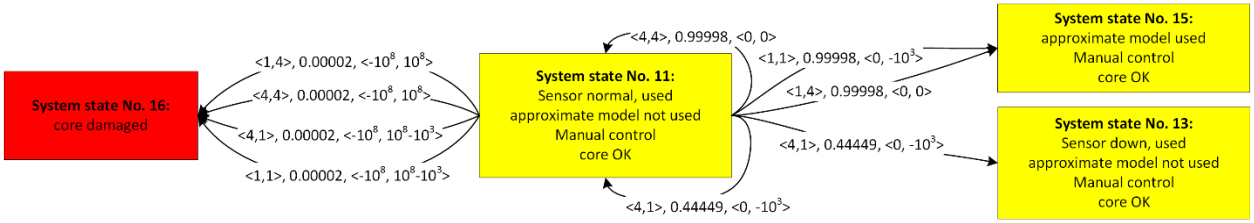


Figure 3-5 Illustration of the resulting MDP model.

3.4.3 Results

The ε-Nash Equilibrium is obtained using the YALMIP toolbox, with $\varepsilon = 4.11 \times 10^{-4}$. This means that the obtained equilibrium is very close to the Nash Equilibrium. The equilibrium policy and the corresponding state values as calculated in Equation 3-1 for the defender and the attacker are presented in Table 3-3. As introduced in Section 3.2, a policy is a probability distribution over the available actions (e.g., (0.58, 0.42, 0) for actions 1, 2, and 4 for the defender at state 1). The reward for the defender is the negative of the loss and in the case study the best case for the defender is there is no loss, so the rewards for the defender in Table 3-3 are always negative.

Table 3-3 Equilibrium policy and state values for the defender and the attacker.

State	Defender's policy				Attacker's policy				Defender's	Attacker's
No.	1	2	3	4	1	2	3	4	value	value
1	0.58	0.42	'x'	0.00	0.47	0.53	'x'	0.00	-8424.73	6757.69
2	1.00	0.00	'x'	0.00	'x'	1.00	'x'	0.00	-19751.51	18321.21
3	0.00	1.00	'x'	0.00	1.00	'x'	'x'	0.00	-20521.25	18635.89
4	0.50	'x'	0.50	0.00	0.50	'x'	0.50	0.00	-9332.46	8014.54
5	1.00	0.00	'x'	0.00	'x'	'x'	'x'	1.00	-22551.00	22254.34
6	1.00	'x'	0.00	0.00	'x'	'x'	1.00	0.00	-20731.26	19519.33
7	'x'	0.66	'x'	0.34	'x'	0.35	'x'	0.65	-9932.43	9126.75
8	0.00	'x'	1.00	0.00	1.00	'x'	'x'	0.00	-20731.26	19519.33
9	'x'	1.00	'x'	0.00	'x'	'x'	'x'	1.00	-17931.80	17507.95
10	0.50	'x'	0.50	0.00	'x'	'x'	'x'	1.00	-23530.72	23530.72
11	0.67	'x'	'x'	0.33	0.50	'x'	'x'	0.50	-11332.28	10726.71
12	'x'	'x'	0.67	0.33	'x'	'x'	0.50	0.50	-11332.28	10726.71
13	1.00	'x'	'x'	0.00	'x'	'x'	'x'	1.00	-19331.53	19331.53
14	'x'	'x'	1.00	0.00	'x'	'x'	'x'	1.00	-19331.53	19331.53
15	'x'	'x'	'x'	1.00	'x'	'x'	'x'	1.00	-13332.09	13332.09
16	'x'	'x'	'x'	1.00	'x'	'x'	'x'	1.00	0.00	0.00

* 'x' means that this action is not available at the state.

3.4.4 Discussions

From Table 3-3, we can see that when the system is in the initial state (state 1), the defender's optimal response strategy is to take actions 1 (i.e., use approximate model), 2 (i.e., use backup computer), and 4 (i.e., no action) with probabilities 0.58, 0.42, and 0, respectively. So this situation favors action 1. However, this does not mean that the defender should always take this action, because this will encourage the attacker to deviate from his/her equilibrium policy. At state 2, where the sensors are compromised and all the other components are normal, the optimal response policy is to take action 1, which is reasonable. At state 5,

where both sensors and the main computer are compromised while being used for control, the optimal policy is to take actions 1 and 2 with probabilities 1 and 0, respectively. This can be explained by the fact that taking action 2 will still expose the backup computer to the attacker in the next time step, while taking action 1 will lead to the use of the approximate model, which is immune to cyber-attacks. The equilibrium policy at other states can be analyzed in the same way. The policy obtained can be used to guide the defender's response to cyber-attacks under various situations.

The state values presented in Table 3-3 also provide us with meaningful insights into the expected long-term cumulative reward for each player if they adopt the equilibrium policy. We can see that at the initial state, the value for the defender is -8424.73, which is not too bad compared with the values at other states. This is because at this state the defender can take a number of actions to mitigate the consequence. The value at state 4 is also high (i.e., the amount of loss is small) because all the components being used are normal and no disturbance is introduced into the system by using either the approximate model or manual control. The value at state 16 is 0 because the core has already been damaged, so there is no additional cost starting from this state. These are the cases where the values are high. There are several states where the values are particularly small (i.e., the amount of loss is large). States 10 and 5 have the smallest values. The results are reasonable because at these two states, both sensors and the backup computer (state 10) or the main computer (state 5) are compromised while being used for control. As a comparison, the values at the states where only one component is compromised while being used (e.g., state 8 vs. state 10, state 3 vs. state 5) are higher.

In summary, the equilibrium policy tells us (i.e., the defender) what response strategy should be taken at each state, and the values tell us how much of reward we can expect starting from each state.

3.5 Conclusion

When cyber-attacks on a nuclear power plant are detected, the operators need to take actions to mitigate the potential consequences. However, this is a challenging task for human operators. This is because the operators need to not only consider the effect of their own actions but also the effect of the attacker's actions. Besides, the operators need to minimize the long-term consequence rather than focusing on short-term benefits. To address this problem, in this research we propose a game-theoretic approach. The interaction between the defender (i.e., the operators) and the attacker is modeled as a two-player, nonzero-sum, stochastic game, which is then represented as a competitive MDP. The MDP model consists of the set of system states, the set of actions for each player, the transitions and the probabilities of transitions between system states, the reward function for each player, and the discount factor. To identify system states and state transitions in the MDP model, an approach based on analyzing the interactions between the two players starting from the initial normal system state is proposed. To obtain credible transition probabilities, an approach based on PRA is proposed.

