Bounding Cyber in Design Basis Threat

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

with a

Major in Computer Science

in the

College of Graduate Studies

University of Idaho

by

Jacob S. Benjamin

Major Professor: Michael Haney, Ph.D.

Committee Members: Robert Borrelli, Ph.D.; Constantinos Kolias, Ph.D.;

Marlene Ladendorff, Ph.D.

Department Administrator: Terence Soule, Ph.D.

August 2020

Authorization to Submit Thesis

This dissertation of Jacob S. Benjamin, submitted for the degree of Doctor of Philosophy with a Major in Computer Science and titled "Bounding Cyber in Design Basis Threat," has been reviewed in final form. Permission, as indicated by the signatures and dates below is now granted to submit final copies for the College of Graduate Studies for approval.

Major Professor:	Michael Haney, Ph.D.	Date
Committee Members:	Robert Borrelli, Ph.D.	Date
	Constantinos Kolias, Ph.D.	Date
	Marlene Ladendorff, Ph.D.	Date
Department Administrator:	Terence Soule, Ph.D.	Date

Abstract

The emergence of cyberweapons and the convergence of Information Technology (IT) and Operational Technology (OT), contribute to the exponential growth in the number and sophistication of cyber-attacks, targeting critical infrastructure. The nuclear sector has recognized that it must employ compensating measures in order to ensure its most critical systems can defend, detect, delay, respond, and recover from cyber-attacks. The Nuclear Regulatory Commission (NRC) has included cybersecurity requirements in the Physical Security and Design Basis Threat Orders. Design Basis Threat (DBT) is a profile of the type, composition, and capabilities of an adversary used to design protection systems at nuclear power plants. These prescribed cybersecurity requirements, are an alternate approach to traditional DBT analysis, that even if implemented correctly, may not be sufficient to defend against an Advanced Persistent Threat (APT). The use of a compliance-based approach has left nuclear power plants unable to quantitatively measure their ability to defend against adversaries with cyber capabilities. This research identifies residual cyber risk at nuclear power plants, advocates for the adoption of Software-Defined Networking (SDN) and face recognition technologies at nuclear facilities, and proposes a novel approach to developing cyber DBTs specific to the facility, its material, or adversary activities that can be empirically investigated through a combination of modeling, simulation and live exercises.

Acknowledgements

First and foremost, I would like to thank my major professor, Dr. Michael Haney, for his support and guidance during my graduate studies.

I graciously acknowledge the financial support of Idaho National Laboratory during the majority of my graduate studies at the University of Idaho.

Much of this work would not have been possible without the support of my teammates in the Nuclear Cyber Team at Idaho National Laboratory and Threat Operations Center at Dragos. I would like to thank them for their enthusiastic encouragement and useful critiques of this work.

Dedication

Thank you to my family, for your unconditional and continuous love and support through my academic career. This would not have been remotely possible without you all.

Table of Contents

List of Tables

List of Figures

List of Acronyms

DBT Design Basis Threat

APT Advanced Persistent Threat

NRC Nuclear Regulatory Commission

CDA Critical Digital Asset

IT Information Technology

OT Operational Technology

CSP Cyber Security Plan

NEI Nuclear Energy Institute

NASA National Aeronautics and Space Administration

PLC Programmable Logic Controller

NSA National Security Agency

CIA Central Intelligence Agency

USB Universal Serial Bus

SIS Safety Instrumentation System

SMB Server Message Block

IAEA International Atomic Energy Agency

SDN Software-Defined Networking

AV Antivirus

EICAR The European Institute for Computer Anti-Virus Research

LSB Least Significant Bit

BMP Bitmap

PNG Portable Graphics Format

RGB Red Green Blue

PMMD Portable Media and Mobile Devices

C2 Command-and-Control

TTPs Tactics, Techniques, and Procedures

PRA Probabilistic Risk Assessment

HAZOP Hazard and Operability Study

FMEA Failure Mode and Effects Analysis

FTA Fault Tree Analysis

DBIR Data Breach Investigations Report

CISA Cybersecurity and Infrastructure Security Agency

HCE High Consequence Event

ICS Industrial Control Systems

IP Internet Protocol

MAC Media Access Control

RSTP Rapid Spanning Tree Protocol

ARP Address Resolution Protocol

VLAN Virtual Local Area Network

BDPU Bridge Data Protocol Unit

NIST National Institute of Standards and Technology

DHS Department of Homeland Security

FBI Federal Bureau of Investigation

AC&D Access Control and Detection

LAS Local Alarm Station

CAS Central Alarm Station

NVR Network Video Recorder

FDB Field Distribution Box

PPS Physical Protection System

GPU Graphics Processing Unit

Chapter 1: Introduction

The emergence of cyberweapons [48] [145] and the convergence of IT and OT [55], contribute to the exponential growth in the number and sophistication of cyber-attacks, targeting critical infrastructure [140] [33] [77] [30] [1] [2] [55]. In 2018, the International Atomic Energy Agency (IAEA) reported there were 450 nuclear power reactors in operation in 30 countries and 55 more in various stages of construction [66]. The United States has 99 Active and 18 Decommissioning Power Reactors in 30 states [21]. These 99 reactors generated 807.1 billion kilowatt hours of electricity and provided nearly 20 percent of the nation's electricity [29] [21]. In the United States, nuclear reactors fall under the Nuclear Reactors, Materials, and Waste sector of critical infrastructure, which covers most civilian nuclear infrastructure [21]. The NRC is the government agency, tasked with protecting public health and safety related to nuclear energy. Its functions include overseeing reactor safety and security, administering reactor licensing and renewal, licensing radioactive materials, radionuclide safety, and managing the storage, security, recycling, and disposal of spent fuel [104]. The NRC and its licensees recognized that they must employ compensating measures in order to ensure its most critical systems can defend, detect, delay, respond, and recover from cyber-attacks [91]. In 2002, the NRC added cybersecurity requirements into the Physical Security and Design Basis Threat Orders [102] for power generating nuclear reactors. DBT is the key input nuclear power plants use for the design of systems against acts of radiological sabotage and theft of special nuclear material [102][62]. The NRC expects its licensees, nuclear power plants, to demonstrate that they can defend against the DBT [102].

1.1 Problem

In the nuclear industry, cyber as a capability of an adversary, or the potential impact cyber-attacks may have on nuclear safety and security is an unbounded risk. In current license documents with the regulator, adversary cyber capabilities are not limited or bounded in any way [101]. Nuclear power plants are required to adequately defend against an adversary with any and all cyber capabilities up to and including the DBT [101]. The DBT includes any action that my cause or contribute to radiological sabotage or theft of special nuclear material [102]. Given an cyber-attack scenario, emulating this DBT, nuclear power plants are unable to demonstrate the effectiveness of their cyber protective measures against specific adversary Tactics, Techniques, and Procedures (TTPs) in a quantifiable manner [25][41] [134] [76]. Currently, licensees demonstrate that they can defend against the DBT through the implementation of their Cyber Security Plan (CSP), which includes the application, evaluation, and ongoing maintenance of prescribed cybersecurity controls [91]. This is a generic compliance-based approach for which no research

has shown to be effective in minimizing actual safety and security incidents [76]. Furthermore, these cybersecurity requirements, even if implemented correctly, may not be sufficient to defend against a persistent adversary with advanced cyber capabilities [76] [78] [32]. If this problem remains unsolved, the health and safety of the public may be at risk. Using nuclear energy for power generation is not a passively-safe process and some High Consequence Events (HCEs) have occurred [126]. Nuclear power plants need the ability to identify when a cyber-attack exceeds their protections and DBT, before a HCE occurs. By detecting and identifying a cyber-attack as beyond DBT, before a HCE occurs, plants can alert the regulator and federal agencies for additional support, to limit or prevent the event through a quicker response and recovery.

1.2 Solution

This solution used qualitative research methods to combine theoretical approaches for cybersecurity for nuclear facilities in a new way. It was developed through the analysis and critique of established theories such DBT, defense-in-depth, regulatory compliance, threat-driven, and consequence-driven approaches. It challenged the established theory that traditional DBT techniques could not be adapted for cybersecurity [65] [134] [75]. The proposed novel approach is based upon the concepts of traditional DBT analysis, however it has been augmented to address the complexities and nuances of cybersecurity. The output of this contemporary approach is similar to that of traditional analysis used for noncyber DBTs. More exactly, it is similar in that it will contain a list of potential adversaries and their attributes, characteristics, and possible actions. Also, in that it will include analysis that will determine whether specific adversaries are relevant to potential targets. It is different from the traditional approach in that it is focused on adversarial cyber capabilities, not conventional physical capabilities. The ultimate deliverable of the new approach is a cyber DBT tailored to the facility that can be used as requirements for designing protection measures as well as be empirically evaluated through a performance test, such as threat hunts, penetration tests, modeling, emulations, simulations, live exercises or a combination of these. The cyber DBT is detailed and comprehensive to a level that plant can easily recognize when a cyber-attack is beyond their protection measures.

1.3 Thesis Overview

Chapter 2 is a literature review. It describes the history and current state of cybersecurity at nuclear power plants. It also provides some background on DBT, including designing physical protection systems, developing DBTs, an example DBT, and cyber DBTs. Additionally, this chapter summarizes past significant cyber-attacks and their impacts on critical infrastructure.

Chapter 3 describes an original experiment conducted by the author to identify residual cyber risk within nuclear power plants. The experiment uses quantitative research methods to evaluate whether an adversary could use steganographic techniques to evade detection by the Portable Media and Mobile Devices (PMMD) kiosks used to protect plant systems and networks from adversaries using removable media as an attack pathway.

