Model-Based Cyber-Security Framework for Nuclear Power Plant

by

Junjie Guo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Nuclear Engineering and Radiological Sciences)
in The University of Michigan
2020

Doctoral Committee:

Professor John C. Lee, Chair
Professor William R. Martin
Doctor Athi Varuttamaseni, Brookhaven National Laboratory
Professor Nickolas Vlahopoulos

Junjie Guo

gjunjie@umich.edu

ORCID iD: 0000-0003-4447-1333

To my family

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to Professor John C. Lee. Thank you for your consistent patience in mentoring me throughout the six years. I really appreciate and look up to your dedication and professionalism in nuclear engineering. I have learned so much from you in not only the knowledge itself but also your attitude towards scientific research.

Second, I would like to thank my committee members, Prof. Marin, Dr. Varuttamaseni, and Prof. Vlahopoulos. They provided invaluable suggestions to me during my thesis formation and studies.

This thesis is a result of a series of collaborations. I especially thank Dr. Youngblood at Idaho National Laboratory and Dr. Varuttamaseni at Brookhaven National Laboratory. Thank you for all the insightful conversations with your extensive experiences in the area and reviewing this work with helpful suggestion. I am also very thankful to my colleagues Steven Wacker and Rafael Pires Barbosa. I was pleased to have worked with you on the cyber-security project.

This dissertation would not have been possible without funding from the Nuclear Energy University Program (NEUP) award number DE-NE0008783.

I want to thank all my friends everywhere in this world who have brought so much pleasure and joy to my Ph.D. journey. Some of you were here in Ann Arbor with me while the others were thousands of miles away. You are always there whenever I needed a friend.

Finally, I would never get here without the support of my family. I can't say

thank you enough to my parents Yaode Guo and Jiamin Wang, who taught me how to be a good person, provided invaluable guidance, and always supported me.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES viii

LIST OF TABLES x

LIST OF ABBREVIATIONS xi

ABSTRACT xiv

CHAPTER

1. Introduction 1

1.1 Overview of Cyber-Security for Nuclear Power Plant 1
1.1.1 General Background 1
1.1.2 Review of Previous Work 2
1.1.3 Statement of Problem 4
1.2 Thesis Outline 6

2. Instrumentation and Control System in Nulcear Power Plant 8

2.1 AP1000 I&C System 9
2.1.1 Protection and Safety Monitoring System 12
2.1.2 Plant Control System 16
2.2 Generic Pressurized Water Reactor I&C System 19
2.2.1 Steam Generator Water Level Control System . . . 19
2.2.2 Reactor Protection System 24
2.3 Comparison Between GPWR and AP1000 I&C Systems . . . 27

3. Cyber Attack Scenario Generation 31

3.1 Analysis of Attack Scenarios 32

3.1.1 SAPHIRE Code Introduction for Attack Scenario Study 32
3.1.2 Attack Tree Analysis 36
3.1.3 Attack Tree Generation for AP1000 38
3.2 Multi-Path Event Tree Representation 42
3.2.1 Generation of MPET 42
3.2.2 MPET Representation for Low-Order Trip Scenarios 43
3.3 Susceptibility Analysis 45
3.3.1 Methodology Development 45
3.3.2 Parametric Analysis on Attack Possibility 49

4. RELAP5 Simulation for AP1000 Power Plant 51

4.1 RELAP5 Intrduction 52
4.2 AP1000 Reactor Model Development and Simulation 56

5. Application Programming Interface Development 63

5.1 API for the RELAP5 Code 63
5.1.1 Operation of API 64
5.1.2 API Framework Functions 65
5.1.3 API Testing 67
5.1.4 Testing of Control Functions via API 67
5.2 API for the GPWR 69
5.2.1 Data Extraction 69

6. Reduced Order Steam Geneator Model 70

6.1 Model Order Reduction Techniques 71
6.2 Development of SG Reduced Order Model 72
6.3 Validation of ROM with RELAP5 Simulation 75

7. Kalman Filter Algorithm for Detection 78

7.1 Kalman Filter Algorithm 78
7.2 Application for Diagnosis 82

8. Attack Scenario Simulation Using GPWR 87

8.1 GPWR Simulator Introduction 87
8.1.1 GPWR Operation and Features 88
8.2 Simulation of Cyber Attack on Sensor and Controller 94
8.2.1 Direct Low SG Water Level Trip 94
8.2.2 Direct High SG Water Level Trip 95
8.2.3 Controller Feedback SG Water Level Trip 96

9. Cyber-attack Mitigation with Controller feedback 99

9.1 PI Controller 100
9.1.1 Introduction 100
9.1.2 PI Controller Model Development for Water Level Control 101
9.2 Optimal Signal Insertion Mitigation 104
9.2.1 Attack Signal Analysis 104
9.2.2 Optimal Mitigation Formulation 107
9.2.3 Mitigation Approach Demonstration 110

10. Summary, Conclusions, and Future Work 114

10.1 Summary 114
10.2 Future Work 116

BIBLIOGRAPHY 119

LIST OF FIGURES

Figure

1.1	Overview of cyber-security framework	6
2.1	AP1000 I&C architecture	10
2.2	Protection and safety monitoring system	13
2.3	Loss of heat sink functional diagram	15
2.4	Cross section detailing the main SG	21
2.5	Main FW control system control diagram	23
2.6	Reactor protection system layout	25
2.7	Low-low SG level trip circuitry	28
2.8	Low SG level with a feedwater flow mismatch trip	29
3.1	Example of attack tree.	37
3.2	Top-level control logic structure for the PLS.	39
3.3	Attack tree for reactor trip due to low SG level.	40
3.4	Top-level logic for trip generation in PMS.	41
3.5	Example of MPET structure.	43
3.6	MP-ET for attacks on the PLS causing low SG water level end state.	44
3.7	MP-ET for attacks on the PLS causing high SG water level end state.	44
4.1	NSSS nodalization diagram for PWR plant.	57
4.2	PHTS and passive safety systems of AP1000	58
4.3	Reactor vessel nodalization diagram.	60
4.4	SG nodalization diagram.	61
4.5	Temperature distribution along the core.	62
5.1	Framework architecture.	64
5.2	API operations with RELAP5 as the Simulation Module.	66
5.3	Dynamic changes in the RELAP5 simulation.	68
6.1	Water-level comparison between RELAP5 and ROM.	76
6.2	Static water level comparison between RELAP5 and ROM.	77
7.1	Flow of information for the Kalman filter.	82
7.2	Kalman filter demonstration.	84
7.3	Illustration of Kalman filtering for SG parameter space.	85
7.4	2-D projection of SG parameter space.	85
8.1	GPWR flowchart.	89
8.2	Control room overview.	90

8.3 C1 panel example 1. 92
8.4 C1 panel example 2. 93
8.5 Low SG water level trip example. 95
8.6 High SG water level trip example. 96
8.7 Controller Feedback on high SG Water Level Trip. 97
8.8 Controller Feedback on low SG Water Level Trip. 97
9.1 PI controller flowchart. 101
9.2 P controller performance. 102
9.3 PI controller performance. 103
9.4 PI controller performance. 104
9.5 Step function attack signal: Case A. 105
9.6 Step function attack signal: Case B. 106
9.7 Ramp function attack signal: Case C. 107
9.8 Controller input - channel 3 water level sensor. 111
9.9 Controller output feedback. 112
9.10 Different mitigation implementation time. 113

LIST OF TABLES

Table

2.1	Primary AP1000 I&C systems capable of executing reactor trip.	11
2.2	Primary AP1000 reactor trip functions.	14
2.3	Primary PLS control systems and responsibilities.	17
2.4	Sensors in SGWLC system.	20
2.5	GPWR reactor trip functions.	26
3.1	Logic symbols in fault tree.	33
3.2	List of components in PLS.	38
3.3	Attack Possibility For FI trip signal.	47
3.4	Susceptibility for low-order reactor trip attacks, reference case.	48
3.5	Susceptibility for low-order reactor trip attacks, parametric case 1.	49
3.6	Susceptibility for low-order reactor trip attacks, parametric case 2.	50
4.1	AP1000 nodalization details.	59
4.2	AP1000 RELAP5 Simulation.	61
5.1	Setup and operational function descriptions.	66
5.2	Step operation function descriptions.	66
5.3	Breakdown of API run times for the RELAP5 SG model.	67
6.1	Parameters for the reference case.	75
6.2	Comparison of RELAP5 results and ROM calculations.	77

LIST OF ABBREVIATIONS

API Application Programming Interface

CDA Critical Digital Assets

CMA Common Mode Attack

CMT Core Makeup Tank

CPS Cyber-Physical System

CRDMs Control Rod Drive Mechanisms

DAS Diverse Actuation System

DDS Data Display and Processing System

DoS Denial-of-Service

ESF Engineered Safety Feature

ESFAS Engineered Safety Features Actuation System

FCS Feedwater Control System

FDI Fault Detection and Isolation

FI Feedwater Isolation

FW Feedwater

GPWR Generic Pressurized Water Reactor

HAGs Hybrid Attack Graphs

I&C Instrumentation and Control

LOHS Loss of Heat Sink

NIS Nuclear Instrument System

MOR Model Order Reduction

MPET Multi-Path Event Tree

NR Narrow Range

NPPs Nuclear Power Plants

NRC Nuclear Regulatory Commission

OCS Operation and Control Centers System

P&ID Piping and Instrumentation Diagram

PI Proportional Integral

PIC Process Instrumentation Cabinet

PLS Plant Control System

PMS Protection and Safety Monitoring System

PRHR Passive Residual Heat Removal

PWR Pressurized Water Reactor

PZRLC Pressurizer Level Control

PZRPC Pressurizer Pressure Control

RCS Reactor Coolant System

ROM Reduced Order Model

RPS Reactor Protection System

RODCS Rod Control System

SA Single Attack

SDCS Steam Dump Control

SG Steam Generator

SGWLC Steam Generator Water Level Control

SSPS Solid State Protection System

UV Under-Voltage

WR Wide Range

ABSTRACT

A model-based cyber-security framework has been developed to address the new challenges of cyber threats due to the increasing implementation of digital components in the instrumentation and control (I&C) system of modern nuclear power plants. The framework is developed to detect intrusions to pressurized water reactor (PWR) systems that could result in unnecessary reactor shutdown events due to out-of-range water levels of steam generators.

The generation of potential attack scenarios demonstrated a process for identifying the most susceptible attack pathways and components in the I&C system. It starts with identifying two key I&C divisions of the modern AP1000 design related to the reactor trip functions, protection and safety monitoring system, and plant control system. The attack tree analysis is performed on the steam generator (SG) water level control system using the SAPHIRE 8.0.9 code. To quantify the system susceptibility to cyber-attack events, causing reactor trips, we propose sensitivity metrics to identify the low-order sets of components that may be compromised and the degree of perturbations needed for each component. The multi-path event tree (MPET) structures are developed to efficiently and intuitively display a large number of dominant or risk-significant attack scenarios instead of the traditional event trees representing minimal cut sets.

A reduced order model (ROM) has been developed to efficiently represent the SG dynamics and facilitate the detection of potential cyber-attacks. The dynamic ROM is built on the energy balance equation for a single vertical boiling channel

approximating a U-tube steam generator. The ROM provides an essential relationship connecting the reactor power, water level, and feedwater flow rate. An application programming interface (API) for the I&C systems serving as the interface between the RELAP5 system code and the ROM has been developed.

A Kalman filtering based detection method has been proposed, providing optimal tracking of SG water level combining the uncertain simulation results with the observation data subject to statistical fluctuations. An observed plant state with significant deviation from the optimal system projection could then indicate potential intrusions into the system. Finally, a mitigation strategy considering the controller feedback is proposed to avoid the reactor trip due to attack on SG water level sensors. The worst-case attack within this issue space is defined, and the maximum delay time allowed for the mitigation is obtained.

CHAPTER 1

Introduction

1.1 Overview of Cyber-Security for Nuclear Power Plant

1.1.1 General Background

Many of the existing nuclear power plants (NPPs) in the world are nearing or at the midpoint of their design life. At the same time, there have been significant advances in electronics, computers, and networks, which have been incorporated into the currently available digital instrumentation and control (I&C) hardware and software. Advanced digital I&C systems have been used extensively in many other industries. In recent years, digital I&C systems have been developed and installed in new and operating NPPs to address obsolescence issues with analog components. The digital technologies can provide far more functionality than their analog counterparts, but the potential of cyber-attacks has escalated into a severe threat for NPPs.

In fact, several cases of cyber-attacks on critical controlled plants have already been reported [1]. In 2003, the Microsoft SQL Slammer worm infected the computer network server of the Davis-Besse nuclear power plant in Ohio. This infection increased data traffic in the site network, preventing the availability of the safety parameter display system and plant process computers for 5 hours [2]. A representative cyber-attack on a nuclear facility is Stuxnet, which physically destroyed

the centrifuges of Iran's uranium enrichment facility in 2010. Stuxnet is a purpose-built, technologically sophisticated, precisely engineered, and complex piece of cyber weaponry [3]. In Korea, the computer network of Korea Hydro & Nuclear Power (KHNP) was attacked, and the attacker took the design and manual of a NPP, and personal information of the employees in 2014 [4].

In response, the NRC developed Regulatory Guide 5.71 [2], the Cyber Security Program for Nuclear Facilities, which provides an approach satisfying the requirements of 10CFR73.54. The International Atomic Energy Agency (IAEA) provides technical guidelines for addressing computer security issues and implementing cyber-security programs used to protect nuclear facilities against possible malicious cyber-threats. The IAEA Nuclear Security Series (NSS) 17 addresses the establishment and improvement of programs to protect computer systems, networks, and other digital systems critical for the safe and secure operation of facilities and the prevention of theft sabotage and other malicious acts [5].

1.1.2 Review of Previous Work

More and more research has been conducted on the cyber-security issues for nuclear power plants. All the research could be classified into three categories. The first category is cyber-security assessment and analysis. Lassell [6] developed a methodology for assessing the cyber-security robustness and vulnerability of critical digital assets (CDA) to identify cyber vulnerabilities. A relative risk index is proposed for the attacker capability, attacker intent, and attack opportunity. The methodology is also tested at university research reactors. Song [7] introduced a practical approach for the cyber-security risk assessment of NPP I&C systems by considering the characteristics and lifecycle of these systems, and by focusing on detailed matters that can be considered when NPP I&C system designers and equipment suppliers perform cyber-security risk assessment activities. Yadav [8] presents the application of

traditional probabilistic risk assessment methods for performing prevention analysis of cyber-attack scenarios. The method is demonstrated by using a fault-tree based formulation for a cyber-attack scenario in a water flow-loop comprised of flow controllers and pumps, and controlled via manual controls, wired signals and wireless signals that are susceptible to a cyber-attack. The demonstration successfully illustrates the powerful capabilities of fault-tree-based prevention analysis as a robust, scalable, and efficient technique to achieve acceptable system reliability, based on preventing only a subset of cyber-attacks.

The second one is the attack scenario generation and simulation. Nichols [9] presents a methodology by which the attack graphs can be used to generate attack scenarios for a given system in an automated manner. These scenarios can then be applied to simulations of systems or testbeds to determine how a system will respond to that attack, extract key components for counter-measures, or evaluate the quality of counter-measures built into the system. Hill [10] presents a new approach that combines cyber-physical system (CPS) simulations with an existing tool known as hybrid attack graphs (HAGs) to help measure and validate the security of CPS. Three different attacks, sensor miscalibration, spoofed sensor data, and a denial-of-service (DoS) attack, are analyzed. Hill [11] describes the use of a framework for the simulation of a nuclear reactor control system deployed within a honey-net to capture abnormal network traffic and attacks. It can be used not only to explore different control strategies for a particular system but also to investigate cyber-attacks and their potential impact.

Finally, a great deal of research focus on the detection and mitigation methods for cyber-attacks [12]. Hwang [13] divided the fault detection, isolation, and reconfiguration (FDIR) methods into the fault detection and isolation (FDI) step, and the controller reconfiguration step. For FDI, various model-based techniques to generate residuals and statistical techniques of testing the residuals are discussed. This

is followed by various techniques to implemente a reconfigurable control strategy in response to faults. Rosich [14] presents a worst-case attack scenario analysis utilizing optimization techniques and a novel approach for detecting attacks on the controller, considering the second derivative of prediction into the model. The significance of this detecting technique is that no specific controller knowledge is necessary. Hence, the vulnerability of the detector can be reduced since no reconfiguration is required. Chamanbaz [15] proposed a novel approach to co-design controller and attack detector using elements from model predictive control for nonlinear cyber-physical systems affected by false data injection attack. He augments the predictive model controller with an additional constraint requiring the system's future trajectory to remain in some time-invariant neighborhood of a properly designed reference trajectory.

Game theory has been introduced into the nuclear security topic. Kim [16] explores a game-theoretic modeling approach examining how physical protection system functions when attacked by an intelligent adversary in league with an insider at a hypothetical nuclear facility. The game-theoretic approach has the advantage of modeling an adversary who has an intention and complete knowledge of the facility. The game-theoretic model efficiently finds optimal equilibrium in the hypothetical game played by the defender and adversary. Do [17] reviews the existing game-theoretic approaches for cyber-security and privacy issues. The research regarding three main applications, cyber-physical security, communication security, and privacy, is selected to demonstrate the utilization of game theory. The authors present the game models, features, and solutions of the selected works and describe the advantages and limitations from design to implementation of the defense mechanisms.

1.1.3 Statement of Problem

With the investigation of the cyber-attack cases and the review of the research papers, we decided to concentrate on potential cyber-attacks resulting in unscheduled

reactor trips. Cyber-attack in a nuclear power plant is not likely to cause very severe damage to the power plant, such as the large break loss of coolant accident we usually study in NPP risk analysis. Reactor trips are part of the reactor protection system, which protects the reactor from design basis accidents. However, it provides a weak point for cyber-attacks to trip a reactor through a number of available pathways. Unscheduled reactor trips could make the power plant unavailable for a period of time and increase the operation and maintenance cost of a nuclear power plant. If we can avoid unnecessary reactor trips, the nuclear power plant economy can be improved.

The AP1000 and Generic Pressurized Water Reactor (GPWR) designs have been selected as the reactor models for this study. Since most of the NPPs built in the U.S. are PWR plant, this type of reactor should be our focus for the cyber-security study. The GPWR is a generic model that can represent most of the PWRs currently in operation. The AP1000 reactor is a generation III+ reactor that includes a large scale use of digital I&C components, and represents the most recent Westinghouse design certified by the U.S. Nuclear Regulatory Commission (NRC).

To narrow down the scope of the investigation, we decided to place our emphasis on attack scenarios that could result in an automated reactor trip initiated by SG levels being out-of-range. Among all 11 reactor trip signals in the AP1000 design, there are four directly related to the SG. An SG is also the connection between the primary and secondary loop, maintaining effective heat removal for nuclear power plants. Hence, SG level-focused attack scenarios for SGs form a distinct point of focus for our cyber-security study.

The overall framework comprising (a) attack scenario generation, (b) system modeling, (3) attack detection, and (d) mitigation action that we have developed is summarized in Figure. 1.1. Various components in the framework will be described in detail in subsequent chapters.

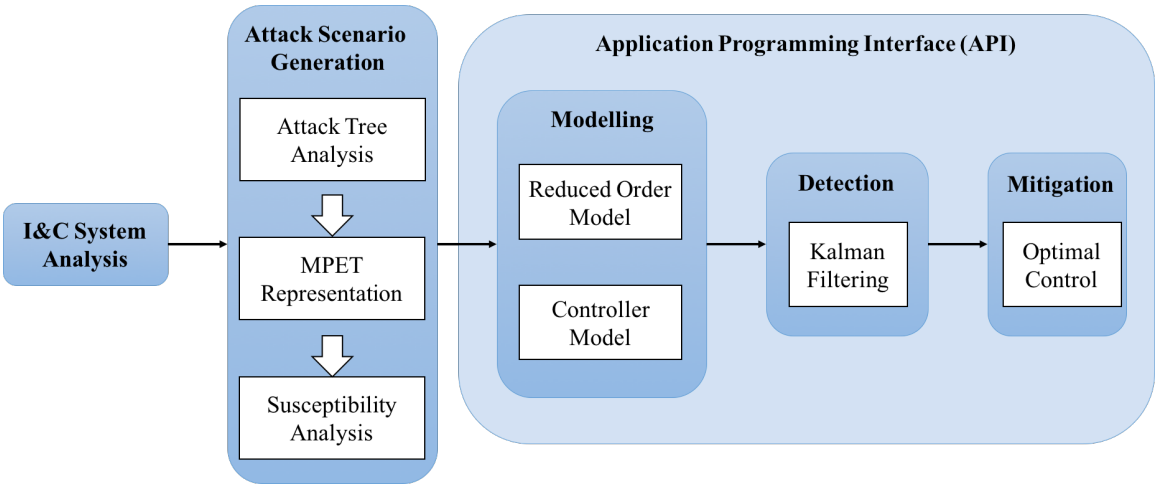

Figure 1.1: Overview of cyber-security framework

1.2 Thesis Outline

Chapter 2 starts with a review of the I&C system in a nuclear power plant, especially for the AP1000 [18] and GPWR [19] reactor models. The two main systems related to the unscheduled reactor trip resulting from cyber-attacks are plant control system and plant protection and monitoring system. The characteristics and functions of these two systems are introduced and discussed.

Chapter 3 deals with the cyber-attack scenarios generation for the SG water level control system. The attack tree analysis has been completed through the SAPHIRE code [20]. The minimal cut sets are obtained and represented in the MPET structure. A new susceptibility analysis is proposed, which can simplify the possible attack scenarios by identifying low-order sets. A parametric study on different attack possibility combinations is discussed.

Chapter 4 introduces the RELAP5 thermal-hydraulic simulation software [21], developed for the NRC for simulation of hydraulic and thermal transients in both nuclear and non-nuclear systems. Details on the RELAP5 software are discussed, together with an AP1000 reactor model.

Chapter 5 discusses the application programming interface (API) developed for the

whole cyber-security framework, which served as the platform for modeling, detection, and mitigation of potential cyber-attack incidents. Details on the operation and function of the API are explained. Test cases are performed, which validates the functionality of the API.

Chapter 6 introduces a reduced-order model for the steam generator in the nuclear power plant. It provides a critical relationship between reactor power, SG water level, and feedwater flow rate. The derivation of the model based on the energy balance equation is presented. Several cases comparing with the results from RELAP5 simulation are also performed validating the correctness and accuracy of the model.

Chapter 7 presents a diagnosis approach based on the Kalman filter algorithm, which takes both the measurement fluctuation and simulation uncertainty into consideration, providing an optimal estimate. The diagnosis approach can be used to detect significant deviation on key parameters resulting from the cyber-attack.

Chapter 8 analyzes various attack scenarios simulated using the GPWR Simulator. The function, feature, and operation of GPWR are introduced. Various attack scenarios on the steam generator control system are simulated, which provide a better understanding on how the system responds to specific attack scenarios.

Chapter 9 provides a new optimal mitigation approach to avoid reactor trips considering the controller feedback in the SG water level control system. The mitigation method inserts a signal in the opposite direction into the system, which can counteract the attack signal's influence. The maximum delay time allowed for the attack detection and mitigation action is presented.

The final chapter summarizes the whole cyber-attack framework and proposes various possible directions for future work.

CHAPTER 2

Instrumentation and Control System in Nulcear Power Plant

It is important to study the I&C system of a typical PWR system and model the detail for analysis of cyber-attack events. In probabilistic risk analysis, random failure studies are performed focusing on the physical component and system in the nuclear power plant, such as pump failure or pipe break. However, the target of cyber-attacks will likely be the I&C system, which has digital devices or connections to the network. With the goal of avoiding unscheduled reactor trips due to potential intrusions into the SG operation, this chapter addresses the I&C system involved in SG operation and reactor trip events. The study of the I&C system could not only identify the plant parameters of interest that can lead to a specific reactor trip function but also recognize the components associated with SG operation and reactor trip. Two reactor designs have been studied, AP1000 and GPWR, representing the advanced Generation III+ PWR design and the current Generation II PWR design, respectively. The investigation of the I&C system in this chapter provides us the knowledge of the I&C system functions in the nuclear power plant, which has laid a solid foundation for the cyber-attack scenario generation later.

2.1 AP1000 I&C System

AP1000 [22, 18] is two-loop PWR nuclear power plant designed by Westinghouse Electric Company. The AP1000 plant builds and improves upon the established technology of major components used in current Westinghouse-designed plants. These components include:

- Steam generators
- Digital instrumentation and controls
- Fuel
- Pressurizers
- Reactor vessels

The established design of the AP1000 plant offers three distinct advantages over other designs: unequaled safety, economic competitiveness and improved and more efficient operations. It is recognized as one of the most advanced nuclear power plants design currently operating around the world. Hence, the design of the I&C system for AP1000 can represent the newest in class, which is a good demonstration for modern PWR I&C systems.

The I&C systems presented in this chapter provide protection against unsafe reactor operation during steady-state and transient power operations. They initiate selected protective functions to mitigate the consequences of design basis events. The AP1000 I&C architecture is illustrated in Figure 2.1 [23]. The figure shows two major sections separated by the real-time data network.

The lower portion of the figure includes plant protection, control, and monitoring functions. It performs the reactor trip functions, the engineered safety features (ESF) actuation functions, and the qualified data processing (QDPS) functions. The I&C

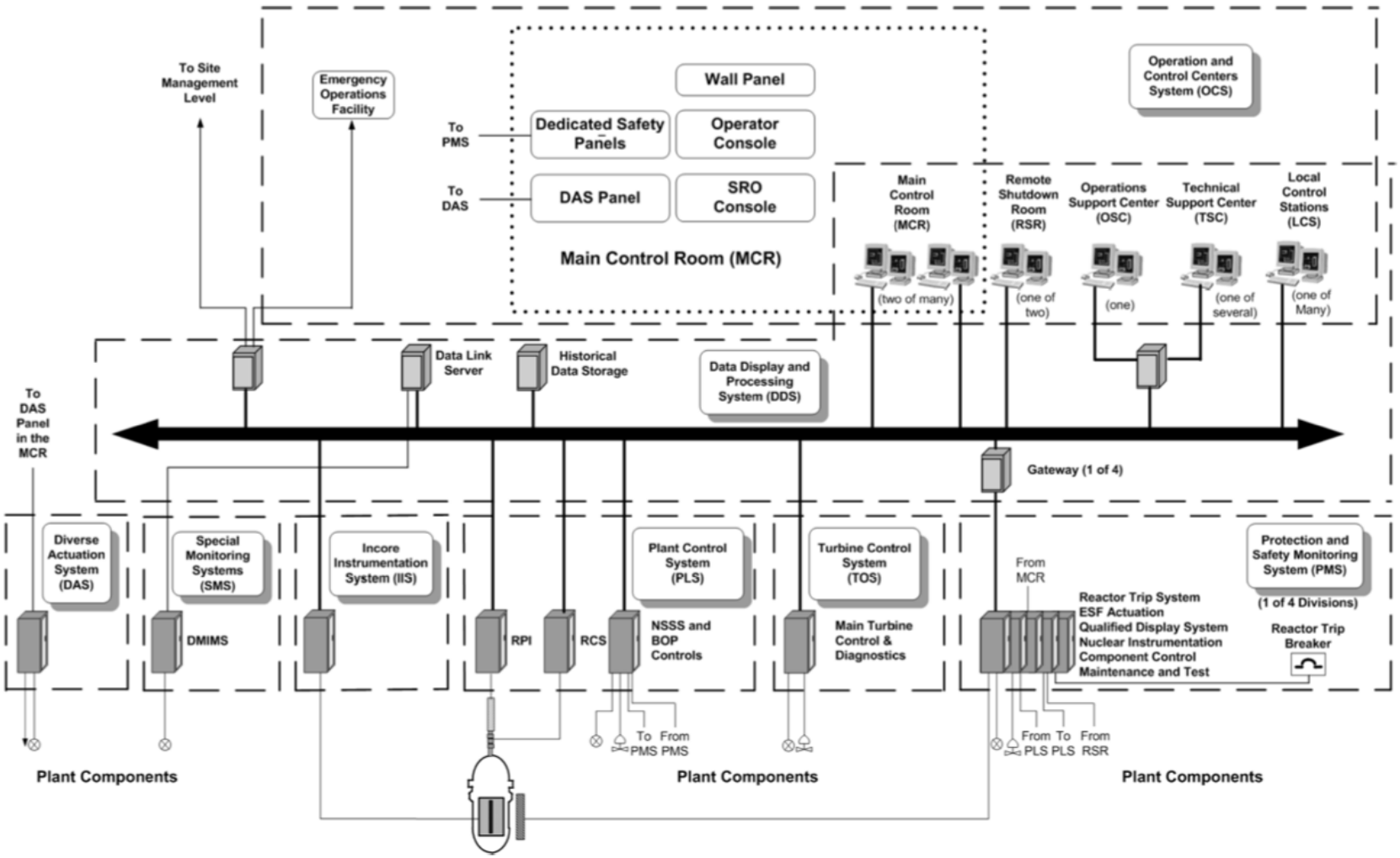

Figure 2.1: AP1000 I&C architecture. *Source:* [18].

equipment performing reactor trip and ESF actuation functions, their related sensors, and the reactor trip switchgear are, for the most part, fourfold redundant. The non-safety-related real-time data network, which horizontally divides Figure 2.1, is a high-speed, redundant communications network that links systems of importance to the operator. The upper portion of the figure depicts the control rooms and data display and processing system.

With many systems of the AP1000 I&C architecture [23] focusing on preventing unscheduled reactor trips, we note that five systems can generate a reactor trip signal. Table 2.1 describes the systems and the role that they play in reactor control and safety.

Table 2.1: Primary AP1000 I&C systems capable of executing reactor trip.

System	**Description**
Protection and Safety Monitoring System (PMS)	Initiates reactor trip and actuation of ESFs in response to plant conditions by monitoring process instrumentation and provides safety-related displays.
Plant Control System (PLS)	Provides automated and manual plant control of nonsafety-related plant components during normal and emergency plant operations.
Diverse Actuation System (DAS)	Performs secondary nonsafety-related operations separate and independent to the PMS, initiates reactor trips, actuates select ESFs, provides secondary interfaces for manual reactor trip, and displays for select plant parameters.
Data Display and Processing System (DDS)	Provides nonsafety-related alarms and displays, performs analysis of plant data, logging of plant data including storage and retrieval, and operational support for plant personnel.
Operation and Control Centers System (OCS)	Developed and implemented based upon a human factors engineering (HFE) program. Includes the main control room, remote shutdown workstation, the local control stations, and the associated workstations for each of these centers.

These five systems can contribute to the generation of the majority of trip signals for the AP1000 plants. Our research focuses on the attacks that may result in automatic reactor trip initiation. The Operation and Control Centers System (OCS), the Diverse Actuation System (DAS) and the Data Display and Processing System (DDS) usually require some form of human interaction for a trip to occur. These systems are therefore not the focus of this study, which leaves the Protection and Safety Monitoring System (PMS) and Plant Control System (PLS) as the primary I&C systems capable of actuating the reactor trip.

2.1.1 Protection and Safety Monitoring System

The protection and safety monitoring system provides the detection of off-nominal conditions and the actuation of appropriate safety-related functions necessary to achieve and maintain the plant in a safe shutdown condition. The PMS has four reactor trip and engineered safety feature (ESF) actuation divisions, and two divisions of safety-related post-accident parameter displays. The functional arrangement of the PMS is depicted in Figure 2.2.

The reactor trip system keeps the reactor within the safe region by shutting down the reactor whenever safety limits are approached. The reactor trip is a protective function performed by the protection and safety monitoring system when it anticipates a parameter's approach to its safety limit. Reactor shutdown occurs when electrical power is removed from the rod drive mechanism coils, allowing the rods to fall by gravity into the reactor core. The plant protection subsystems maintain surveillance of key process variables directly related to equipment mechanical limitations (such as pressure) and variables that directly affect the reactor's heat transfer capability (such as flow and temperature).