As a case study, the proposed game-theoretic approach is applied to a simplified digital feedwater control system. The system is defined, the MDP model is developed, and the equilibrium policy and state values for both players are obtained using dynamic programming. An examination of the equilibrium policy and state values found that the results are consistent with our intuition and/or a comprehensive analysis. The obtained policy will provide the defender with guidance in responding to cyber-attacks, in particular in complex situations where there are tradeoffs between the available actions. The state values will provide the defender with an expected long-term cumulative reward starting from each system state.

In future research, the case study will be extended to a more complex nuclear power plant system. The other types of games, e.g., Bayesian games where each player has incomplete information on the other player's available strategies or rewards, will be considered to obtain a more realistic modeling and analysis, which leads to more credible suggestions to the defender in responding to cyber-attacks. Methods for obtaining credible rewards will also be investigated.

4 Semi-Markov Game with Finite Horizon for Attack Response

This work is presented in a detailed manner in the paper titled "Finite-horizon Semi-Markov Game for Time-sensitive Attack Response and Probabilistic Risk Assessment in Nuclear Power Plants", which has been submitted to the journal of Reliability Engineering and System Safety as a copyrighted document.

Digital systems have been more widely deployed in nuclear power plants. Accompanied with the safety and efficiency benefits brought by such systems are the concerns about the physical consequences that can be potentially caused by cyber attacks on these systems. Particularly, since the Stuxnet attack on Iran's uranium enrichment facilities [14], cybersecurity has drawn increasing attention in the nuclear industry. Various measures have been deployed in nuclear power plants to strengthen the cyber-security posture, including measures for preventing cyber attacks on nuclear power plants from occurring [17] and measures for detecting cyber attacks in a timely manner [18]. To further improve the cybersecurity of nuclear power plants, it is also important for the plant operator to adequately respond to a cyber attack once one is detected. However, in responding to a cyber attack the plant operator may be faced with the following challenges. First, a cyber attack can progress very fast and therefore requires the plant operator to respond in a timely manner. Considering the complexity of a nuclear plant, the operator's knowledge limitations, and the high pressure imposed on the operator, it is likely that his/her response will be delayed, a suboptimal decision is made, or even an error is committed. Second, besides considering his or her own actions, the operator also needs to consider the attacker's actions, since both players' action affect the progression of the event and therefore the consequence of the event. Third, instead of taking myopic actions to maximize the short-term benefit, the operator should take actions to maximize the long-term benefit. Therefore, it is important to develop adequate plans to support the plant operator in responding to a cyber attack, the necessity of which has been emphasized by the U.S. Nuclear Regulatory Commission, the Nuclear Energy Institute, the National Institute of Standards and Technology, and the Department of Homeland Security. In addition to developing an optimal cyber-attack response strategy, it is also informative to assess the risk to nuclear power plants due to a cyber attack.

In this research, we develop a finite-horizon semi-Markov game for time-sensitive attack response and probabilistic risk assessment. Compared with existing studies on cyber-attack response and cybersecurity risk assessment, we have made the following contributions in this research.

For cyber-attack response, we have developed a generic model for describing a two-player, finite-horizon, general-sum semi-Markov game, and the method for deriving the optimal cyber-attack response strategy based on game theory and the generic model. In particular, we have modeled the interactions between the defender/operator and the attacker in continuous time, instead of discrete stages as in existing studies. The arbitrary distribution for sojourn time between two states in a transition, which is one of the features of a semi-Markov game, provides more flexibility in modeling the defender-attacker interactions. We have also considered a finite-horizon game, instead of an infinite-horizon game. It is more reasonable to restrict the game to a finite amount of time, since after a certain amount of time the game will end because the nuclear plant may have transitioned to a stable and safe state or adequate countermeasures may have been taken to terminate the game. It is worth noting that a finite-horizon game renders the response of the defender time sensitive or non-stationary. In addition, we have developed a recursive method for identifying system states, and a method by integrating probabilistic risk assessment techniques [19], i.e. event tree and fault tree, for quantifying state transition probabilities. System states and state transition probabilities are two essential elements in the semi-Markov game model. Finally, we have developed a dynamic program to derive the mixed-strategy Nash equilibrium of the finite-horizon semi-Markov game, which informs the optimal cyber-attack response strategy.

For probabilistic risk assessment, besides the advantages provided by the generic model for finite-horizon semi-Markov games, we have developed an analytical algorithm and a Monte Carlo simulation-based algorithm for deriving three risk metrics. These are the probability of the first arrival time at the undesirable states, the probability of arriving at the undesirable states before or at a specified time, and the probability distribution of system states at any time. One promising feature of the two algorithms is that they support a persistent update of the risk metrics as the game progresses from the beginning to the end of the finite horizon.

For demonstration, we have applied the proposed method to a generic digital feedwater control system [6] that can be found in pressurized water reactors. For this system, we consider sensors for monitoring the system state, an approximate model that is used as a backup of the sensors [15], a main computer for system control in normal operation, a backup computer that takes over system control in case of the main computer failure or compromise, and operator manual control in case of both computer failures. Following the proposed method, we developed the semi-Markov game model, which consists of sixteen system states. For cyber-attack response and probabilistic risk assessment, we consider two cases for comparison. In the first case, we assume that the game has just begun and we are at the beginning of the game, and in the second case, we assume that the game has progressed to a time point close to the end of the game. For both cases, we derived the mixed-strategy Nash equilibrium of the game, which informs the optimal action for the defender at each system state and any time point in the finite horizon of 24 hours, and the state value which is the maximum payoff that can be achieved at each system state. We also derived the three risk metrics at each system state and any time point in the finite horizon. Besides, we compared the mixed strategy for the defender obtained based on the proposed method with a baseline response strategy, which is stationary and rule-based and similar to current practices in nuclear power plants. The result verifies that the risk is reduced when the mixed strategy is adopted compared to the one when the baseline strategy is used.