Chapter 4 is a novel solution developed using qualitative research methods to combine theoretical approaches for cybersecurity for nuclear facilities in a new way. It was created through the analysis and critique of established theories such DBT, defense-in-depth, regulatory compliance, threat-driven, and consequence-driven approaches. It challenges the established theory that traditional DBT techniques could not be adapted for cybersecurity [65] [134] [75]. The approach results in a detailed cyber DBT, tailored to the facility that can be empirically tested and verified. It bounds cyber in DBT by enabling licensees to better quantify and evaluate their performance against specific adversary cyber TTPs.

Chapter 5 defines "inherently secure" and posits that nuclear power plants can significantly improve the security of their OT environments by transitioning to "inherently secure" technologies, such as SDN. SDN limits the impact of cyber-attacks and thus the scope of cyber in DBT, by directly addressing one of the most common adversary tactics, lateral movement.

Chapter 6 proposes that nuclear power plants utilize face recognition technology as a measure to combat risk associated with the insider threat. It proposes the integration of face recognition system technology with the existing physical security infrastructure and equipment including a high level system design. The proposed system may provide support in the mitigation of the risk posed by insiders by passively auditing and validating access to vital areas and enforcing the two-person rule. Augmenting the insider mitigation program for earlier detection of tampering, helps bound cyber in DBT by directly addressing another common adversary tactic, the use of an insider.

Chapter 7 summarizes and concludes the thesis.

Chapter 2: Literature Review

Cyber hygiene is helpful for warding off online ankle biters and if done perfectly in a Utopian world, might thwart 95% of attackers. But in the real world, virtually all places, it registers as barely a speed bump for sophisticated attackers aiming at a particular target.
-Michael Assante

2.1 Cybersecurity at Nuclear Power Plants

The NRC views nuclear security as a balancing of risks and costs, with the understanding that achieving a "zero" level of risk is impossible [73]. Cybersecurity risk mitigation for nuclear power plants began in 2002 and 2003, when the NRC included cybersecurity requirements in the Physical Security and Design Basis Threat Orders [102].

Voluntary Cyber Program

In 2005, the NRC supported a voluntary cybersecurity program for those in the nuclear industry who wished to participate. In the absence of further guidance from the NRC, the Nuclear Energy Institute (NEI), developed and published NEI 04-04, *Cyber Security Program for Power Reactors*. This document contained an acceptable approach to developing a cybersecurity program and was endorsed by the NRC in December of 2005 [52]. NEI is a policy organization whose members include companies that own or operate nuclear power plants, reactor designers, engineering firms, etc. They develop policy on key legislative and regulatory issues affecting the nuclear energy sector [52]. Most plants participated in the voluntary program, likely as an effort to avoid further cybersecurity regulation [52][9].

The Cyber Rule

In March 2009, Code of Federal Regulations Title 10, Section 73.54, *Protection of digital computer and communication systems and networks*, known colloquially as, "The Cyber Rule", was released. This rule required that each licensed nuclear power plant submit a CSP that satisfied the stated requirements for review by the Commission [101]. NEI issued the sixth revision of NEI 08-09, *Cyber Security Plan for Nuclear Power Reactors*, and in May 2010, it was endorsed by the NRC for compliance with 10 CFR 73.54 [3] [52]. The NRC set cybersecurity milestones related to the implementation of the operating power reactor cybersecurity plans [52]. These milestones were split into two phases. Phase one consisted of seven interim cybersecurity milestones and phase two was full program implementation, also called milestone 8 [9] [52]. The NRC's two phase approach allowed the nuclear industry to gradually increase

their cybersecurity posture while focusing on the highest priority actions first [9]. The NRC used limited scope onsite inspections to provide implementation oversight of the seven interim cybersecurity milestones [9]. After December 31, 2017, the milestone 8 deadline, additional full scope inspections occurred over the next two years [9].

The seven interim cybersecurity milestones are [9]:

1. establish a multi-disciplinary cyber assessment team,
2. identify critical digital assets,
3. establish a defensive architecture and isolate as many critical assets as possible,
4. control portable media and devices,
5. enhance insider mitigation strategies,
6. establish security controls for the most significant components, and
7. ensure ongoing monitoring and assessment of controls is established.

Milestone 8, full program implementation, included any requirements listed in the CSP that were not in addressed in the interim seven milestones. These requirements were mostly programmatic and administrative in nature, such as modifying policies and procedures [91]. The most significant requirement in terms of work and time, in scope of milestone 8, is the completion of cybersecurity assessments, for all Critical Digital Assets (CDAs), against the cybersecurity controls listed in the appendices D and E of NEI 08-09 [9]. The cybersecurity controls in NEI 08-09 are based on National Institute of Standards and Technology (NIST) Special Publications (SP) 800-53, *Risk Management Framework*, and SP 800-82, *Guide to Industrial Control Systems Security* [3]. The scope of this work was burdensome that NEI developed additional guidance, NEI 13-10, *Cyber Security Controls Assessments* to help streamline the assessment process [52]. The methodology augments NEI 08-09 by including a graded-approach for applying the cybersecurity controls based on the importance and complexity of the CDA [93]. The guidance in NEI 13-10 was endorsed by the NRC in 2015 and plants could utilize the guidance without changing their CSP. The industry estimated that 60% of CDAs could take advantage of the optimizations of NEI 13-10. The completion of any engineering modifications needed to address gaps identified during the assessments was also in scope of the full program implementation deadline [91] [9]. This requirement was particularly demanding as engineering modifications at nuclear power plants can take up to several years to complete [1][2].

2.2 Design Basis Threat

The DBT concept was introduced in 1979 by the NRC in Title 10, Code of Federal Regulations, Part 73. The regulation defines DBT as a, "profile of the type, composition, and capabilities of an adversary" [102]. The NRC expects its nuclear licensees to demonstrate that they can defend against the DBT [102].The IAEA offers a similar, but more detailed, definition in its own publication, *Development, Use and Maintenance of the Design Basis Threat*. DBT is the, "State's description of a representative set of attributes and characteristics of adversaries, based upon (but not necessarily limited to) a threat assessment, which the State has decided to use as a basis for the design and evaluation of a physical protection system" [62]. Nuclear power plants use DBT as the key input for designing systems to protect against acts of radiological sabotage, as well as to prevent the theft of special nuclear material [102][62].

Designing Physical Protection Systems

The design for Physical Protection Systems (PPS) takes into account basic principles of, "defense-in-depth, minimum consequence of component failure, balanced protection and graded protection in accordance with the significance or potential radiological consequences" [61]. Primary design considerations include detection, delay, response, and deterrence measures for the mitigation of an active insider [61] [120]. Physical protection sensors can be characterized in terms of the performance in the following: detection probability, false alarm rate, and susceptibility to environmental factors adverse effect [61][120]. Physical barriers, such as walls, doors, and vehicle barriers can be characterized in terms of delay, the time it takes for an adversary to overcome them [120]. The goal of the PPS designer is to interweave the detection and delay measures in such a way that allows the reaction force to interrupt and neutralize the adversary within a designated time frame [120]. This approach allows physical protection specialists to predict, control, and quantify the performance of the physical protection systems [61] [120].

Developing DBTs

DBTs are developed by analyzing credible threat intelligence and threat information as well as past nuclear security events [65]. The output is a list of potential adversaries and their attributes, characteristics, and possible actions [65]. The analysis determines whether specific adversaries are relevant to potential targets and ultimately, results in DBTs specific to the facility, its material, or adversary activities [65]. Furthermore, a nuclear power plant's PPS is designed and evaluated on the basis of the DBT [65] [61]. The DBTs contain detailed and quantitative data which can be compared against the PPS [61]. The detailed quantitative description of the DBT is almost always classified [61] [120]. An example

of a detailed DBT description was found in the course material provided at by the IAEA in their DBT Workshops and is provided below [61]. Further analysis of this example DBT and how it translates to physical protection measures was found in research by Tudor Radulescu is provided below the example [120].

Example DBT

> *Attempt of theft of a significant amount of NM (e.g. 10Kg of Pu) by a group of 6 outsiders equipped with 10 Kg TNT explosive, automatic weapons (including light infantry weapons) and specific commercially available intrusion tools. They have a comprehensive knowledge of the facility and associated PP measures. Willing to die or to kill. No collusion with insider.*
>
> *Based on such quantitative information, the physical protection measures can be designed in such a way that:*
>
> - *the protected targets (vital areas) comprise any location that hosts nuclear materials (plutonium) in significant quantities;*
> - *the intrusion detection systems are suitable to detect military trained intruders, with high mobility / equipped with light baggage (few Kg of explosives and infantry weapons), with comprehensive knowledge of the facility, of the vulnerabilities of the detection systems and physical barriers and with tools to sabotage an intrusion detection system;*
> - *the physical barriers on any possible adversaries paths can withstand attacks with explosives with cumulative quantities of 10 Kg;*
> - *the physical barriers offer significant delay to commercially available intrusion tools, in such a way that the delay is more than the time required by the reaction force to intercept the attackers;*
> - *the reaction force is sized in such a way that they have a neutralization probability higher than 90% against a team of 6 attackers armed with lights infantry weapons and with the willing to have up to 5 members killed in order to attain the mission goal.*

Cyber DBTs

The NRC and the nuclear industry used the alternate approach of prescriptive regulatory requirements for cyber DBT [101]. However, there is still a clear link between the cybersecurity requirements and traditional DBT. Milestone 6 is the cybersecurity requirement that is most closely aligned with traditional DBT. Licensees were required to identify, document, and implement cybersecurity controls for a special

subset of their CDAs, called target set CDAs [100]. Target set CDAs are those that could adversely impact the design function of physical security target set equipment. The intent of this milestone was for licensees to provide a high degree of protection against cyber and blended attacks that could lead to radiological sabotage [100]. NEI 08-09 does not use a graded approach, meaning that, other than being identified and remediated first, target set CDAs are not subject to additional protection measures or cybersecurity controls than less critical CDAs [91] [9]. This was later fixed with NEI 13-1 which applied a graded approach to the assessment process by removing controls from less critical, indirect, CDAs. Milestone 5, the enhancement of insider mitigation strategies to include observation for the detection of obvious cybersecurity tampering of CDAs, is in most cases, only implemented on target set CDAs [52] [9]. Plants modified existing insider mitigation rounds, already being performed by security forces, to include the installation and validation of tamper tape on target set CDAs in addition to looking for obvious indications of cyber-related tampering, such as inserted Universal Serial Bus devices (USBs) [100] [9].