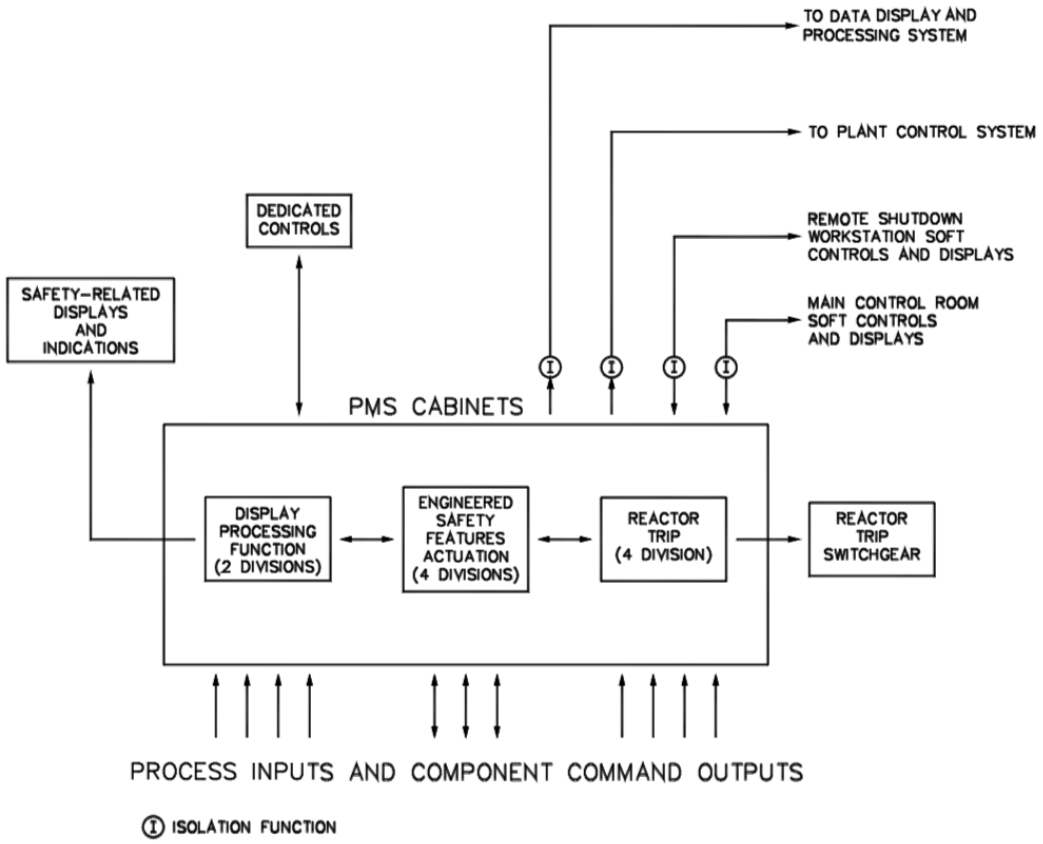

Figure 2.2: Protection and safety monitoring system. *Source:* [23].

Four redundant measurements, using four separate sensors, are provided for each variable used for the reactor trip. Analog signals are converted to digital form by analog-to-digital converters within the PMS. Signal conditioning is applied to selected inputs following the conversion to digital form. Following necessary calculations and processing, the measurements are compared against the applicable setpoint for that variable. A partial trip signal for a parameter is generated if one channel indicates measurement exceeding its predetermined or calculated limit. The processing of variables for the reactor trip is identical in each of the four redundant divisions of the protection system. Each division sends its partial trip status to each of the other three divisions over isolated data links. Each division is capable of generating a reactor trip signal if two or more redundant channels of a single variable are in a

partial trip state.

The reactor trip signal from each of the four divisions of the PMS is sent to the corresponding reactor trip switchgear breakers. Each of the four reactor trip actuation divisions consists of two reactor trip circuit breakers. The reactor is tripped when two or more actuation divisions output a reactor trip signal. This automatic trip demand initiates the following two actions. It de-energizes the under-voltage trip attachments on the reactor trip breakers, and it energizes the shunt trip devices on the reactor trip breakers. Either action causes the breakers to trip. Opening the appropriate trip breakers removes power to the rod drive mechanism coils, allowing the rods to fall into the core. This rapid negative reactivity insertion causes the reactor to shut down.

The main reactor trip functions in the AP1000 can be grouped around certain operational protections. These groupings form eleven main reactor trip functions of interest, summarized in Table 2.2.

Table 2.2: Primary AP1000 reactor trip functions.

Nuclear startup	Automatic depressurization system actuation
Nuclear overpower	Passive residual heat removal actuation
Core heat removal	Core makeup tank injection
Primary overpressure	Safeguards actuation
Loss of heat sink	Manual reactor trip
Feedwater isolation	–

There are four reactor trips related to the steam generator water level among these eleven trip functions:

- Loss of Heat Sink Trip
- Feedwater Isolation Trip
- Core Makeup Tank Injection Trip
- Reactor Trip on Passive Residual Heat Removal System Actuation

Loss of heat sink trip protects the reactor from loss of heat sink in the event of a loss of feedwater to the steam generators. The reactor is tripped when two out of the four water level sensors in any steam generator produce signals below the setpoint value as shown in Figure. 2.3.

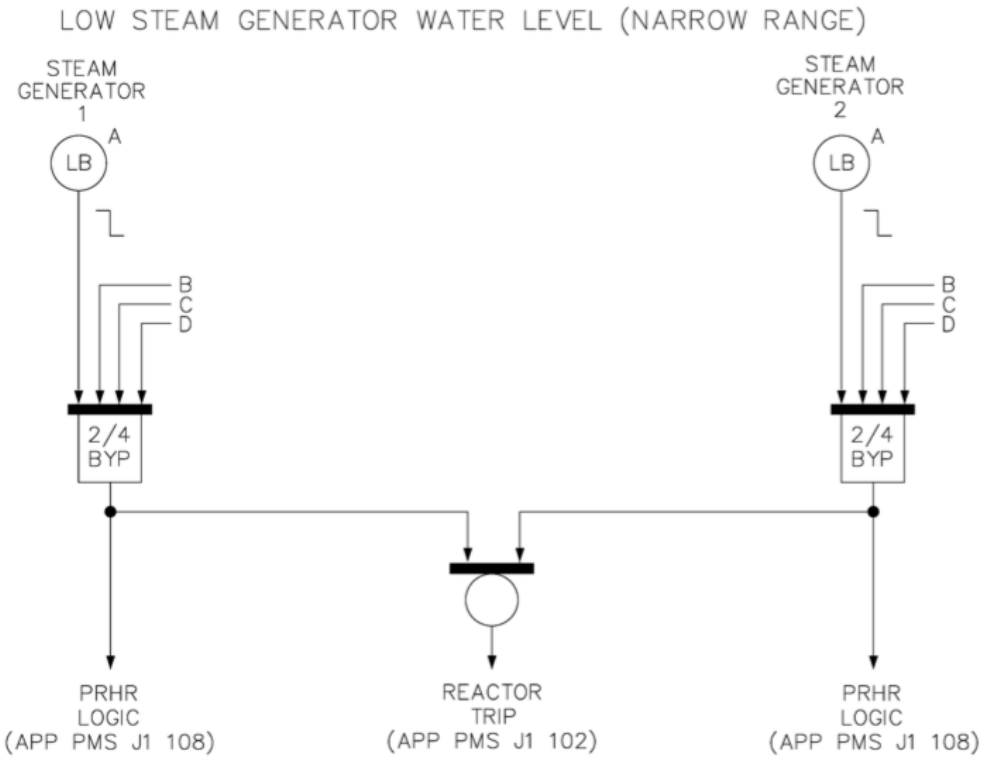

Figure 2.3: Loss of heat sink functional diagram. *Source:* [18].

A feedwater isolation trip is an anticipatory trip based on the expectation that a reactor trip would occur after steam generator feedwater is isolated. The trip is initiated if any steam generator water level exceeds the High-2 setpoint. Core Makeup Tank injection trip is initiated if core makeup injection occurs either automatically or manually. Low wide-range steam generator level coincident with high hot-leg temperature is one condition that can generate signals to align the core makeup tanks for injection, which involves the steam generator water level. The passive residual heat removal (PRHR) system actuation trip is initiated when the PRHR

system's discharge valves come off their fully shut seat, allowing flow through the PRHR heat exchanger. Low narrow-range steam generator level coincident with low startup feedwater flow is one of the conditions to generate a signal to align the PRHR heat exchanger to passively remove core heat.

2.1.2 Plant Control System

The PLS provides the functions necessary for the plant's normal operation from cold shutdown through full power, by controlling non-safety-related components in the plant operated from the main control room or remote shutdown workstation. The PLS contains non-safety-related control and instrumentation equipment to change reactor power, control pressurizer pressure and level, control feedwater flow, and perform other plant functions associated with power generation. The PLS accomplishes these functions by using the following: rod control, pressurizer pressure and level control, steam generator water level control, steam dump (turbine bypass) control, and rapid power reduction. The PLS provides automatic regulation of reactor and other key system parameters in response to changes in operating limits (load changes) and acts to maximize margins to plant safety limits and maximize the plant transient performance.

The PLS includes the equipment from the process sensor input circuitry to the modulating and non modulating control outputs and the digital signals to other plant systems. Modulating control devices include valve positioners, pump speed controllers, and the control rod equipment. Non-modulating devices include motor starters for motor-operated valves and pumps, breakers for heaters, and solenoids for actuation of air-operated valves. The PLS cabinets contain the process sensor inputs and the modulating and non-modulating outputs, and also includes equipment to monitor and control the control rods. The functions performed by the plant control system are listed in Table 2.3 [18].

Table 2.3: Primary PLS control systems and responsibilities.

System	Functional Description
Reactor Power Control System	Coordinates the responses of various reactivity control mechanisms, load follow operations, load regulation/frequency control, and axial nuclear power distribution control.
Rod Control System	Maintains power and reactor coolant temperature, without challenges to protection systems, during normal operating transients in conjunction with the reactor power control system.
Pressurizer Pressure Control	Maintains or restores the pressurizer pressure to the nominal operating value following normal transients. Reacts to avoid challenges to the protection systems during transients.
Pressurizer Water Level Control	Establishes, and maintains pressurizer water level to its operating region as a function of reactor coolant system temperature to minimize charging and letdown requirements. No challenges to the protection system result from normal operational transients.
Feedwater Control System	Maintains the steam generator water level at a predetermined setpoint during steady-state operation. It also maintains the water level within operating limits during normal transient operation. The feedwater control system restores normal water level following a unit trip.
Steam Dump Control	Reacts to prevent a reactor trip following a sudden loss of electrical load. Removes stored energy and residual heat following a trip so that the plant can be brought to equilibrium no-load conditions without actuation of the steam generator safety valves. Also provides for maintaining the plant at no- or low-load conditions to facilitate controlled cooldown.
Rapid Power Reduction	For large, rapid load rejections (turbine trip or grid disconnect from 50-percent power or greater) a rapid nuclear power cutback is implemented. Results in a reduction of thermal power to a level that can be handled by the steam dump system.
Defense-In-Depth Control	Provides control of systems performing defense-in-depth functions.

The AP1000 control systems share a common hardware design and implementation philosophy. They are also functionally integrated to enhance responsiveness during plant transients. Specific design requirements have been imposed that limit the impact of individual equipment failures. The function of the plant control system is performed by several major system assemblies, including distributed controllers, signal selector algorithms, and real-time data network.

The distributed controller performs system-level and component-level control calculations, provides the capability for an operator interface to the controlled components, transmits control signals to discrete, modulating, and networked interfaced control components, and provides plant status and plant parameter information to the real-time data network. The distributed controllers receive process inputs and implement the system-level logic and control algorithms appropriate for the plant

operating mode. Control functions are distributed across multiple distributed controllers so that single failures within a controller do not degrade the performance of control functions performed by other controllers. The major control functions implemented in different distributed controllers include reactor power control, feedwater control, pressurizer control, and turbine control.

Signal selector algorithms provide the plant control system with the ability to obtain inputs from the PMS. The signal selector algorithms select those PMS signals representing the actual status of the plant and reject erroneous signals. Therefore, the control system does not cause an unsafe control action to occur even if one of four redundant protection channels is degraded by random failure simultaneous with another of the four channels bypassed for test or maintenance. Each signal selector algorithm receives data from each of the redundant divisions of the PMS. The data are received from each division through an isolation device and the signal selector algorithms provide validated process values to the PLS. They also provide the validation status, the average of the valid process values, the number of valid process values, and alarms. For the logic values received from the PMS, such as permissives, two-out-of-four (2/4) voting is used to provide a valid logic value to the PLS. The signal selection algorithm is executed in the PLS, and the results are not available to PMS or DAS. Therefore, PMS and DAS performance, controls, and displays are independent of the signal selector algorithm.

The real-time data network is a redundant data highway supporting both periodic and aperiodic data transfers of non-safety-related signals and data. Periodic transfers consist of process data that are broadcast over the network at fixed intervals and are available to all destinations. Aperiodic data transfer is generally used for messages or file transfers. The real-time data network provides communications among the distributed controllers, the PMS gateways, the incore instrumentation, and the special monitoring system.

2.2 Generic Pressurized Water Reactor I&C System

The Generic Pressurized Water Reactor (GPWR) [19] is the generic reactor model in the GPWR nuclear plant simulator newly installed at the University of Michigan. It represents a three-loop Westinghouse PWR system. Similar to the AP1000 design, there are two systems mainly related to our cyber-attack study: the nuclear control I&C system and the reactor protection system (RPS).

There are five major control systems relating to the reactor and steam generation. These control systems are (1) rod control system (RODCS), (2) pressurizer level control (PZRLC), (3)pressurizer pressure control (PZRPC), (4) steam generator water level control (SGWLC), and (5) steam dump control (SDCS). Each system uses a predetermined program to perform its control function. For a control system to operate, it must rely on inputs for various parameters that have been or may be affected by the control process.The SGWLC system, involving both primary and secondary components, is introduced next with our three-category classification method on I&C system components.

2.2.1 Steam Generator Water Level Control System

The SGWLC system provides automatic control of the steam generator water level over the entire range of power operation, which is a focus of our study. Stable and reliable control is achieved through the use of two subsystems: the Feedwater (FW) Control System, which controls the main feedwater valves during power operation, and the Feedwater Bypass Control System, which controls the FW bypass valves at low power levels. The SGWLC is designed to adjust the FW flow to maintain a programmed level of 57 % in the SGs during normal plant operation and plant transients. The SGWLC system components are categorized as either sensor, control logic processor, or actuation system.

2.2.1.1 Sensors

In order for the SGs to maintain the proper heat sink, the secondary-side feedwater to the SGs is automatically adjusted through continuous control based on three physical parameters: (a) narrow range (NR) and wide range (WR) water level, (b) FW flow rate, and (c) steam flow rate. For a Generation II Westinghouse PWR plant which is represented in the GPWR, Table 2.4 lists the specific sensors, and their physical location illustrated Figure 2.4 [24]. The locations for the NR and WR gauges and steam flow gauges are highlighted together with the feedwater inlet nozzle and steam separation equipment.

Among the sensors, the level sensor is the most important one. Five water level sensors are used on each SG. Four sensors are protection grade, narrow-range level transmitters. The four narrow-range level channels are used for protection, control, and indication.

Table 2.4: Sensors in SGWLC system.

Sensor	**Description**	**Location**
NR Level	Differential pressure cell	Upper 240 inch of SG head
WL Level	Differential pressure cell	Total 610 inch from tube to mid-deck
FW Flow	Venturi	FW inlet line
Steam Flow	Venturi	SG outlet steam flow restrictor

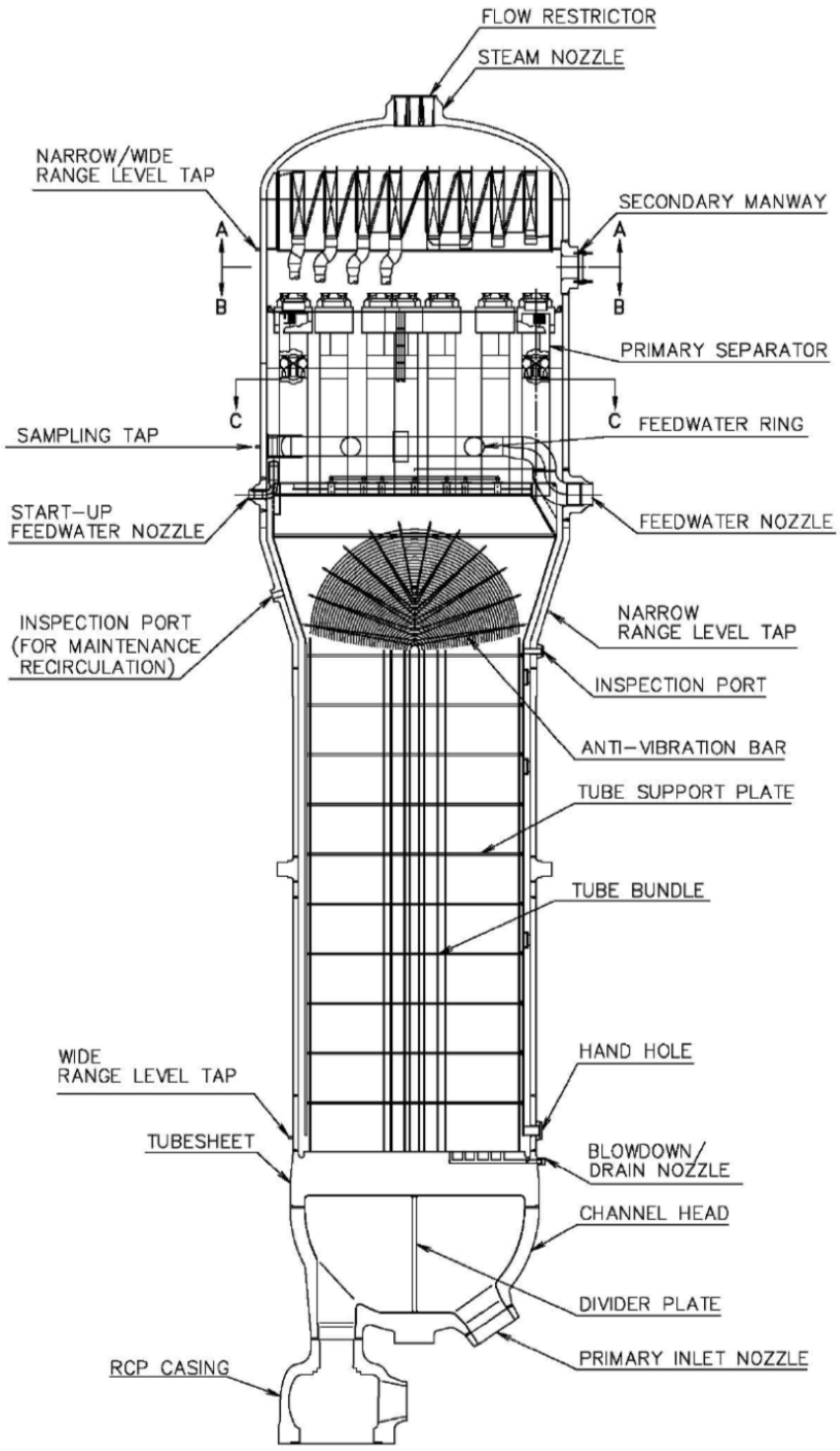

Figure 2.4: Cross section detailing the main SG.*Source:* [25].

2.2.1.2 PI Controller for SG water level

Automatic SG water level control is the result of either the FW Control System of the FW Bypass Control System acting to maintain programmed SG level from 0 to 100 percent power. This control is accomplished through a proportional integral (PI) controller that senses steam flow, feed flow, and SG level. Details for the main FW control and corresponding data flow are illustrated in Figure 2.5 [25].

The steam flow is sensed and corrected for density by a steam pressure detector. The resulting steam flow signal is fed to a summer. A flow error signal is produced by subtracting the feed flow signal from the steam flow signal. The flow error signal goes to a PI controller.

The actual SG level is sent through a lag circuit to dampen out natural oscillations in the level signal. A level error signal is produced by subtracting the actual level from the program level. The level error signal is sent to the PI controller. The level and flow error signals are added to produce a total error signal. This total error signal is the output of the PI controller when it is in AUTO. The operator can manually control the output of the controller. The output then goes to the I/P converter to position the FW control valve.

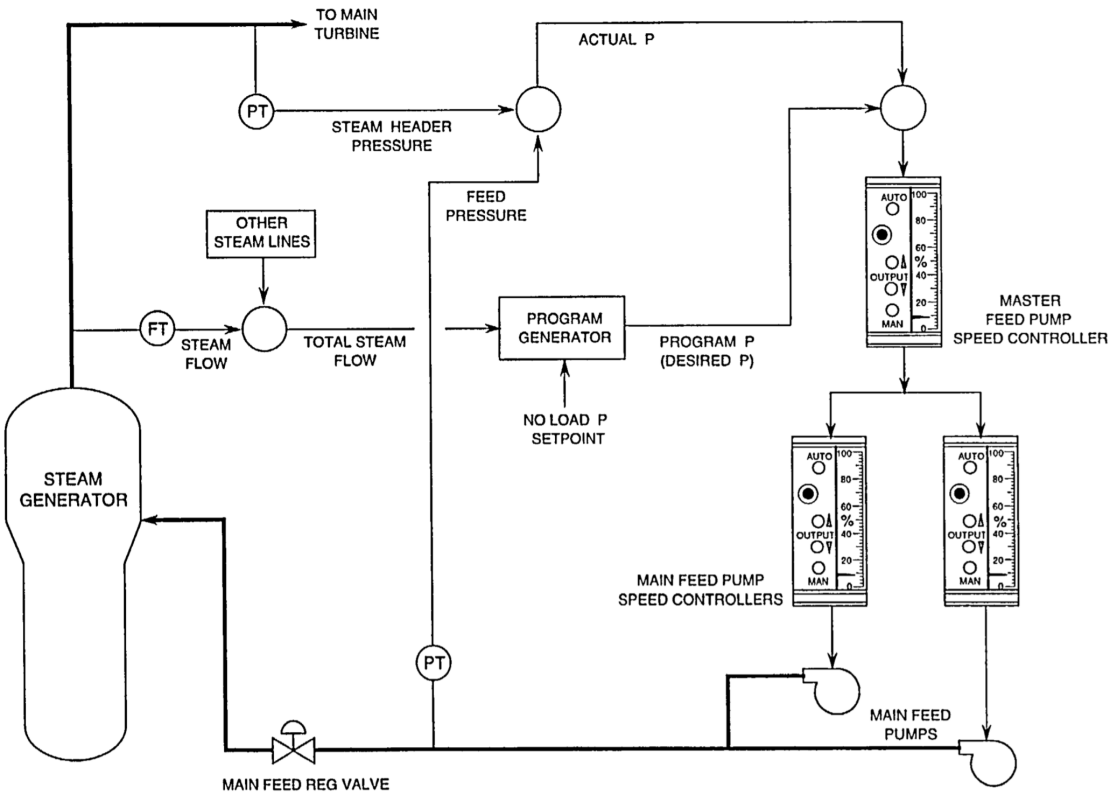

Figure 2.5: Main FW control system control diagram.*Source:* [25].

2.2.1.3 Actuators

As shown in Figure 2.5, the SG water level is controlled via the feedwater regulating valve (FRV) located between feedwater pump and SG. The FW control valves are sized to operate in automatic in the range of 8 to 100 percent reactor power. In this range, the FW control valves are designed to have a nearly linear relationship between valve position and FW flow rate. The FW control valve is sized at 12 inches to support feed flow in the range of 8 to 100 percent power. The FW bypass control valve is a 3-inch valve that permits automatic operation at a low power level from 0 to 8 percent reactor power, which is not a focus of our research.

2.2.2 Reactor Protection System

The purpose of the Reactor Protection System (RPS) is to process input signals from selected plant parameters and send a reactor trip signal to the reactor trip breakers when abnormal plant conditions exist. The purpose of the Engineered Safety Features Actuation System (ESFAS) is to process input signals from selected plant parameters and send an actuation signal to the ESF equipment when abnormal plant conditions exist. When a reactor trip signal is generated, the RPS shuts down the reactor by opening the reactor trip breakers. The trip and actuation signals for both the RPS and ESFAS are processed through the Solid State Protection System (SSPS) cabinets.

2.2.2.1 SSPS Cabinets

The SSPS is a dual train redundant system, consisting of two four-bay cabinets, one single bay control board demultiplexer cabinet, and a computer mounted demultiplexer assembly. Each of the four-bay cabinets is composed of an input relay bay, a logic bay, and two output relay bays. Figure 2.6 shows the simplified interface diagram of the SSPS.

The SSPS input relays receive inputs from Process Instrumentation Cabinet (PIC) bistables, Nuclear Instrument System (NIS) bistables, and field contacts. When a trip or actuation signal is received from one of the bistable or field contact inputs, the associated input relay de-energizes, closing a contact to provide the trip or actuation signal to the logic cabinet. The SSPS logic bay receives input from the input relays and determines if the required actuations logic (coincidence) is met for a reactor trip or ESF signal. If the required coincidence is met, the logic bay initiates the reactor trip or ESF signal. The reactor trip breakers and bypass breakers are designed to open on an automatic or manual trip signal, which will interrupt power from the rod control system to the control rod drive mechanisms (CRDMs), causing the rods to

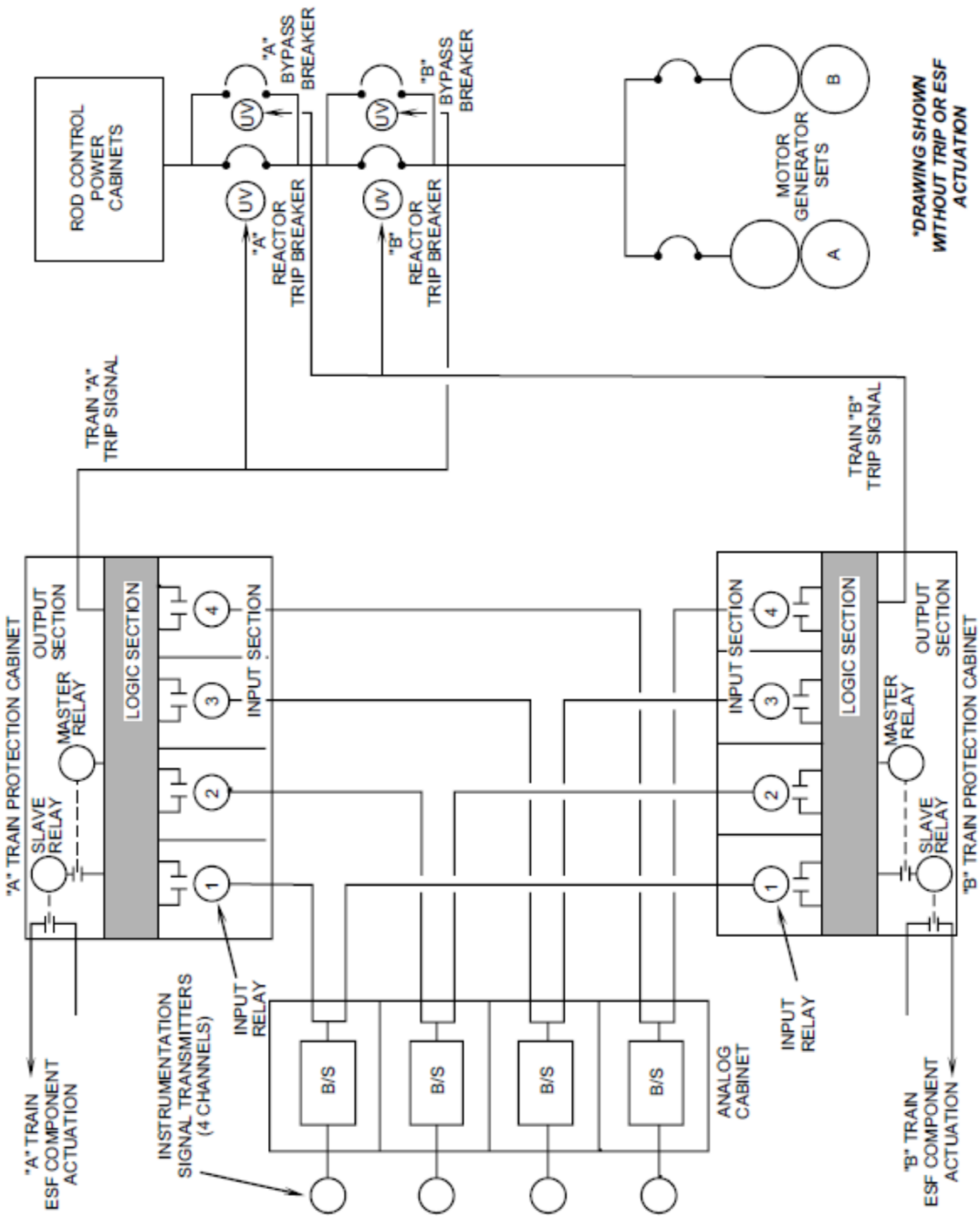

Figure 2.6: Reactor protection system layout. *Source:* [25].

fall to the bottom of the core. Reactor trip breakers A and B are normally shut and are in series such that opening either breaker will cause the rods to drop into the core. Each reactor trip and bypass breaker has redundant trip coils: under-voltage (UV) trip coil and shunt trip coil. The UV coil is maintained energized by the output of the SSPS logic bay when a trip signal is not active. When SSPS initiates a reactor trip, the power to the UV coil is removed, causing the reactor trip breaker to open. The shunt trip coil, normally de-energized, is energized when a reactor trip is initiated.

2.2.2.2 GPWR reactor trip functions

The RPS automatically keeps reactor operation within a safe region by shutting down the reactor whenever the limits of the region are exceeded. The safe operating region is defined by several considerations, such as mechanical/hydraulic limitations on equipment and heat transfer phenomena. For the GPWR system, reactor trips will be initiated for the 18 events listed in Table 2.5. Included in the list are two SG related trip functions featuring low-low water level and low SG level-low feedwater flow events.

Table 2.5: GPWR reactor trip functions.

Manual	Source range high flux
Intermediate range high flux	Power range high flux
Power range high neutron flux rate	Overtemperature ΔT
Overpower ΔT	Low primary coolant flow
RCP bus undervoltage	RCP bus underfrequency
Pressurizer high pressure	Pressurizer low pressure
Pressurizer high water level	SG low-low water level
Low SG level/low feedwater flow	Safety Injection signal
Turbine trip	SSPS General warning alarm

The purpose of the SG low-low water level trip is to protect the reactor by preventing operation without adequate heat removal capability. A loss of SG heat sink will lead to a reactor coolant system (RCS) overtemperature and overpressure excursion, and could eventually lead to a loss of core cooling capability. This circuit trips

the reactor if two out of three level indicators of any SG indicate below the low-low trip setpoint of 25 percent NR SG level. The trip logic is shown in Figure 2.7 with three level sensors for each SG.

A low SG level with feedwater flow mismatch trip is included to enhance the reactor trip system's overall reliability. It is redundant to the SG low-low level trip. The trip is actuated when the steam flow from one SG exceeds the feedwater flow to the same SG by 40% of rated steam flow coincident with a low level (25% NR) in the same SG level. For each SG, there are two SG level circuits, and two feedwater flow mismatch circuits. One of two of the circuits at the setpoint in any SG will generate a reactor trip. The actual circuity is illustrated in Figure 2.8.

In summary, there is a four-channel sensor system used in the I&C system of GPWR to monitor each SG's water level. One of them (channel three) serves as the controller input for the SG water level control system. Three of them (channels one, two, and three) serve as monitors to initiate the low water level reactor trip. All four channels supply the signal to initiate a high steam generator level trip.

2.3 Comparison Between GPWR and AP1000 I&C Systems

The I&C system of a nuclear power plant functions as the "nerve system" of the plant. We have made a considerable effort to study the I&C systems in both the GPWR and AP1000 designs. There are many similarities, as well as some differences between the GPWR and AP1000 I&C systems. Both systems provide operators with critical safety information on plant operation, allowing operators to control various plant safety systems during routine operations, and automatically protect the reactor core during potential accident events.