5 Prototype Display Design

This section provides a summary of the accomplishments that fulfill the completion of the milestone to "develop a display prototype" as defined in ***Task 2.4*** of the original proposal for the "Support for Reactor Operators in Case of Cyber-Security Threats" a Department of Energy (DOE) Office of Nuclear Energy (NE) Nuclear Energy University Program (NEUP) funded project. The project identified the need to provide nuclear power plant operators with information about the cyber security state of the plant. To those ends several methods have been combined to provide the plant with information. A small task in the project was to develop the means to provide that information in a concise and clear manner that would provide context for the components or systems that have been identified as impacted. To those ends a functional design has been completed which is able to provide an interface that accomplishes the design goals of an interface that captures the following relevant communication and control attributes of a nuclear power plant:

- System architecture
  - Components
  - Communication interconnections and interfaces
- Configuration information with drill down to details as provided
  - Settings
  - Software/Firmware revisions with whitelist confirmation
  - Supply chain information as provided
- Consume and display results from cyber security tools created as part of this project
  - Machine learning results
  - Malware signature detection
  - Game theoretical results and recommendations
  - Other applications available today or in the future
- Other possible benefits
  - Provide a nexus for in context forensics logs and other information—not part of the current project.

5.1 Background

Given the existence and installation of tools to detect and evaluate the cyber security state of a nuclear power plant (NPP), a need to provide that information in a manner that can be interpreted by personnel at the NPP exists. Tools for this purpose have been researched with the development of visual approaches for various control systems. Visualization tools have previously been developed at Idaho National Laboratory (INL) that have the ability to give geographic context [20]. While geographic context, like that in Figure 5-1, may be important, a logical system interface display provides the context of the interconnection between components.

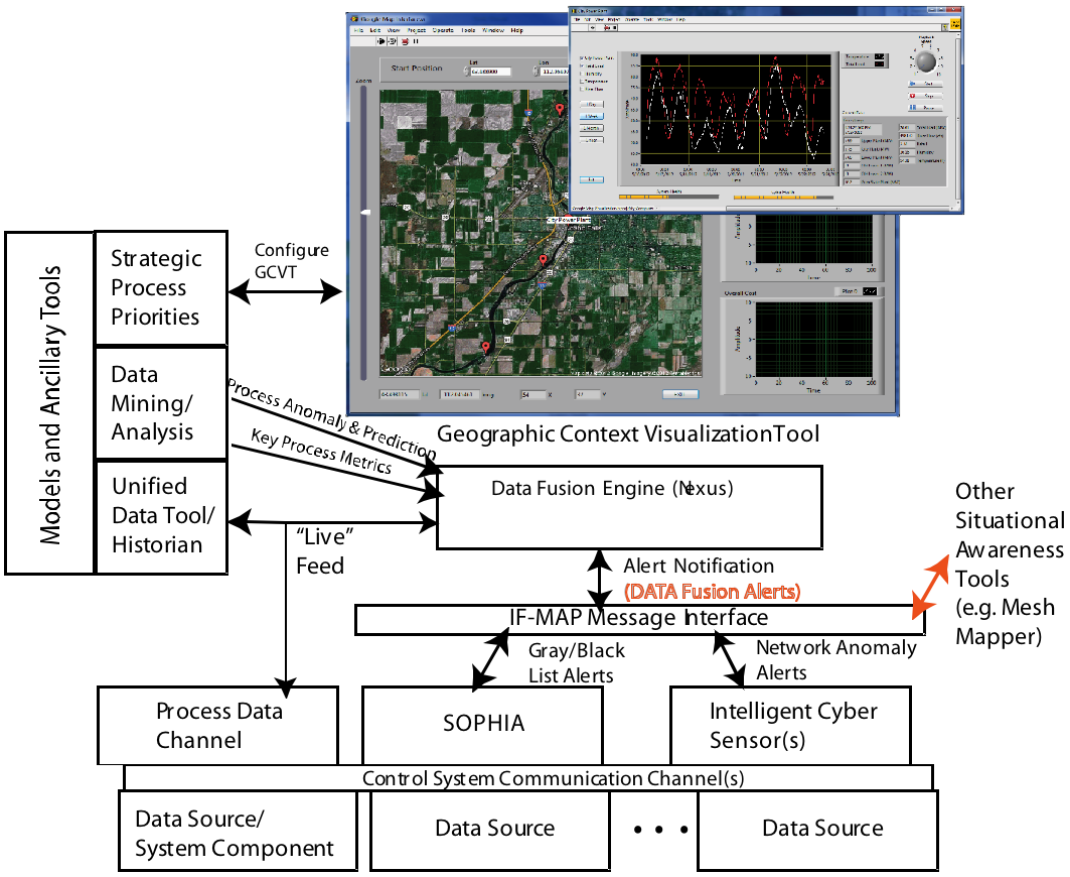


Figure 5-1: An example of a geographic context display [20].

Other displays concepts, created at INL that combine a more illustrative depiction of the state of the system have been developed, as in Figure 5-2. The concept of simply displaying the physical and cyber health of the system along with other pertinent information appeared to be a reasonable starting point for the development of a prototype that would provide the NPP with the information determined by security toolsets. Whereas, in the grid example of Figure 5-2. The relative physical location of the components and topology of the interconnect that determines the physics of the system is important; the current project focuses on the digital assets and the connections between those assets. There would be utility in providing both context in a simplified form; however, the scope and budget for this project required a pragmatic down select to show the logical blocks. The prototype design, therefore, focused on these logical blocks and the interconnections.