2.3 Emergence of Cyberwarfare

The term "cyber" was first popularized in the 1940s by a mathematician named Norbert Wiener [94]. Over the next several decades the term "cyber" became a prefix meaning related to the Internet or digital technology. For example, the word "cyberpunk", describes a genre of storytelling that centers around the Internet culture or the word "cyberbullying", which describes bullying that occurs through digital means, such as email, messaging, or social media [94]. Eventually the prefix "cyber" was added to the word warfare to describe actions a nation state has taken against another nation state using digital technology. Cyberwarfare and its definition is a heavily debated topic among experts and currently there is not a consensus or agreed upon definition [90]. For the purpose of this dissertation, cyberwarfare will be used as a generic term used to describe offensive actions that leverage digital technologies to diminish or destroy an adversary capability, especially capabilities associated with defense, infrastructure, or traditional warfare. The next few sections of this chapter review the emergence of cyberwarfare, by discussing past examples of cyberwarfare and cyberweapons.

> *Cyber weapons are, in a way, the perfect weapons. They get the job done, they are cost-effective, and they are deniable.* -Mikko Hypponen

Titan Rain

For a variety of reasons, there is not a clear, unanimously agreed upon list of instances of cyberwarfare. The earliest, documented instance of a cyber-attack, that meets the definition introduced in this chapter is Titan Rain. Titan Rain was a series of coordinated attacks on computer systems belonging to the United

States government starting as early as 2003 [43]. While the Chinese government denies involvement, evidence suggest that the attacks originated in Guangdong, China [130]. The attackers were able to gain access to many defense contractor networks including those at Lockheed Martin, Sandia National Laboratories, Redstone Arsenal, and National Aeronautics and Space Administration (NASA) [43]. Titan Rain is believed to be one of the earliest instances of state-sponsored APT, a term, originating from the United States Air Force in 2006, referring to group, such as a government, with both the capability and the intent to target, persistently and effectively, a specific entity [12]. Although no sensitive or classified information was reported stolen, Titan Rain caused friction between the United States and Chinese governments [38].

Estonia

On April 27, 2007, there were a series of cyber-attacks on a multitude of Estonian assets. The attackers targeted websites of Estonian organizations, including their parliament, banks, ministries, newspapers, and broadcasters [132]. Most of the attacks were distributed denial-of-service attacks ranging from single individuals using various methods like ping floods to expensive rentals of botnets usually used for spam distribution [6]. The attacks were amid a disagreement between Estonia and Russia and are commonly believed to be a direct result of the disagreements [132]. Experts believe the attacks to be at least in part state-sponsored as the efforts of the attacks exceeded the skills of individual activists or organized crime [6]. Specifically, the attacks required the cooperation of a state-owned large telecommunications company [39]. Further evidence of state involvement is the targeting of critical systems whose network Internet Protocol (IP) addresses would not be generally known [6]. These critical systems are used for telephony and financial transaction processing [6]. It has been estimated that, at the time it occurred, it may have been the second-largest instance of state-sponsored cyberwarfare, following Titan Rain [39].

2.4 Events in the Nuclear Sector

Stuxnet

The first and likely the most famous cyberweapon is Stuxnet. Identified in 2010, it targeted specific Industrial Control Systems (ICS), exploited several zero-day vulnerabilities, and succeeded in destroying a large number of centrifuges at the Natanz Nuclear Enrichment Facility, in Iran [145]. Stuxnet targets Programmable Logic Controllers (PLCs). The PLCs at Natanz, were used to control the centrifuges responsible for separating nuclear material [89]. Once infecting the systems, Stuxnet caused the fast-spinning centrifuges to tear themselves apart. Experts have since analyzed Stuxnet and documented that

it has three modules: a worm that executes all routines related to the main payload of the attack; a link file that automatically executes the propagated copies of the worm; and a rootkit component responsible for hiding all malicious files and processes, preventing detection of the presence of Stuxnet [136]. The facility in Natanz believed the common ICS myth that their systems were protected from cyber-attacks because the systems were air-gapped, or physically isolated from other networks, including the internet [16] [36]. It is believed that Stuxnet was introduced to the Natanz environment via an infected USB drive, thereby jumping its air-gap [16]. Expert analysis determined that once loaded, the worm then propagates across the network, scanning for Siemens Step7 software on computers used for controlling PLCs [44]. In the event the worm cannot find the software or the PLC, it becomes dormant [44] [16]. If it finds both the software and PLC, it installs the rootkit module onto the PLC and Step7 software [44] [16]. It also modifies the code in Step7 to send unexpected commands to the PLC, all while displaying, a false loop of normal values to the operator's console [54]. U.S. General Keith B. Alexander stated, "he and his cyber warriors have already launched their first attack. The cyberweapon that came to be known as Stuxnet was created and built by the National Security Agency (NSA) in partnership with the Central Intelligence Agency (CIA) and Israeli intelligence in the mid-2000s" [5]. Stuxnet while not the first instance of cyberwarfare, was the known first instance of a cyberweapon. This weapon was used successfully to diminish an adversary's capability of enriching uranium, commonly used for nuclear weapons [145].

Wolf Creek

A breach of the business network of Wolf Creek Nuclear Operating Corporation in Burlington, Kansas was first made public in the summer of 2017 [78]. The spokeswoman for Wolf Creek, maintains that despite the intrusion, "the safety and controls systems for the reactor were never at risk", because vital plant components are not connected to business networks or the internet [78]. She continued saying that the plant, "runs on an analog system and operates as an 'island' that cannot be remotely hacked" [78]. However, as seen with Stuxnex, adversaries are capable of jumping the air-gap. This stance of air-gaps and data diodes negating all cyber risk for OT is a commonly held myth [36]. The Federal Bureau of Investigation (FBI) and Department of Homeland Security (DHS) have issued alerts [19] alleging the Russia was behind the breach. According to the New York Times, many cybersecurity experts saw this event as a signal that Russia is positioning itself to disrupt the United States' critical facilities in the event of a conflict [115] [140].

2.5 Events in the Energy Sector

Industroyer

Since Stuxnet, an alarming number of cyberweapons have been developed and released with varying degrees of impact [77]. Many of these weapons, such as Havex, BlackEnergy, and Industroyer specifically target the energoy sector and or the electric grid [77]. Industroyer and CrashOverride are two names given to the same malware that was employed in the December 17th, 2016 cyber-attack on transmission substations in Kiev, Ukraine [147]. The result of the cyber-attack was an undesired impact to Ukrainian electric grid operations, specifically, a loss of power to one fifth of the city of Kiev, for the period of one hour [147]. The malware was discovered by Slovak internet security company, ESET [77]. Industroyer is unique as it was the first ever malware framework designed and deployed to attack electric grids [17]. The malware is a modular framework consisting of an initial backdoor, a loader module, and several supporting and payload modules [17][77]. The creators of Industroyer, perhaps inspired by Stuxnet, chose to understand and codify their knowledge of the industrial process to disrupt operations [77]. Analysis by industry experts at Dragos, found that many of the capabilities of this cyberweapon were not used in the attack, indicating that perhaps this particular attack may have been intended to be more of a proof of concept, than outright warfare [77]. The analysis also highlighted that several characteristics of Industroyer indicate a significant step forward in the evolution of the cyberweapon tradecraft [77]. The characteristics include its modularity, scalability, and most notably, its codification of tactics learned from a previous attack on the Ukraine one year prior, in December 2015 [77].

Shamoon

In 2012, the Iranian adversary group known as the “Cutting Sword of Justice” deployed a destructive malware known as Shamoon in a cyber-attack campaign against multiple oil companies in the Middle East [114]. The malware is reportedly responsible for damaging over 35,000 workstations and causing a week long work delay [114]. Shamoon was designed to spread across networks and make workstations unusable by erasing hard drives and overwriting them with a corrupted image [129] [7]. The malware was modular, with three distinct components, the Dropper, the Wiper, and the Reporter [79] [13]. Shamoon exploited 32-bit NT kernel versions of Microsoft Windows, but could detect and load a 64-bit version, if necessary [79]. Shamoon confirmed concerns that Iran had learned from the effectiveness of Stuxnet and developed similar techniques to target their enemies [146].

Triton

The evolution of sophisticated cyberweapons deployed against the energy sector continued to advance after Industroyer, with Triton. Trisis and Triton are two names given to the same malware, that was discovered in November 2017, after it was deployed against a petroleum and natural gas utility in the Middle East [30]. The malware, like Stuxnet, was tailored to attack ICS. Specifically, Triton targets Schneider Electric's Triconex Safety Instrumentation System (SIS) and it enables the replacement of logic in the final control elements [30]. A SIS, like these Triconex, are responsible for maintaining safe conditions, in the event other equipment or process failures occur [30]. They often operate independently of normal process control logic systems and are focused on detecting and preventing dangerous physical events [30]. Examples of uses may include stopping rotating machinery when a dangerous condition is detected or stopping inflow or heating of gasses when a dangerous temperature, pressure, or other potentially life-threatening condition exists [30]. Experts do not currently know what the specific safety implications of Triton would be in a production environment. However, alterations to logic on the final control element imply that there could be a risk to operational safety [30]. Initially, Russia was suspected to be responsible for the malware, but it was not conclusively determined. However, recently, the cybersecurity firm, FireEye, uncovered evidence that tracked the origin of the malware to a Russian government-owned technical research institute in Moscow [68]. FireEye managed to do this by examining how the attackers may have gained access to critical components needed to build the Triton attack framework [68]. Fortunately, Triton is not a highly scalable or easily replicated attack, because each SIS is unique [30]. However, Triton does mark another step forward in the evolution of cyberweapon tradecraft, as it outlined a success path for adversaries to potentially increase the damage from their attacks by succeeding in diminishing or destroying the safety protections of a physical process [30].