AP1000 is, however, a newer design, which has highly integrated digital I&C designs for safe and efficient operation. The overall architecture is clear, and the subsystem modules are well divided and classified. AP1000 also introduced several

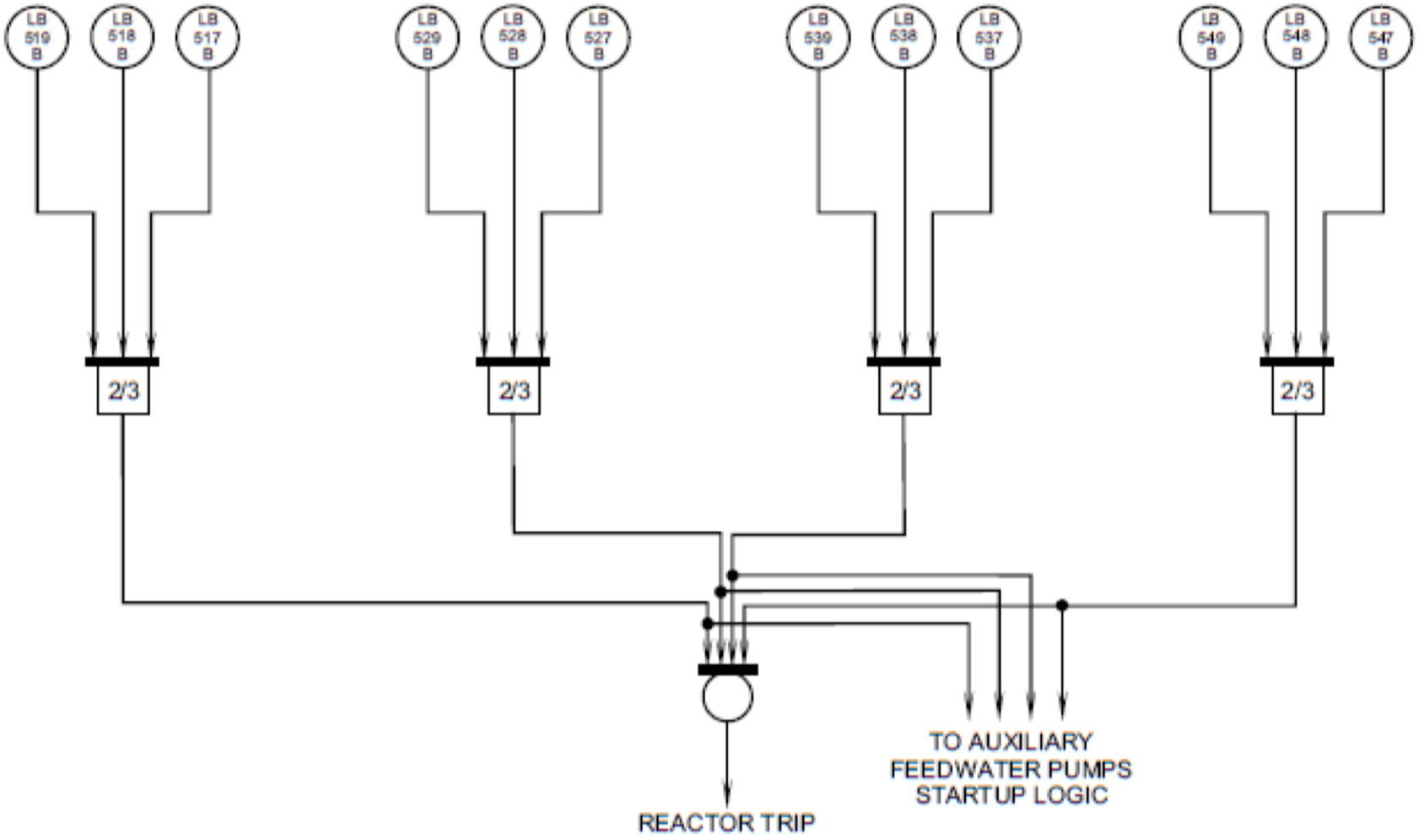

Figure 2.7: Low-low SG level trip circuitry.*Source:* [25].

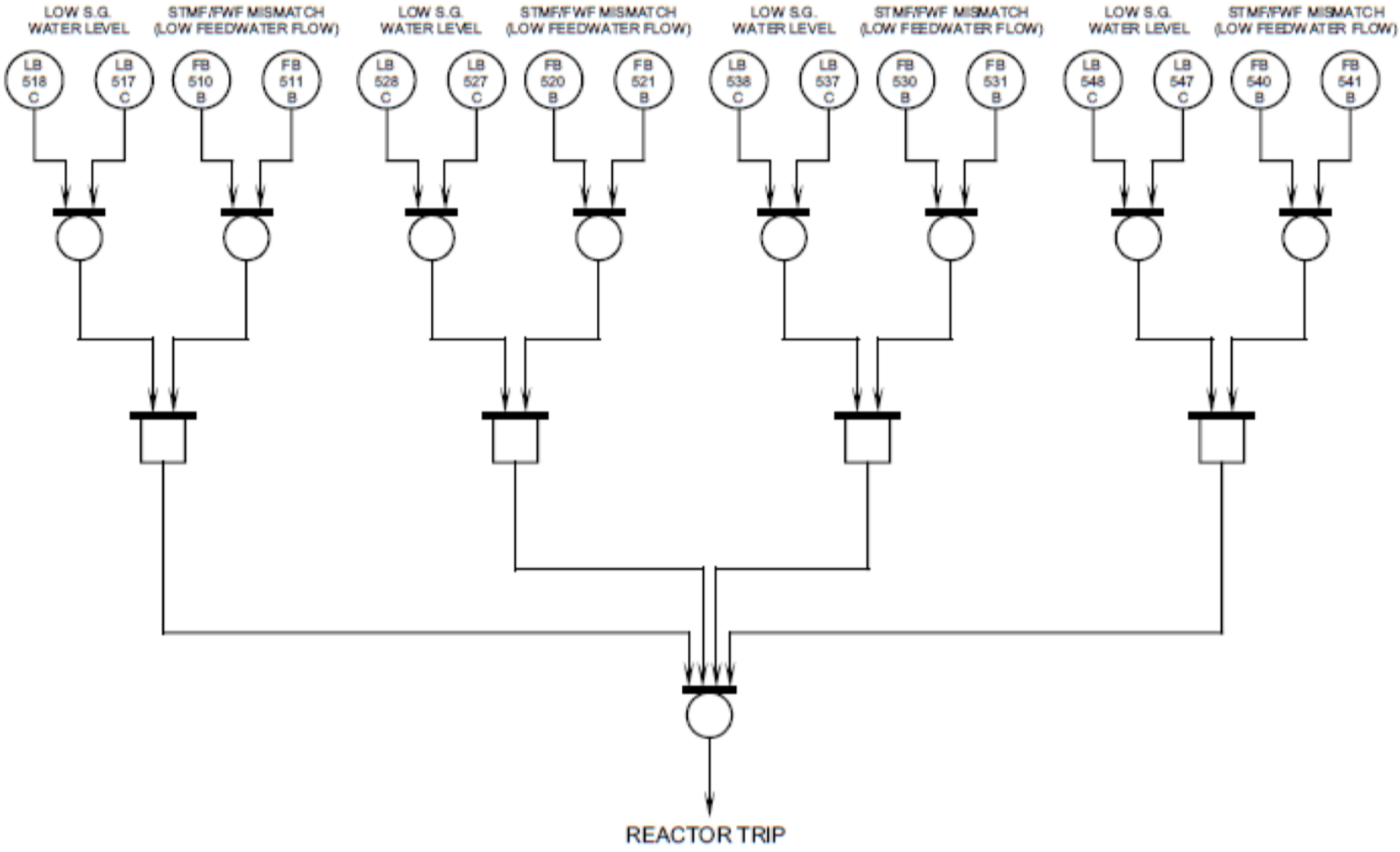

Figure 2.8: Low SG level with a feedwater flow mismatch trip.*Source:* [25].

new passive safety cooling systems, including Passive Safety Injection, PRHR and Passive Containment Cooling Systems. These new systems also brought new reactor trip functions for the protection and monitoring system, such as PRHR actuation and core makeup tank injection trips as summarized in Table. 2.2.

CHAPTER 3

Cyber Attack Scenario Generation

With the knowledge gained of the I&C system in AP1000, we now begin the task to find out how a cyber-attack can sabotage the I&C system. We develop a methodology to quantify the most susceptible attack pathways and identify at-risk components that may result in an unscheduled shutdown due to a cyber-attack on the I&C system. Through our investigation, we note that four reactor trips originate from automatic control of SG water level outside an acceptable range in a total of 11 reactor trip functions. Thus, SG water-level focused attacks serve as a distinct point of focus for our cyber-attack scenario study.

An attack tree analysis has been performed in our attack tree generation and logic representation using SAPHIRE software. A multi-path event tree (MPET) structure is developed to display the dominant attack scenarios obtained from the analysis result. A new sensitivity metric is proposed to identify the low-order sets of components that may be compromised, and the degree of perturbations needed for each component. This approach allows us to rank all potential attack scenarios and identify the most significant ones with high levels of susceptibility.

3.1 Analysis of Attack Scenarios

In our study of the I&C system in Chapter 2, we recall that there are two primary I&C system responses for the generation of automatic reactor trips for the AP1000 plant. All the trip signals originate in the PMS, which contains the local coincidence logic that decides whether a trip is needed. This logic takes the real-time sensor values for all kinds of reactor state parameters, which are a part of PLS. The cyber-attack on PMS can result in a reactor trip directly. However, the digital components in PMS have a higher security level and are not accessible to the attacker. The components in PLS operating on the real-time network do not have a security level as high as PMS. These components could be more accessible for the attacker. Hence, in this study, we focus on attacks originating in the PLS.

For attack scenarios generation and analysis, we further classify and define the components in the I&C system into three modules: a sensor module, a controller module, an actuator module. The sensor module is defined as the detector, sensor, or indicator measuring system parameters or indicating system state. The controller module is defined as the components that process the signals from the sensor module, e.g. PI controller, bistable processor station, and local coincidence logic. The actuator module is defined as the components receiving signals from the controller module and interact directly with the physical process, e.g. a regulating valve or pump. Cyber-attack scenarios can initiate at any of these modules; hence, our cyber-attack scenarios are generated in these three modules.

3.1.1 SAPHIRE Code Introduction for Attack Scenario Study

The SAPHIRE code [20] is developed for performing a complete probabilistic risk assessment (PRA) for complex engineering systems. SAPHIRE can be used to model a complex system's response to initiating events and follow through the system evolution, and obtain outcome frequencies (or probabilities). Specifically, for nuclear

power plant applications, SAPHIRE 8 can identify important contributors to core damage (Level 1 PRA) and containment failure during a severe accident, which leads to releases (Level 2 PRA). It can be used for a PRA where the reactor is at full power, low power, or at shutdown conditions.

SAPHIRE contains editors or options for creating event trees and fault trees, defining accident sequences and basic event failure data, solving system fault trees and accident sequence event trees, quantifying cut sets, performing sensitivity and uncertainty analyses, documenting the results, and generating reports.

3.1.1.1 Fault Tree Analysis

A fault tree [26] model consists of a top event (usually defined by a heading in an event tree) and a connecting logic structure that models the combinations of events that take place to result in the undesired top event. In SAPHIRE, a fault tree generally represents a failure model. The fault tree logic structure can consist of any combination of the logic symbols, listed in Table. 3.1.

Table 3.1: Logic symbols in fault tree.

Symbol	**Description**
Basic event	A simple failure or fault. It may be a hardware failure, a human error, or an adverse condition.
AND gate	The logic operation for this gate requires all inputs into the AND gate must occur for failure to occur.
OR gate	The logic operation for this gate requires only one of the total number of inputs into the OR gate to occur for failure to occur.
N/M gate	The logic operation for this gate requires that N of the M inputs into the gate must occur for failure to occur.
TRANSFER gate	The transfer gate indicates that logic is continued from some other location.

3.1.1.2 Determination of Minimal Cut Sets

Once we have built a fault tree, we could determine the minimal cut sets of this fault tree [27]. The fault tree consists of many levels of basic events and sub-events linked together by AND gates and OR gates. Certain optimization functions are performed on the fault tree logic before it is processed. Next, we need to identify the cut sets based on the fault tree structure. The cut-set identification algorithm will locate the uppermost gate, the top gate, and perform a top-down analysis. Each gate's logic is recursively replaced with its input events until the resulting logic is in terms of basic events only. This results in a list of conjunctions of basic events. Each event intersection is a cut set of the fault tree and identifies a set of events that will cause the function modeled by the fault tree to occur. The list of cut sets identifies the logical combinations of events that will cause the top event to occur. The cut sets may need further reduction, which is applied to obtain a simpler collection of cut sets. In particular, the cut sets generated should be minimal, the list should not be simplifiable, and the cut set has no other cut sets as a subset.

For example, if $A \cap B \cap C$ causes the top event to occur, then $A \cap B \cap C$ is a cut set. If $A \cap B$ is also a cut set, then $A \cap B \cap C$ is not minimal, and it is discarded from the list. If neither A alone nor B alone cause the top event to occur, $A \cap B$ is a minimal cut set, and it is retained in the list.

3.1.1.3 Quantification for Probabilities and Frequencies

Probability is the only satisfactory way to quantify our uncertainty about an uncertain event E. Some basic probability relationships are described here [28, 29]. The probability of the union of n events is

$$P(A_1 \cup A_2 \cup ...A_n) = \sum P(A_i) - \sum_{i<j} P(A_i A_j) + ... + (-1)^n P(A_1 A_2 ... A_n). \quad (3.1)$$

The probability of the intersection of n events is

$$P(A_1A_2...A_n) = P(A_n|A_1A_2...A_{n-1})P(A_2|A_1)P(A_1). \tag{3.2}$$

The probability of the intersection of n events when the events are statistically independent is

$$P(A_1A_2...A_n) = P(A_1)P(A_2)...P(A_n). \tag{3.3}$$

These basic rules for the probability can be used for quantifying minimal cut sets and fault trees [30, 31]. The individual cut set probabilities are determined by multiplying the probabilities of the applicable basic events, e.g. for the probability C_i of cut set i is given by

$$C_i = q_1q_2...q_n \tag{3.4}$$

where

$$q_i = \text{ probability of the k-th basic event in the i-th cut set.}$$

The fault tree quantification process is performed in two steps: (1) calculation of individual cut set probabilities and (2) combining the cut set probabilities. The exact probability of the union of the cut sets can be found, in principle, by Eq.(3.1), where each A_i is a cut set.

Considering the comprehensive functionality of SAPHIRE, we have used the code to build the attack trees and generate minimal cuts for potential cyber-attack scenarios. The SAPHIRE analysis focusing on unscheduled reactor trips due to the perturbation in the steam generator system is discussed in Section 3.1.2.

3.1.2 Attack Tree Analysis

Attack trees are conceptual diagrams showing how an asset, or target, might be attacked [32]. They have been used in a variety of applications. Especially in the field of information technology, they have been used to describe threats to computer systems and possible attacks to realize those threats. For our study, we use the attack tree to analyze the possible attack targets and scenarios for cyber-attack events.

Attack trees are hierarchical, graphical diagrams. Like fault trees, the diagrams are usually drawn inverted, with the root node at the top of the tree and branches descending from the root. The top or root node represents the attacker's overall goal. The nodes at the lowest levels of the tree represent the activities performed by the attacker. Nodes between the leaf nodes and the root node depict intermediate states or attacker sub-goals. The gate structure is similar to the OR and AND gate we introduced in Section 3.1.1.3. The cut sets generated from the tree are known as attack scenarios. Attack trees can become large and complex, especially when dealing with specific attacks. A full attack tree may contain hundreds or thousands of different scenarios, all leading to the attack's completion.

An example of a tree describing attacks on a hypothetical nuclear plant's cooling systems is shown in Figure 3.1. The target of the attack is damage to cooling pumps. There are two AND nodes and three OR nodes. For example, the sabotage pump node is an OR node. There are two leaf nodes under this node. This means that we could either cause electrical damage or blow up the pumps to sabotage pumps.

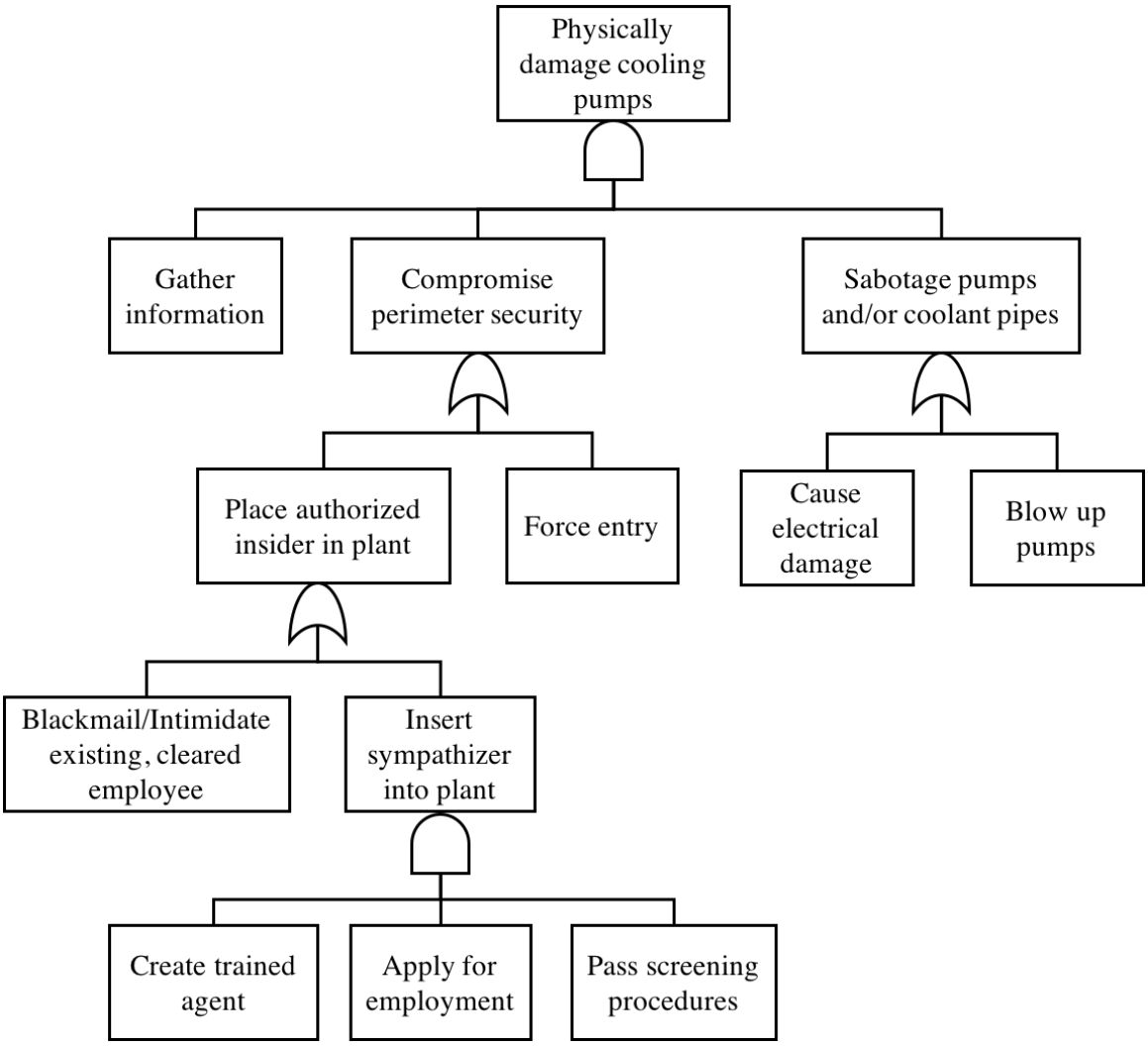

Figure 3.1: Example of attack tree.

Once we have completed the structure of the attack tree, we can assign values to the various leaf nodes, then make calculations about the nodes. There are several ways to assign values. Assigning measures to represent "possible" and "impossible" steps to the nodes is one way to look at the tree. Any Boolean value can be assigned to the leaf nodes and then propagated up the tree structure in the same manner. It is also possible to assign continuous values to nodes. There are many other possible continuous node values, including the probability of success of a given attack and the likelihood that an attacker will try a given attack.

To create an attack tree, we need to identify the attack goals first. Each goal forms

a separate tree, although they might share subtrees and nodes. Then, we should think of all attacks against each goal and add them to the tree. Once we have the attack tree and have evaluated all node values, the attack tree can be used to make security decisions. Attack trees provide a formal methodology for analyzing the security of systems and subsystems [33, 34].

3.1.3 Attack Tree Generation for AP1000

With the observation that all attacks could originate in the PLS, attack tree and minimal cut sets are generated with the logic structure and knowledge of the SGWLC system. An attack on a single module or some combinations of the three modules could lead to the unintended generation of reactor trip signals in the PMS. A list of all the components in PLS is summarized in Tabel 3.2. Figure 3.2 illustrates the logic structure for PLS control components and their relationship to the PMS structure for a trip signal generation due to an out-of-range SG water level caused by an attack on the PLS in terms of sensors, controllers, and actuators.

Table 3.2: List of components in PLS.

Component	**Number**	**Component**	**Number**
Narrow range level sensor	4	Wide range level sensor	4
Steam pressure sensor	4	Steam flow sensor	2
FW flow sensor	2	Hot leg temperature sensor	4
SG level control program	1	FW flow PI controller	1
FW Regulating valve	1	Pump	2

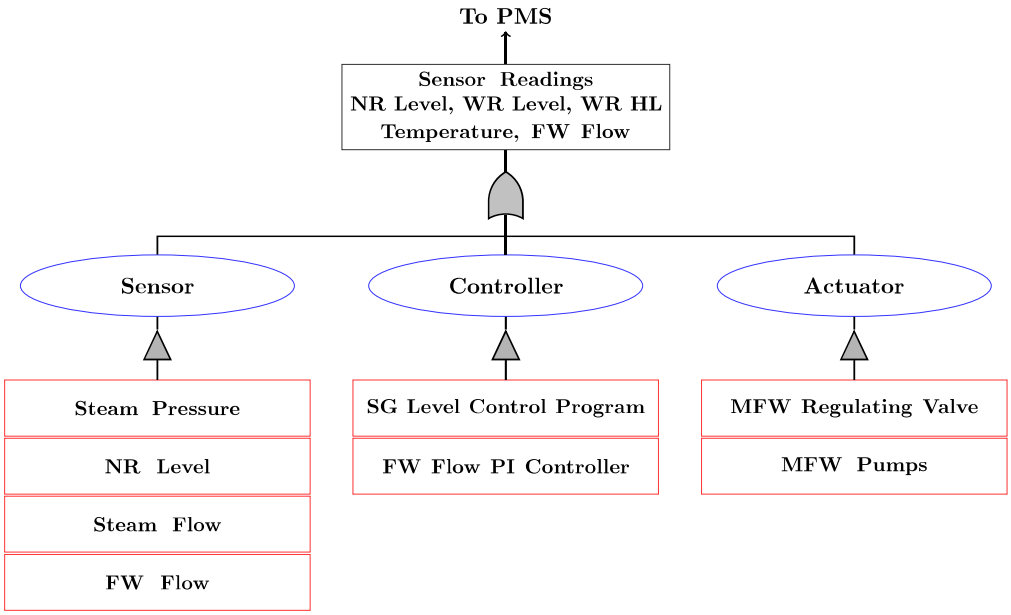

Figure 3.2: Top-level control logic structure for the PLS.

From these structures, we can construct an event tree using the SAPHIRE code. A part of the PLS attack tree built using the SAPHIRE code is shown in Figure. 3.3.

We may then generate minimal cut sets with SAPHIRE for the attack tree. An attack probability must be assigned to each component to calculate the minimal cut set probability and rank them. Since the probability of an attack on a single component is difficult to quantify, we start by creating two main groups of components. The first is comprised of primary control components, including NR level sensors, steam pressure sensors, all control logic structures, and the MFW regulating valve, all of which are assigned the *attack possibility* metric $\lambda_j = 0.1$. The second group is made up of secondary control components, including FW flow sensors, steam flow sensors, and the MFW pump, all of which are assigned $\lambda_j = 0.05$. This metric is mostly a placeholder at this stage; it does not mean an attack will occur at these sites or that they are the most accessible sites to attack, but it does focus the discussion at this stage.

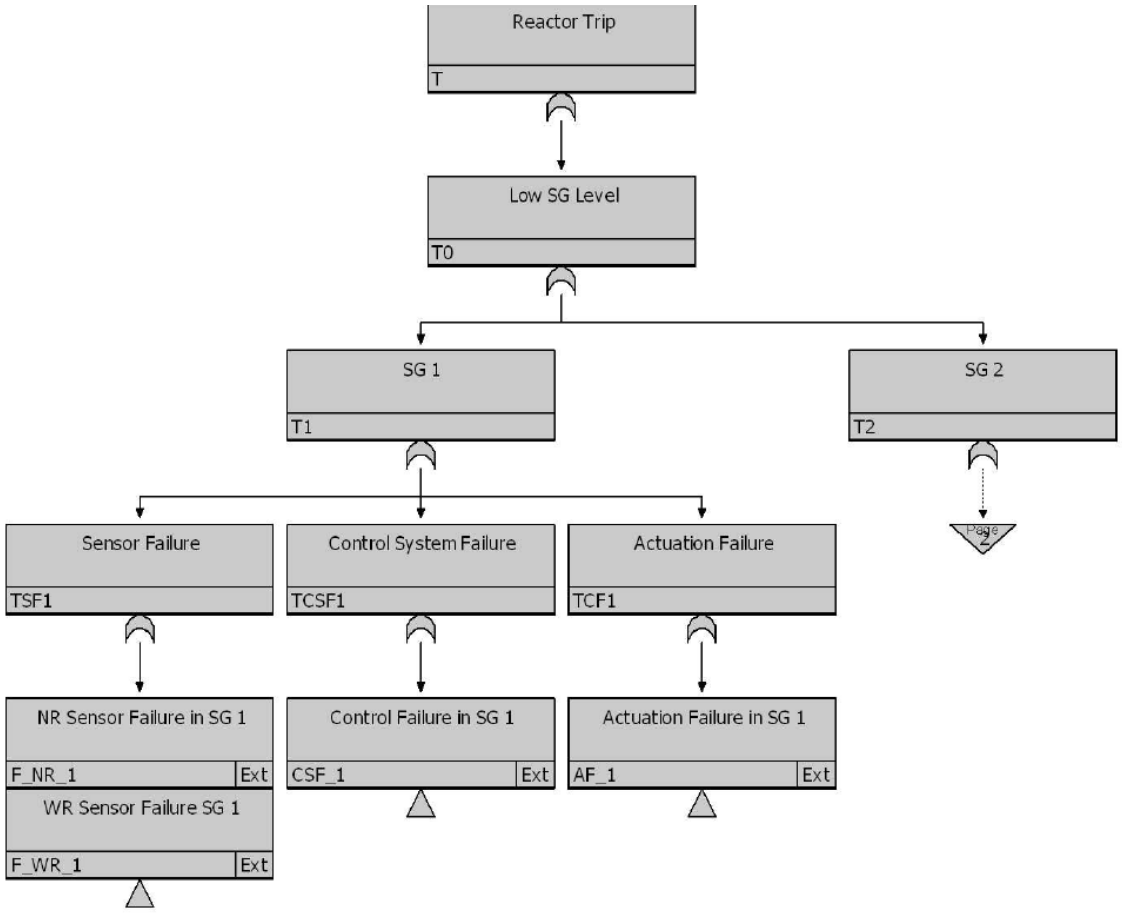

Figure 3.3: Attack tree for reactor trip due to low SG level.

There are several hundreds cut sets. After reduction, we obtain 54 minimal cut sets. The minimal cut-set list can be very long, especially when there are many combinations of basic events. With a properly developed Multi-Path Event Tree (MPET) representation, we may visually display the combination of attacks on components that could result in top events.

For an attack focused on causing reactor trip due to out-of-range SG water level, the resulting trip signal will come from one of two main trip signal classes: (a) the generic trip signals feedwater isolation (FI) and Loss of Heat Sink (LOHS) on high and low SG water level, respectively, or (b) the ESF-related trip signals for PRHR actuation and Core Makeup Tank (CMT) injection. In this structure, the generic trip signals are the primary mechanisms for protecting the reactor from any number

of issues related to improper interaction between the primary and secondary sides. Alternately, ESFs can function as either a redundancy feature, in case of failure in the generic trip signals, or as a safeguard against unanticipated plant conditions. Because ESF actuation usually results in a more drastic response, e.g., the CMT injecting borated coolant into the core, they typically carry more logical checks than their generic trip signal counterparts.

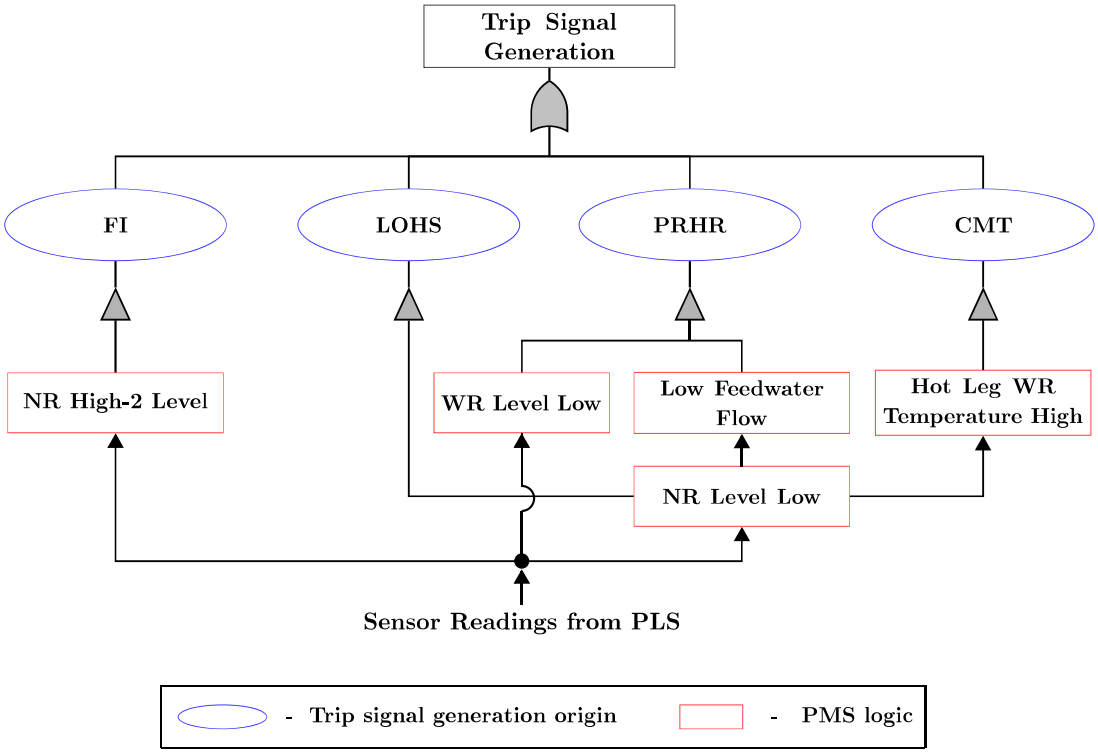

Figure 3.4: Top-level logic for trip generation in PMS.

From the PLS logic structure for automated and manual SG level control, we are able to relate the control functions and mechanisms to the sensors of interest to the PMS. Building on this understanding, a logic-tree structure can be constructed for pathways to trip signal generation in the PMS due to the out-of-range SG level, as illustrated in Figure 3.4.

3.2 Multi-Path Event Tree Representation

Identification of cut sets and minimal cut sets is one of the most important qualitative analyses of an event\fault tree. A cut set is a set of basic events whose occurrence ensures that the top event occurs. A minimal cut set is a cut set that cannot be reduced without losing its status as a cut set, which can also be defined as a group of sets consisting of the smallest combinations of basic events that result in the occurrence of the top event. For a complex system or network, the analysis results invariably result in a large number of minimal cut sets, which is not intuitive and hard to interpret. Hence, an MPET structure is developed to efficiently and intuitively display a large number of dominant or risk-significant attack scenarios instead of the traditional event trees representing minimal cut sets.

3.2.1 Generation of MPET

The MPET formulation uses a Boolean logic structure to graphically represent the minimal cut sets in terms of basic events in a succinct manner [35, 36]. With the experience of complex system analysis, we discover that some of the minimal cut sets will contribute to the majority of the result. These dominant minimal cut sets are always sharing similar basic events. Hence, we proposed a methodology to graphically representing the minimal cut sets based on the characteristics we discovered.

To generate the MPET, we developed a methodology to follow, and a simple example is presented to demonstrate the methodology. In this simple example, there are three components or stages of system evolution, X_1, X_2, and X_3. Under X_1, there are three basic events X_{11}, X_{12}, and X_{13}. Under X_2, there are two basic events X_{21} and X_{22}. Under X_3, there are three basic events X_{31}, X_{32}, and X_{33}. The basic events or states $\{X_{ij}\}$ have an associated quantification $\{P_{ij}\}$. Hence, for a complex system, we first need to divide the system into subgroups or cluster the basic events into functional groups from the minimal cut set results. A Boolean logic notation is

introduced into graphically connecting these basic events and generating the MPET diagram. The MPET structure for our example is illustrated in Figure 3.5. The MPET graphically represents a total of 18 minimal cut sets into a simple structure for rapid analysis and evaluation of end states.

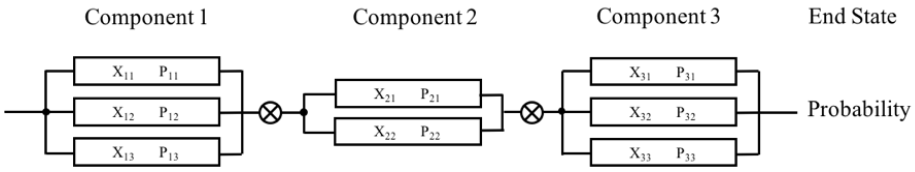

Figure 3.5: Example of MPET structure.