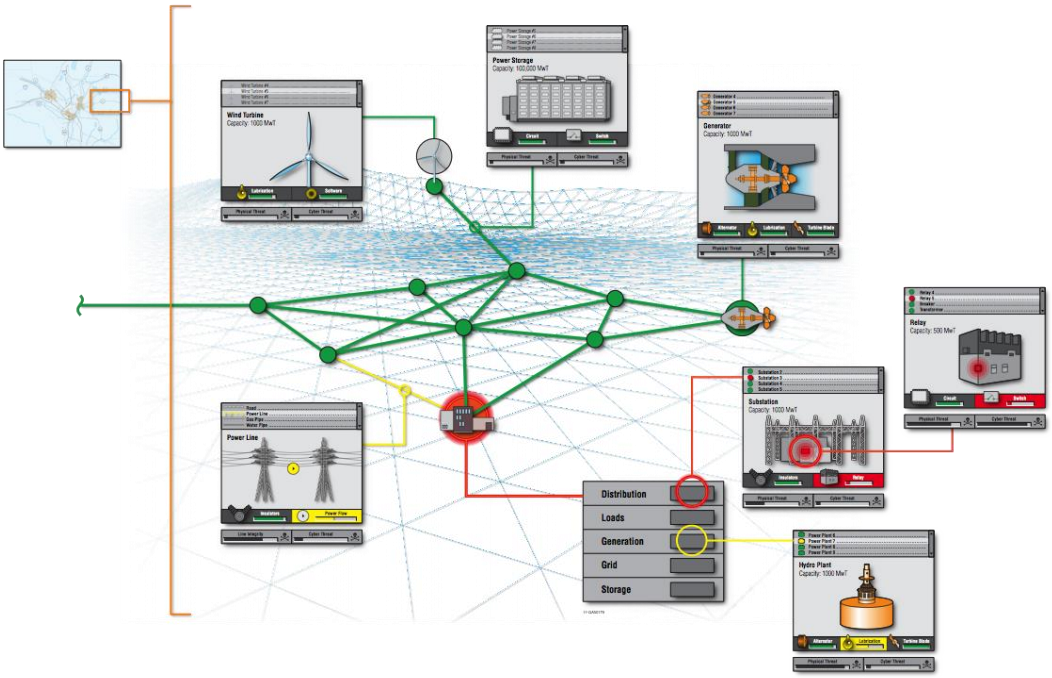


Figure 5-2: Concept of a power and electricity display that contains less specific geography. Credit Rieger/McJunkin INL concept design—unpublished.

5.2 Design

A generic diagram of a NPP control system was used as a beginning basis for the display [21]. The initial concept that was presented to the project team is shown in Figure 5-3. A goal of the software design was to make the display configurable to adapt to represent the current fleet as well for the potential of a modernized NPP or new designs. Through the course of the project discovery of the types of information that would be available for visualization occurred. Since development of the prototype and the cyber security tools were happening in parallel, the display input methods were designed with a generic file up system to which either the tools can adapt or in the future a new standard input can be developed.

The project team agreed upon the following design goals:

- System architecture
  - Components
  - Communication interconnections and interfaces
- Configuration information with drill down to details as provided
  - Settings
  - Software/Firmware revisions with whitelist confirmation
  - Supply chain information as provided
- Consume and display results from cyber security tools created as part of this project
  - Machine learning results
  - Malware signature detection

- Game theoretical results and recommendations
- Other applications available today or in the future

- Other possible benefits
  - Provide a nexus for in context forensics logs and other information—not part of the current project.

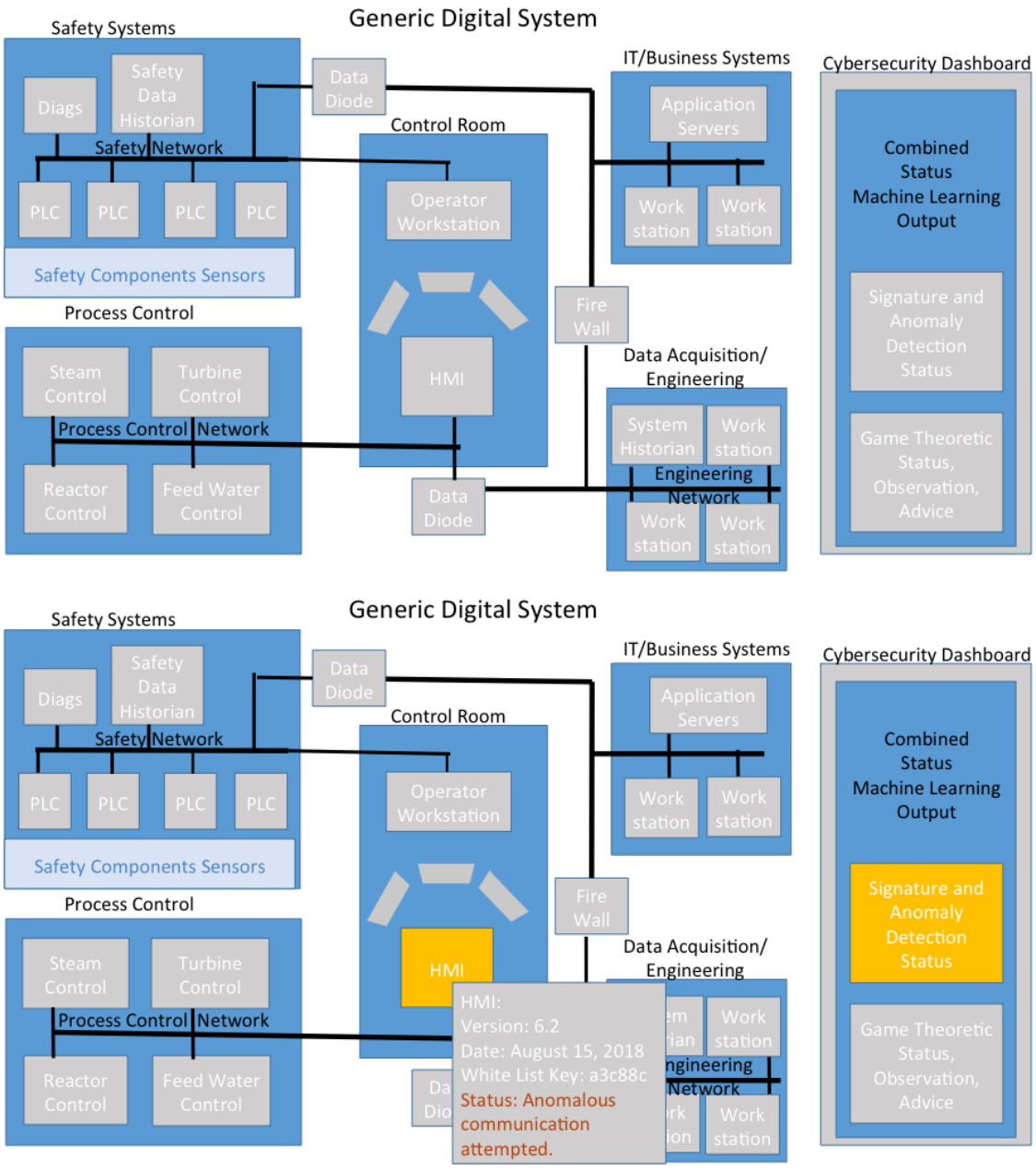


Figure 5-3: Display concept representing generic digital systems in NPP organized by security level groupings showing lines of digital communications (top) and the indicators for the locations of concern with drill downs to available details (bottom).