2.6 Events in Other Sectors

This chapter focused on the Energy and Nuclear sectors, but consequences to other sectors, such as systems used for healthcare or transportation, can be equally life-threatening to the public. Similar OT is prevalent in these other sectors and they are vulnerable to cyber-attacks [55] [31] [32]. OT in any sector is vulnerable, because they are rarely, if ever patched [55] [31]. Recent malware incidents, such as WannaCry, Petya, and NotPetya exploited these unpatched systems and significantly impacted these lesser discussed sectors [131] [32]. The damage they caused demonstrated the physical effects and severe consequences, even indiscriminate cyber-attacks, can cause OT equipment residing in critical infrastructure [30] [32] [131]. A targeted or coordinated attack on these systems from an APT or other sophisticated cyber

adversary, would likely result in exponentially greater devastating impacts. Two events from other sectors are discussed below, one that impacted several sectors and one from the Defense Industrial Base sector.

Eternal Blue

Eternal Blue was an exploit developed by the NSA that was leaked by the Shadow Brokers hacker group on April 2017 [131]. The exploit was used as part of several ransomware attacks and banking Trojans [48]. Eternal Blue exploits a vulnerability in Microsoft's implementation of the Server Message Block (SMB) protocol [42]. The NSA had this exploit for five years prior to alerting Microsoft, after the breach was discovered, but before the exploits were leaked [116]. Microsoft released a patch for the vulnerabilities in March 2017 and two months later the exploits were used in the WannaCry ransomware attacks [131]. The impact of WannaCry was so rampant in life-critical systems, like those deployed in hospitals, that Microsoft broke precedent and released emergency patches for many of its end of life operating systems, including Windows XP [82]. Eternal Blue demonstrated to the world the dangers of stockpiling vulnerabilities.

Iranian Missile Systems

In June 2019, the United States launched a cyber-attack on Iranian weapons systems. The attack was successful in disabling computer systems that controlled Iran's rocket and missile launchers [119]. Tensions between the United States and Iran have escalated since the United States withdrew from the 2015 nuclear deal with Iran [88]. Since then, the United States has increased sanctions against Iran and Iran has attack two oil tankers in the strategic Strait of Hormuz and a United States surveillance drone [88]. It was the latest incident, with the drone, that prompted a response from United States. However, according to the President of United States, the cyber-attack was used in lieu of a convention strike [88]. This choice of a cyberweapon instead of a conventional weapon, signifies another step in the evolution of cyberwarfare. The ability to accomplish a significant goal against an adversary, such as disabling or destroying their offensive weapon systems, with little or no casualties, is a bigger display of power than conventional warfare.

2.7 Conclusion

This chapter documented the emergence of cyberwarfare and established a definition for it as offensive actions using digital technology. Next, it analyzed several past instances of cyberwarfare including cyber-attacks on the Energy and Nuclear sectors of critical infrastructure. These attacks demonstrated the severe impact and effectiveness of nation-states targeting these important, yet antiquated, systems.

Eternal Blue demonstrated the danger of stock-piling exploits for use as cyberweapons, and the cyber-attack retaliation for the Iranian attack on the U.S. drone displayed cyber dominance. Technology has become embedded in nearly every aspect of modern society, including warfare. Cyberwarfare has emerged, evolved, and will continue to be an increasingly important aspect of warfare tactics.

Chapter 3: Identification of a Residual Cyber Risk at Nuclear Power Plants

The deficiencies in the existing methods of cyber defense have been increasingly exposed as state-sponsored and state-run attacks have become more frequent and use more sophisticated and extensive resources. -Richard J. Danzig

3.1 Introduction

Steganography

Steganography is a means of secret communication and has existed in various forms for over two thousand years [10]. It is often referred to as the art of, "covered or hidden writing" and its earliest usage dates back to the ancient Greeks, during their battles against Xerxes [72]. Prior to the battle, a warning message was sent to the Spartans of the incoming Persian invasion [72]. The communication was able to avoid Persian detection, because the message was carved into the wooden backing of wax tablets. These tablets were commonly used, as a re-writable surface, so their presence was not suspicious. Because the wax covered the message, the Greeks were able to receive the message before Xerxes invaded [118]. With the emergence of computers and technology, steganography has have evolved to incorporate digital techniques for hiding messages. It is often confused with encryption, because they are both used for private communications, but the concepts are actually quite different. Encryption is used to communicate privately by obscuring the message, so that it is incomprehensible to anyone but the intended party [10]. Steganography is about covert communication. It does not obscure the message, so much as it obscures the communication itself. Today, there are hundreds of digital steganography applications easily available on the Internet [10]. The most common legitimate use of digital steganography is the embedding of watermarks to combat intellectual property piracy [70] [26]. Conversely, according to the National Institute of Justice , "One of the most common illicit uses for steganography is for the possession and storage of child pornography images. However, steganography can also be used to commit fraud, terrorist activities and other illegal acts" [95]. Steganography is especially dangerous, because it can easily avoid detection by traditional digital security software, like anti-virus [139]. Steganography allows seemingly innocuous files, such as pictures of animals, to act as carriers for embedded messages, data, or even files. Upon initial inspection, these images would appear unaltered and offer no indication of the information hidden within them [69]. However, it is important to note that images are not the

only files that can act as a carrier. Steganography applications exist that use a variety of mediums including, video files, audio files, and Voice over Internet Protocol (VoIP) network traffic as carriers for hiding data [10]. However, studies have shown that steganography used in conjunction with malware could become an offensive weapon [69]. For example, botnets have been found that embedded their code to contact a Command-and-Control (C2) host in memes [142]. Once connected to the C2 host, the malware then downloads additional code [142]. Thankfully, significant research has been done in the area of detecting steganography [10]. Many free, open source, or commercial software exist to detect steganography applications and carrier files [10] [141]. However, these applications are considered a niche technology, usually within the digital forensics field, and are not typically included in traditional malware software suites [141].

PMMD Cyber Risk Mitigation at Nuclear Power Plants

A common way of introducing external data sources to isolated environments is with PMMD [9] [52] [33]. However, anytime PMMD accesses these critical systems and networks, they risk exposure to malware or cyber-attacks [92] [33] [31]. Milestone 4 of the cybersecurity program implementation, the control of PMMD, is meant to mitigate this risk [9] [52]. Nuclear power plants installed PMMD kiosks at their facilities to meet this milestone [112][9][52]. PMMD kiosks such as those offered by OPSWAT [112][110] and Tresys [133], are small, free-standing physical structures, containing a computer and a touchscreen that allow users to insert portable media and scan for malware. They come with a variety of features including secure erase, secure transfer, digital signature validation, the ability to handle over twenty media types, and boast up to 64 malware detection engines [112] [110] [133]. Due to industry guidance and NRC oversight the PMMD program has been implemented pretty consistently across the United States' nuclear fleet [91] [9]. Most plants have multiple kiosks deployed across the site at strategic locations [111]. Commonly chosen locations include the Maintenance and Test Equipment Tool Room and Work Control Center, because authorized PMMD is checked out of the inventory in the tool room and work orders are processed and approved in the Work Control Center. Workers interacting with critical systems are guaranteed to pass through those two locations [111]. By procedural adherence, all PMMD used on CDAs are scanned and cleared by these kiosks [91] [9] [52]. The intent is for the PMMD program, in which the kiosks are the cornerstone of, to provide protection against media-borne threats, such as BadUSB, data-borne attacks, such as viruses and Trojans, and supply chain threats, such as files being tampered within transit [112] [9] [52].

Residual Risk

The definition of risk varies depending up on the field of expertise [117], but risk is usually defined as consequence of an undesired event, multiplied by the probability of the event's occurrence [71] [18]. Residual risk refers to the amount of risk remaining after risks have been addressed either by controls or other mitigations [87] [18]. As stated in Section 2.1, the NRC views nuclear security as a balancing of risks and costs, with the understanding that achieving a "zero" level of risk is impossible [73]. However, the NRC glossary defines risk-informed decision making as, "an approach to regulatory decision making, in which insights from probabilistic risk assessment are considered with other engineering insights" [97] [117]. In the context of cybersecurity and nuclear power reactors, residual cyber risk can be defined as the amount of risk remaining, after the licensees have completed the full implementation of their cybersecurity programs.

Problem

The nuclear power industry uses PMMD kiosks as its main technical cybersecurity control for the mitigation of the risk posed by their continued use of PMMD. Steganography is a sophisticated technique to hide data, including hiding malicious code from Antivirus (AV) software. The author hypothesizes that this technique could be used by cyber adversaries to avoid detection by the PMMD kiosks and introduce malware to isolated plant systems and networks. For example, an adversary could use a basic steganographic technique, such as Least Significant Bit (LSB) insertion, to hide malicious code within a carrier file, such as an image, on PMMD, and extract it later, once it is past the kiosk's defenses, such as at its run time on a CDA. An experiment was created to test this hypothesis and if proven true, then residual risk remains for nuclear power plants, in regards to PMMD introducing malware to segregated plant systems and network.