Our MPET representation and associated end-state frequency offer several advantages compared to the traditional event tree method. In particular, we can identify the state and time evolution of each component, resulting in either success, degradation, or failure end-states. We are also able to distinguish the end states for different scenarios. With a properly developed MPET, reactor operators can determine the end state of each scenario more efficiently than conventional ETs. Hence, the operators can take proper actions faster for safely managing attack events, thereby avoiding reactor trips. This MPET structure would also allow convenient modifications to the cut sets and represent suspected cyber-threats for system analysis, evaluation, and verification.

3.2.2 MPET Representation for Low-Order Trip Scenarios

We have created two main groupings of attacks: the first aiming to attack components in such a manner that SG water level is decreased to the point at which a LOHS-, PRHR-, or CMT-based trip is actuated, and the second where the SG water level is increased to the point that the logic for an FI trip is met.

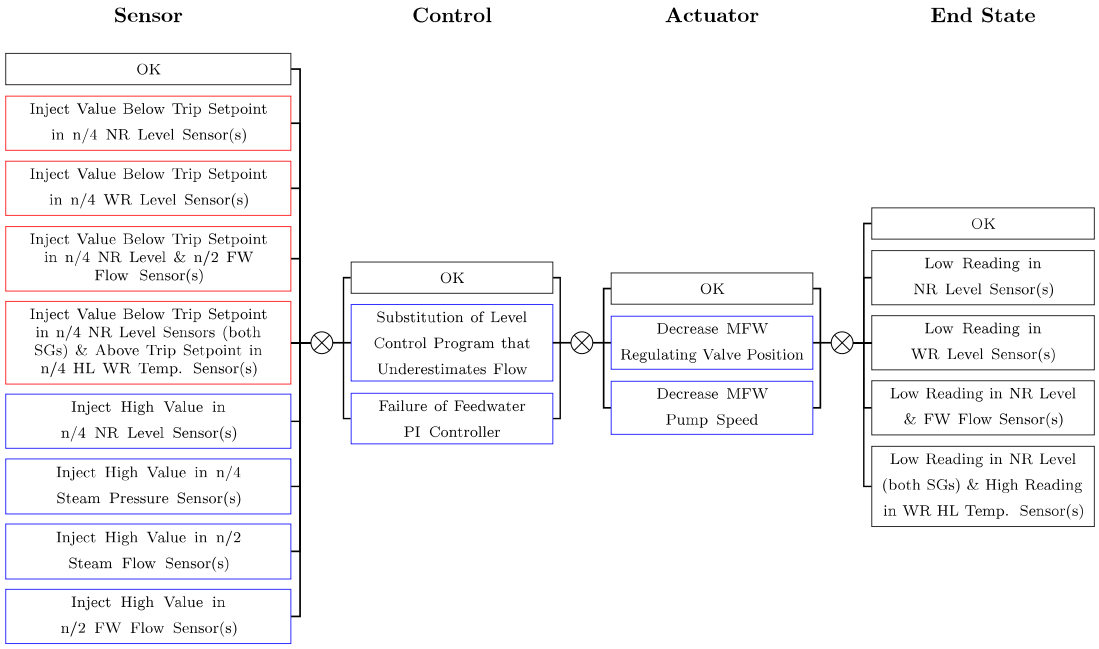

Figure 3.6: MP-ET for attacks on the PLS causing low SG water level end state.

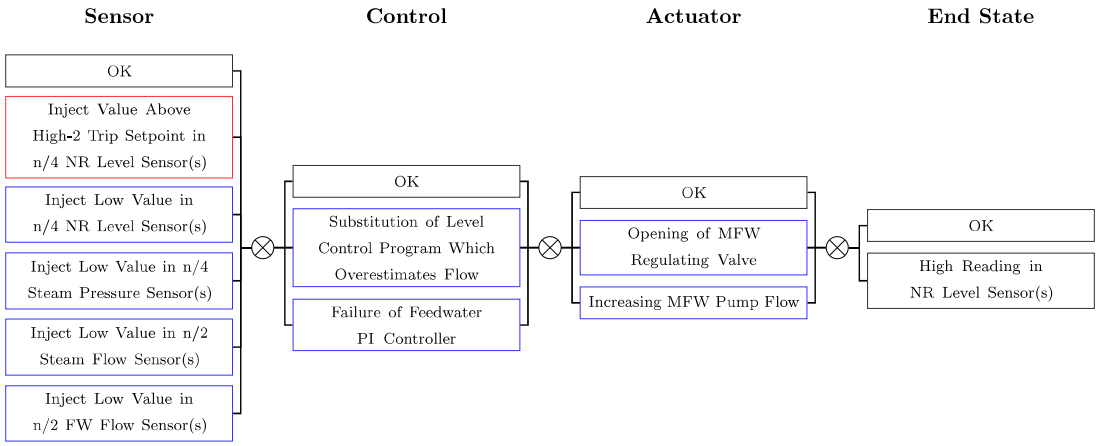

Figure 3.7: MP-ET for attacks on the PLS causing high SG water level end state.

Under the structure we have defined, two main methods of attack arise (1) common mode attack (CMA) on multiple sensors that immediately meets a condition for a trip in the PMS without the need for changes to occur in the physical system, or (2) single attack (SA) or CMA that is below the trip setpoint and uses normal feedwater control system (FCS) control behavior to drive the physical system outside the normal operating envelope, eventually causing reactor trip due to out-of-range

values. Figures 3.6 and 3.7 show MPET representations of the potential minimal cut sets for the failure of SG water-level control, resulting in low or high SG water levels, respectively, that lead to a spurious reactor trip. In these figures, the first method for the attack is bordered in red, and the second is bordered in blue.

The MPETs summarized in Figures 3.6 and 3.7 represent all possible attack scenarios that can result in spurious reactor trips brought about by improper control of the SG water level. Under this structure, the power of the MPET can be seen through the representation of the hundreds of potential cut sets in a half-page graphic.

3.3 Susceptibility Analysis

Further analysis can simplify the possible attack scenarios by identifying low-order sets requiring the lowest number of distinct actions, and the smallest system deviation needed. To accomplish this, metrics bound by system architecture are introduced to analyze system susceptibility to a given attack, rather than relying on assumptions about the attacker or attack feasibility. From these metrics, calculations and analyses are performed using the SAPHIRE 8.0.9 code.

3.3.1 Methodology Development

An attack tree structure is used to determine an estimate for the susceptibility to a given attack. To relate attacks on the PLS to the required system change needed for the PMS to actuate a trip, the *sensitivity metric* η_i for sensor i used in the PMS trip logic is defined as

$$\eta_i = 1 - \frac{|N_i - T_i|}{N_i}, \tag{3.5}$$

where N_i is the average (or setpoint) value for the sensor at full-power operation, and T_i is the value at which the sensor meets the condition(s) for reactor trip. Thus, attack pathways that require the least deviation from their normal operating values are

considered more likely to initiate a trip. The attack possibility λ_j is also raised to the power of the number n of components compromised by the attack, thus favoring sets that require fewer components to determine the *susceptibility* χ_{ij} to attack scenario j causing a trip via sensor i, as in

$$\chi_{ij} = \eta_i \times \lambda_j^n. \tag{3.6}$$

The formulation of Eqs. 3.5 and 3.6 may be illustrated for a spurious trip on a high SG water level end state, as outlined in Figure 3.7. An attack targeting the sensor group could compromise NR level sensor feedback, $\lambda_j = 0.1$, by injecting a false low value into the feedback division, $n = 1$, so that the system increases FW flow for a long enough time period and the remaining NR level sensors meet the conditions for FI, $\eta_i = 0.63$, resulting in an attack *susceptibility* $\chi_{ij} = 0.063$. The attack possibility for FI trip signal is summarized in Table 3.3. The table lists the attack possibility for each scenario. An analysis of low-order PLS-PMS attack scenarios is summarized in Table 3.4, with attacks grouped according to the fundamental control structure they target, and $\chi_i = \sum_j \chi_{ij}$ covering all potential attack scenarios $j = 1, \ldots, G$. Preliminary investigations showed that the large number of components required for an attack obtaining trip via CMT significantly lowered the susceptibility contribution compared to other sets, and are therefore not included in the tabulation.

Table 3.3: Attack Possibility For FI trip signal.

Trip Signal	PMS Trip Logic Sensor	Attack Possibility	Attack Description
FI	n/4 NR Level Sensors High	5.000E-02	Inject false low value in NR level sensor feedback division to increase SG level
		5.000E-02	Inject false low value in steam pressure sensor feedback division to increase SG level
		2.500E-03	Inject false low value in two of four NR level sensors to increase SG Level
		2.500E-03	Inject false low value in two of four steam pressure sensors to increase SG Level
		5.000E-02	Attack increases position of MFW regulating valve to increase SG level
		5.000E-02	Attack substitutes false SG level control program which overestimates flow to increase SG level
		5.000E-02	Attack tampers with PI level controller to overestimate flow and increase SG level
		1.000E-01	Inject false low value in steam flow feedback division sensor to increase SG level
		1.000E-01	Inject false low value in FW flow feedback division sensor to increase SG level
		1.000E-01	Attack increases MFW pump speed to increase SG level
		1.250E-04	Inject false low value in three of four NR level sensors to increase SG Level
		1.250E-04	Inject false low value in three of four steam pressure sensors to increase SG level
		1.000E-02	Inject false low value in two of two FW flow sensors to increase SG level
		1.000E-02	Inject false low value in two of two steam flow sensors to increase SG level
		6.250E-06	Inject false low value in four of four NR level sensors to increase SG level
		6.250E-06	Inject false low value in four of four steam pressure sensors to increase SG level
		2.500E-03	Common mode attack on 2/4 NR level sensors cause trip on high level
		1.250E-04	Common mode attack on 3/4 NR level sensors cause trip on high level
		6.250E-06	4/4 NR level sensors cause trip on high level due to common mode attack

Table 3.4: Susceptibility for low-order reactor trip attacks, reference case.

System State	Trip Signal	PMS Trip Logic (End State)	η_i	Attack Group (G)	$\sum_j^G \lambda_j^n$	χ_i (%)
SG Level High	FI	High Level Reading in NR Level Sensors	0.63	Sensor	0.34	17.5
				Control	0.20	10.4
				Actuator	0.15	7.8
SG Level Low	LOHS	Low Level Reading in NR Level Sensors	0.44	Sensor	0.34	12.2
				Control	0.20	7.2
				Actuator	0.15	5.4
	PRHR	Low Revel Reading in WR Level Sensors	0.44	Sensor	0.34	12.2
				Control	0.20	7.2
				Actuator	0.15	5.4
		Low Level Reading in NR Level & Low Flow Reading in FW Flow Sensor(s)	0.27	Sensor	0.33	7.2
				Control	0.20	4.4
				Actuator	0.15	3.3

The attack scenario analysis showed that most high-susceptibility pathways consist of attacks on primary control components and sensor feedback divisions, and require that only one 2/4 bypass condition be met. The differences between the nominal value for sensors and the corresponding trip value has a relatively large impact on the susceptibility rankings, as demonstrated by the proximity between the nominal SG NR water level, $N_i = 57$ %, and the trip pathway through FI for high NR water level, $T_i = 78\%$, compared to the LOHS path for low NR water level, $T_i = 25\%$. Except for the WR level trip path in PRHR, the large number of components required for most ESF trip pathways had significantly lower susceptibilities than their generic counterparts. While this is reassuring from a broad prevention standpoint, as ESF actuation is typically a more severe response than normal trips, these scenarios should probably be weighted heavier due to potentially being more desirable targets.

We have developed an approach that highlights susceptible pathways for spurious reactor trip comprising the least number of components and the least deviation from the nominal system state. We note that groups with large numbers of components and high redundancy, e.g., sensors, are more susceptible to increased opportunity for attacks capable of causing a spurious reactor trip. These findings suggest that

sensors and control logic structures are a distinct focus for further inquiries into attack detection and mitigation methods for NPP I&C systems.

3.3.2 Parametric Analysis on Attack Possibility

As discussed in Section 3.3.1, the probability of an attack on a single component is very difficult to quantify. In the demonstration, we arbitrarily assumed the primary control components have the attack possibility metric $\lambda_j = 0.1$, and the secondary control components have the attack possibility metric $\lambda_j = 0.05$. In this subsection, we will discuss how the attack possibility metric will influence the attack scenario susceptibility analysis. During the analysis, the sensitivity metric of different reactor trip functions remains the same since the system itself defines it.

For parametric case 1, we use the same attack possibility metric $\lambda_j = 0.1$ for both primary and secondary control components. The result of the susceptibility analysis of this situation is summarized in Table 3.5.

Table 3.5: Susceptibility for low-order reactor trip attacks, parametric case 1.

System State	**Trip Signal**	**PMS Trip Logic (End State)**	η_i	**Attack Group (G)**	$\sum_j^G \lambda_j^n$	χ_i (%)
SG Level High	FI	High Level Reading in NR Level Sensors	0.63	Sensor	0.45	18.9
				Control	0.20	8.4
				Actuator	0.20	8.4
SG Level Low	LOHS	Low Level Reading in NR Level Sensors	0.44	Sensor	0.45	13.1
				Control	0.20	5.8
				Actuator	0.20	5.8
	PRHR	Low Revel Reading in WR Level Sensors	0.44	Sensor	0.25	13.1
				Control	0.20	5.8
				Actuator	0.20	5.8
		Low Level Reading in NR Level & Low Flow Reading in FW Flow Sensor(s)	0.27	Sensor	0.44	7.8
				Control	0.20	3.5
				Actuator	0.20	3.5

The total attack possibility increases due to the increase in the attack possibility of secondary control components. The possibility and the distribution for each group had changed as well. The sensor and actuator group has a higher share in the overall

susceptibility. This could result from the sensor and actuator are mostly being located on the secondary side. The overall distribution does not change too much, compared to the last demonstration. Hence, the qualitative conclusions for the initial analysis still remain for this case.

If the attacker thinks that secondary control components are easier to attack, the secondary control possibility could be higher than the primary ones. For parametric case 2, we assign the attack possibility metric $\lambda_j = 0.05$ for the primary control components, and $\lambda_j = 0.1$ for the secondary control components. The result of the susceptibility analysis for the parametric case 2 is summarized in Table 3.6.

Table 3.6: Susceptibility for low-order reactor trip attacks, parametric case 2.

System State	**Trip Signal**	**PMS Trip Logic (End State)**	η_i	**Attack Group (G)**	$\sum_j^G \lambda_j^n$	χ_i (%)
SG Level High	FI	High Level Reading in NR Level Sensors	0.63	Sensor	0.33	20.2
				Control	0.10	6.2
				Actuator	0.15	9.2
SG Level Low	LOHS	Low Level Reading in NR Level Sensors	0.44	Sensor	0.33	14.0
				Control	0.10	4.3
				Actuator	0.15	6.4
	PRHR	Low Revel Reading in WR Level Sensors	0.44	Sensor	0.33	14.0
				Control	0.10	4.3
				Actuator	0.15	6.4
		Low Level Reading in NR Level & Low Flow Reading in FW Flow Sensor(s)	0.27	Sensor	0.33	8.5
				Control	0.10	2.6
				Actuator	0.15	3.9

The total attack possibility decreases a little bit. It is because there are more control components on the primary side. Compared to the reference case, the overall distribution for attack possibility among groups and trip logic does not change too much. Hence, the qualitative conclusion still holds for this case as well.

CHAPTER 4

RELAP5 Simulation for AP1000 Power Plant

After the generation of the cyber-attack scenarios, we turn to the development of the capability to simulate the dominant scenarios. For this kind of simulation, we need simulation capable of both the thermal-hydraulic transients and the I&C components, e.g. sensor and controller. The RELAP5 code has been chosen for the thermal-hydraulic transient simulation. The RELAP5 code has been developed for best-estimate transient simulations of light water reactor coolant systems during postulated accidents. The code models the coupled behavior of the reactor coolant system and the core for loss-of-coolant accidents and operational transients, such as anticipated transients without scram, loss of offsite power, loss of feedwater, and loss of flow. A generic modeling approach is used that permits simulating a variety of thermal-hydraulic systems. Control system and secondary system components are included to permit modeling of plant controls, turbines, condensers, and secondary feedwater systems. Based on these advantages of RELAP5, we have used RELAP5 as the simulation software for our initial cyber-attack study. The transient simulation capability can be extended to the cyber-attack simulation as well. We begin with a basic introduction to RELAP5. An AP1000 nuclear power plant model has been developed in RELAP5 for the reduced-order model development and cyber-attack analysis. The simulation results and validation cases are also presented in this chapter.

4.1 RELAP5 Intrduction

The RELAP5/MOD3 code is based on a nonhomogeneous and nonequilibrium model for the two-phase system. It is solved by a fast, partially implicit numerical scheme to permit the economic calculation of system transients [21]. The RELAP5 code can provide important first-order effects for accurate prediction of system transients, but that is sufficiently simple and cost-effective so that parametric or sensitivity studies are also possible.

The RELAP5 thermal-hydraulic model solves eight field equations for eight primary dependent variables. The primary dependent variables are pressure p, phasic specific internal energies U_g, U_f, vapor volume fraction (void fraction) α_g, phasic velocities v_g, v_f, noncondensable quality X_n, and boron density ρ_b. The independent variables are time t and distance x. The basic two-fluid differential equations that form the basis for the hydrodynamic model are presented.

Let us start with the single-phase fluid mass conservation equation [37] with density ρ and fluid velocity $\mathbf{v}$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}), \tag{4.1}$$

the momentum conservation equation in terms of hydrostatic pressure p and acceleration of gravity $\mathbf{g}$, and the viscous momentum transport represented through shear stress tensor $\boldsymbol{\tau}$

$$\frac{\partial}{\partial t}\rho \mathbf{v} = -\nabla \cdot (\rho \mathbf{v}\mathbf{v}) - \nabla \cdot \boldsymbol{\tau} - \nabla p + \rho \mathbf{g}, \tag{4.2}$$

and the energy conservation equation with internal energy density U, heat flux $\mathbf{q}$ and source term S

$$\frac{\partial}{\partial t}\rho U + \nabla \cdot (\rho U \mathbf{v}) = -\nabla \cdot \mathbf{q} - p\nabla \cdot \mathbf{v} + \mathbf{v} \cdot (\nabla \cdot \boldsymbol{\tau}) - \nabla \cdot (\boldsymbol{\tau} \cdot \mathbf{v}) + S. \tag{4.3}$$

The phasic equations for the two-fluid model may be formally derived by averaging fluid conservation equations over a differential volume associated with each phase along the flow channel, as well as over time. The channel total cross-sectional area is A. The fractional cross-sectional area occupied by the vapor and liquid phase with the vapor fraction α_g and liquid fraction α_f

$$A_g = \alpha_g A, A_f = \alpha_f A. \tag{4.4}$$

The basic field equations for the two-fluid nonequilibrium model consist of two phasic continuity equations, two phasic momentum equations, and two phasic energy equations [38]. With ρ_g and ρ_f for the vapor and liquid densities, respectively, we obtain the phasic continuity equations

$$\frac{\partial}{\partial t}(\alpha_g \rho_g) + \frac{\partial}{\partial x}(\alpha_g \rho_g v_g) = \Gamma, \tag{4.5}$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \frac{\partial}{\partial x}(\alpha_f \rho_f v_f) = -\Gamma, \tag{4.6}$$

where Γ represents the vapor generation rate comprising mass transfer at the vapor-liquid interface in the bulk fluid and near the channel wall.

The phasic conservation of momentum equations are presented in terms of momenta per unit volume using the phasic primitive variables v_g and v_f. The momentum effects are secondary to mass and energy conservation in reactor safety analysis and a less exact formulation (compared to mass and energy conservation) is acceptable. Hence, the momentum equation comes with the following simplifications: the Reynolds stresses, covariance terms, phasic viscous stresses, and interfacial momentum storage are neglected, the phasic pressures are assumed equal, and the interfacial pressures are assumed equal to the phasic pressures. The momentum equations for both the vapor phase and liquid phases are combined with the corresponding conti-

nuity equations, respectively, to yield

$$\alpha_g \rho_g \frac{\partial v_g}{\partial t} + \frac{1}{2}\alpha_g \rho_g \frac{\partial v_g^2}{\partial x} = -\alpha_g \frac{\partial p}{\partial x} + \alpha_g \rho_g g - F_{wg} - F_{ig} + \Gamma(v_i - v_g) \tag{4.7}$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \frac{1}{2}\alpha_f \rho_f \frac{\partial v_f^2}{\partial x} = -\alpha_f \frac{\partial p}{\partial x} + \alpha_f \rho_f g - F_{wf} - F_{if} + \Gamma(v_i - v_f) \tag{4.8}$$

where the force terms on the RHS of Equations are

F_{wg}, F_{wf} = wall friction forces on vapor and liquid, respectively,

F_{ig}, F_{if} = interface friction forces on vapor and liquid, respectively,

and the interfacial momentum transfers associated with vapor generation are explicitly written in terms of the interface speed v_i and vapor generation rate Γ. The RELAP5 formulation includes additional terms associated virtual mass effects representing bubble motions in the channel, which are not included in here [39].

The energy equations are written in terms of the internal energy U,

$$\frac{\partial}{\partial t}(\alpha_g \rho_g U_g) + \frac{\partial}{\partial x}(\alpha_g \rho_g U_g v_g) = -p\frac{\partial \alpha_g}{\partial t} - p\frac{\partial}{\partial x}(\alpha_g v_g) + + Q_{wg} + Q_{ig} + Q_{fg} - \Gamma h_g, \tag{4.9}$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f U_f) + \frac{\partial}{\partial x}(\alpha_f \rho_f U_f v_f) = -p\frac{\partial \alpha_f}{\partial t} - p\frac{\partial}{\partial x}(\alpha_f v_f) + Q_{wf} + Q_{if} + Q_{ff} - \Gamma h_f, \tag{4.10}$$

where the energy source terms on the RHS represent

Q_{wg}, Q_{wf} = wall heat fluxes for vapor and liquid, respectively,

Q_{fg}, Q_{ff} = wall friction and pump effects for vapor and liquid, respectively,

Q_{ig}, Q_{if} = interface heat fluxes for vapor and liquid, respectively,

and the bulk energy transfers associated with the vapor generation are explicitly represented with the vapor enthalpy h_g, liquid enthalpy h_f, and Γ. The RELAP5 formulation includes additional terms associated with bulk interface mass transfer and thermal boundary layer.

The interface between the phases is assumed to have no volume, the sum of the transfer rates of mass, momentum, and energy should vanish at the interface. The interface condition for the mass transfer rate Γ has already been accounted for explicitly in deriving the continuity equations. For the momentum equations and energy equations the interface condition yields

$$\Gamma = -\frac{F_{ig} + F_{if}}{v_g - v_l} = -\frac{Q_{ig} + Q_{if}}{v_g - v_l}. \tag{4.11}$$

It is also noted that

$$F_{wg} + F_{Wf} = F_w, \tag{4.12}$$

$$Q_{wg} + Q_{wf} = Q_w. \tag{4.13}$$

The difference equations are based on the concept of a control volume (or mesh cell) in which mass and energy are conserved by equating accumulation to the rate of influx through the cell boundaries. A semi-implicit numerical solution scheme is used to solve the finite-difference equations [40, 41]. The equation of state $\rho = f(h, p)$ connecting water density ρ to enthalpy h and pressure p is obtained with the ASME

steam table to evaluate phasic densities ρ_g and ρ_f and other properties of water. The constitutive relations are defined through flow regimes and flow-regime-related models. Four flow regime maps are used: a horizontal map for flow in pipes, a vertical map for flow in pipes and bundles, a high mixing map for flow in pumps, and an engineered cementitious composite (ECC) mixer map [42, 43]. A point reactor kinetics equation represents the reactor core power.

The code includes many generic component models from which general systems can be simulated. The component models include pumps, valves, pipes, heat-releasing or absorbing structures, reactor point kinetics, electric heaters, jet pumps, turbines, separators, accumulators, and control system components. Thus, the code provides the ability to model the details of the entire nuclear steam supply system (NSSS) illustrated by the nodalization diagram for a PWR plant in Fig. 4.1. The diagram illustrates a reactor core with coolant channels, two steam generators with two cold legs and one hot leg, and a pressurizer, together with various emergency core cooling systems. The designation and connection of control volumes and junctions are also illustrated in the diagram.

4.2 AP1000 Reactor Model Development and Simulation

The AP1000 is a 2-loop pressurized water reactor (PWR) with a power rating of 3415 MWt [18]. Its core design is very similar to a 3-loop PWR containing 157 fuel assemblies and a 14-ft active core length. Two hundred sixty-four fuel rods in a 17 $\times$ 17 square array constitute a fuel assembly. The remaining 25 positions contain the guide tubes and a central instrument tube. The AP1000 design has two Delta 125 type steam generators [24], each rated at 1707.5 MWt. Each pump has a height of 6.73 m and a power input of 6.6 MW and a rated speed of 188.5 rad/s and a flow rate of 4.97 m^3/s. The pressurizer in the AP1000 is larger than in earlier PWRs with a vessel volume of 59.46 m^3 (47.6% water and 52.4% steam) and length of 12.77

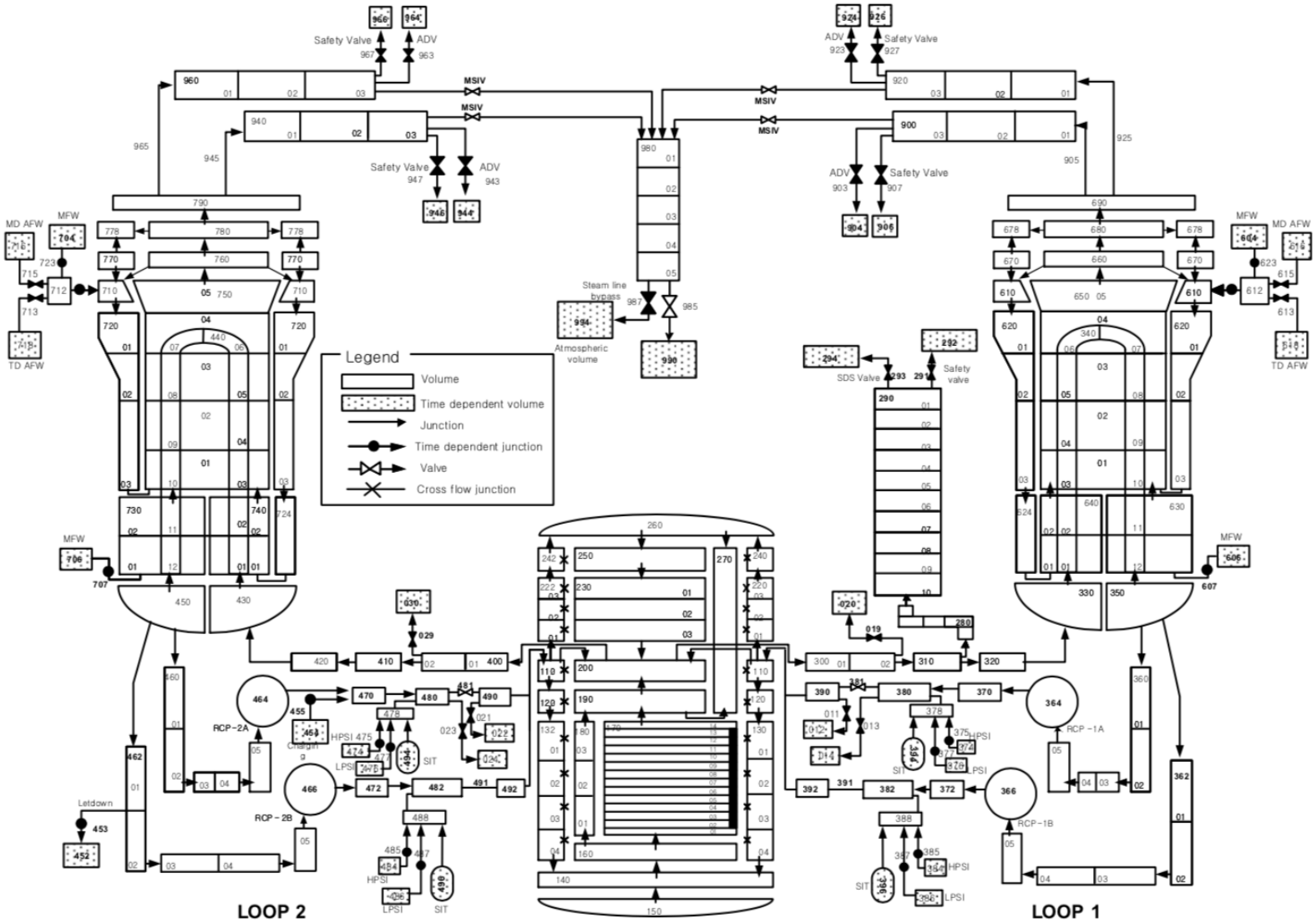

Figure 4.1: NSSS nodalization diagram for PWR plant.

m [44]. The most advantage AP1000 has is that all safety systems are based on natural circulation. Figure. 4.2 shows a schematic of the primary heat transport system (PHTS) and the passive safety systems for this design [22].

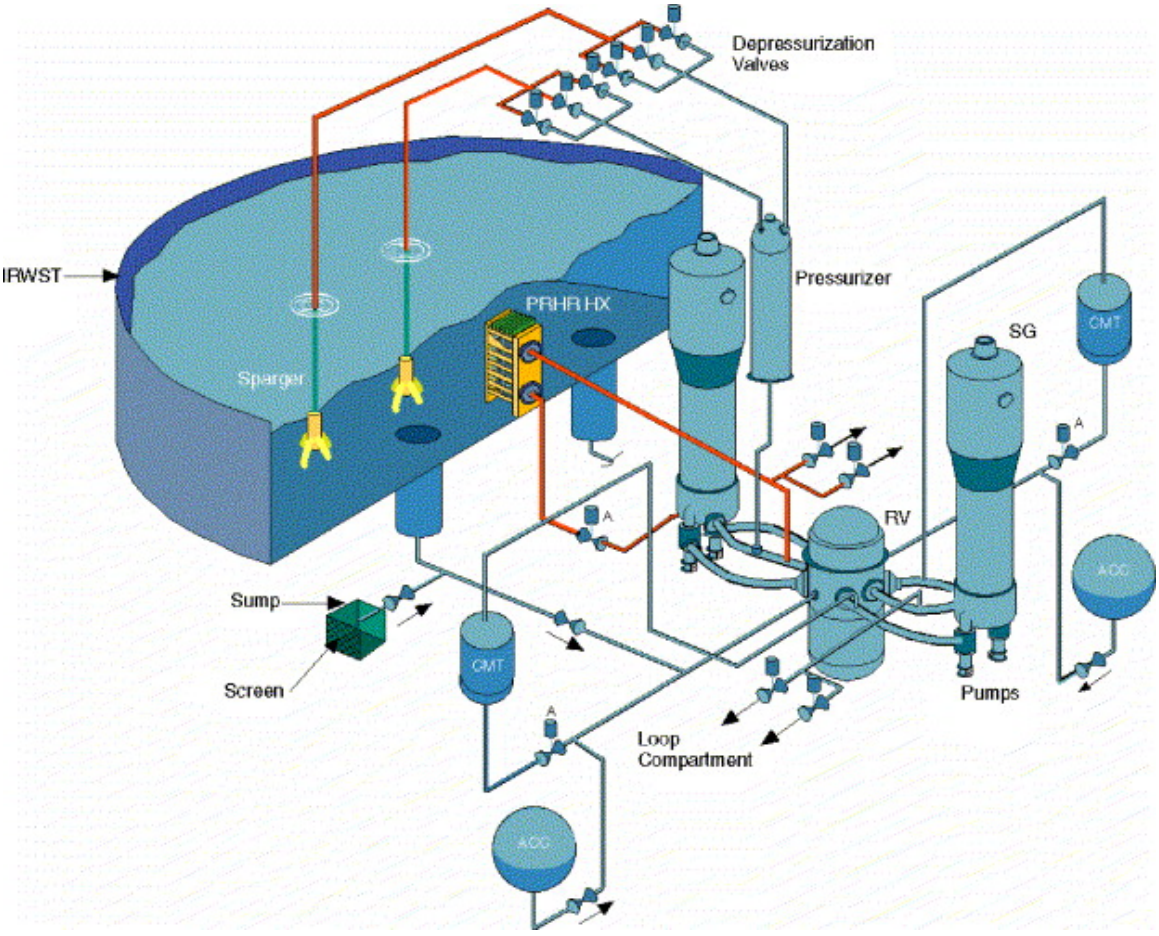

Figure 4.2: PHTS and passive safety systems of AP1000. *Source:* [22]

An input deck has been created to simulate the system state of AP1000. This AP1000 model is used to develop a reduced-order SG model and provide the measurements in the implementation of the Kalman filter algorithm for attack detection. The input deck is modified from that for a PWR with an AP1000 passive core cooling system input deck used for analyzing small-break loss-of-coolant accidents (SBLOCAs). An AP1000 model nodalization details are summrized in Table 4.1.