5.3 Software Description

The prototype is a dynamic graphical user interface (GUI) based visualization of cyber-attack data for a control system. There are three main components to the system. First the GUI is defined through Xtensible Markup Language (XML). Second there is a file based input system that will provide real time updates to the GUI. Third, there is a logging and forensics system that will allow for review of system events. The prototype is part of a system that will allow a user to visualize the events that have been affecting the system. With the XML definitions this singular application is able to be integrated into many different applications. Figure 5-4 shows the visualization tool GUI. The dashboard can be configured to provide additional summary information from cyber security tools including recommendations to the operators, if available.

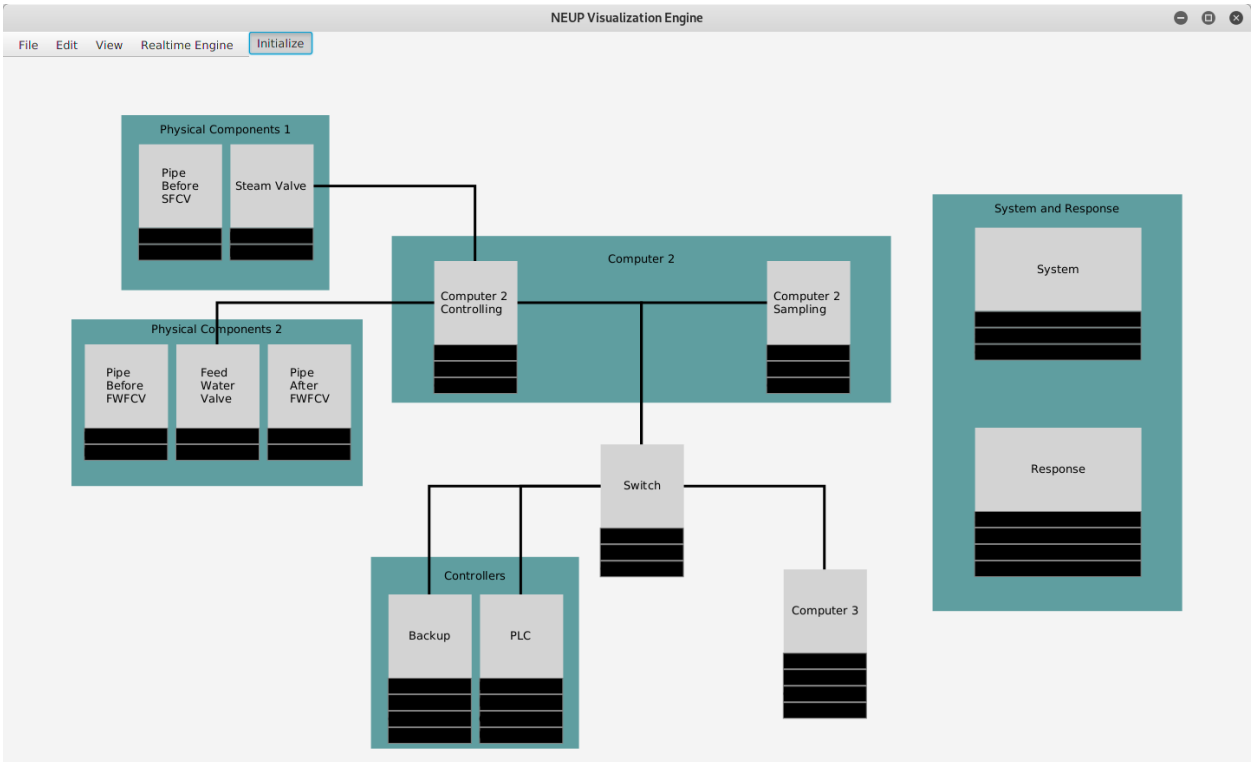
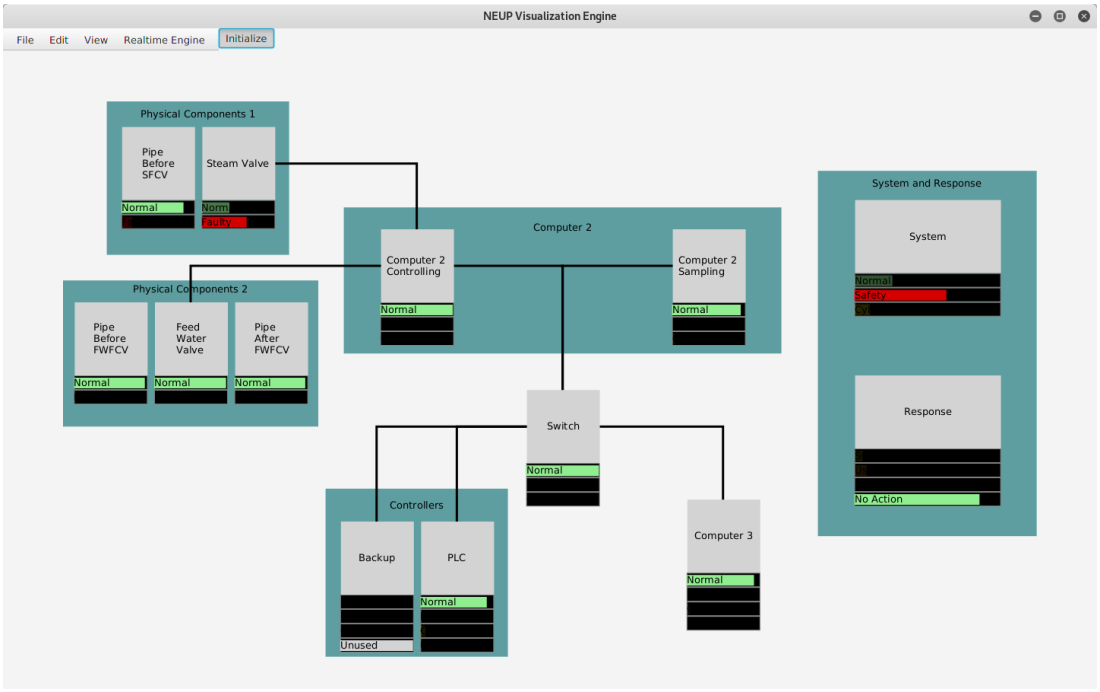
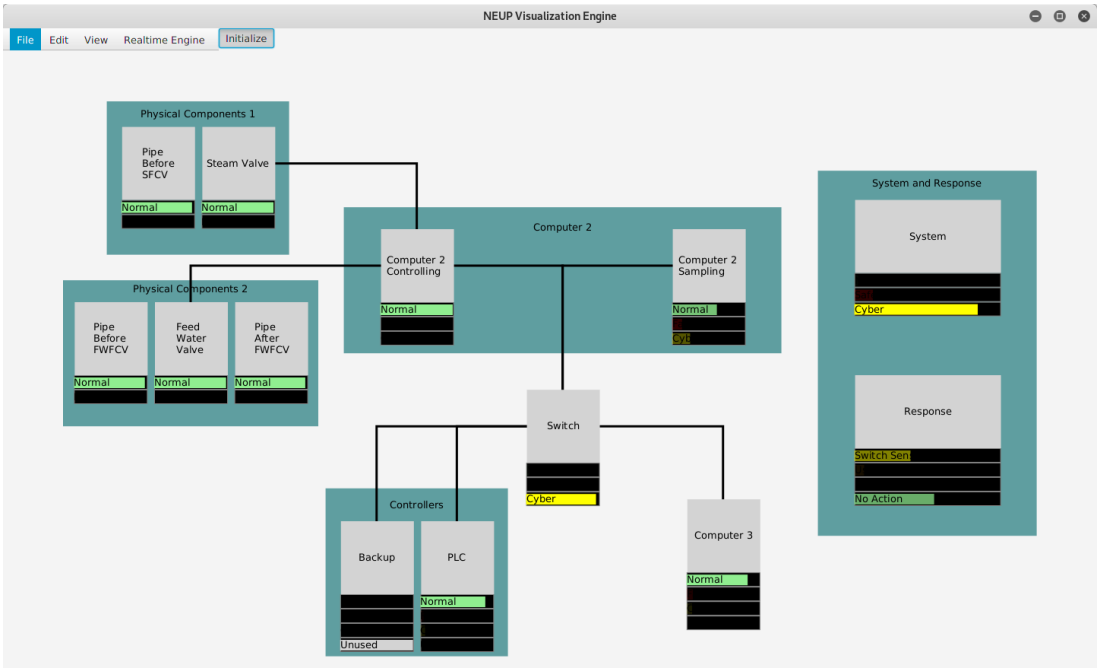