3.2 Method and Materials

The European Institute for Computer Anti-Virus Research (EICAR) Test File

EICAR developed a benign file to test the response of AV software [40]. This file allows people to test and validate anti-malware products without having to use a real computer virus [40] [122]. This is a preferred method over using real malware because of the possibility of causing real damage [40] [122]. Most anti-malware products, with the known exception of Malwarebytes [80], respond to the file as if it were a

real virus, although the description in the report usually has an obvious name, such as "EICAR-AV-Test" [40]. The file is a legitimate DOS program, and when ran, produces the message "EICAR-STANDARD-ANTIVIRUS-TEST-FILE!" [40]. The file consists entirely of printable ASCII characters, so that it can easily be created with a regular text editor [40]. EICAR provides four versions of the file. The first, eicar.com, contains the ASCII string as described in Figure 3.1. The second file, eicar.com.txt, is a copy of this file with a different filename. The third version contains the test file inside a zip archive. Most anti-malware products will detect a malware inside an archive file [40] [122]. The last version is a zip archive containing the third file. This file can be used to determine whether the anti-malware solution is checking archive files within archive files [40].

```
X5O!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-FILE!$H+H*
```

Figure 3.1: EICAR Test File

LSB Technique for Image Steganography

The most common carrier type for steganography is image files [81]. Images are comprised of pixels that are represented by a 24-bitmap value, consisting of three 8-bit bytes that represent the color of the pixel in terms of Red Green Blue (RGB) values ranging from 0 to 255 [47] [58]. These values are normally represented in hexadecimal (Base-16) format as opposed to the standard Base-10 format for numbers [58]. A pixel with an RGB value of FF 9B 00 (255, 165, 0) is represented as orange [72]. It has a red value of 255(FF), a green value of 165 (9B), and a blue value of 0 [72]. Most most image steganography algorithms, utilize LSB embedding mechanism [81]. LSB embedding is the simplest approach to hiding data within an image file [72]. It works by exploiting the fact that changing the least significant bit of each of the RGB value of an image would produce only a, "minor change in the intensity of the color represented by the pixel and this change is not perceptible to the human eye" [81]. Some LSB algorithms modify randomly chosen pixels, while others only use those located in certain areas of the image [10]. When using LSB algorithms, it is best to keep size of the payload less than one-eighth of the size of the carrier file to ensure the changes are not visually discernible [24].

LSB Python Package

Stego-LSB is a python package created by Ryan Gibson [51]. It requires Python version 3.6 or higher and is freely available under the MIT license [51]. Stego-LSB uses a LSB algorithm to hide a file in the color information of an RGB image, specifically Bitmap (BMP) or PNG file types [51]. For each color channel (R,G,B) in each pixel of the image, the Python module overwrites the least significant bits of the

color value with the data from a selected payload file [51]. In order to make recovering the payload data easier, Stego-LSB also hides the file size of the payload file in the first few color channels of the carrier image file [51]. The module also comes with StegDetect, which provides one method for detecting simple LSB steganography in images [51].

Image Carrier File

Stego-LSB requires a BMP or PNG image file to be the carrier medium. A PNG image obtained from Google image search of a lion was used for this study. The image is 186KB in size and is illustrated in Figure 3.2.

Figure 3.2: Carrier PNG File

3.3 Experiment

Experiment Overview

The purpose of this study was to answer the question of whether PMMD kiosks, characterized in section 3.1, can detect malware hidden, using the steganographic technique of LSB Insertion, described in section 3.2, residing on portable media. The experiment uses quantitative research methods to answer this question. The author created a limited test set of images using the steganographic applications and techniques described in Section 3.2. The test set included one hidden payload, the EICAR test file, described in Section 3.2 and depicted in Figure 3.1. The payload was consistent in terms of file size and file type throughout the study. The control group of this experiment is the normal, unconcealed

EICAR test file on an USB drive. The independent variable is the state of the malware payload, hidden or unhidden. The dependent variable is the detection result of the file after being scanned by the kiosk.

Experiment Limitations

The experiment was limited to image steganography. It did not include other types of steganography such as audio, video, network, etc. Image steganography was chosen because it is the most common form of steganography [72]. The author assumed that if the PMMD kiosks detected steganography at all, they would detect the most common medium. The experiment was limited in the types of malware hidden with the steganography. The only type of malware used in the experiment is the benign malware developed by EICAR. This limitation was chosen on purpose, because the author does not own the PMMD kiosks and did not want to assume any legal responsibilities for introducing real malware to the kiosk or the kiosk's owner's environment. The EICAR test file is a perfect substitute for any type of malware, because every AV engine, with the known exception of MalwareBytes, treats the file as if it were real malware. The author addresses this assumption in the experiment by first showing that both the PMMD kiosk and VirusTotal detect the EICAR test file as malicious every time it is found.

Experiment Execution

The researcher copied the EICAR test file, described in section 3.2 on to an USB drive, henceforth named USB-1. Then the researcher used the commands in Figure 3.3, leveraging the Python package described in Section 3.2, to embed a copy of the test file into an image file and copied that file to an USB drive, henceforth named USB-2. The researcher then inserted USB-1 into a PMMD kiosk and documented the detection results. Next, the researcher inserted USB-2 into the same PMMD kiosk and documented the detection results. The researcher also scanned all files using the VirusTotal website to provide additional context [138].

```
$ stegolsb steglsb -a -i input_image.png -s input_file.zip -n 2
# OR
$ stegolsb steglsb -h -i input_image.png -s input_file.zip -o steg.png -n 2 -c 1
# OR
$ stegolsb steglsb -r -i steg.png -o output_file.zip -n 2
```

Figure 3.3: Stego-LSB Commands

3.4 Results and Analysis

Results

Table 3.1: Experiment Results

Filename	***Kiosk***	***Payload***	***USB #***	***AV Engines***	***Detection***	***Correct Result***
eicar.com	OPSWAT	Unhidden	1	4	Infected	Yes
lion.png	OPSWAT	Hidden	2	4	Clean	**No**
lion.png	OPSWAT	N/A	3	4	Clean	Yes
eicar.com	VirusTotal	Unhidden	N/A	59	Infected	Yes
lion.png	VirusTotal	Hidden	N/A	59	Clean	**No**
lion.png	VirusTotal	N/A	N/A	59	Clean	Yes

Analysis

As shown in Table 1, the PMMD kiosk correctly categorized the unconcealed EICAR test file and the original lion.png, as expected. However, when the EICAR test file was hidden inside lion.png, using LSB image steganography, the kiosk failed to identify the file as malicious. The results were consistent when using VirusTotal, confirming the root cause of the issue is with the AV engines themselves and not necessarily the kiosks. Regardless, the study confirms that steganographic techniques avoid detection by the PMMD kiosks. This short fall could be leveraged by cyber adversaries. Some APTs have already been observed using steganography in malware campaigns [107]. More advanced steganography techniques exist using a variety of mediums and have an even lower detection rate than image basic image steganography [10]. The results of the study confirm residual cyber risk remains for nuclear power plants relying on PMMD kiosks to prevent the introduction of malware to isolated plant systems and networks from PMMD.

3.5 Conclusion

Nuclear power plants use PMMD to facilitate necessary data flow activities such as software updates, reporting, and audits in and out of their air-gapped networks. Some APTs have been observed using steganography in malware campaigns. This study evaluated the effectiveness of PMMD kiosks against steganographic hiding techniques. The kiosks could not detect malware hidden using the most common and basic steganographic techniques. The results of the experiment confirm the existence of residual cyber in nuclear power plants, regarding their PMMD program, as it relies on these PMMD kiosks to

prevent the introduction of malware to their isolated plant systems and networks from PMMD.

Chapter 4: A Novel Approach to Creating DBTs

No amount of spending on defenses will shield you completely from hackers. It's time for another approach. -Andy Bochman

4.1 Challenges with Developing Cyber DBTs

For cyber, most countries opted for the alternate approach of prescriptive regulatory requirements instead of traditional DBT analysis [65]. This is likely due to the challenges associated with using traditional DBT analysis for cyber capabilities. The challenges include the rapid evolution of the cyber threat landscape, inaccurate modeling of cyber-initiated events, and the concept of mal-operation.

The cyber threat landscape, including adversary TTPs, changes at a faster rate than physical security threats [78] [61] [120] [76] [53]. The update cycle of DBTs is unable to keep pace with the ever-changing cyber threat landscape [61] [120] [78] [33] [76] [53]. Regulators can issue and update cybersecurity controls for nuclear plants to implement, but sophisticated cyber adversaries will continue to develop new techniques to bypass those security controls [144]. This cycle often results in a constant arms race of attack and defend, attack and defend [144]. The nuclear industry has already expended significant resources in this costly cyber-arms race [53] [9] [52]. Comparing the budgets of nation-state cyber adversaries [46] with the budgets of the nuclear power plants [9] suggests that the costs of the cyber-arms race is to the advantage of the adversaries and not nuclear power plants [78] [115] [15].

Hazard analysis modeling methodologies such as Probabilistic Risk Assessment (PRA), Hazard and Operability Study (HAZOP), Failure Mode and Effects Analysis (FMEA), and Fault Tree Analysis (FTA) are unable to accurately model cyber-attacks [144] [117] [41] [134]. There are many reasons these well established methodologies cannot be easily or accurately be applied to cyber-attacks. The most obvious reason being that cyber-attacks are out of scope of their original purpose and design [144] [117], and thus they are likely to be lacking the necessary inputs. Hazard analysis modeling methodologies focus on equipment failures or human error as initiating events for a hazard, and not cyber-induced failures [144] [117] [41] [134]. Perhaps, the most significant reason these methodologies cannot be used accurately, is that they cannot differentiate between indiscriminate and targeted cyber-attacks [144] [117] [41] [134]. As detailed in Chapter **??**, a review of the modern cyber threat landscape, including the emergence of cyberwarfare and the development of cyberweapons, revealed that if targeted by an advanced cyber adversary, all organizations will be likely be compromised [32] [33] [31] [144].