Table 4.1: AP1000 nodalization details.

Component number	**Volumes**	**Type**	**Description**
335	6	Pipe	Core channel
315	8	Annulus	Downcomer
320	6	Pipe	Downcomer bypass
350	4	Pipe	Upper plenum
356	3	Pipe	Upper head
100, 104, 200, 204	2	Pipe	Hot legs
118, 114, 210, 214	1	Pipe	Cold legs
107, 109, 207, 209	1	Single volume	Hot and cold plenum of SG
108, 208	8	Pipe	U-tubes of SG
113, 209	1	Pump	Pumps
176, 276	6	Pipe	Downcomers of SG secondary side
170, 270	6	Pipe	Boiler plus riser
171, 271	1	Separator	Separator
150	6	Pipe	Pressurizer
196, 191, 296, 291	1	Pipe	Core Makeup Tank
290, 190	1	Accumulator	Accumulator
167	6	Pipe	IRWST
166	8	Pipe	PRHR Heat Exchanger

RELAP5 pipe components with equivalent flow areas have been used to model the reactor core, divided into six nodes. The fuel is modeled as a 1D RELAP5 heat structure. The fuel rod is discretized into 17 nodes. The first six nodes represent the fuel, the next two nodes are the fuel gap, and the rest is the fuel cladding. The temperature-dependent thermal conductivity data and volumetric heat capacity data are defined for these compositions. The axial power profile is input for the fuel model. The downcomer region is modeled as an annulus with eight nodes. The two hot legs are represented using pipe components. The nodalization diagram for reactor vessel is presented in Figure 4.3.

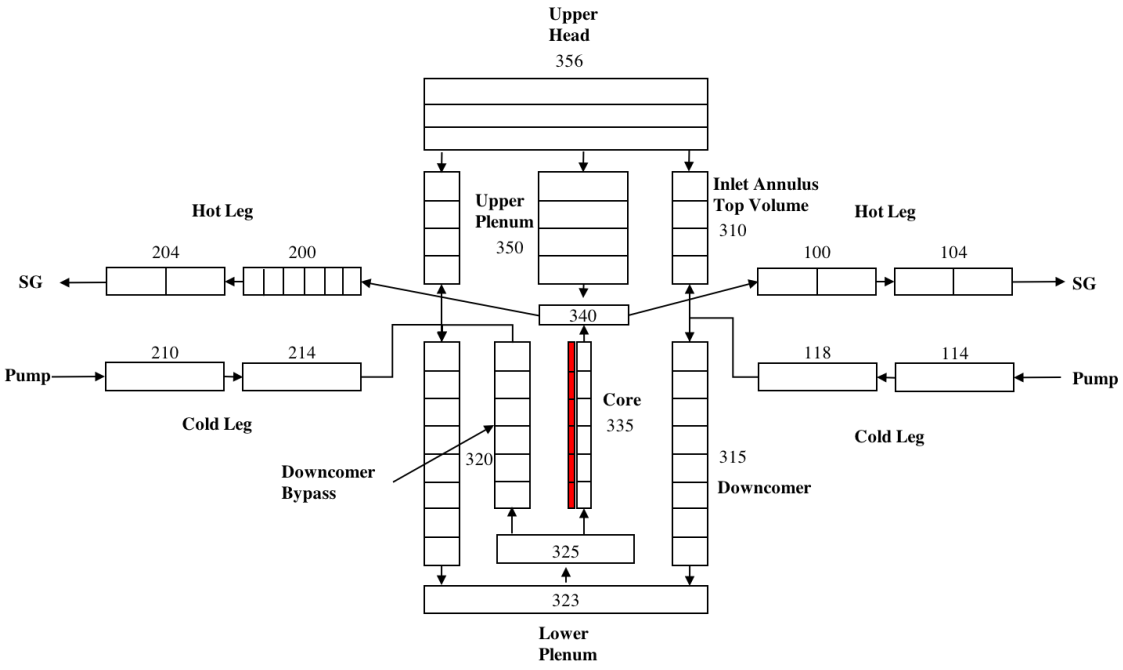

Figure 4.3: Reactor vessel nodalization diagram.

Since the SG is the focus of our study, Figure. 4.4 presents a nodalization diagram for the AP1000 SG. The SG primary side is modeled as a single U-tube with eight volumes with an equivalent flow area and volume for AP1000. The secondary side is modeled using a RELAP5 pipe component with six vertical volumes. The specific steam separator model in RELAP5 has been selected for the separator. The dryer and dome are modeled as a branch and a single volume, respectively. The main feedwater and auxiliary feedwater are modeled using a time-dependent junction and time-dependent volume, controlled through the input deck. A valve is used as the main steam isolation valve (MSIV), which separates the SG from the steam line to the turbine. The turbine is not modeled here. Instead, a time-dependent volume is used. All the passive core cooling system (PCCS) including the core make-up tanks, the accumulators, the automatic depressurization system, the in-containment refueling water storage tank, and the PRHR heat exchanger are all modeled. However, it is not the focus of our cyber-attack simulation study. We will not discuss the details of these components.

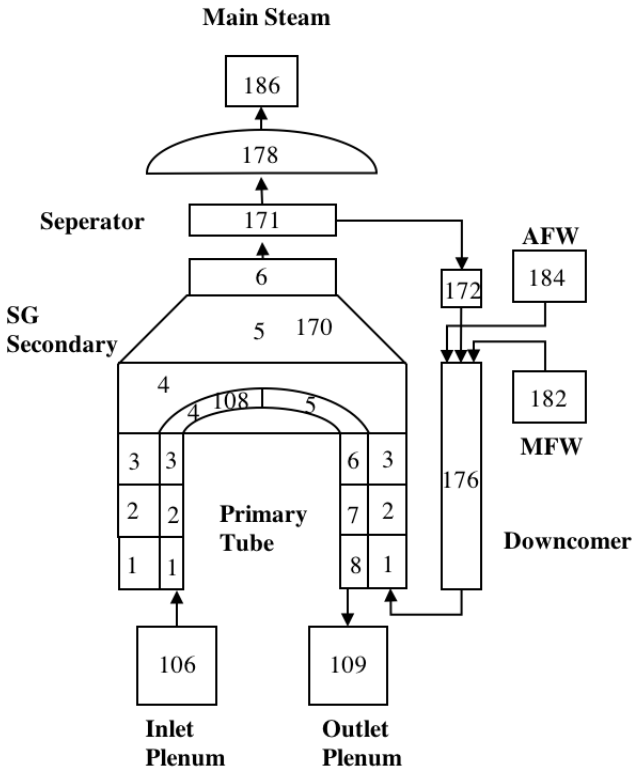

Figure 4.4: SG nodalization diagram.

Some simulation results are presented below. The steady-state conditions after running a null transient for 2000 s are reported in Table‘4.2. They closely match the design data. Figure. 4.5 presents the temperature rise across the core.

Table 4.2: AP1000 RELAP5 Simulation.

Parameter	**RELAP5**	**Design**
Core thermal power (MWt)	3415	3415
Core inlet flow rate (Mg/s)	14.778	15.17
Core outlet flow rate (Mg/s) (loop 1/2)	7.38/7.41	7.59
Vessel inlet temperature (K)	549.8	553.8
Vessel outlet temperature (K)	591.8	594.2
SG secondary feedwater temperature (K)	499.7	499.7
SG secondary steam flow (kg/s)	942.6	943.0

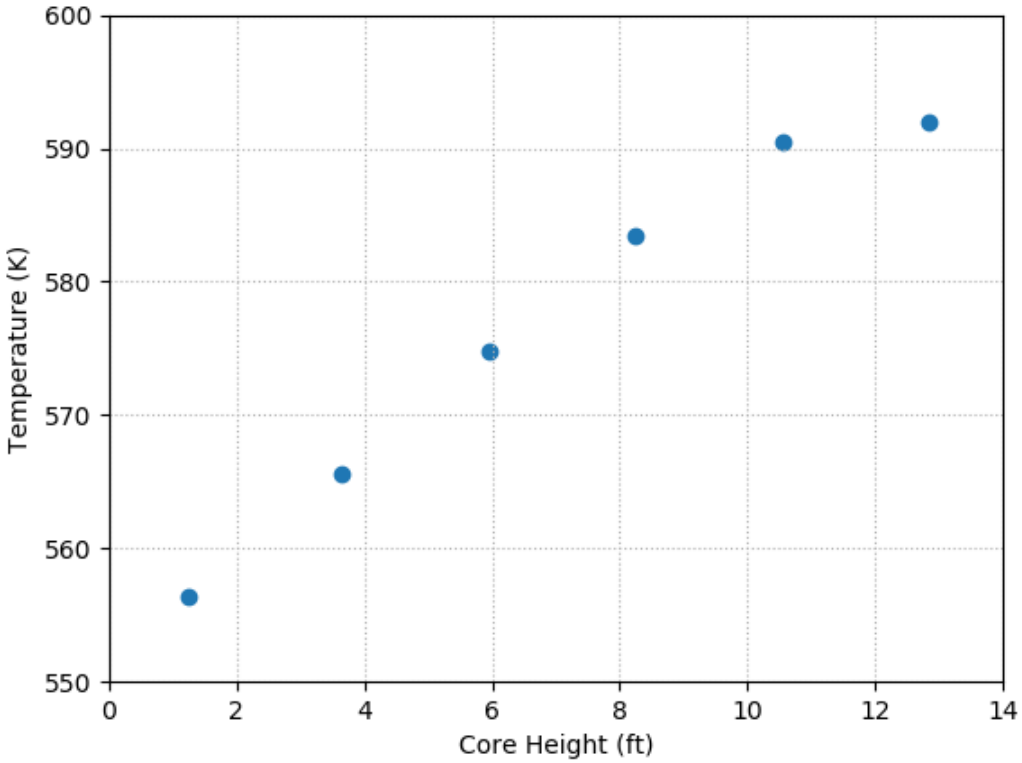

Figure 4.5: Temperature distribution along the core.

We note that there are small discrepancies between our RELAP5 calculation and the AP1000 design values for various parameters. However, it still satisfies the requirement for our demonstration of the reduced-order model and Kalman filtering based detection approach.

CHAPTER 5

Application Programming Interface Development

As stated in Chapter 4, the cyber-attack scenario simulation requires more than the capability of thermal-hydraulic simulation software. The cyber-attack scenarios involve the I&C system component and the evolution of the I&C system component status should be presented. Hence, an API for the I&C systems is developed as a platform to handle the cyber-attack scenario modeling, detection, and mitigation. The API is programmed in Python [45], which is an interpreted, high-level, general-purpose scripting language. The API can be divided into two separate parts: the first part deals with the RELAP5 software and ROM; the second part deals with the GPWR simulator. The API RELAP5 part can implement the control model of the system and allow the application of Kalman filter, providing optimal estimates of the system parameters. It relies on a ROM to provide simulation results for the system evolution, while observed surrogate data are extracted from RELAP5. The GPWR part can deal with the automatic control into the simulation and the selection and plotting of output data. More functions for GPWR are still underdevelopment.

5.1 API for the RELAP5 Code

To work with RELAP5 in the overall framework, an API has been developed to handle the interface with the evaluation module (EM), extract the physical quantities

simulated by RELAP5, and convert them into the appropriate format for the EM. The EM receives RELAP5 data representing the system, the ROM-simulated data, and control the RELAP5 simulation, providing it with control-related information generated by the EM, as illustrated in Figure 5.1.

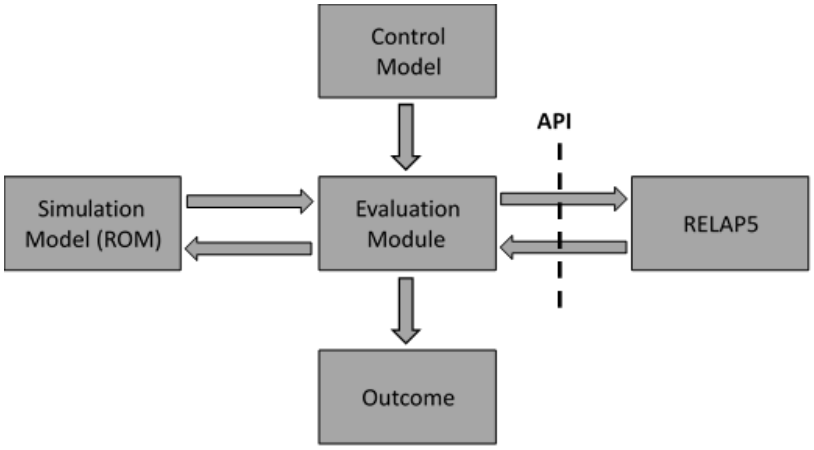

Figure 5.1: Framework architecture.

The EM evaluates the observed data and the simulated data generated by the ROM in a Kalman filter, providing the optimal estimate for the system parameters. It also houses the implementation of the control system model and manages both the RELAP5 and ROM simulations. This architecture allows the use of methods not implemented in RELAP5, e.g., Kalman filter or other such methods that the designer or analyst might implement in the future.

5.1.1 Operation of API

The API extracts all control-related information from the RELAP5 simulation and delivers it to the EM in a convenient form; it also takes all the control-related information generated or modified by the control model implemented in the EM and converts it to the RELAP5 input format. This is performed via two operations: setup and step. The approach to connect the API to RELAP5 is to act upon its input and output files. This approach has the advantage of not requiring access or modification

to the RELAP5 source code, guaranteeing full modularity and transferability between systems. The API is written in Python 3.7 to allow cross-systems portability, making use of its standard library and other open-source libraries.

The first operation executed by the API is the *setup* operation, which reads the original RELAP5 input file for the simulation and generates all the control objects used in the simulation, e.g., options controls, trip functions, and hydrodynamic components. It makes all the control objects available to the EM and sets up the simulation's initial environment. This operation is not mandatory, as all the relevant setup information can be directly entered into the EM, but it enables the use of preexisting RELAP5 input decks in the overall toolkit.

In the *step* operation, the EM sends a modified RELAP5 input file accounting for any changes in the state of the simulation. At every timestep, the EM directly inputs the value of the control variables used in RELAP5 by changing the input file to be used for that timestep. Then, the API will read the RELAP5 output file for that timestep, extract the physical quantities of the system, and present them to the EM, which then uses that information in its control model to generate new values of the control variables for the next timestep, until the simulation ends.

5.1.2 API Framework Functions

The two API operations, setup and step, are summarized in Figure 5.2. The setup operation is implemented using the functions `read_r5_input` and `build_objects`, described in Table 5.1. The step operation is implemented using the functions `write_r5_input` and `read _r5_output`, described in Table 5.2.

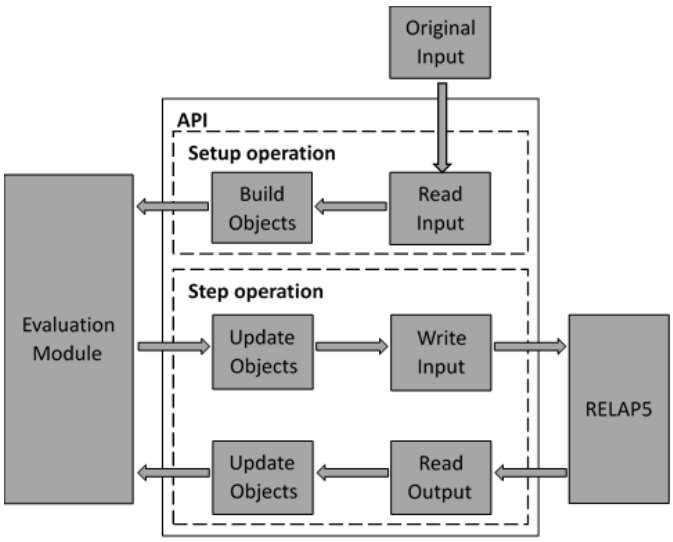

Figure 5.2: API operations with RELAP5 as the Simulation Module.

Table 5.1: Setup and operational function descriptions.

Function	**Description**
`read_r5_input`	Reads the input file, parses its lines, and creates a dictionary with a list of strings containing the arguments of each RELAP5 card, indexed by the card numbers.
`build_objects`	Takes each entry of the dictionary created by `read_r5_input`, assigns it to a class, and creates a components dictionary with all the control-related objects built, indexed by the cards numbers.
`build_trip_signals`	Returns a dictionary of trip signals used in RELAP5 implementation of the control model.
`build_trip_variables`	Returns a dictionary relating each trip signal to the variables used in its definition.
`build_minor_edits`	Reads the trip variable dictionary and appends the components dictionary with the necessary minor edits to retrieve information on trip variables.

Table 5.2: Step operation function descriptions.

Function	**Description**
`write_r5_input`	Writes a RELAP5 input file based on the components dictionary updated by the EM. Each component class has a specific method implemented for writing a list of strings with its features according to RELAP5 input syntax.
`read_r5_output`	Retrieves in the RELAP5 output file the relevant information about the objects in the components dictionary and updates the control objects.

The API models the control objects in classes, one for each type of control-related components that RELAP5 uses, such as timestep controls, trips, single junctions,

valves, and heat structures. This process occurs at every timestep of the simulation, allowing the EM to fully control the simulation run by RELAP5 and implement the intended control model.

5.1.3 API Testing

The preliminary API testing protocol was built over functioning PWR input files for RELAP5, representing both steady-state and transient conditions. Input files generated via the setup operation of the API, based on the original functioning input files, were executed by RELAP5, producing output identical to the original input deck.

An additional API verification was also performed. The FW mass flow rate was perturbed through the API and compared to the same perturbation accomplished by manually changing the input file, yielding identical results. Table 5.3 compares the API operation time with the RELAP5 CPU time, indicating that, for this simple problem, the API operation time is less than 1.0 s for two simple cases with different FW flowrates W_s.

Table 5.3: Breakdown of API run times for the RELAP5 SG model.

Simulation Time (s)	**CPU Time (s)**					
	$W_s = 802.6$ kg·s^{-1}			$W_s = 700.0$ kg·s^{-1}		
	RELAP5	**API**	**Total**	**RELAP5**	**API**	**Total**
25	11.033	0.859	11.892	14.206	0.878	15.084
50	21.144	0.924	22.068	20.838	0.868	21.706
75	31.186	0.949	32.135	31.267	0.866	32.133
100	39.984	0.836	40.784	41.492	0.843	42.335
125	52.276	0.850	53.126	52.272	0.895	53.167
150	62.283	0.848	63.131	64.441	0.864	65.305

5.1.4 Testing of Control Functions via API

One of the main functionalities of the API is to allow dynamic changes in RELAP5 simulations. In the first dynamic test case, the simulation started with $W_s = 802.6$

kg·s^{-1} for SG unit 1, followed by an abrupt change to 700 kg·s^{-1} when the simulation reached 100 s, then changed back to the initial value at 250 s. In the second scenario, the simulation started with $W_s = 700$ kg·s^{-1} for SG unit 1, and at 250 s was linearly increased to 802.6 kg·s^{-1} until 350 s, and then extends until the simulation reaches 400 s.

With enough time for the system to enter a steady state after the changes, the SG water level should approximate that calculated without dynamic interference from the API. Figure 5.3 indicates that the initial results are identical to the ones previously calculated. However, the transient converges to the same state after the changes are made, provided sufficient time is allowed to achieve an equilibrium.

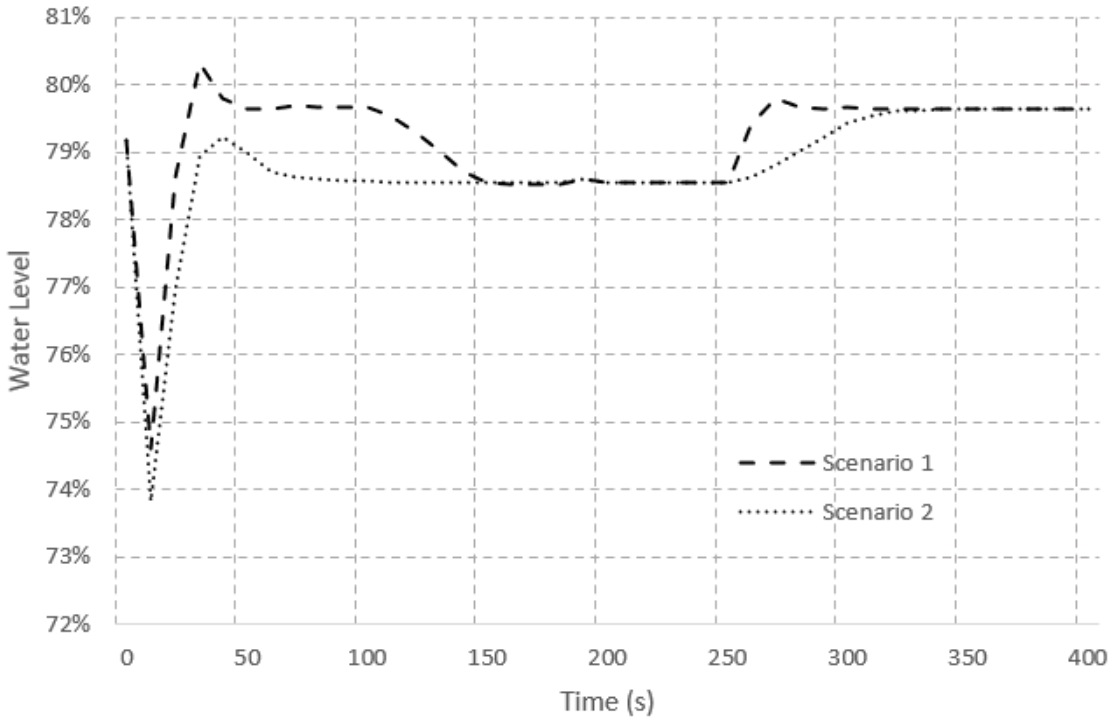

Figure 5.3: Dynamic changes in the RELAP5 simulation.

5.2 API for the GPWR

The GPWR simulator is a comprehensive simulator software with lots of useful functions already built in it. To simulate our cyber-attack scenarios automatically with GPWR and extract data for analysis, the API could be used effectively, although the API for GPWR is at an early stage of development. The design framework should be able to embed the API into the software, which can control the simulator and interact with the detection and mitigation approaches during the simulation.

5.2.1 Data Extraction

GPWR can save the trend data into a CSV format file. However, we need to import the data into a Python script with a well-organized data structure for future detection and mitigation studies. The API can read the output file from GPWR and let the user select and save the required data into the data frame in Python. Since the API does not connect directly to the source code at this point in our study, there is some limitation for it. It cannot provide real-time data extraction capability while the simulation is running.

The data analysis library pandas [46] is used. First, the API reads the CSV file from the GPWR simulator. Then, delete the irrelevant part of the output and put the relevant data into the pandas data frame. The API will use the console to present the name of each column. The user can type the parameters to be plotted in the console. Finally, the API will print out the plot and save the data file into a new data frame structure file for use with detection and mitigation studies.

CHAPTER 6

Reduced Order Steam Geneator Model

RELAP5 simulation can provide us with accurate results for various thermal hydraulic transient simulations. However, the run time could be an issue when we want to have online or real-time detection and mitigation. Hence, we have developed a reduced-order model (ROM) to resolve this problem. ROMs are simplifications of high-fidelity, complex models and capture the behavior of these source models so that engineers or analysts can quickly study a system's dominant effects using minimal computational resources. There are various complex mathematics models for the SG [47, 48]. In order to provide the system state and facilitate the detection of potential cyber-attacks, we developed a quantitative reduced-order model for the SG, which can represent the SG dynamics and provide the water level of SG given time-dependent data for the reactor power and feedwater flow rate. The model is built on the first-principle energy balance equation. It integrates through a boiling channel, which is coupled to the primary coolant channel by representing the heat flux via the average temperature of the primary coolant flow. To simplify the geometry, a single vertical boiling channel is considered to approximately represent the combination of co- and counter-current heat transfer processes inherent in a U-tube SG. All geometric and physical characteristics of the SG are obtained from the RELAP5 simulation of an AP1000 reactor model.

6.1 Model Order Reduction Techniques

There are several definitions of Model Order Reduction (MOR). Originally, MOR was developed in the area of systems and control theory, which studies properties of dynamical systems in the application for reducing their complexity, while preserving their input-output behavior as much as possible [49, 50]. Numerical mathematicians have also taken up the field. Nowadays, model order reduction is a flourishing field of research, both in systems and control theory and in numerical analysis. This simplification is needed in order to perform simulations within an acceptable amount of time and limited storage capacity, but with a reliable outcome. Model order reduction finds applications within all fields involving mathematical modeling, and many reviews exist for the topics relevant to electronics, fluid and structural mechanics. In our case, we would like to have online predictions of the system behavior with acceptable computational speed to be able to perform detection or mitigation related to cyber-attack scenarios.

The reduced-order model tries to capture the essential features of a structure quickly. Contemporary model order reduction techniques can be broadly classified into four classes: orthogonal decomposition methods [51], balancing method, and simplified physics [52] or operation-based reduction methods [53]. The first three methods fall into the category of projection-based reduction approaches. Projection-based reduction relies on the projection of either the model equations or the solution onto a basis of reduced dimensionality.

The simplified physics approach is used here for our steam generator problem. It can be described as analogous to the traditional mathematical modeling approach, in which a less complicated description of a system is constructed based on assumptions and simplifications using physical insights or otherwise derived information.

6.2 Development of SG Reduced Order Model

The first-principle energy balance equation is the basis of the reduced-order SG model. Recall the single-phase energy conservation equation Eq. (4.3). If fluid enthalpy $h = U + pV$ is introduced, and assume further that the kinetic energy, viscous heating, and gravitational potential energy are negligibly small, we obtain

$$\frac{\partial}{\partial t}(\rho h) = -\nabla \cdot (\rho h \mathbf{v}) - \nabla \cdot \mathbf{q} + \frac{\partial p}{\partial t} + S. \tag{6.1}$$

For a vertical channel in our model, we average over the channel with cross-section area A, wetted perimeter M, and mass velocity G, which yields

$$\frac{\partial}{\partial t}(\rho h) = -\frac{\partial}{\partial t}(Gh) + \frac{M q_s}{A} + \frac{\partial p}{\partial t} + S, \tag{6.2}$$

where the wall heat flux q_s is defined to be positive for heat flux into the secondary channel.

For a SG of length H, cross-sectional area A, and FW mass flowrate $W_s(t)$, the flow is divided into single- and two-phase regions separated by the water level at $z_0(t)$. The heat flux $q_s(z,t)$ is represented in terms of the fluid temperature $T_s(z,t)$, primary coolant temperature T_p, and effective heat transfer coefficient U_1 as

$$q_s(z,t) = U_1[T_p - T_s(z,t)]. \tag{6.3}$$

The primary side temperature is represented as the average of the primary inlet temperature $T_{p,in}$ and outlet temperature $T_{p,out}$

$$T_p = \frac{T_{p,in} + T_{p,out}}{2}, \tag{6.4}$$

assumed to be either constant or slowly varying for the relatively slow SG transients

of interest. For the single-phase region $z(t) < z_0(t)$, the energy balance can be derived from Eq. (6.2) with fluid density ρ_s and heat capacity C_s,

$$A\rho_s C_s \frac{\partial T_s(z,t)}{\partial t} = -W_s(t) C_s \frac{\partial T_s(z,t)}{\partial z} + MU_1[T_p - T_s(z,t)]. \tag{6.5}$$

The partial differential equation involving both time t and channel length z can not be solved analytically. Hence, we need to discretize Eq. (6.5) in t and z, then numerically solve for $T_s(z,t)$. For the purpose of discretizing the equation in time with the notation $T_s(z,t_n)$ and time step $\Delta t = t_n - t_{n-1}$, we can obtain

$$A\rho_s C_s \frac{T_s(z,t_n) - T_s(z,t_{n-1})}{\Delta t} = -W_s(t_n) C_s \frac{\partial T_s(z,t_n)}{\partial z} + MU_1[T_p - T_s(z,t_n)]. \tag{6.6}$$

Then, we can further diecretize the equation in vertical direction z with $\Delta z = z_m - z_{m-1}$. The equation becomes

$$\begin{aligned} A\rho_s C_s \frac{T_s(z_m,t_n) - T_s(z_m,t_{n-1})}{\Delta t} = & - W_s(t_n) C_s \frac{T_s(z_m,t_n) - T_s(z_{m-1},t_n)}{\Delta z} \\ & + MU_1[T_p - T_s(z_m,t_n)]. \end{aligned} \tag{6.7}$$

To solve this finite-difference equation, we can first solve it along the vertical direction at each time step, then using Crank-Nicholson [54] scheme updates $T_s(z,t_n)$ at each time step. Since the time step Δt for the time-dependent feedwater flow rate simulated with RELAP5 is 1.0 second, we use the same time step here for our differential equation. Since the water level changes as a function of time, we use a dynamic spatial mesh. We divided the vertical length from the bottom to the water level $z_0(t_n)$ into 10 equal intervals. We use $T_s(z_m,t_n)$ to calculate $T_s(z_m,t_{n+1})$ through Eq. (6.7). Then, compare the temperature at $T_s(z_0,t_{n+1})$ with the saturated temperature T_{sat}. If it is higher than the saturated temperature

$$T_s(z_0, t_{n+1}) > T_{sat}, \tag{6.8}$$

we decrease the mesh size at t_n, and get a new temperature distribution $T_s(z_m, t_n)$. If

$$T_s(z_0, t_{n+1}) < T_{sat}, \tag{6.9}$$

we increase the mesh size at t_n, and get a new temperature distribution $T_s(z_m, t_n)$. Repeat these steps until $T_s(z_0, t_{n+1})$ converges to T_{sat}.

For a steady-state flow, Eq. (6.5) reduces to

$$W_s C_s \frac{dT_s(z)}{dz} = M q_s(z) = M U_1 [T_p - T_s(z)], \tag{6.10}$$

which can be readily solved for $T_s(z)$ in terms of the secondary inlet temperature $T_{s,in}$. We integrate Eq. (6.10) over $[0, z]$, and obtain

$$T_s(z) = T_p + (T_{s,in} - T_p) \exp\left(-\frac{\gamma_1}{W_s C_s} z\right), \quad \gamma_1 = M U_1. \tag{6.11}$$

The water level z_0 may be determined by setting $T_s(z_0)$ to the saturation temperature T_{sat} via a homogeneous equilibrium model [37]

$$z_0 = \frac{W_s C_s}{\gamma_1} \ln\left(\frac{T_p - T_{s,in}}{T_p - T_{\text{sat}}}\right). \tag{6.12}$$

The effective heat transfer coefficient of U_1 and wetted perimeter M is obtained via the reference case of RELAP5 simulation. Then, the water level could be quickly calculated through the model. For the single-phase heat transfer coefficient U_1, we assume that the Dittus-Boelter correlation [37] for the Nusselt number is applicable, indicating a 0.8^{th}-power dependence on the mass flow rate W_s of feedwater.

For the two-phase region $z > z_0$ with $T_s(z)$ remaining at the saturation temperature T_{sat}, the steady-state energy balance is represented in terms of the FW enthalpy

$h_s(z)$ with effective two-phase heat transfer coefficient U_2

$$W_s \frac{dh_s(z)}{dz} = MU_2[T_p - T_{\text{sat}}], \tag{6.13}$$

yielding an expression for the exit quality x_e in terms of the latent heat of vaporization h_{fg}

$$x_e = \frac{1}{h_{fg}} \frac{\gamma_2}{W_s}(T_p - T_{\text{sat}})(H - z_0), \quad \gamma_2 = MU_2. \tag{6.14}$$

The two-phase coefficient γ_2 should conceptually include the effects of steam separation equipment in the SG and the exit quality x_e is assumed to be 100% in our approximate model.

The overall energy balance between the primary and secondary sides of the SG requires integration of the heat flux over the whole length H of the SG, which includes the single-phase heat transfer rate together with the two-phase region. The total heat transfer rate into the SG is equal to the total power produced in the core

$$P_{SG} = W_s[C_s(T_{\text{sat}} - T_{s,in}) + x_e h_{fg}]. \tag{6.15}$$

6.3 Validation of ROM with RELAP5 Simulation

To set up the reduced order SG model, we performed a RELAP5 simulation for AP1000 reactor as the reference case. The relevant system parameters are summarized in Table 6.1 at full power $P = 3.42$ GWt and the FW flowrate $W_s = 942.6$ kg·s^{-1}.

Table 6.1: Parameters for the reference case.