Figure 5-4: Screen Capture of the Display

The GUI system represents the real time information display. There are four main components to the GUI system, control indicators, titles, groupings, and interconnections. Control indicators are the most basic and important element in the system; these represent a physical, or objects. In the case of NPP, these are typically digital assets and communication devices between digital assets. The system will update the colors and text of these objects when more information becomes available. Titles are just simple text to be used for any purpose required. Groupings are rectangles that reference many different control indicators. The groupings are representative of a room or a building but could also represent virtual organizational structures such as virtual networks or security lines. Interconnections represent the connection, physical and/or digital, between different control indicators. The components drawn are all defined through XML. These XML files can be loaded and unloaded as required by the system.



(a) Screen Capture of the Steam Valve Failure Scenario



(b) Screen Capture of the DOS Attack Scenario

Figure 5-5: Screen Capture of the Display at two different times showing modification of the status of the blocks.

The visualization prototype is capable of handling all different cyber and safety probabilistic data. This is represented as different length bars beneath the different controls. This allows for an at-a-glance view of the different possible problems affecting the system. An input standard has been defined to input individual component or overall status and events to the system. Two screen captures at different times, shown in

Figure 5-5, illustrate the dynamics of the display as new updates are provided to an input directory. Both the length of the three bars below each element and the opacity of the bar indicate the relative magnitude of cyber security, safety, and unknown. This is intended to provide a stronger stimulus to the operator when events are more certain. If all is normal the indicators fade into the background.

The operator also may click on any element of interest and see more detailed information in a "drill-down" dialog box, as shown in Figure 5-6. There the description of the component and the numeric status of the cyber security tools are shown with any text that has been provided to communicate details of an event or recommendations to the operators.

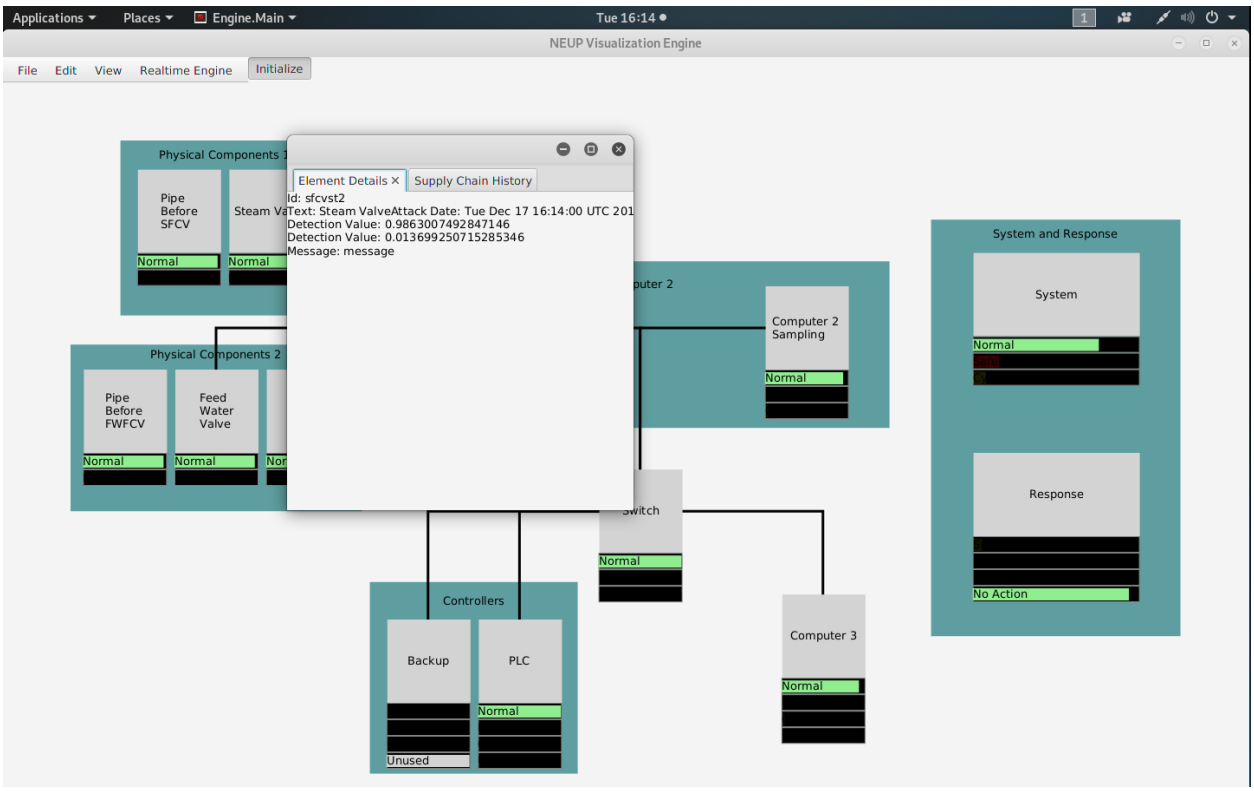


Figure 5-6: Screen Capture of the Display illustrating a drill down to the details of a component and the status of the cyber security tools.