The third challenge is centered on the concept of mal-operation. Cyber-attacks can use the functionality of a trusted system to perform operations outside of the intended design and without the operator's

knowledge [113] [144] [77]. Safety assessments are concerned primarily with the physical domain and cyber risk assessments are focused on the digital domain [55] [127]. The challenge with assessing cyber risk at nuclear power plants is the physical and the digital domains are intertwined and often have significant overlap [55] [127]. In ICSs, like those used at nuclear power plants, cyber-attacks can cause serious impacts in the physical domain [113] [8] [145]. Cyber adversaries can, "bypass or manipulate traditionally engineered safety barriers and present false information, invalidating the fundamental basis of a safety analysis" [144], as seen in the Stuxnet, Industroyer, and Triton cyber events [30] [77] [145]. Adding to the challenge, these systems were designed to meet engineering requirements, like functionality and failure mode analysis, not security requirements and thus often use antiquated and insecure protocols [2] [1] [117]. For these environments, cyber protections should extend beyond typical IT controls such as AV. The protections must prevent mal-operations such as an adversary issuing legitimate commands in a malicious manner or stopping legitimate commands from operators or controllers from reaching their objectives [144].

The challenges of an ever-changing threat landscape, the inability to use hazard analysis modeling, and the concept of mal-operation certainly contribute to the reason many countries chose the alternate approach of cybersecurity requirements instead traditional DBT analysis. The proposed approach in the following section addresses and overcomes these challenges for applying traditional DBT analysis to cyber DBTs.

4.2 Methodology

Traditional DBTs are developed using a quantitative approach that includes analyzing credible threat intelligence and threat information, past nuclear security events, and site-specific configurations of equipment to determine pertinent adversaries that are relevant to the facility, its systems, or material [65]. This approach can be extended to developing cyber DBTs, by addressing and overcoming the challenges list in Section 4.1 of this chapter. The desired outcome is one or more cyber DBT tailored to the plant, with quantitative data and unambiguous analyses that can be tested and verified.

Past Cyber Events

The logical starting place for identifying credible threats is to review and analyze past cyber events. This review can start as broad as all cyber events, but the goal is to gather threat information relevant to the nuclear facility. The Verizon Data Breach Investigations Report (DBIR) provides perspectives on general cyber threats, not specific to nuclear [137]. The author recommends reviewing any cyber events that are specific to the nuclear and energy sectors of critical infrastructure first and then expanding to

other sectors, if desired. The inclusion of the energy sector is suggested because of the similar equipment and mission objectives. The Nuclear Threat Initiative [105] has references to cyber events at nuclear facilities dating back to 1990. The author has summarized many of the most recent events in both sectors in Chapter 2 of this dissertation. Dragos, Inc. provides year in review reports similar to the Verizon DBIR, but specific to ICSs, like those in use by the nuclear and energy sectors [33] [32] [31].

Credible Threat Intelligence

Amy Bejtlich, the Director of Threat Intelligence Analysis at Dragos, defined threat intelligence as, "actionable knowledge and insight about adversaries and their malicious activities that improves visibility, enables defenders to reduce harm to their organizations, and drives better decision-making about adversaries and their malicious behaviors" [14]. When evaluating threat intelligence, analysts should rate intelligence based on whether it is complete, accurate, relevant, and timely [14] [33]. Completeness makes sure the data provides sufficient detail to enable proper response [14]. Accuracy of the data reduces mistakes and increases the impact of the threat intelligence [14]. Relevance ensures the intelligence addresses threats pertinent to an organization in a consumable manner [14]. Timeliness is scored based on whether the data is delivered quickly enough to reduce adversary dwell time or the defender's time to recovery [14]. These criteria are how analysts ensure the intelligence reduces harm by providing context to the threat and informing on action or non-action [14] [33]. Threat intelligence can framed as strategic, operational, or tactical based on its intended audience [14] [33]. If strategic, the intelligence may provide business context, strategic impact, inform risk management strategies [14]. For operational audiences, the intelligence may support remediation, threat hunting, detection, budget decisions, or collection management [14]. Lastly, if tactical, the intelligence may provide technical indicators or threat behavior analytics [14] [33]. For example the same threat intelligence can be crafted to produce technical reports, executive insights, advisories and alerts, or even machine indicators [14].

Threat intelligence for ICS can come from many sources including government agencies like the Cybersecurity and Infrastructure Security Agency (CISA) [20], or private firms like Dragos, Inc. [35] or FireEye [45]. These services can provide valuable threat information about adversaries such as their TTPs as well as indicators or artifacts. Tracking TTPs, instead of just indicators like tools, artifacts, domains, IP addresses, and hashes, etc., allows defenders to better defend and track adversaries even while they constantly evolve [11]. It is trivial for an adversary to slightly modify code so that it generates a different hash, and even easier to switch IP addresses, but changing TTPs requires significant work [11]. Documenting and tracking TTPs requires a common language or lexicon for mapping cyber capabilities of adversaries or documentation of cyber-attacks. Two of the most common frameworks for this purpose

are MITRE's ATT&CK, Enterprise [85] and ICS [84], and the Director of National Intelligence's Common Cyber Threat Framework [27]. The use of frameworks allows defenders to better track, link, and trend adversary TTPs while compiling information from multiple sources [59]. CISA has recently started provided their alerts using MITRE's ATT&CK framework. For an example, see [22] for the CISA alert about ransomware impacting pipeline operations.

Threat intelligence can provide comprehensive information about adversary behaviors and targeting that can help inform proactive defense [33]. More explicitly, it helps defenders scope and scale protection, detection, and response activities [33] [14]. The value of threat intelligence is correlated with how actionable and digestable it is, reinforcing the need to provide it using a common lexicon [14] [59].

Site Specific Targeting

The IAEA recommends that DBTs be specific to the facility, its material, or adversary activities [61]. A key concept of DBT development is identifying those few critical functions that are relied upon to prevent radiological sabotage or theft of special nuclear materials [102]. These functions, sometimes called the Crown Jewels [34], will likely be different for every facility, and therefore, to be most effective, it should be conducted for each individual facility. There are existing methodologies describing ways of identifying systems or components that can contribute to HCEs such as Idaho National Laboratory's Consequence-driven Cyber-informed Engineering [49] and Dragos' Crown Jewel Analysis [34]. These two methodologies stand out from others like NEI 10-04 [92] and MIT [128], because they are adversarial-minded in their targeting and assessment scoping. Key areas of focus for targeting should include physical infrastructure and interdependencies, instances of horizontal application of technology, and reliance on automation and control capabilities [49]. Determining realistic impacts of potential consequences are best calculated with a cross-disciplinary experts including cybersecurity, engineering, and physical security [49] [91]. For nuclear power plants many of these consequences and impacts have already been documented in Final Safety Analysis Reports, Physical Security Plans, Technical Specifications, and failure scenarios. The identification of potential targets and their consequences specific to the facility allows one to filter and refine the credible threat intelligence and past cyber events to only those that are relevant to the facility, material, or adversary.

Relevance Determination Process

The relevance determination process is how one refines the abundance of information from various sources down to just those that are relevant to the facility. Past cyber events, credible threat intelligence, and site specific targets are the inputs for the relevance determination process. The various inputs are

combined and analyzed to determine overlaps. Specifically, an overlap in adversaries that have motivation to target the facility, equipment at the facility that can have a HCE, adversaries with the capabilities to cause the HCE in the system, and any past experience that shows motivation, opportunity, capability of specific adversaries or HCEs. The intersecting information is then drafted into the common lexicon, such as ATT&CK, in such a way that it is both actionable and tailored to the facility.

4.3 Results

In this example, a nuclear power plant uses a Triconex SIS in one of their systems that performs safety-related functions. A consequence-based target identification, as described in Section 4.2, was performed and determined that the Triconex SIS can contribute to a HCE, if compromised. The cross-disciplinary team identified two possible impacts or consequences from SIS manipulation: plant shutdown and an unsafe physical condition resulting in physical damage to the environment. A review of past cyber events discovered Triton, malware deployed against a petroleum and natural gas utility in the Middle East that compromised a Triconex SIS [30] [68]. Credible threat intelligence from Dragos World View [35] included an adversary group, Xenotime, that has targeted and compromised SIS in an oil & gas facility. The group also targets electric sectors within the Middle East, Europe, and North America [37] [86]. The MITRE ICS ATT&CK website included common TTPs used by Xenotime, specifically, Drive-By Compromise [T817], External Remote Services [T822], Valid Accounts [T859], and Supply Chain Compromise [T862], in addition to the custom developed tools and ICS-tailored malware [S0013] [86].

Below is a cyber DBT developed using the example above the proposed approach in Section 4.2:

> *Attempt to cause physical damage to safety instrumentation systems [T880] by a nation-state cyber adversary. The adversary has been known to use Drive-By Compromise [T817], External Remote Services [T822], Valid Accounts [T859], and Supply Chain Compromise [T862], in addition to the custom developed tools and ICS-tailored malware [S0013]. They have destructive capabilities, understand process implications, and have specific knowledge of industrial control systems. Willing to cause physical harm or kill. No collusion with insider.*

4.4 Analysis

The resulting cyber DBT, shown in section 4.3, contains detailed, relevant information tailored to the facility, material, or adversary. It specifies that the targets comprise of SISs, the adversary has destructive capabilities, understands process implications, and possesses specific knowledge of ICS. Furthermore, the

content of the cyber DBT is actionable. It is digestible in such way that defenders can map the adversary TTPs to applicable mitigations and layer detection and protection mechanisms along the Cyber Kill Chain [60] and or the ICS Cyber Kill Chain [4]. This cyber DBT enumerates the following adversary TTPs: Drive-By Compromise [T817], External Remote Services [T822], Valid Accounts [T859], and Supply Chain Compromise [T862], and Triton [S0013]. Figure 4.1 shows these TTPs in the form of an attack tree with an adversary goal of SIS exploitation.