Parameter		**Parameter**	
H (m)	13.78	T_{sat} (K)	535.7
W_s (kg·s^{-1})	942.6	C_s (kJ·kg^{-1}K^{-1})	4.99
$T_{p,in}$ (K)	549.8	z_o (m)	4.81
$T_{p,out}$ (K)	591.8	h_{fg} (MJ·kg^{-1})	1.65
$T_{s,in}$ (K)	499.7	-	-

The ROM is validated by running the same scenario as RELAP5 and comparing the water level variation over time. A disturbance is introduced at the beginning of the simulation, resulting in variations of the water level. A new equilibrium level is reached without a significant time delay. Results from RELAP5 and the static and dynamic ROMs are compared in Figure 6.1. The blue line is the RELAP5 result, compared with the orange and green lines representing the static and dynamic ROM results, respectively. We observe that the time-dependent ROM follows the RELAP5 variations reasonably well, apart from a small discrepancy in the asymptotic or equilibrium water level. This difference is primarily due to approximations in the heat transfer coefficient model. The water level obtained from the static ROM result of Eq. (6.12) agrees well with the asymptotic dynamic ROM result, as it should.

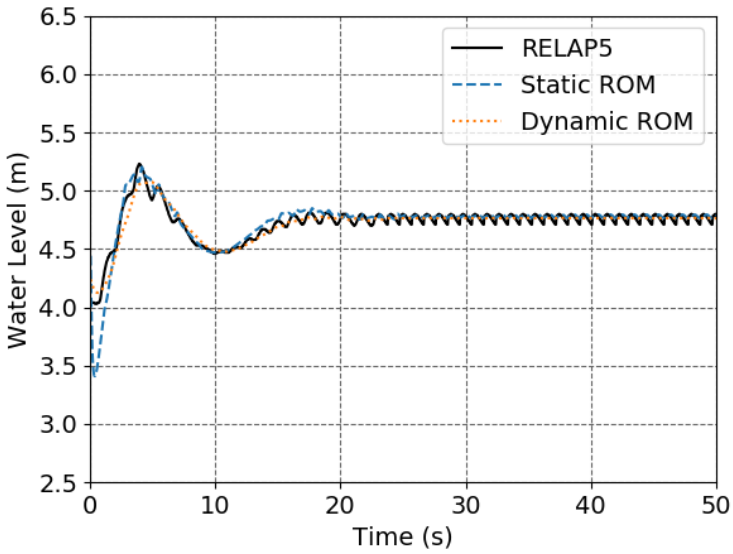

Figure 6.1: Water-level comparison between RELAP5 and ROM.

Three additional cases are simulated, with the FW flowrate varied at the beginning of the transient to represent different disturbances to the SG. Figure 6.2 illustrates the comparison of the water level between the RELAP5 simulations and static ROM estimates. Numerical data corresponding to the comparison in Figure 6.2 are summarized in Table 6.2.

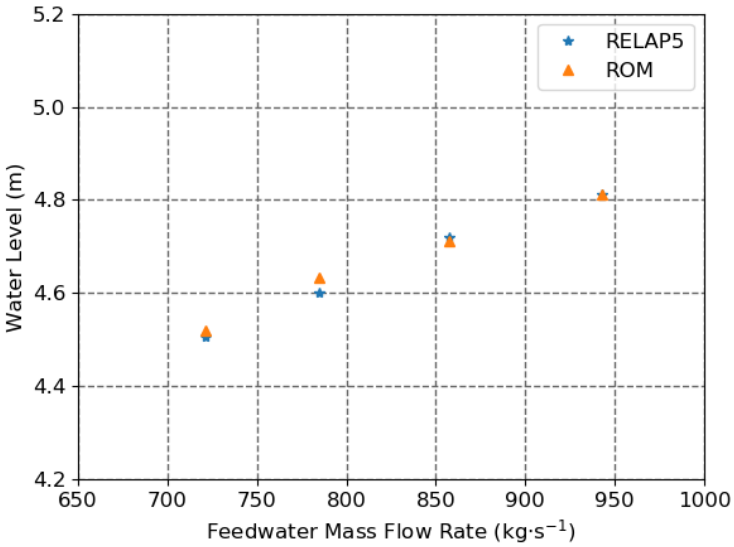

Figure 6.2: Static water level comparison between RELAP5 and ROM.

Table 6.2: Comparison of RELAP5 results and ROM calculations.

Case	W_s (kg·s^{-1})	Power(GWt)		γ_1 (MW·m^{-1}K^{-1})		γ_2 (MW·m^{-1}K^{-1})		z_0 (m)	
		RELAP5	**ROM**	**RELAP5**	**ROM**	**RELAP5**	**ROM**	**RELAP5**	**ROM**
1	942.6	3.42	3.44	0.681	0.681	4.90	4.90	4.81	4.81
2	857.3	3.20	3.16	0.630	0.632	4.41	4.46	4.72	4.71
3	784.7	3.00	2.92	0.593	0.588	4.00	4.10	4.60	4.63
4	721.2	2.80	2.70	0.552	0.550	3.59	3.72	4.51	4.52

The validation cases indicate that the ROM provides sufficiently accurate predictions of the total thermal power and water level as a function of the feedwater flow rate W_s. Hence, we expect the ROM to be used effectively in the future development of the detection algorithms.

CHAPTER 7

Kalman Filter Algorithm for Detection

Simulation of the cyber-attack scenarios is not the end of the research. We want to go further to detect the occurrence of cyber-attacks in the I&C system. The direct measurement from the system will suffer from noise, which cannot be effectively used for detection. Kalman filter [55] is an algorithm that uses a series of measurements, containing statistical noise and other inaccuracies. It produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone. Hence, we would like to use this characteristic of the Kalman filter to develop a detection approach for cyber-attack problems. We can combine the measurement from the nuclear power plant with our real-time ROM to provide an optimal estimate of a dynamical system subject to modeling uncertainties and inherent observation errors. Any observation of the plant parameters indicating a significant deviation from the optimal Kalman filter estimates can signal a potential cyber-attack in the system.

7.1 Kalman Filter Algorithm

Kalman filtering is an algorithm that provides estimates of some unknown variables given the measurements observed over time. Kalman filtering has demonstrated its usefulness in various applications, such as guidance, navigation, and control [56].

There is various research discussing applications of Kalman filter in the detection problems in other areas [57, 58]. Kalman filters have relatively simple forms and require small computational power, which can satisfy the real-time requirement in our cyber-attack detection.

The Kalman filter is a minimum-variance parameter estimation algorithm that generates an optimal estimate of system state vector $x(t)$ given observation vector $y(t)$, accounting for modeling uncertainties for $x(t)$ and statistical fluctuations in $y(t)$. The optimal estimate $\hat{x}(t)$ is obtained so that the covariance of the system estimation is minimized. Consider a dynamical system represented by $x(t)$ subject to white Gaussian noise vector $w(t)$ with covariance Q,

$$\frac{dx(t)}{dt} = F(t)x(t) + w(t), \langle w(t) \rangle = 0, \left\langle w(t)w^T(t') \right\rangle = Q\delta(t-t'), \tag{7.1}$$

where $x(t)$ is determined indirectly through observation $y(t)$ subject to white Gaussian noise vector $v(t)$ with covariance R,

$$\frac{dt(t)}{dt} = M(t)x(t) + v(t), \langle v(t) \rangle = 0, \left\langle v(t)v^T(t') \right\rangle = R\delta(t-t'), \tag{7.2}$$

The optimal system estimate $\hat{x}(t)$may be considered a statistical expectation of the exact system state $x(t)$ given observation $y(t)$,

$$\hat{x}(t) = \langle x(t)|y(t) \rangle , \tag{7.3}$$

such that the covariance matrix

$$P(t) = \left\langle [x(t) - \hat{x}(t)] [y(t) - \hat{x}(t)]^T \right\rangle \tag{7.4}$$

is minimized. For system diagnosis application, the basic formualation for a discretized linear Kalman filter is summarized. The state transition matrix is defined

over the time interval $[t_{k-1}, t_k]$,

$$\Phi(k|K-1) = \exp\left[\int_{t_{k-1}}^{t_k} F(t)dt\right], \tag{7.5}$$

so that Eq. (7.1) can be discretized as

$$x(k) = \Phi(k|k-1)x(k-1) + w(k) = \Phi x(k-1) + w, \tag{7.6}$$

and the measurement equation Eq. (7.2) is similarly discretized,

$$y(k) = M(k)x(k) + v(k) = Mx(k) + v. \tag{7.7}$$

The covariance matrix of Eq. (7.4) may be written for time steps $k-1$ and k as

$$P(k-1) = \left\langle [x(k-1) - \hat{x}(k-1)] \left[y(k-1) - \hat{x}(k-1)\right]^T \right\rangle, \tag{7.8}$$

$$P(k) = \left\langle [x(k) - \hat{x}(k)] \left[y(k) - \hat{x}(k)\right]^T \right\rangle. \tag{7.9}$$

The Kalman filter is a recursive estimator, which means the estimate for the current state needs the estimated state from the previous time step and the current measurement. We can also conceptualize it as two steps: predict and update. Before the measurement at time step k is taken, the prior state and variance are estimated

$$\hat{x}^-(k) = \hat{x}^-(k|k-1) = \Phi\hat{x}(k-1), \tag{7.10}$$

$$P^-(k) = \Phi P^-(k-1)\Phi^T + Q(k). \tag{7.11}$$

In the update step of the Kalman filter, after a new measurement is taken at step k, the objective is to add to the prior estimate of Eq. 7.10 a term proportional to the measurement residual,

$$\xi(k) = y(k) - M\hat{x}^{-}(k), \tag{7.12}$$

so that the resulting posterior estimate

$$\hat{x}(k) = \hat{x}^{+}(k) = \hat{x}^{-}(k) + K\left[y(k) - M\hat{x}^{-}(k)\right] \tag{7.13}$$

minimizes the estimation error

$$\epsilon(k) = x(k) - \hat{x}^{(}k) \tag{7.14}$$

or equivalently the covariance $P(k)$ of Eq. (7.9). An alternate form of measurement residual $\xi(k)$ can be obrained via Eq. (7.6) and Eq. (7.7),

$$\xi(k) = Mx(k) + v(k) - M\hat{x}^{-}(k) = M\left[\Phi x(k-1) + w\right] + v(k) - M\hat{x}^{-}(k), \tag{7.15}$$

which can be rewritten with $\epsilon(k)$ at time step $k-1$:

$$\xi(k) = M\left[\Phi\epsilon(k-1) + w\right] + v. \tag{7.16}$$

Substituting the terms in Eq. 7.14 provides a more useful form:

$$\epsilon(k) = (I - KM)[\Phi\epsilon(k-1) + w] - Kv. \tag{7.17}$$

The posterior estimate of the covariance matrix becomes

$$P(k) = P^{+}(k) = (I - KM)P^{-}(k)(I - KM)^{T} + KRK^{T} \tag{7.18}$$

Minimization of the posterior covariance matrix $P(k)$ may be accomplished by taking a derivative of the trace of $P(k)$ with respect to K and setting it to zero, which finally

provide the Kalman gain matrix at time step k:

$$K(k) = P^{-}(k)M^{T}\left[MP^{-}(k)M^{T} + R\right]^{-1}. \tag{7.19}$$

Then, we can yield an simpler form of the posterior covariance matrix,

$$P(k) = P^{+}(k) = (I - KM)P^{-}(k). \tag{7.20}$$

The flow of information for the Kalman filter algorithm is illustrated in Figure. 7.1.

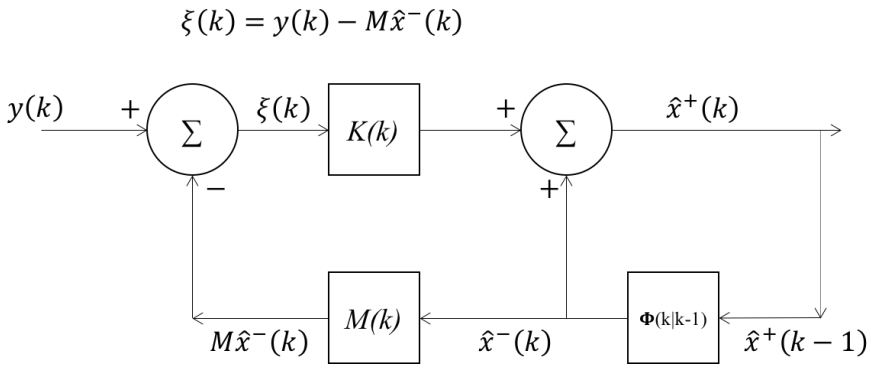

Figure 7.1: Flow of information for the Kalman filter.

7.2 Application for Diagnosis

We now build a diagnosis approach based on Kalman filtering, with the measurement data and simulation results to obtain an optimal estimation of the system state. An observed plant state having a significant deviation from the optimal system estimate could then indicate potential intrusion into the system. In our demonstration, using the ROM presented in Chpater 6, we obtain an optimal estimate of the water level z_0 in the SG. In this case, we treat the RELAP5 result as the observation data subject to some statistical fluctuations and the ROM as the simulation subject to uncertainties. RELAP5 is a deterministic model, so a normal distribution noise is added to the RELAP5 results to represent realistic observation data for our demonstration

calculations.

A transient case involving a SG water level variation is simulated to demonstrate the applicability of Kalman filtering. The system state $x(k)$ is the water level from ROM calculation, which is divided into a constant term and a dynamic term

$$x(k) = \begin{bmatrix} z_{0,\mathrm{ROM}}(k) \\ z_{0,\mathrm{RELAP5}}(k) - z_{0,\mathrm{ROM}}(k) \end{bmatrix}, \tag{7.21}$$

and the measurement $y(k)$ is the simulation data from RELAP5. The state transition matrix and the measurement matrix are

$$\Phi = \begin{bmatrix} 1 & 0 \\ 0 & e^{-\Delta t} \end{bmatrix}, \tag{7.22}$$

$$M = \begin{bmatrix} 1 & 1 \end{bmatrix}. \tag{7.23}$$

The result is shown in Figure 7.2, where both the ROM and RELAP5 data are assumed subject to 1% uncertainty or noise. The feedwater flow rate is modified at the beginning of the simulation, resulting in the SG water level change. The solid line in the figure represents the optimal estimate obtained through the process. The noise from both the measurement and the simulator data are taken into consideration. Hence, we can use Kalman filtering as a diagnosis method regardless of the noise issue.

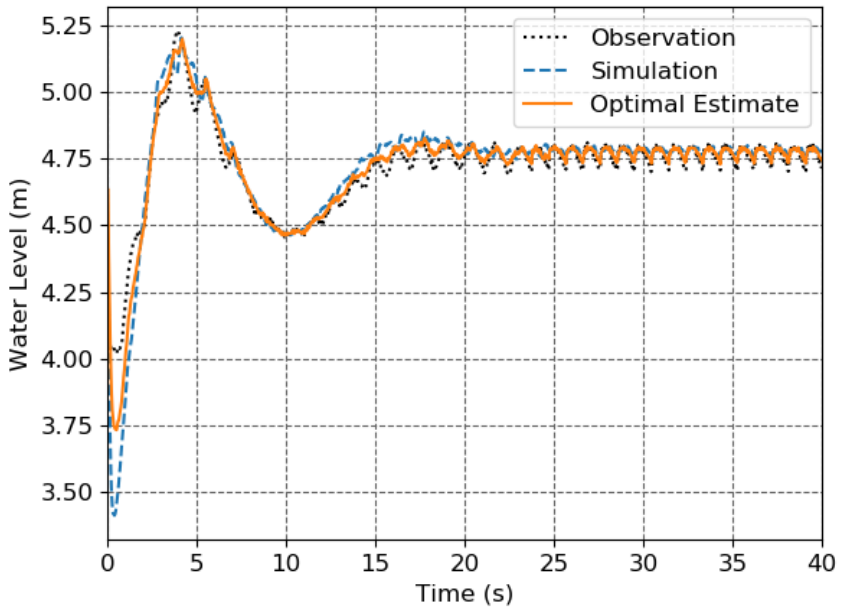

Figure 7.2: Kalman filter demonstration.

With the unique relationship connecting water level z_0, FW flow rate W_s, and reactor power level P, a 3-D plot with these three SG parameters is shown in Figure 7.3. The best estimate from the Kalman filter and the combined error from observation and simulation are plotted in the figure. Any observation of the plant parameters indicating a significant deviation from the Kalman filter estimates in the 3-D space could signal a potential cyber-attack on the SG system. This Kalman filter approach can be expanded to other key parameters as well for the cyber-security diagnosis in general.

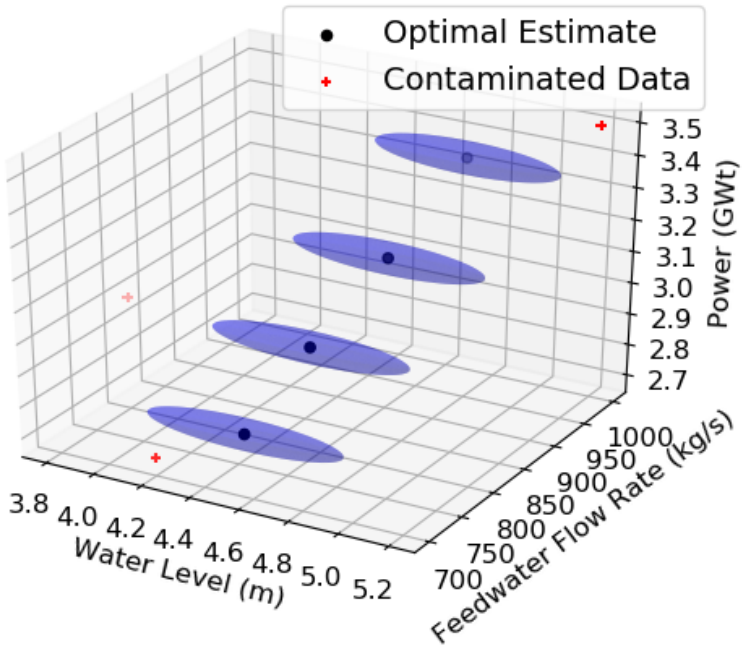

Figure 7.3: Illustration of Kalman filtering for SG parameter space.

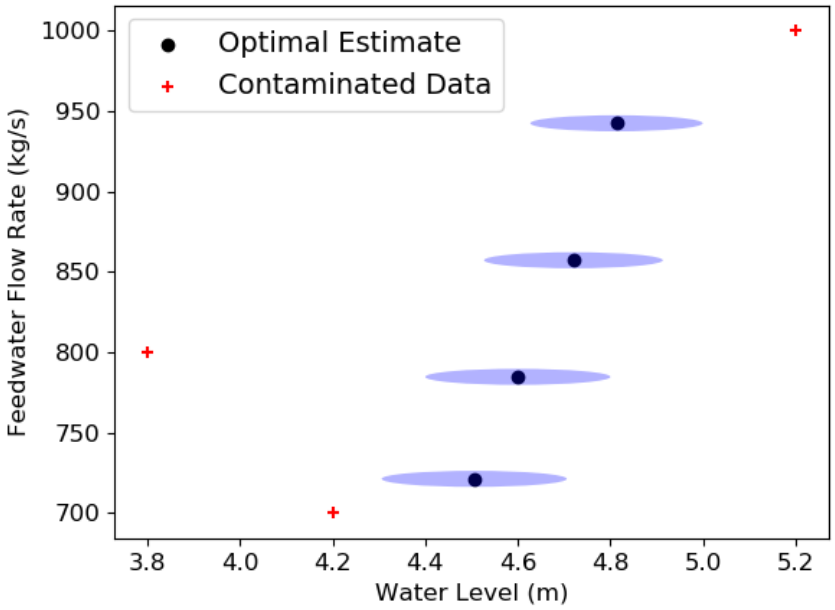

Figure 7.4: 2-D projection of SG parameter space.

A 2-D projection to the SG water level and feedwater flow rate phase has a clear visualization of the relationship between the optimal estimate and the contaminated data. In Figure 7.4, we can observe that the contaminated data falls out of the optimal estimate error range.

CHAPTER 8

Attack Scenario Simulation Using GPWR

The GPWR simulator is newly installed at the University of Michigan, which provides a comprehensive simulation capability, including the control room panel and some I&C system. It has been applied in various research areas related to human factors [59, 60] and cyber threats [61]. We use GPWR to perform more detailed simulations on attack scenarios. The GPWR simulator is introduced in this chapter primarily related to its operation and function. The dominant cyber-attack scenarios we identified in Chapter 3 are also simulated though the GPWR simulator. The results are presented and analyzed, providing us new insight into the controller involved cyber-attack.

8.1 GPWR Simulator Introduction

Generic Pressurized Water Reactor (GPWR) is a full-scope nuclear power plant simulator acceptable for U.S. NRC licensed operator training and requalification. It includes high fidelity models that allow full plant operation, including Normal Operations, Abnormal Operations, and Emergency Operations as required by ANS-3.5. The simulator response has been validated against actual operating plant data. It is a comprehensive package with several components handling different tasks. The core platform is SimExec with the real-time executive environment for all other compo-

nents.

GPWR includes the RETACT model, a thermal-hydraulic engineering analysis tool in real-time representing the primary system. The nuclear reactor kinetics is calculated by the REMARK model, a real-time multi-group advanced reactor kinetics software, with the JTOPMERET model representing the balance of plant (BOP). It provides real-time execution of nuclear steam supply system (NSSS) and BOP simulations with real-time plant response, audible alarms, switch positioning, and relay activation. There is also a control panel for the operator to manipulate. We can access all relevant models, including gauges, recorders, controllers, actuators, and the built-in piping and instrumentation diagram (P&ID) provides dynamic interaction during a simulation. The simulator also comes with the source code in Fortran, which allows system modifications in the future.

8.1.1 GPWR Operation and Features

An overview of GPWR operation flowchart is shown in Figure 8.1. There are two routes to manage the GPWR operation. One is to monitor and control through JADE operator station (JOS), the other is to simulate through JADE Graphical Simulator.

The JOS is for displaying soft panel drawings. Using JOS, we can manipulate the devices, and monitor the responses through the graphical models. Figure 8.2 is a control room panel overview. There are nine primary panels at the front of the control room and a couple of accessory panels on the side, which is the typical design for the NPP control room.

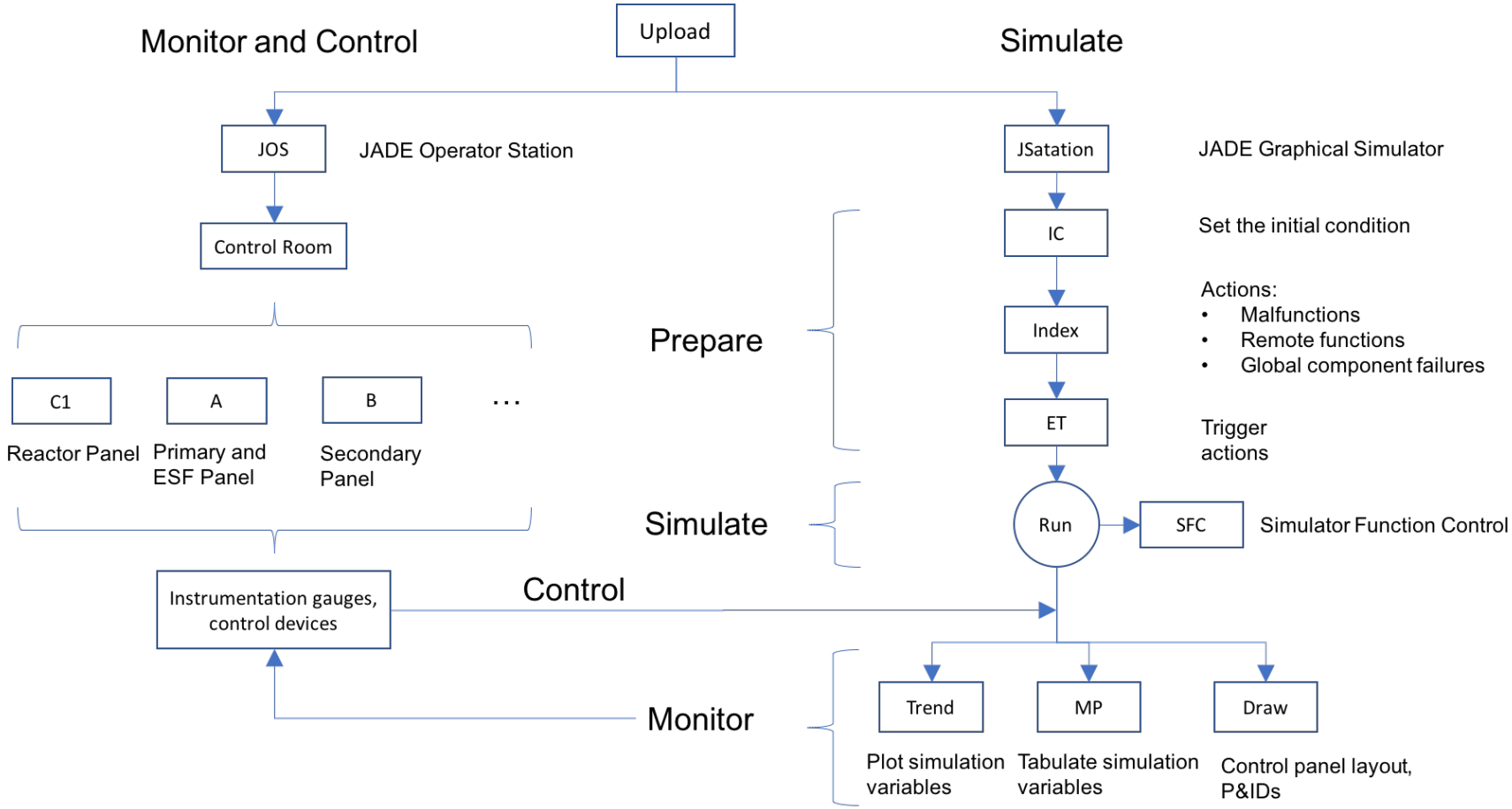

Figure 8.1: GPWR flowchart.

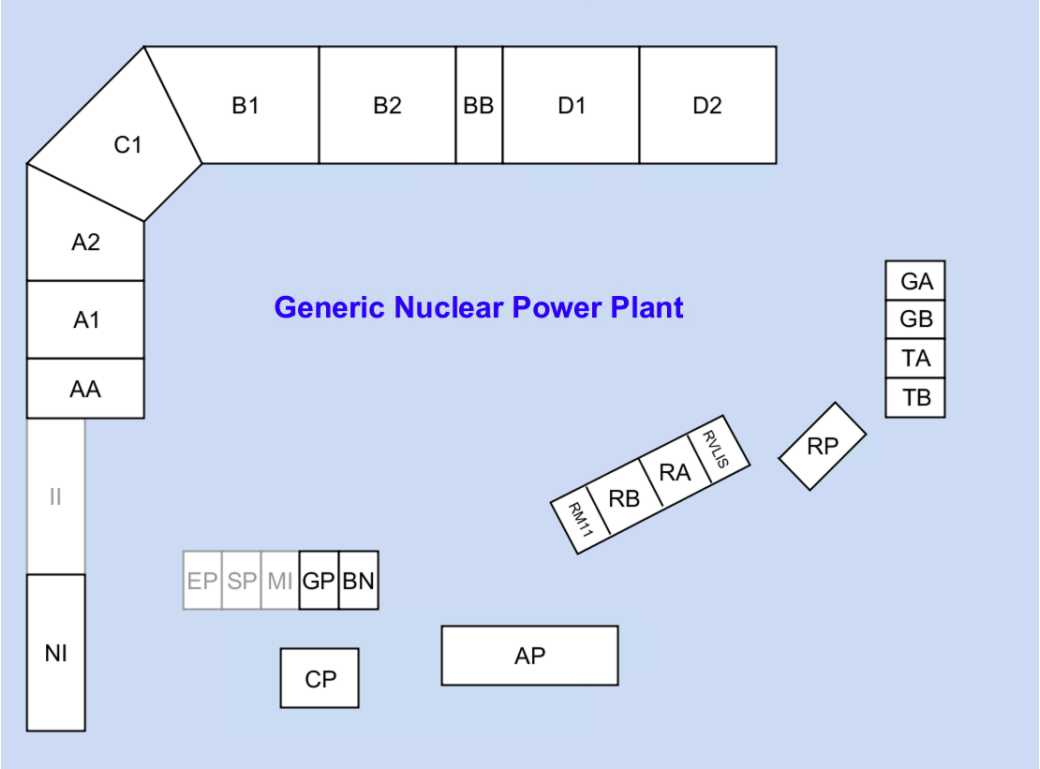

Figure 8.2: Control room overview.

The nine main control panels can be divided into four sections. Section A covers primary loop control and ESF actuation. Section B covers secondary loop control. Section C is the main panel for the overall control of the reactor. Section D covers the heating, ventilation, and air conditioning system.

Parts of panel C1 are captured in Figures 8.3 and 8.4. Panel C1 is the main panel, including most of the important parameters and control. Figure 8.3 shows the alarm system. We observe the status of reactor trip functions from this section. The block at the top left corner is the alarm board for the reactor coolant system (RCS). There are various RCS alarms, such as the RCS loop A/B/C low flow alert, the RCS loop A/B/C low average temperature alert. The gauges at the bottom left corner are the temperature indicators for all three loops. More gauges displaying critical components statuses, such as the reactor coolant pump, neutron detector, appear in this panel. Figure 8.4 shows the control rod step counters and reactor trip switch.

The control rod can be inserted or withdrawn manually from this panel, as well as the manual reactor trip. The control rod and shut down rod indicators show the positions of control rod banks and shut down rod banks. Any rod bank can be maneuvered through the rod motion switch with the rod bank selector. All instrumentation gauges and control devices are accessible from JOS, allowing for control and monitoring of the simulation process from the operator perspective.

Another path is to manage the reactor operation from the JStation platform. Among the important features in JStation are three functions mostly related to cyber-attack simulations: (a) the action index function (b) the dynamic piping and instrument diagram (P&ID), and (c) the trend function. The action index function activates malfunctions, remote functions, I/O overrides, global component failures, external parameters, fixed-parameter overrides, and annunciator overrides to manipulate the simulator.

The P&IDs are dynamic, displaying current parameter values and status of equipment. They can also provide access to actions associated with the components and equipment. There are hundreds of P&ID diagrams covering multiple systems in the GPWR simulator. For the steam generator low-level trip system, the P&ID can display each SG water level sensor's status during the simulation and the logic of how the SG sensor will initiate the reactor trip. The status of the SG water level sensor can also be manipulated through this dynamic P&ID, which is a useful tool for our cyber-attack simulations.

The trend function can plot, collect, and save the values from the simulation for performance analysis. We can set up a list of parameters for monitoring and collect from more than 1200 parameters in the nuclear plant model. With the powerful simulator in hand, we could do cyber-attack scenarios simulation based on our analysis before.

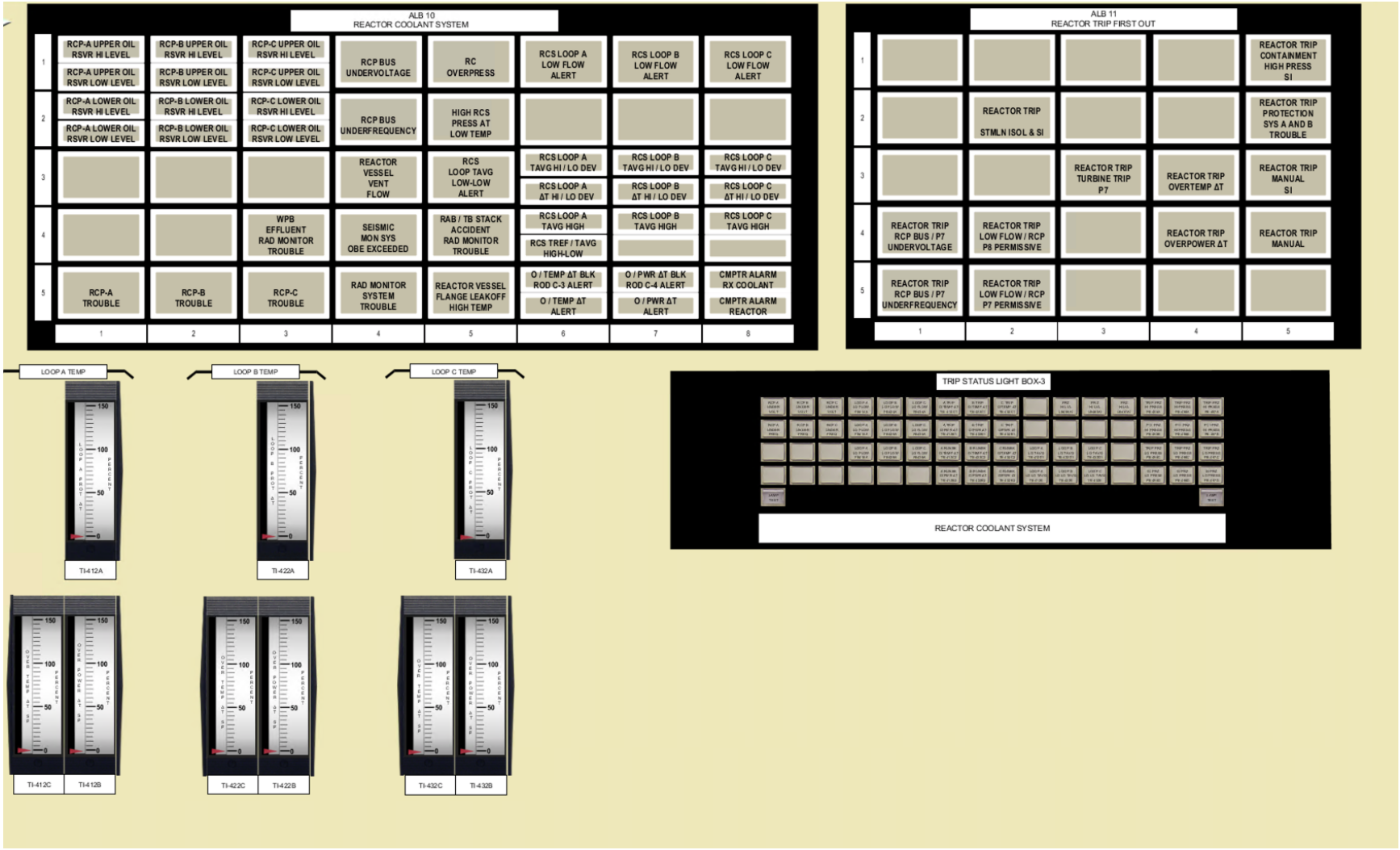

Figure 8.3: C1 panel example 1.