5.4 Conclusion and Path Forward

The prototype display developed for the Support for Reactor Operators in Case of Cyber-Security Threats project provides a basic feature set to enable visualization of the advanced cyber-security tools developed throughout the project in the context of the control and communications system of a NPP. The prototype will be configured to put into use in the remainder of the project as demonstrations and experiments continue to be conducted. The team in the remainder of the period of performance can evaluate the utility of the prototype. At this time, the design team is considering the appropriate manner to make the prototype available beyond the team. Possibilities include: providing the code base as open source and exploring possibilities of commercializing with partners. Key steps remain in determining the ultimate utility of providing this information in the control room or other parts of the NPP facility. Opportunities to evaluate and improve the prototype include study of human factors in design and in use with operators. Advances in the aesthetic and utility will be considered based on the evaluation of the team. Appropriate opportunities to publish the outcomes of the evaluation will be pursued with the project team and others (specifically human factors researchers and/or others with operational expertize). Additionally, the prototype is flexible enough to take input from other sources and could be coupled with other operational tools.

Acknowledgement

This research was performed using funding received from the DOE Office of Nuclear Energy's Nuclear Energy University Program.

References

[1] Zhu B, Sastry S. SCADA-specific Intrusion Detection / Prevention Systems : A Survey and Taxonomy. Proceedings of the 1st Workshop on Secure Control Systems (SCS), 2010.

[2] Giraldo J, Urbina D, Cardenas A, Valente J, Faisal M, Ruths J, et al. A Survey of Physics-Based Attack Detection in Cyber-Physical Systems. ACM Comput Surv 2018;51:76:1–76:36. https://doi.org/10.1145/3203245.

[3] Vasilomanolakis E, Srinivasa S, Cordero CG, Mühlhäuser M. Multi-stage attack detection and signature generation with ICS honeypots. NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium, 2016, p. 1227–32. https://doi.org/10.1109/NOMS.2016.7502992.

[4] Ramotsoela DT, Hancke GP, Abu-Mahfouz AM. Attack detection in water distribution systems using machine learning. Human-Centric Computing and Information Sciences 2019;9:13. https://doi.org/10.1186/s13673-019-0175-8.

[5] Forbes J, Huang T, Kanazawa K, Russell S. The BATmobile: Towards a Bayesian Automated Taxi. Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 1995, p. 1878–1885.

[6] Aldemir T, Stovsky MP, Kirschenbaum J, Mandelli D, Bucci P, Mangan LA, et al. Dynamic Reliability Modeling of Digital Instrumentation and Control Systems for Nuclear Reactor Probabilistic Risk Assessments. Washington, DC: U.S. Nuclear Regulatory Commission; 2007.

[7] Varuttamaseni A, Bari RA, Youngblood R. Construction of a Cyber Attack Model for Nuclear Power Plants, San Francisco, CA, USA: 2017.

[8] Raghavan TES, Filar JA. Algorithms for stochastic games — A survey. ZOR - Methods and Models of Operations Research 1991;35:437–72. https://doi.org/10.1007/BF01415989.

[9] Huang L, Chen J, Zhu Q. A Factored MDP Approach to Optimal Mechanism Design for Resilient Large-Scale Interdependent Critical Infrastructures. 2017 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), 2017, p. 1–6. https://doi.org/10.1109/MSCPES.2017.8064531.

[10] Huang L, Chen J, Zhu Q. A Large-Scale Markov Game Approach to Dynamic Protection of Interdependent Infrastructure Networks. In: Rass S, An B, Kiekintveld C, Fang F, Schauer S, editors. Decision and Game Theory for Security, Springer International Publishing; 2017, p. 357–76.

[11] Lye K, Wing JM. Game strategies in network security. IJIS 2005;4:71–86. https://doi.org/10.1007/s10207-004-0060-x.

[12] Poolsappasit N, Dewri R, Ray I. Dynamic Security Risk Management Using Bayesian Attack Graphs. IEEE Transactions on Dependable and Secure Computing 2012;9:61–74. https://doi.org/10.1109/TDSC.2011.34.

[13] Liu Y, Ning P, Reiter MK. False Data Injection Attacks Against State Estimation in Electric Power Grids. ACM Trans Inf Syst Secur 2011;14:13:1–13:33. https://doi.org/10.1145/1952982.1952995.

[14] Falliere N, Murchu LO, Chien E. W32.Stuxnet Dossier. Symantec Corp.; 2011.

[15] Cárdenas AA, Amin S, Lin Z-S, Huang Y-L, Huang C-Y, Sastry S. Attacks Against Process Control Systems: Risk Assessment, Detection, and Response. Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, New York, NY, USA: ACM; 2011, p. 355–366. https://doi.org/10.1145/1966913.1966959.

[16] Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants. Appendix V: Quantitative Results of Accident Sequences. U.S. Nuclear Regulatory Commission; 1975.

[17] Stouffer K, Pillitteri V, Lightman S, Abrams M, Hahn A. Guide to Industrial Control Systems (ICS) Security. National Institute of Standards and Technology; 2015. https://doi.org/10.6028/NIST.SP.800-82r2.

[18] Scarfone K, Mell P. Guide to Intrusion Detection and Prevention Systems (IDPS). National Institute of Standards and Technology; 2007. https://doi.org/10.6028/NIST.SP.800-94.