Each adversary TPP has one or more associated mitigations within the ATT&CK framework vetted by the security community [85]. Using the MITRE Enterprise [85] and ICS ATT&CK [84] websites, one can easily identify these mitigations for each of the five TTPs used by Xenotime, the example adversary used in this chapter. There are eleven unique mitigations associated with these five adversary TTPs [86]. Figure 4.2 illustrates the mapping of these mitigations to the identified adversary TTPs. In the figure, the TTPs are in denoted by red font and arrows are drawn to the associated mitigations, written in green font. Figure 4.3 illustrates the mitigations, applied to the attack tree from Figure 4.1 with multiple layers of defenses (mitigations) placed between the adversary and their target, the SIS. It emphasizes the preferred approach of applying defensive measures, using defense-in-depth, to ensure no single point of failure exists. In addition to delaying the adversary, each barrier allows the defender an opportunity to detect the adversary and activate appropriate response and recovery measures.

There are thirteen mitigations in Figure 4.3 between the adversary and the SIS. Thirteen is a relatively low number in comparison to the hundreds of controls in listed in NEI 08-09 [91], especially when one considers that two of the thirteen are duplicates. Additionally, many of these mitigations can be clearly linked to common cybersecurity controls already required by the NRC. For example, network architecture described in NEI 08-09 appendix A [91], if implemented correctly, addresses three of the mitigations: Filter Network Traffic, Network Segmentation, and Limit Access Over the Network. Six others, External Remote Services, Set Antivirus, User Account Management, Password Policies, Update Software, and Vulnerability Scanning all are listed in NEI 08-09 Appendices D and E [91] and by rule should already be addressed by the plant. Only three mitigations, Restrict Web-Based Content, Multi-factor Authentication, and Keep Physical Key Switch in Run Mode are not explicitly listed in NEI 08-09. Restrict Web-Based Content [M1021] and Multi-factor Authentication [M1032] are common controls used in corporate IT environments and should also be considered for use within OT environments [85]. The last mitigation, Keep Physical Key Switch in Run Mode, is not cybersecurity control listed in the ATT&CK framework or NEI 08-09. It is a physical mitigation, detailed in the analysis of the Triton malware, done by Dragos [30], which was reviewed during the Credible Threat Intelligence and Past Cyber Event phases of this approach. This simple and cheap mitigation prevents malware like Triton, from modifying the