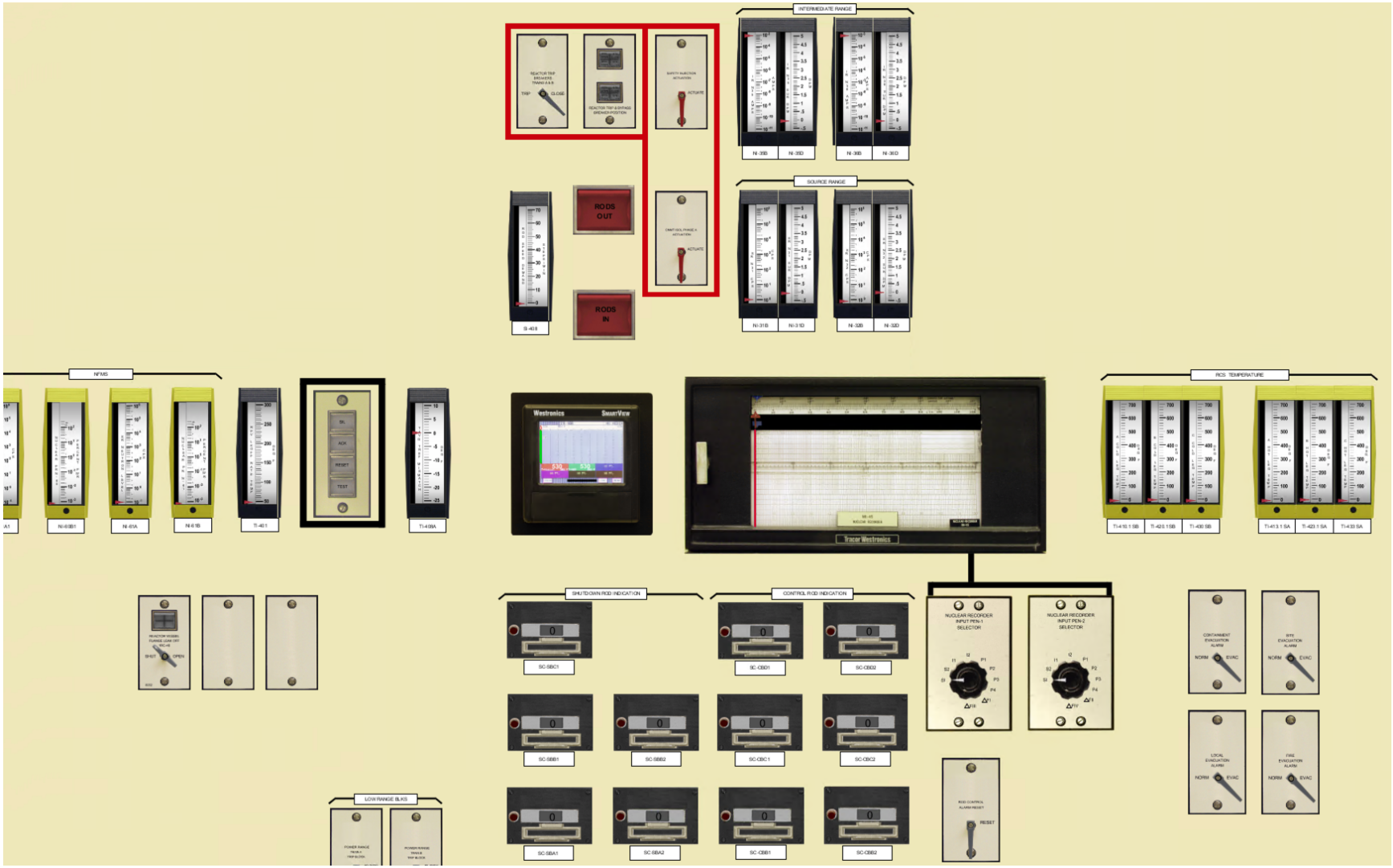

Figure 8.4: C1 panel example 2.

8.2 Simulation of Cyber Attack on Sensor and Controller

As discussed in Chapter 2, there is a four-channel sensor system used in the I&C system of GPWR to monitor each SG water level. One of them (Channel 3) serves as the controller input for the SG water level control system. Three of them (Channel 1, 2, and 3) function as monitors to initiate the low water level reactor trip. All four channels supply the signal to initiate a high steam generator level trip. Based on this design, we may consider a couple of different attack modes that can result in a reactor trip.

8.2.1 Direct Low SG Water Level Trip

Figure 8.5 demonstrates a direct low SG water level trip, which illustrates a corruption of channel 2 at 15 s. It is set to a lower value of 20%, which is below the set point for the reactor trip. However, a single channel corruption will not initiate a reactor trip or result in any change in the system. Following that, channel 4 is corrupted at 35 s and set to the same value as channel 2. Since channel 4 is not in the logic of a low-level steam generator trip, the corruption of these two channels will not result in a reactor trip. Eventually, channel 1 is corrupted at 50 s, which satisfies the low-level SG trip condition. If two out of three channels among channels 1, 2, and 3 are below the set point, the reactor is tripped. Finally, the reactor is tripped immediately right after the channel 1 suffers corruption as expected.

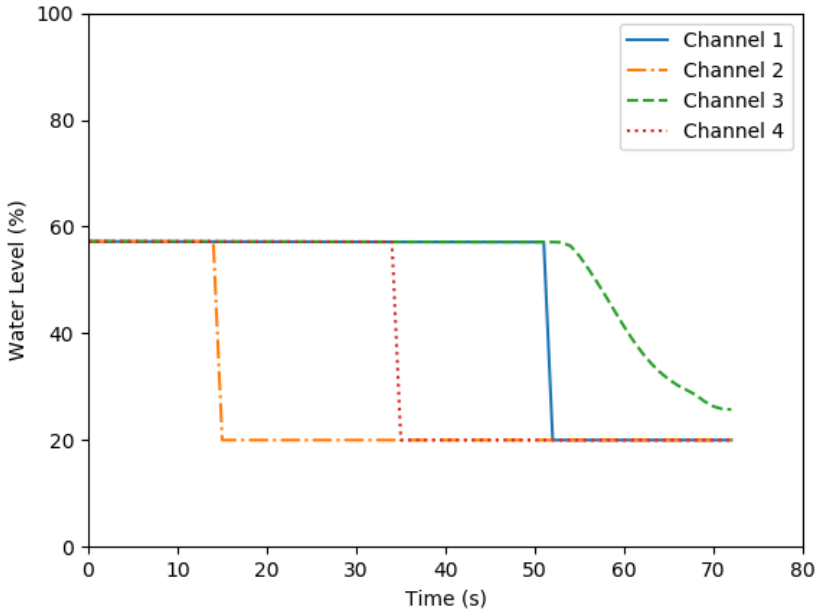

Figure 8.5: Low SG water level trip example.

8.2.2 Direct High SG Water Level Trip

Unlike the condition for low SG water level trip, high SG water level trip requires that any two out of all four channels are above the set point. Figure 8.6 demonstrates a direct high SG water level trip. We observe that a single channel corruption will not result in a reactor trip, but any additional channel corruption will trip the reactor as shown.

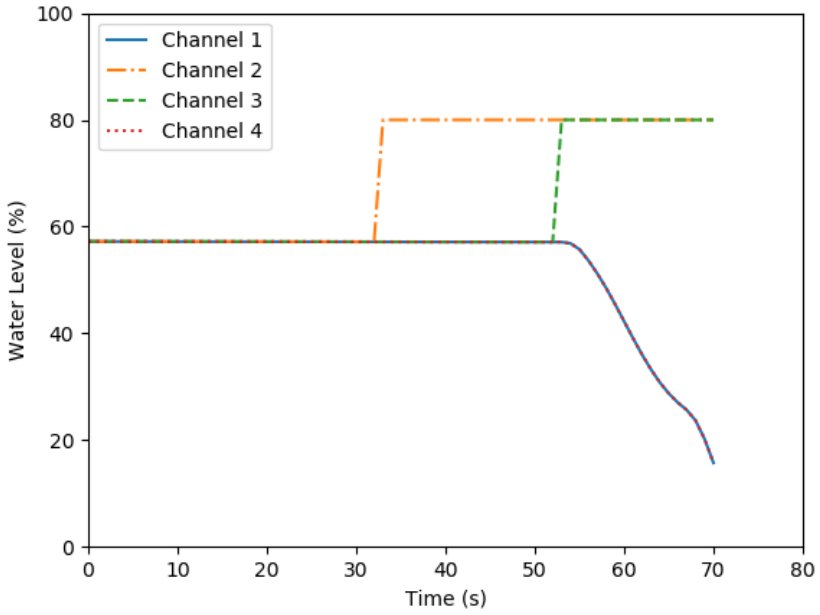

Figure 8.6: High SG water level trip example.

In the two cases considered involving direct low SG and high SG level trips, the SG state is not actually perturbed from normal operation state. The cyber-attack would cause corruption and perturbation only in the I&C system control logic, which nonetheless results in a reactor trip.

8.2.3 Controller Feedback SG Water Level Trip

We now turn our attention to the perturbation that would involve the SG controller. Channel 3 supplies the water level signal to the water level control system. Hence, the disturbance on channel 3, which will not trip the reactor immediately, but will feed the wrong signal to the feedback control loop. In time as the PI controller comes into action, the system output could be perturbed, which results in the change of other water level channels and a reactor trip.

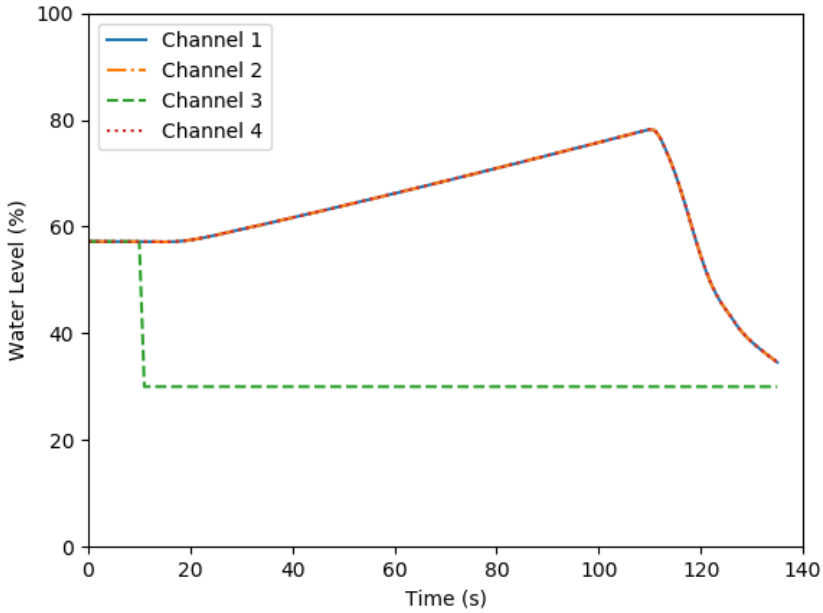

Figure 8.7: Controller Feedback on high SG Water Level Trip.

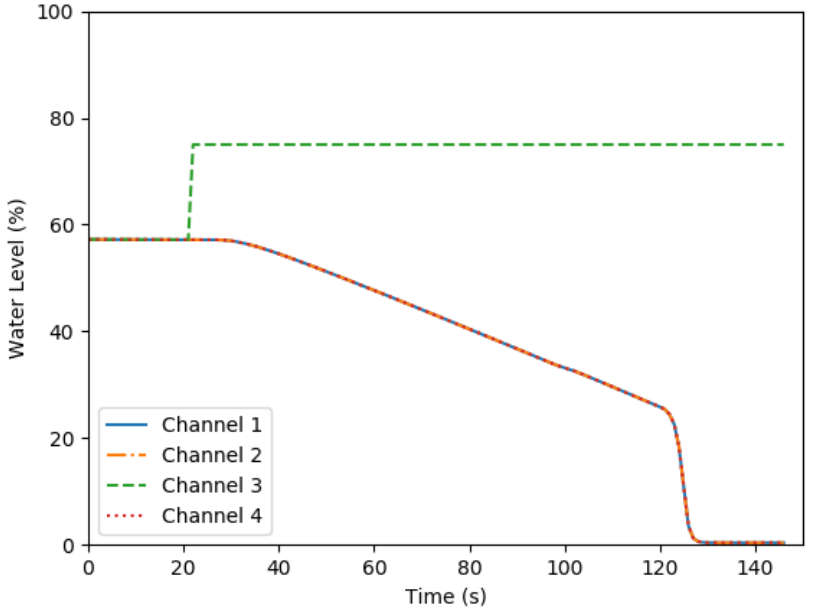

Figure 8.8: Controller Feedback on low SG Water Level Trip.

Figures 8.7 and 8.8 demonstrate the phenomenon we described above. In Figure 8.7, the attacker spoofs the channel 3 water level sensor and sets it to a different

value, which does not reflect the system state for the SG. The controller will take this information to adjust the regulating valve to compensate for this low-level sensor value. This results in a change in the system state and the increase of the output water level, shown in the other three channels level sensor. This process takes some time until the un-attacked channels reach the high level set point. Finally, it will initiate a reactor trip. Figure 8.8 illustrates a similar situation, but with a perturbation introduced to channel 3 sensor, causing the controller to reduce the actual water level in the remaining three channels and eventually resulting actually in a low SG level and a reactor trip.

In this kind of controller feedback trip, the spoofed value does not need to deviate significantly from the true system. A small deviation can still result in a reactor trip. However, the larger the perturbation on the controller input is, the faster the reactor will trip. Since this type of attack can change the actual system state, which is more severe than the scenarios considered for Figure 8.5 and 8.6, we would like to study the controller mechanism and propose an approach to mitigate this kind of attack in Chapter 9.

CHAPTER 9

Cyber-attack Mitigation with Controller feedback

Once the deviation from normal operating state due to cyber-attack is detected, we need to consider a mitigation method that can prevent the unnecessary reactor trip resulting from the attack. As we discussed in Chapter 8, some of the SG sensor attacks do not result in a reactor trip directly. Only a small deviation for the controller input from the spoofed sensor will be fed into the controller resulting in a compensation in the opposite direction for the water level. Rather than directly setting the water level to trip the reactor immediately in PMS, this type of attack requires only a small deviation and a period of time to trip the reactor. This period provides the defender some time to mitigate the reactor trip. Hence, this type of attack is our focus on the mitigation method. In this type of scenario, the PI controller plays an essential role in the steam generator water level control system. Hence, we start this chapter with a study on the PI controller. Then, several possible attack signals are examined, and the worst-case attack scenarios under this particular circumstance are considered. Finally, an optimal mitigation method is introduced. The mechanism behind the formulation is derived and the maximum delay time allowing for mitigation response is obtained.

9.1 PI Controller

9.1.1 Introduction

PI Control is one of the most popular control algorithms used in the industry, given its simplicity and effectiveness [62]. It can regulate flow, temperature, pressure, level, and many other industrial process variables [63, 64].

Its output is made up of the sum of the proportional and integral control actions. The proportional term is the main driving force in the controller. It changes the controller output which is proportional to the error or difference between the desired point and actual state, with the proportional gain providing a fast response. For example, if the error is large and positive, the control output will be proportionately large and positive, taking into account the gain factor K_c. The integral term is to increment or decrement the controller output over time to reduce the error. Given enough time, the integral action will drive the controller output until the error is zero. For example, suppose there is a residual error after the application of proportional control. In that case, the integral term seeks to eliminate the residual error by adding a control effect due to the historic cumulative value of the error. When the error is eliminated, the integral term will cease to grow. This will result in the proportional effect diminishing as the error decreases, but this is compensated for by the growing integral effect.

The setpoint (SP) is the target value, and process variable (PV) is the measured value that may deviate from the desired value. The error from the setpoint is the difference between the SP and PV is defined as

$$e(t) = SP - PV. \tag{9.1}$$

The value of the controller output $u(t)$ is fed into the system as the manipulated

variable input,

$$u(t) = u_{bias} + K_c e(t) + \frac{K}{\tau} \int_0^t e(t)dt. \tag{9.2}$$

The u_{bias} term is a constant that is typically set to the value of $u(t)$ when the controller is first switched from manual to automatic mode. The two tuning values for a PI controller are the controller gain K_c and the integral time constant τ. The FI controller logic is shown in Figure 5.

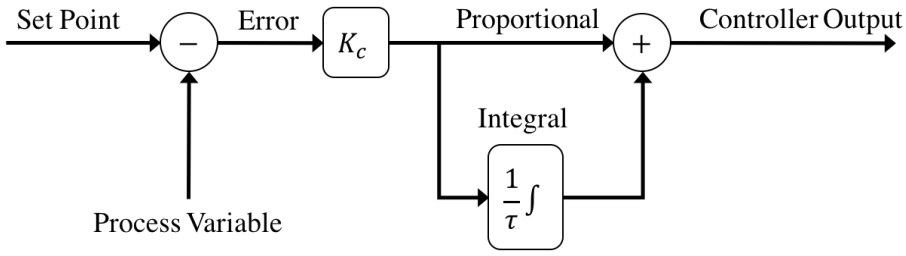

Figure 9.1: PI controller flowchart.

Digital controllers are implemented with discrete sampling periods, and a discrete form of the PI equation is needed to approximate the integral of the error. This modification replaces the continuous integration with a summation of the error and uses Δt as the time between sampling intervals and n_t as the number of sampling intervals:

$$u(t) = u_{bias} + K_c e(t) + \frac{K_c}{\tau} \sum_{i=0}^{n_t} e_i(t)\Delta t. \tag{9.3}$$

9.1.2 PI Controller Model Development for Water Level Control

An only proportional controller model is developed first to control the water level in the steam generator. The result is shown in Figure 9.2.

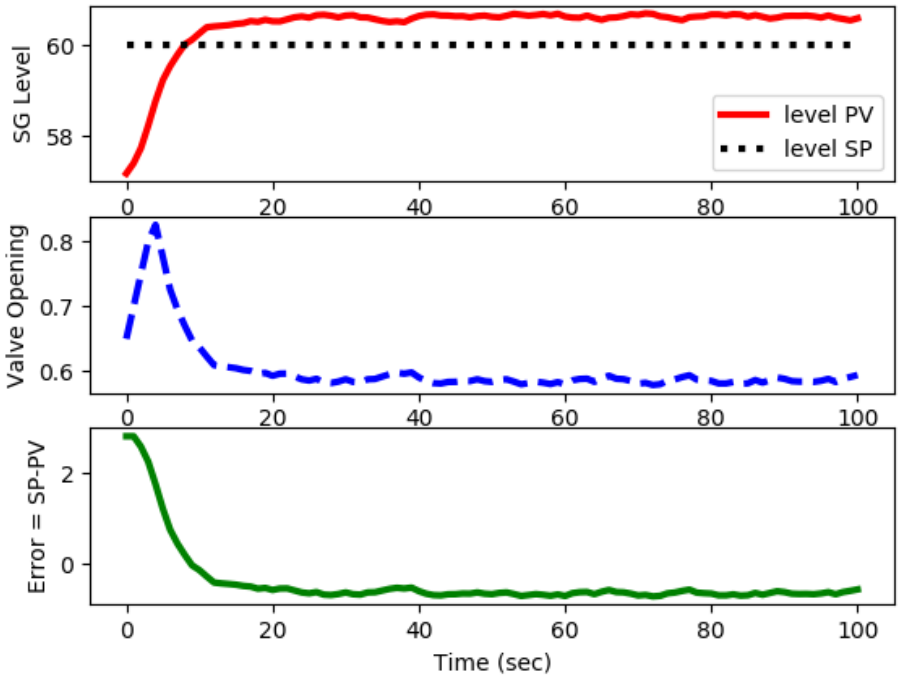

Figure 9.2: P controller performance.

The reactor is operating at full power. The initial water level is 57%, and the valve position is 0.65. We adjust the setpoint for the controller from 57% to 60%. The measurement noise of the water level sensor has also been taken into consideration. The proportional controller cannot control the water level as desired. There is always an offset error between the setpoint and the present value, which cannot be eliminated by proportional only control.

Then, a PI controller model is developed in Python to represent the controller behavior in the SGWLC system. The PI controller takes the water level as the control value and the regulating valve position as the controller action. The action of the combined proportional and integral maneuvers drives the water level to the designed setpoint as illustrated in Figure 9.3.

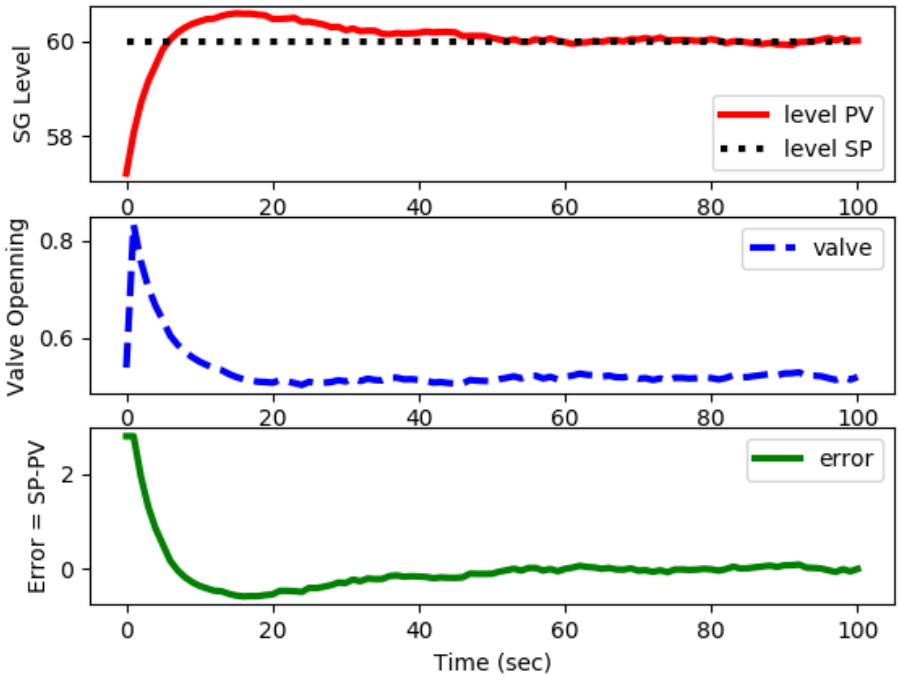

Figure 9.3: PI controller performance.

The parameters and the relationship between the SG level and the regulating valve opening are obtained from the GPWR simulation. The reactor is operating at full power. The initial water level is 57%, and the valve position is 0.65. We adjust the set point for the controller from 57% to 60%. The measurement noise of the water level sensor has also been taken into consideration. The controller parameters are tuned through multiple experiments [65, 66]. We can observe that the SG water level has a small overshoot initially and finally arrives at the setpoint without any error in about 60 s of simulation, which validates the function of the PI controller.

The second validation case accounts for the controller model and the GPWR simulation result as illustrated in Figure 9.4. In this case, we tried out several different combinations of the tuning parameters to fit the GPWR simulation result. From the figure, the controller model behavior is fairly close to the GPWR simulation result.

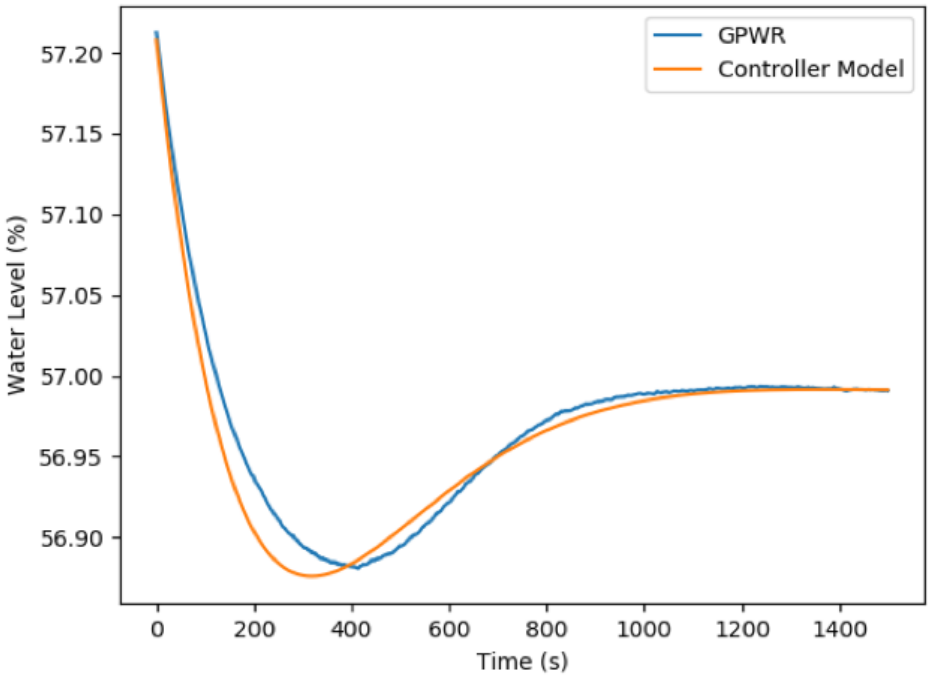

Figure 9.4: PI controller performance.

9.2 Optimal Signal Insertion Mitigation

As our cyber-attack simulation from the Chapter 8 illustrates, the cyber-attack on the feedback controller has a severer effect than the other types of attack. This observation leads us to propose a mitigation strategy that could counter-act the spoofed signal.

9.2.1 Attack Signal Analysis

Under regular operation, the water level should be maintained constant and stable at the setpoint. Two types of attack signals have been tested with the PI controller, the step function, and the ramp function. It is assumed that the attacker has spoofed the signal into the controller by setting a higher water level. Hence, this would eventually result in a low-level SG trip.

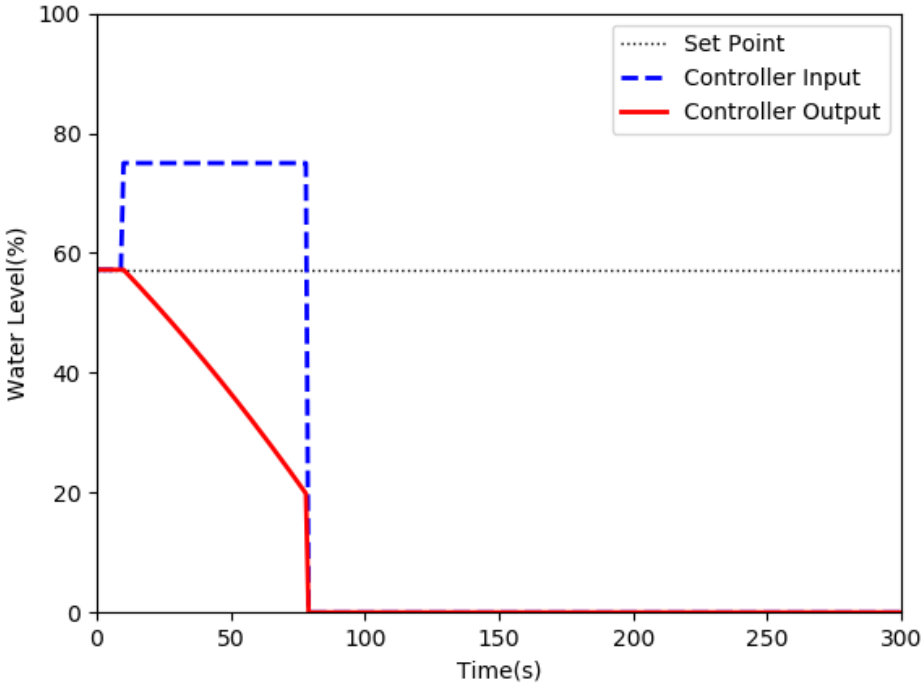

Figure 9.5: Step function attack signal: Case A.

Case A: We consider a step function perturbation, shown in Figure 9.5. The controller input is set to be higher than the actual water level at 75%, so the controller will adjust the feedwater flow rate to compensate for this high water level, which results in the decrease of the water level. After a while, it will trip the reactor through a low steam generator water level. Without mitigation, the water level will reach the low water level trip point around 68 s after the attack.

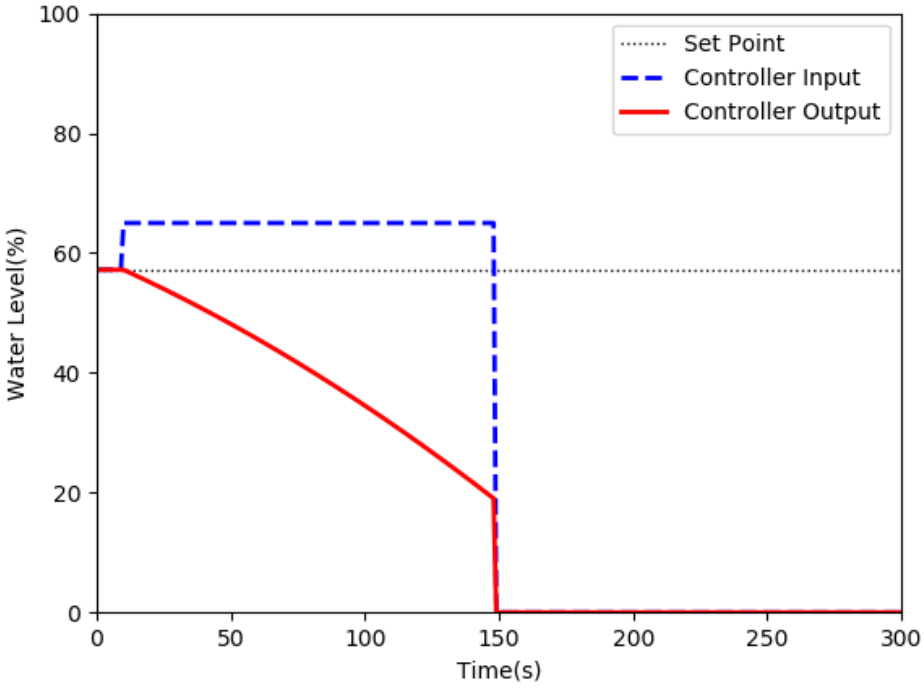

Figure 9.6: Step function attack signal: Case B.

Case B: Another step function insertion similar to case A is considered but with the controller input is set to 65%. The controller will still adjust the feedwater flow rate to compensate for this high water level value. Since the deviation between the spoofed value and the setpoint is smaller than case A, it takes longer to trip the reactor than case A. Without mitigation, the water level will reach the low water level trip point of 20% in 140 s.