[19] Keller W, Modarres M. A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: a tribute to the late Professor Norman Carl Rasmussen. Reliability Engineering & System Safety 2005;89:271–85. https://doi.org/10.1016/j.ress.2004.08.022.
[20] Rueff G, Wheeler B, Vollmer T, McJunkin T, Erbes R. INL Control System Situational Awareness Technology Annual Report 2012. Idaho National Laboratory (INL); 2012. https://doi.org/10.2172/1060987.
[21] Secure Network Design Techniques for Safety System Applications at Nuclear Power Plants A Letter Report to the US NRC. 2010.

List of Publications

Book chapter

SpringerBrief book on "Game and Decision-Theoretic Methods for Mitigating Cyber Threats in Nuclear Power Plants" under preparation.

C. Smidts, X. Diao, P. Vaddi. "Next Generation Network Architecture and Autonomous Cyber Defense," in Michael Haney, Indrajit Ray, Craig Rieger and Quanyan Zhu, eds., Industrial Control Systems Security and Resiliency: Practice and Theory, Springer 2019.

P. Chen, S. Choudhury, L. Rodrigiez, A. Hero and I. Ray. “Towards Cyber-Resiliency Metrics for Action Recommendations Against Lateral Movement Attacks,” in Michael Haney, Indrajit Ray, Craig Rieger and Quanyan Zhu, eds., Industrial Control Systems Security and Resiliency: Practice and Theory, Springer 2019.

L. Huang, Q. Zhu, "Strategic Learning for Active, Adaptive, and Autonomous Cyber Defense", a Springer book chapter of ARO Invitational Workshop on Autonomous Adaptive Cyber Systems.

L. Huang, Q. Zhu, "Deception and Counter-deception Bayesian Game: Adaptive Defense Strategies Against Advanced Persistent Threats for Cyber-physical Systems," a book chapter in the Cyber Deception, E. Al-Shaer, K. Hamlen, J. Wei, and C. Wang (Eds.), Springer, 2018.

Peer-reviewed journal articles

Y. Zhao, L. Huang, C. Smidts, and Q. Zhu. “Finite-horizon Semi-Markov Game for Time-sensitive Attack Response and Probabilistic Risk Assessment in Nuclear Power Plants,” *Revision submitted to* Reliability Engineering and System Safety.

P. Vaddi, M. Pietrykowski, D. Kar, X. Diao, Y. Zhao, T. Mabry, I. Ray, C. Smidts. “Dynamic Bayesian Networks based Abnormal Event Classifier for Nuclear Power Plants in case of Cyber Security Threats,” *will be submitted to* Progress in Nuclear Energy.

Y. Zhao, and C. Smidts. “A control-theoretic approach to detecting and distinguishing replay attacks from other anomalies in nuclear power plants,” *submitted to* Progress in Nuclear Energy.

L. Huang, and Q. Zhu, 2020. “A dynamic games approach to proactive defense strategies against Advanced Persistent Threats in cyber-physical systems,” *Computers & Security*, *89*, p.101660.

Y. Zhao, X. Diao, J. Huang, C. Smidts. "Automated Identification of Causal Relationships in Nuclear Power Plant Event Reports," Nuclear Technology, vol. 205 (issue 8), pp. 1021-1034, 2019.

Q. Zhu and S. Rass, “On Multi-Phase and Multi-Stage Game-Theoretic Modeling of Advanced Persistent Threats,” in IEEE Access, vol. 6, pp. 13958-13971, 2018.

Peer-reviewed conference papers

Y. Zhao, L. Huang, C. Smidts, Q. Zhu, "A Game Theoretic Approach for Responding to Cyber-attacks on Nuclear Power Plants," NPIC&HMIT 2019.

P. Vaddi, C. Smidts, Y. Zhao, X. Diao. "Dynamic Bayesian Networks Based Event-Classification in Support for Reactor Operators in Case of Cyber-Security Threats," NPIC&HMIT 2019.

B. Bezawada, R. Indrajit, and T. Kushagra. "AGBuilder: An AI Tool for Automated Attack Graph Building, Analysis, and Refinement." *IFIP Annual Conference on Data and Applications Security and Privacy*. Springer, Cham, 2019.

L. Huang, and Q. Zhu. "Adaptive honeypot engagement through reinforcement learning of semi-markov decision processes." *International Conference on Decision and Game Theory for Security*. Springer, Cham, 2019.

L. Huang and Q. Zhu, "Distributed and Optimal Resilient Control of Large-Scale Interdependent Critical Infrastructures," Winter Simulation Conference (WSC) 2018, Gothenburg, Sweden, December 9-12, 2018.

L. Huang, Q. Zhu, "Analysis and Computation of Adaptive Defense Strategies against Advanced Persistent Threats for Cyber-physical Systems," in International Conference on Decision and Game Theory for Security, Seattle, WA, USA, October 29 - 31, 2018.

D. Mulamba, A. Amarnath, B. Bezawada, and I. Ray. "A Secure Hash Commitment Approach for Moving Target Defense of Security-critical Services." In *Proceedings of the 5th ACM Workshop on Moving Target Defense*, pp. 59-68. ACM, 2018.

B. Bezawada, M. Bachani, J. Peterson, H. Shirazi, Indrakshi Ray, and Indrajit Ray. "Behavioral Fingerprinting of IoT Devices." In *Proceedings of the 2018 Workshop on Attacks and Solutions in Hardware Security*, pp. 41-50. ACM, 2018.

Y. Zhao, X. Diao, C. Smidts. "Preliminary Study of Automated Analysis of Event Reports from Nuclear Power Plants Based on Natural Language Processing Techniques," The Probabilistic Safety Assessment and Management conference 2018 (PSAM 14), Los Angeles, California, USA, Sep. 16-21, 2018.

Dieudonne Mulamba, Indrajit Ray, and Indrakshi Ray. "On Sybil Classification in Online Social Networks Using Only Structural Features," 16th Annual Conference on Privacy, Security and Trust August 28-30, 2018, Belfast, Northern Ireland, United Kingdom.

L. Huang and Q. Zhu, "Adaptive Strategic Cyber Defense for Advanced Persistent Threats in Critical Infrastructure Networks", Critical Infrastructure Network Security (CINS), ACM SIGMETRICS 2018, Irvine, California, USA on June 18, 2018.

C. Smidts, Y. Zhao, X. Diao, I. Ray, J. Hollern, Q. Zhu, T. McJunkin. "Support for Reactor Operators in Case of Cyber-Security Threats," ANS Winter Meeting & Expo 2017, Washington, D.C., Oct. 29-Nov. 2, 2017.