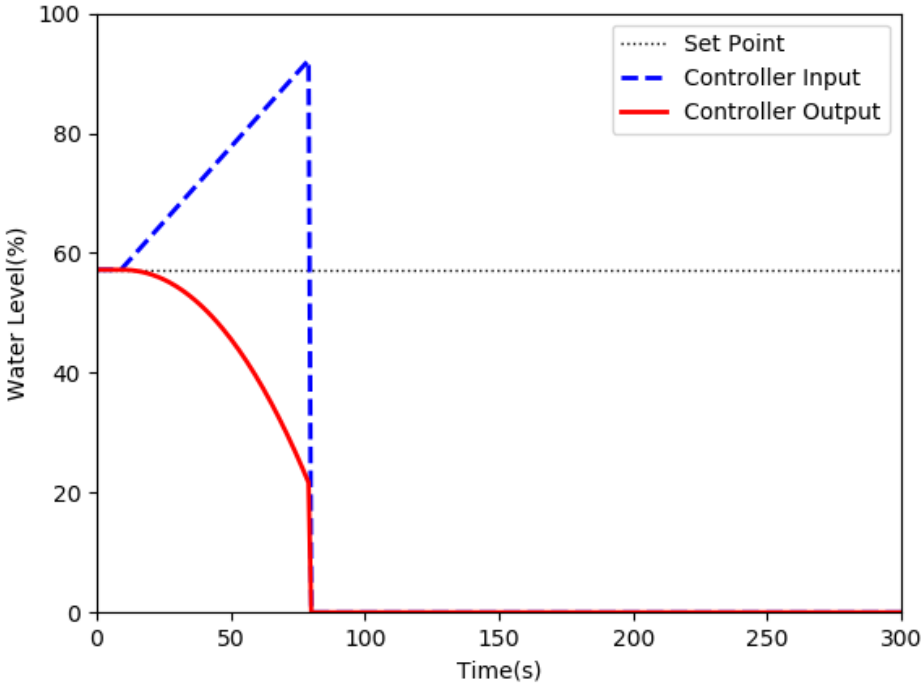

Figure 9.7: Ramp function attack signal: Case C.

Case C: A ramp function perturbation is considered as illustrated in Figure 9.7 where the attacker gradually increases the controller input. The time to trip depends on the slope of the ramp function. In this case, the reactor gets tripped in around 70 seconds.

9.2.2 Optimal Mitigation Formulation

Some sensor corruptions can most easily be mitigated by selecting an uncorrupted sensor as the controller input from the control room by the operator. However, sometimes the attacker may take over the controller resulting in the unavailability of this switch mitigation. Under the circumstance, the attack signals are discussed in Section 9.2.1 can be mitigated by inserting signal into the controller in the opposite direction, which can cancel out the attacker's spoof and bring the system back to normal. There are four channels of water level sensors. One of the four signals is fed

into the controller as the input $z_{in}(t)$. The other three measuring the actual water level can be collapsed into one parameter equaling the controller output $z_{out}(t)$. The setpoint of the water level is set as z_0. For regular operation, the relationship between the parameters is

$$z_{out}(t) = z_{in}(t) = z_0. \tag{9.4}$$

The attacker inserts a spoofed value $a(t)$ into the controller input at time t,

$$z_{c,in}(t) = z_0 + a(t) \tag{9.5}$$

The difference $a(t)$ between the setpoint and spoofed value feeds into the controller to adjust the feedwater flow to compensate for it. In an ideal world, assume that we can detect the cyber-attack immediately after it happens. Then, we can insert a mitigation signal $m(t) = -a(t)$ into the controller input to cancel $a(t)$ and restore the channel to normal setpoint z_0. For a practical mitigation approach, we assume a time delay of Δt between the attack and the implementation of the mitigation. Our objective for the optimal mitigation algorithm is to maintain the SG water level not to exceed the trip setpoints z^+_{trip} and z^-_{trip}. Hence, we would like to find the maximum time allowed for the implementation of the mitigation action.

We assume that the attacker inserts only a step function into the water level data to the controller. Other functions can be extended with similar derivation. The worst attack scenario under this assumption is that the attacker inserts the maximum possible signal z^+_m or the minimum possible signal z^-_m into the controller. The controller will then drive the regulating valve and compensate for this error with the maximum speed, which would be the worst attack scenario. The maximum possible attack scenario for z^+_m would lead to the controller input

$$a(t) = \left(z_m^+ - z_0\right) u(t), \tag{9.6}$$

where $u(t)$ is the Heaviside step function representing the attack initiated at $t = 0$. This will be fed into the controller, and the output of the controller will be

$$\begin{aligned} z_{a,out} =& K_c \left[(z_m^+ - z_0)u(t) + \frac{1}{\tau} \int_0^t (z_m^+ - z_0)u(t')dt' \right] \\ =& K_c(z_m^+ - z_0) \left(1 + \frac{t}{\tau}\right) u(t). \end{aligned} \tag{9.7}$$

We now insert the mitigation signal $m(t)$ into the controller input

$$m(t) = \left(z_0 - z_m^+\right) u(t - \Delta t), \tag{9.8}$$

where Δt is the time delay. We can calculate the output of the controller from this mitigation signal:

$$\begin{aligned} z_{m,out} =& K_c \left[\left(z_0 - z_m^+\right) u(t - \Delta t) + \frac{1}{\tau} \int_0^t \left(z_0 - z_m^+\right) u(t' - \Delta t)dt' \right] \\ =& K_c(z_0 - z_m^+) \left(1 + \frac{t - \Delta t}{\tau}\right) u(t - \Delta t). \end{aligned} \tag{9.9}$$

The sum of these two signals $a(t)$ and $m(t)$ is the signal the controller now receives:

$$\begin{aligned} z_{in}(t) = a(t) + m(t) &= \left(z_m^+ - z_0\right) u(t) + \left(z_0 - z_m^+\right) u(t - \Delta t), \\ &= (z_m^+ - z_0 t) \left[u(t) - u(t - \Delta t)\right]. \end{aligned} \tag{9.10}$$

The controller output is

$$\begin{aligned} z_{out}(t) &= z_{a,out}(t) + z_{m,out}(t) \\ &= K_c(z_m^+ - z_0)\left[\left(1+\frac{t}{\tau}\right)u(t) - \left(1+\frac{t-\Delta t}{\tau}\right)u(t-\Delta t)\right] \\ &= K_c(z_m^+ - z_0)\left[\left(1+\frac{t}{\tau}\right)[u(t)-u(t-\Delta t)] + \frac{\Delta t}{\tau}u(t-\Delta t)\right]. \end{aligned} \tag{9.11}$$

The output $z_{out}(t)$ should not exceed the reactor trip set point z_{trip}^-, which results in

$$K_c(z_m^+ - z_0)\left[\left(1+\frac{t}{\tau}\right)[u(t)-u(t-\Delta t)] + \frac{\Delta t}{\tau}u(t-\Delta t)\right] < (z_0 - z_{trip}^-), \tag{9.12}$$

for $t \geq 0$. We may simplify Eq. (9.12) to derive the maximum delay time allowed

$$\Delta t_m < \frac{\tau}{K_c}\frac{z_0 - z_{trip}^-}{z_m^+ - z_0} \text{ for } t \geq \Delta t_m. \tag{9.13}$$

Similarly, we could derive the inequality for the minimum possible signal

$$\Delta t_m < \frac{\tau}{K_c}\frac{z_{trip}^+ - z_0}{z_0 - z_m^-} \text{ for } t \geq \Delta t_m. \tag{9.14}$$

The smaller of the two time dealys from Eqs. (9.13) and (9.14) can be set as the maximum time delay to be allowed, which is also the latest time for the operator to mitigate the worst-case attack.

9.2.3 Mitigation Approach Demonstration

An example illustrates the effectiveness of the mitigation method. A comparison has been performed between attack trajectory and mitigation trajectory. The attacker spoofed the controller input signal to 70%, as shown in Figure. 9.8. The orange line is the attack trajectory. Suppose we do not respond to the attack with some mitigation approaches. The controller will adjust the feedwater flow rate to compensate for the

incorrect water level signal. The actual SG water level will decrease and reach the trip setpoint and trip the reactor at 130 s after the attack.

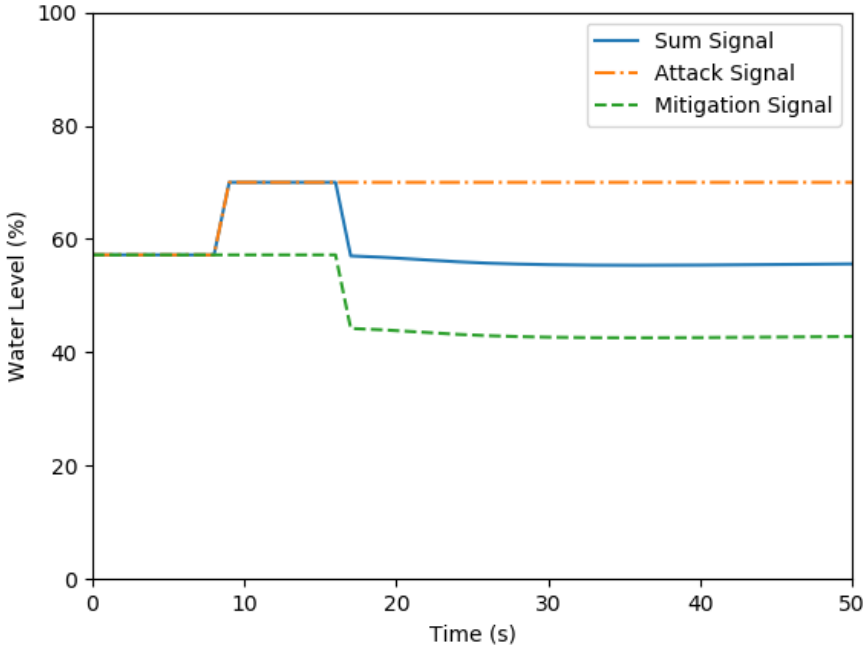

Figure 9.8: Controller input - channel 3 water level sensor.

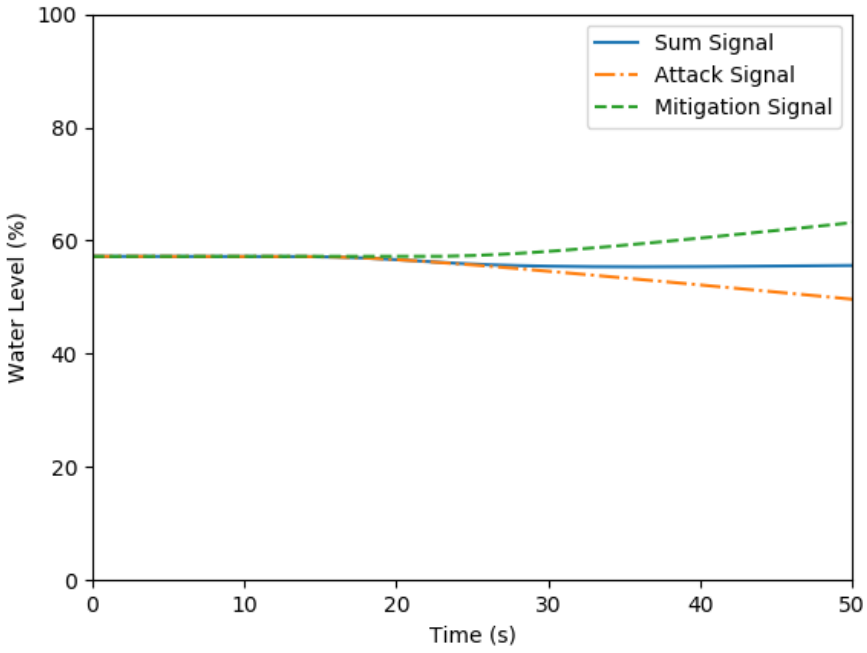

Figure 9.9: Controller output feedback.

If the mitigation is applied to the system, the trajectory will evolve like the blue line in Figures 9.8 and 9.9. The mitigation is implemented after 10 s of the attack initiation. We note that the controller input and the actual water level will quickly adjust to normal and avoid the reactor trip.

For the maximum delay time allowed, the worst-case attack for a low SG water level trip is shown to set the controller input as 100%. With the PI controller paramter $K_c = 0.26$ and $\tau = 25s$, the maximum delay time Δt_m can be calculated through Eq. (9.13)

$$\Delta t_m = \frac{\tau}{K_c} \frac{z_0 - \bar{z_{trip}}}{z_m^+ - z_0} = \frac{25}{0.26} \cdot \frac{0.57 - 0.20}{1.00 - 0.57} = 82.7 \text{ s.} \tag{9.15}$$

Two different implementation times of mitigation are shown in Figure. 9.10. The blue line is the mitigation applied within the maximum time delay allowed. The orange line is the mitigation applied out of the maximum time delay allowed.

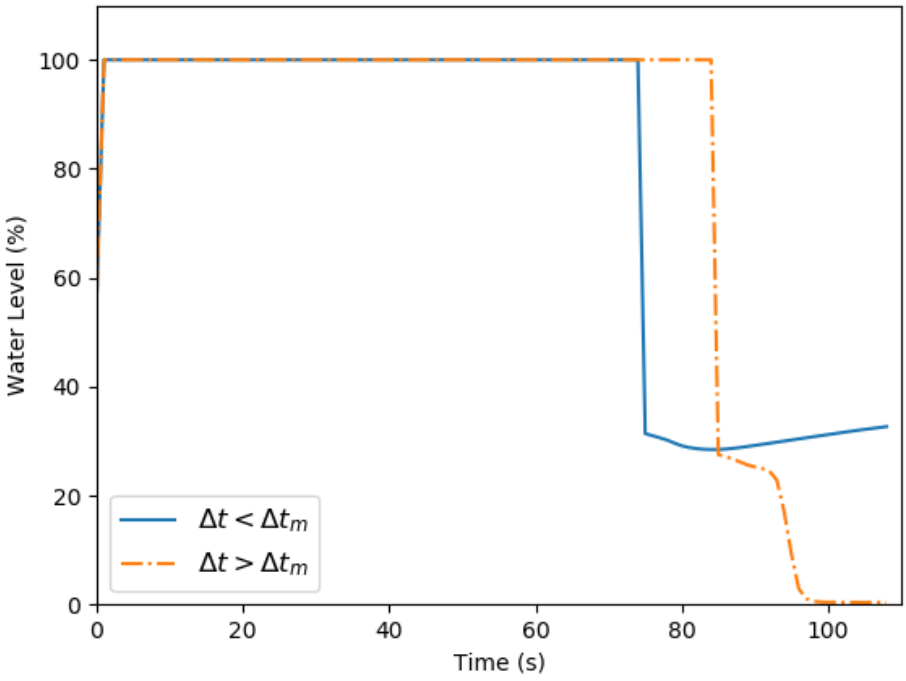

Figure 9.10: Different mitigation implementation time.

Figure. 9.10 shows the result for different mitigation time. We observe that if the mitigation is applied within the maximum delay time allowed, i.e. $\Delta t < \Delta t_m$, the reactor will not be tripped. The mitigation will finally drive the reactor back into the set point. If the mitigation is applied later than the maximum delay time, i.e. $\Delta t > \Delta t_m$, it will no longer be useful. The reactor will still be tripped.

CHAPTER 10

Summary, Conclusions, and Future Work

10.1 Summary

The study in this thesis focuses on the investigation and development of a cyber-security framework for the I&C system in nuclear power plants. The research is motivated by the ongoing effort to respond to the cyber-security issues emerging from the increasing implementation of digital components in the I&C system of modern nuclear power plants. The model-based cyber-security framework provides a systemic approach covering the analysis of the I&C system in nuclear power plants and potential cyber-attack scenarios as well as modeling and simulation of the potential cyber-attack scenarios. The detection and mitigation methods are also included in the framework for potential cyber-intrusions.

The model-based cyber-security framework is developed to detect and cope with the intrusions and disruptions to the steam generator water level control system, which could result in unnecessary reactor shutdown events. Five modules are contained in the framework. The framework starts with an investigation on the I&C system for the operating and advanced PWR plant. We have placed our emphasis on the steam generator water level control system and the trips from the SG water level being out-of-range. PMS and PLS are the two systems most related to the operation control and reactor trip function of the AP1000 design. Similar studies have been

performed for the GPWR system.

The SAPHIRE code is used to build the attack trees for both systems. Based on the attack trees, we are able to generate the minimal cut sets representing the attack scenarios. The MP-ET structure is proved to be a powerful tool for visually representing a large number of attack scenarios in an easily digestible manner, especially for common cause failure or common mode attack. To quantify the susceptibility of cyber-attack scenarios causing reactor trips, we propose sensitivity metrics to identify the low-order sets of components that may be compromised and the degree of perturbations needed for each component. This approach allows us to rank all potential attack scenarios and identify the most significant ones. Various attack possibility combinations are performed using these metrics. The high-susceptibility pathways for cyber-attack scenarios consist of primary control components, sensor feedback divisions, and trip logic paths with minimal inter-component dependencies.

For the modeling and simulation tasks in the framework, we first developed an API serving as the platform for RELAP5, GPWR, and the reduced order model we built. It is developed in Python with two primary operations, setup, and step. The API can execute the RELAP5 code, accurately transfer data between coupled codes, and perform calculations in a reasonable time frame. Moreover, the API can accommodate the application of the Kalman filter and associated detection methods. A reduced-order model for the steam generator has been developed, representing the dynamics of the steam generator. The ROM provides sufficiently accurate predictions of the water level and obtains the unique key relationship among the reactor power, SG water level, and FW flow rate. A possible cyber-attack scenario on the I&C system has been simulated through the newly-installed GPWR simulator. We also analyzed a new attack scenario involving controller feedback through the GPWR simulation, which requires only a small deviation to trip the reactor.

New detection and mitigation methods have also been developed. The Kalman

filter algorithm provides optimal tracking of SG water level combining the uncertain simulation results with the observation data subject to statistical fluctuations. An observed plant state with significant deviation from the optimal system projection could then indicate potential intrusions. Unlike traditional detection methods, this method avoids the issue associated with uncertain observation by using the Kalman filter algorithm to combine the simulation and the measurement. The mitigation method proposed can avoid reactor trip by inserting a counter-acting signal to cancel out the attack signal, especially for the case with controller feedback. The worst-case attack under a particular circumstance is analyzed, which could influence the most on the water level control system. The maximum delay time allowed for the defender's action time is also determined through an optimization formulation.

10.2 Future Work

This thesis discusses a model-based cyber-security framework developed to avoid intrusions and disruptions to steam generator systems that could result in unnecessary reactor trip events. The framework could extend to other systems thereby providing us with more insights for cyber-attack on other I&C systems in nuclear power plants. There are a total of 11 reactor trip functions in the AP1000 system. We can perform a similar analysis on other reactor trip functions and plant control systems generating attack scenarios. Our analysis can be extended to other types of reactors as well.

The susceptibility quantification method of the cyber-attack causing reactor trips can rank all the attack scenarios and identify the dominant ones. It combines the sensitivity metrics of the reactor trip signal with the attack possibility of each component. Since we cannot quantify the probability of attack, we have ranked attacks on the basis of how many corruptions are needed to accomplish them, and the magnitude of the spoof required. More effort can be made to obtain more realistic and meaningful values for the estimated possibility. In addition, we could substitute the

attack possibility by attack difficulty or other measures for various components.

The API has mainly focused on the interface between RELAP5 and ROM. It only contains limited GPWR simulator connection to date. With more in-depth research on the source code of GPWR, we could make better use of the API with GPWR. Finally, we could build a uniform platform to extend the GPWR to test and verify the detection and mitigation approaches for cyber-attacks.

We also note a significant potential for alternate mitigation techniques. However, there are still challenges that need to be resolved before the use of the mitigation signal can be advocated for actual use. When is it better to trip the reactor than to continue operation? Does continuing to operate have any implications on safety? Will it exacerbate potential consequences of undetected attacks? A figure of merit can be proposed to compare alternative responses of the operator. Furthermore, a series of dynamic response and mitigation should be implemented instead of one action. Further research is required to definitely answer these questions.

As we discussed, more interaction between attacker and defender could be represented in the response and mitigation strategy. Game theory is the study of mathematical models of strategic interaction among rational decision-makers. Cyber-attackers should have some knowledge of the nuclear system to make smart decisions on the attack components and approach. We could define the payoff and constraint for both players and explore the equilibrium condition between them. If we could incorporate game theory into the cyber-attack mitigation, it could make the mitigation more effective and useful.

Cyber-security has recieved increasing attention even on the design and construction stage recently [67, 68]. Roh [69] proposes a cybersecurity system that can be used in control networks of nuclear power plants that require high reliability. The proposed system consists of Detection on Attacking Control System (DACS), DACS Management Program (DMP) to centrally manage multiple DACS, and Central Monitoring

Server (CMS) to store system logs meeting the requirements of the U.S. Nuclear Regulatory Commission and the Korea Nuclear Cyber Security Regulations. Guo [70] recommends building an active defense system based on trusted computing technology and boundary protection technology to make the Industrial Control System (ICS) in NPP more secure and explore an ICS test platform allowing us to monitor the running state of ICS and to verify any security measures .

Through the cyber-security project, there has been a long standing question in our study: what is the difference between a cyber-attack and a random failure, and how can we differentiate them? It is not an easy question to answer. Maybe a perfect cyber-attack could not be differentiated from a random failure. Alternatively, the difference between them is unimportant, so we could perhaps treat them in the same way. With increased research and study on cyber-security, we hope to develop a better answer to this question.

BIBLIOGRAPHY

[1] D.-Y. Kim, "Cyber security issues imposed on nuclear power plants," *Annals of Nuclear Energy*, **65**, 141–143 (2014).

[2] "Cyber Security Pro grams for Nuclear Facilities," Regulatory Guide 5.71, U.S. Nuclear Regulatory Commission (2010).

[3] S. Collins and S. McCombie, "Stuxnet: the emergence of a new cyber weapon and its implications," *Journal of Policing, Intelligence and Counter Terrorism*, **7**, *1*, 80–91 (2012).

[4] J. Shin, H. Son, and G. Heo, "Cyber security risk evaluation of a nuclear i&c using bn and et," *Nucl. Eng. Technol*, **49**, 517–524 (2017).

[5] "Computer Security at Nuclear Facilities," IAEA Nuclear Security Series No. 17., International Atomic Energy Agency (2011).

[6] S. Lassell, A. Hawari, J. Benjamin, and K. Barnes, "Methodology Development for Cybersecurity Vulnerability Assessment of University Research Reactors," *Trans. Am. Nucl. Soc.*, **118**, 324–326 (2018).

[7] J.-G. Song, J.-W. Lee, C.-K. Lee, K.-C. Kwon, and D.-Y. Lee, "A cyber security risk assessment for the design of I&C systems in nuclear power plants," *Nucl. Eng. Technol*, **44**, 919–928 (2012).

[8] V. Yadav, R. W. Youngblood, K. L. Le Blanc, J. Perschon, and R. Pitcher, "Fault-Tree Based Prevention Analysis of Cyber-Attack Scenarios for PRA Applications," Annual Reliability and Maintainability Symposium, 1, IEEE (2019).

[9] W. Nichols, Z. Hill, P. Hawrylak, J. Hale, and M. Papa, "Automatic Generation of Attack Scripts from Attack Graphs," 1st International Conference on Data Intelligence and Security, 267, IEEE (2018).

[10] Z. Hill, W. M. Nichols, M. Papa, J. C. Hale, and P. J. Hawrylak, "Verifying attack graphs through simulation," Resilience Week, 64, IEEE (2017).

[11] Z. Hill, S. Chen, D. Wall, M. Papa, J. Hale, and P. Hawrylak, "Simulation and analysis framework for cyber-physical systems," Proceedings of the 12th Annual Conference on Cyber and Information Security Research, 1 (2017).

[12] D. Miljković, "Fault detection methods: A literature survey," Proceedings of the 34th international convention MIPRO, 750, IEEE (2011).

[13] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE transactions on control systems technology*, **18**, 636–653 (2009).

[14] A. Rosich, H. Voos, Y. Li, and M. Darouach, "A model predictive approach for cyber-attack detection and mitigation in control systems," 52nd IEEE Conference on Decision and Control, 6621, IEEE (2013).

[15] M. Chamanbaz, F. Dabbene, and R. Bouffanais, "A Physics-Based Attack Detection Technique in Cyber-Physical Systems: A Model Predictive Control Co-Design Approach," IEEE, 2019 Australian & New Zealand Control Conference, 18 (2019).

[16] K.-N. Kim, M.-S. Yim, and E. Schneider, "A study of insider threat in nuclear security analysis using game theoretic modeling," *Annals of Nuclear Energy*, **108**, 301–309 (2017).

[17] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou, and S. S. Iyengar, "Game theory for cyber security and privacy," *ACM Computing Surveys (CSUR)*, **50**, 1–37 (2017).

[18] "Westinghouse AP1000 design control document," ML11171A500, Westinghouse Electric Company (2012).

[19] "Generic Pressurized Water Reactor (GPWR) simulation software," GSE Performance Solutions, Inc. (2019).

[20] "Systems Analysis Program for Hands-On Integrated Reliability Evalua- tions (SAPHIRE), Technical Reference," NUREG/CR-6952, vol. 2, U.S. Nuclear Regulatory Commission (2008).

[21] "RELAP5/MOD3.3 Code Manual, Volume 1: Code Structure, Systems Models, and Solution Methods," NUREG/CR-5535, rev. 1, U.S. Nuclear Regulatory Commission (2019).

[22] T. L. Schulz, "Westinghouse AP1000 advanced passive plant," *Nuclear engineering and design*, **236**, 1547–1557 (2006).

[23] "AP1000 Protection and Safety Monitoring System Architecture Technical Report," rev. 4, WCAP-16675-NP, Westinghouse Electric Company (2010).

[24] R. F. Condrac, R. P. Waszink, and C. G. Pankiewicz-Nohr, "The delta 125 steam generator design for the AP1000," Proceedings of the 2004 international congress on advances in nuclear power plants-ICAPP'04 (2004).

[25] "Westinghouse Technology Systems Manual," ML11223A214, U.S. Nuclear Regulatory Commission (2011).

[26] W.-S. Lee, D. L. Grosh, F. A. Tillman, and C. H. Lie, "Fault Tree Analysis, Methods, and Applications A Review," *IEEE transactions on reliability*, **34**, 194–203 (1985).

[27] J. Vatn, "Finding minimal cut sets in a fault tree," *Reliability Engineering & System Safety*, **36**, 59–62 (1992).

[28] N. D. Singpurwalla, "Foundational issues in reliability and risk analysis," *SIAM Review*, **30**, 264–282 (1988).

[29] H. Leblanc, "The Autonomy of Probability Theory (Notes on Kolmogorov, Rényi, and Popper)," *The British journal for the philosophy of science*, **40**, 167–181 (1989).

[30] W. E. Vesely, F. F. Goldberg, N. H. Roberts, and D. F. Haasl, "Fault tree handbook," NUREG-0492, U.S. Nuclear Regulatory Commission (1981).

[31] J. Fussell, "How to hand-calculate system reliability and safety characteristics," *IEEE Transactions on Reliability*, **24**, 169–174 (1975).

[32] S. Mauw and M. Oostdijk, "Foundations of attack trees," International Conference on Information Security and Cryptology, 186, Springer (2005).

[33] B. Schneier, "Attack trees," *Dr. Dobb's journal*, **24**, 21–29 (1999).

[34] V. Nagaraju, L. Fiondella, and T. Wandji, "A survey of fault and attack tree modeling and analysis for cyber risk management," International Symposium on Technologies for Homeland Security, 1, IEEE (2017).

[35] S. Kim, J. Guo, K.-I. Ahn, J. C. Lee, ET AL., "Construction of Multi-Path Event Tree for Station Blackout Events," *Trans. Am. Nucl. Soc.*, **117**, 951–953 (2017).

[36] K. R. McCarroll and J. C. Lee, "A Simplified Multi-Path Event Tree for the Integral, Inherently-Safe Light Water Reactor (I2S-LWR)," *Trans. Am. Nucl. Soc.*, **115**, 826–829 (2016).

[37] J. C. Lee, *Nuclear Reactor Physics and Engineering*, John Wiley & Sons (2020).

[38] V. H. Ransom and A. Course, "Numerical modeling of two-phase flows," *Cours de l'Ecole d'été d'Analyse Numérique-CEA-INRIA-EDF*, pp. 12–23 (1989).

[39] D. Drew, L. Cheng, and R. Lahey Jr, "The analysis of virtual mass effects in two-phase flow," *International Journal of Multiphase Flow*, **5**, 233–242 (1979).

[40] J. H. Mahaffy, "A stability-enhancing two-step method for fluid flow calculations," *Journal of Computational Physics*, **46**, 329–341 (1982).

[41] J. A. Trapp and R. A. Riemke, "A nearly-implicit hydrodynamic numerical scheme for two-phase flows," *Journal of Computational Physics*, **66**, 62–82 (1986).

[42] Y. Taitel and A. E. Dukler, "A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow," *AIChE Journal*, **22**, 47–55 (1976).

[43] Y. Taitel, D. Bornea, and A. Dukler, "Modelling flow pattern transitions for steady upward gas-liquid flow in vertical tubes," *AIChE Journal*, **26**, 345–354 (1980).

[44] A. Trivedi, C. Allison, A. Khanna, and P. Munshi, "RELAP5/SCDAPSIM model development for AP1000 and verification for large break LOCA," *Nuclear Engineering and Design*, **305**, 222–229 (2016).

[45] G. Van Rossum and F. L. Drake Jr, *Python reference manual*, Centrum voor Wiskunde en Informatica Amsterdam (1995).

[46] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*, O'Reilly Media, Inc. (2012).

[47] J. Prock, "Mathematical modeling of a steam generator for sensor fault detection," *Applied mathematical modelling*, **12**, 581–609 (1988).

[48] D. C. Arwood and T. Kerlin, "A mathematical model for an integral economizer U-tube steam generator," *Nuclear Technology*, **35**, 12–32 (1977).

[49] W. Schilders, *Model Order Reduction: Theory, Research Aspects and Applications*, Springer (2008).

[50] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, *Reduced Order Methods for Modeling and Computational Reduction*, Springer (2014).

[51] S. Boyaval, C. Le Bris, T. Lelievre, Y. Maday, N. C. Nguyen, and A. T. Patera, "Reduced basis techniques for stochastic problems," *Archives of Computational methods in Engineering*, **17**, 435–454 (2010).

[52] P. Benner, S. Gugercin, and K. Willcox, "A survey of projection-based model reduction methods for parametric dynamical systems," *SIAM review*, **57**, 483–531 (2015).

[53] W. H. Schilders, H. A. Van der Vorst, and J. Rommes, *Model order reduction: theory, research aspects and applications*, vol. 13, Springer (2008).

[54] P. Jamet, "Stability and convergence of a generalized Crank-Nicolson scheme on a variable mesh for the heat equation," *SIAM Journal on Numerical Analysis*, **17**, 530–539 (1980).

[55] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, **92**, 401–422 (2004).

[56] H. Musoff and P. Zarchan, *Fundamentals of Kalman filtering: a practical approach*, American Institute of Aeronautics and Astronautics (2009).

[57] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using Kalman filter," *IEEE transactions on control of network systems*, **1**, 370–379 (2014).

[58] J.-A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," IEEE/RSJ International Conference on Intelligent Robots and Systems, 1514, IEEE (2007).

[59] R. Lew, T. Ulrich, and R. Boring, "Simulation Technologies for Integrated Energy Systems Engineering and Operations," International Conference on Applied Human Factors and Ergonomics, 566, Springer (2020).

[60] H. Jokstad, O. Berntsson, R. McDonald, R. Boring, B. Hallbert, and K. Fitzgerald, "Implementation of Software Tools for Hybrid Control Rooms in the Human Systems Simulation Laboratory," INL/EXT-14-33710, Idaho National Laboratory (2014).

[61] S. S. Adams, N. Murchison, and R. J. Bruneau, "Investigating cyber threats in a nuclear power plant." SAND2018-11557C, Sandia National Laboratory (2018).

[62] S. Bennett, "A brief history of automatic control," *IEEE Control Systems Magazine*, **16**, 17–25 (1996).

[63] M. A. Johnson and M. H. Moradi, *PID control*, Springer (2005).

[64] R. C. Panda, *Introduction to PID controllers: theory, tuning and application to frontier areas*, BoD–Books on Demand (2012).

[65] J. Zhong, "PID controller tuning: A short tutorial," Lecture notes, Department of Mechanical Engineering, Purdue University (2006).

[66] A. O'Dwyer, "A summary of PI and PID controller tuning rules for processes with time delay. Part 1: PI controller tuning rules," *IFAC Proceedings Volumes*, **33**, 159–164 (2000).

[67] J.-W. Lee, C.-K. Lee, J.-G. Song, and D.-Y. Lee, "Cyber Security Considerations in the Development of I&C Systems for Nuclear Power Plants," Proceedings of the International Conference on Security and Management, 1, The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp) (2011).

[68] C. Poresky, C. Andreades, J. Kendrick, and P. Peterson, "Cyber security in nuclear power plants: Insights for advanced nuclear technologies," Department of Nuclear Engineering, University of California, Berkeley, Publication UCBTH-17-004 (2017).

[69] J.-H. Roh, S.-K. Lee, C.-W. Son, C. Hwang, and J. Park, "Real-time Network Intrusion Detection System with Supporting Cyber Security Regulations for Nuclear Power Plants," Transactions of the Korean Nuclear Society Virtual Spring Meeting (2020).

[70] Y. Guo, X. Lou, E. Bajramovic, and K. Waedt, "Cybersecurity risk analysis and technical defense architecture: Research of ICS in nuclear power plant construction stage," Proceedings of the 3rd IAEA International Conference on Nuclear Security: Sustaining and Strengthening Efforts (2020).