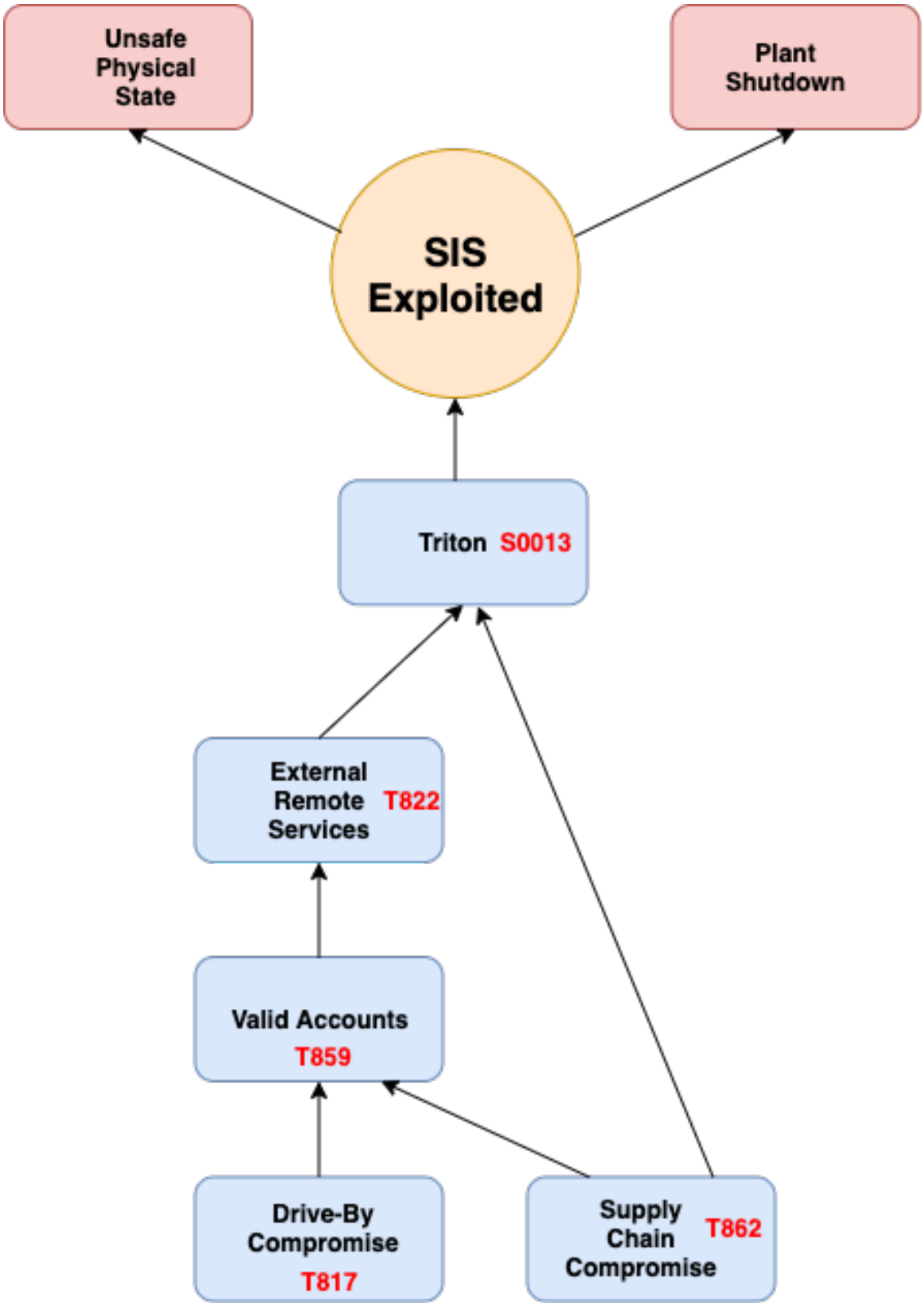

Figure 4.1: Adversary Attack Tree

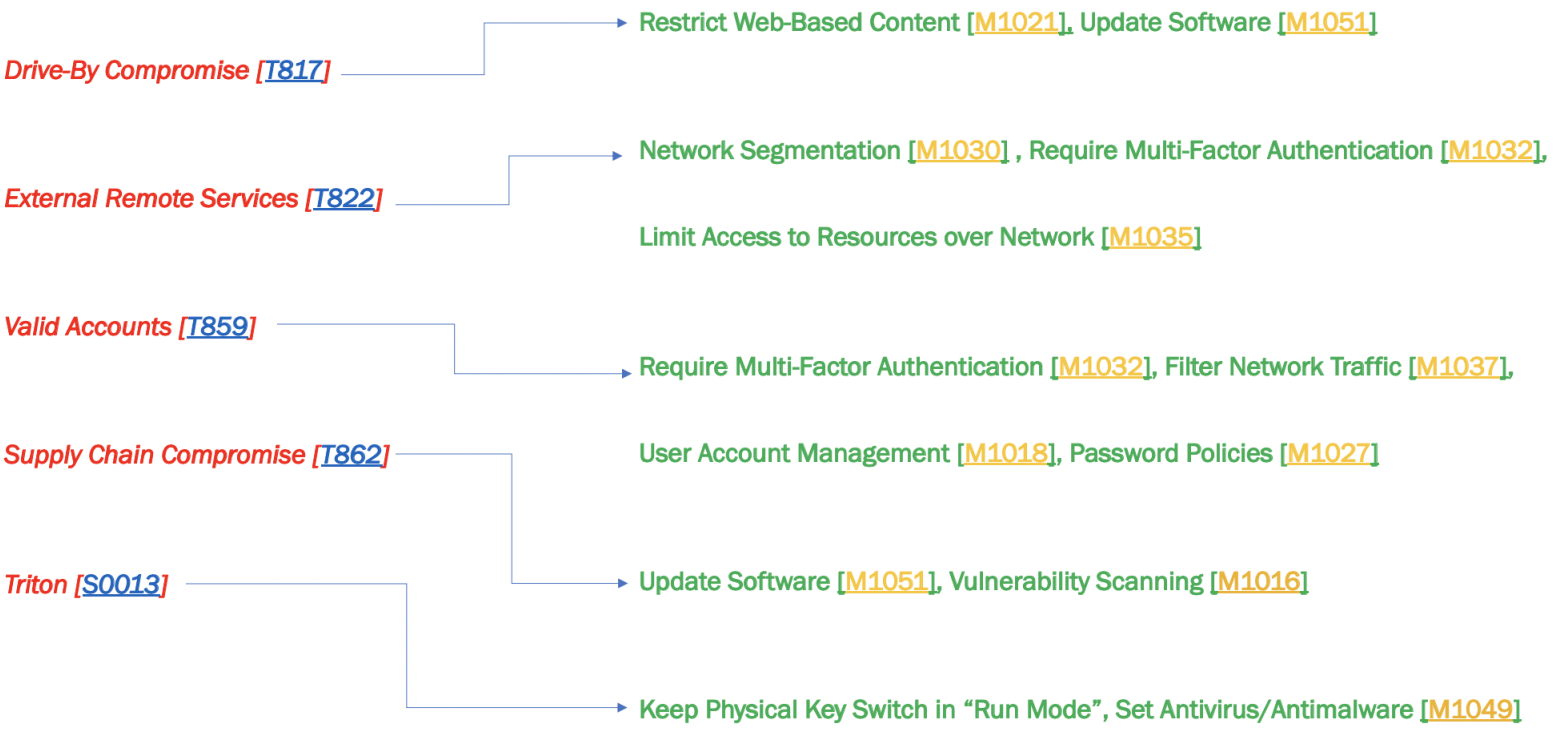

Figure 4.2: Mapped Mitigations

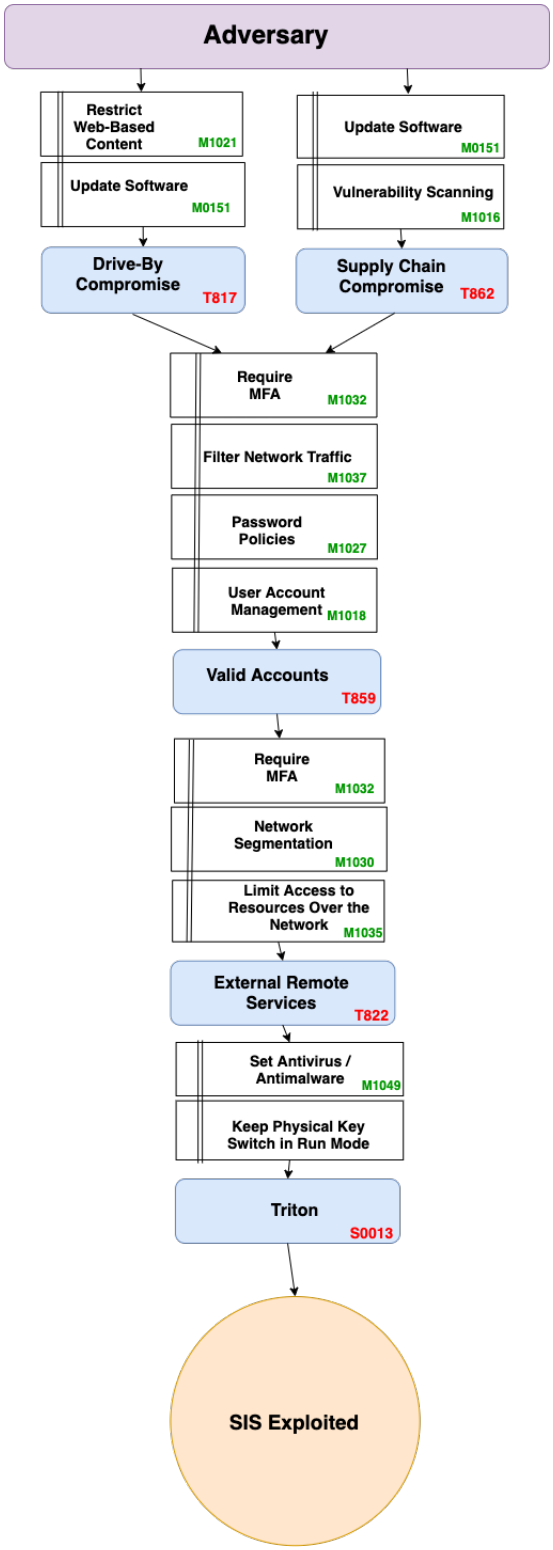

Figure 4.3: Overlapping Defensive Barriers

control logic of in the SIS. This mitigation, while not technical, is the most effective one to prevent the HCE described in the Cyber DBT. An additional potential defensive measure the plant could implement is modifying the milestone 5 rounds, already required by the NRC, to include verifying the position of the physical key on a periodic basis. This example highlighted how powerful mapping adversary tactics to mitigations can be for designing defenses. Engineers can follow this process to create customized and relevant cybersecurity design requirements in engineering modification packages.

4.5 Future Research

The "grand vision" of this approach is to integrate the resulting cyber DBT into a modeling and simulation or emulation framework. In its current state, plants can evaluate their effectiveness against the cyber DBT manually by validating each protection against the identified adversary tactic. The next goal for this research is to automate this evaluation process using adversary emulation software such as Cobalt Strike [23] or Scythe [123]. Both Cobalt Strike and Scythe use the MITRE ATT&CK framework, enabling easier integration with a cyber DBT developed using the approach proposed in this chapter. The researcher's ultimate goal is to create a system that integrates adversary emulation software, such as Scythe, with a nuclear reactor simulator, such as the one from Western Services Corporation [143], to test and evaluate cyber DBTs in an automated fashion. The combined use of adversary emulation and nuclear reactor modeling and simulation would allow plant operators to evaluate scenarios and defensive protections effectively, while maintaining minimal risk to operability and safety.

4.6 Conclusion

This chapter discussed the three main challenges that prevent plants from using traditional DBT analysis for developing cyber DBTss. The challenges included the rapid evolution of cyber capabilities and adversaries, the unpredictable nature of cyber-initiated events, and the concept of mal-operation. This chapter proposed a novel approach that combines threat-driven and consequence-driven cyber risk management practices with traditional DBT analysis to overcome these challenges. The resulting cyber DBT is detailed and customized to the facility, its material, or adversary, with quantitative data and analyses that can be tested and verified.

Chapter 5: Limiting Adversarial Lateral Movement with SDN

The best defense against deliberate acts, such as terrorism, sabotage, vandalism or theft, is the application of inherent security principles to facility design and operation. -Paul BayButt

5.1 Introduction

Until recently, OT networks, such as those in nuclear power plants, were not designed with security as a requirement or priority [2]. Their security relied upon traditional static defense mechanisms such as network air-gaps, obscure protocols and access mechanisms [1]. Research and case studies have shown that these mechanisms are insufficient in preventing targeted attacks from a persistent, well-funded adversary, such as nation-state [33] [32]. OT networks are frequently be referred to as "inherently insecure", meaning they trust data inputs without proper validation or authentication [4] [2] [1]. It is not uncommon for ICS to lack capability or support of even basic security features prevalent in IT environments [2] [1]. The lack of these features and capabilities makes securing nuclear power plants environments from cyber-attacks challenging. In recent years, it became common practice for the nuclear industry to apply traditional IT controls to their OT systems and networks [102] [3] [9] [91]. However, recent research suggest that a greater degree of overall security can be achieved by changing the way engineering systems are designed rather than adding security controls, after the completion of the design phases [2] [1]. The author posits that the best way to achieve high assurance that nuclear power plants are adequately protected against cyber-attacks, is to implement equipment that promotes an "inherently secure" environment. Unfortunately, "inherently secure" does not currently have a universally agreed upon definition. This chapter proposes a definition for "inherently secure" and discusses how SDN can promote an inherently secure environment at nuclear power plants.

5.2 Inherently Secure

A definition for "inherently secure" can be created by leveraging the already well-accepted term "inherently safe". "Inherently safe" is a commonly used expression, within the nuclear industry, to describe the physics of equipment or a process that meet certain criteria involving safety. For example, a process is said to be inherently safe, if it has a low level of danger even if things go wrong [56]. This mindset directly contrasts with processes where a high degree of hazard is controlled by protective systems such as redundancy or engineered controls. Furthermore, as perfect safety cannot be achieved, it is common

practice is to work towards an inherently safer design. “An inherently safer design is one that avoids hazards instead of controlling them, particularly by reducing the amount of hazardous material and the number of hazardous operations in the plant” [56]. The nuclear industry can apply these safety concepts and principles to security and work towards transitioning to an“inherently secure” environment. The author proposes the following definition for “inherently secure”. Equipment can be said to be inherently secure, if it is immune to many of the most commonly known vulnerabilities, such as those in the top 25 items on the Common Weakness Enumeration [83] list, as well as be resilient enough to resist zero-day attacks.

5.3 Software-Defined Networking

SDN technology is an approach to network management that can inherently mitigate many common cyber vulnerabilities [124]. SDN is, “the physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices” [108]. SDN began as an approach to improve network performance and monitoring by making network configuration more dynamic and similar to cloud computing than traditional network management [124]. SDN architecture is directly programmable, agile, centrally-managed, and programmatically configured [108]. SDN works by using flow match rules to whitelist network data flows [124]. Specifically, the “ingressing packets are matched against the ingress port, Ethernet source or destination Media Access Control (MAC) address, Ethertype, Virtual Local Area Network (VLAN) identifier, IP source or destination address, and so on” [124]. Network administrators can then go on to define specific actions for ingressing messages that match the various criteria [124]. SDN administrators can also utilize a set of counters to monitor the ingress and egress of traffic and the overall network health [124]. SDN simplifies the design of secure networks by allowing network administrators to easily create zones within network levels. In addition to providing an additional layer of security, zones simplify the administration, communication, and application of other computer security measures [124].

5.4 Solution

The defensive architecture implemented at a nuclear power plants rely on strong network boundary devices such as data diodes to isolate as many critical assets as possible [9] [91]. However, once on the isolated side, the networks are mostly flat and lack data flow controls to prevent lateral movement within a network level [2] [33]. Loss of functionality or operational concerns are often cited as reasons not to implement horizontal data flow security measures for network levels [52] [9] [32]. SDN can apply horizontal data flow controls throughout the network without operational concerns or functionality losses through

the use of its network flows and traffic engineering and is by design, inherently secure against many common network attacks [124] [121]. Traditional network management techniques utilize features like MAC tables, Rapid Spanning Tree Protocol (RSTP), and cast types for many securities and conveniences [124]. However, these features also make traditional networks vulnerable to cyber-attacks such as, MAC flooding and table poisoning, address resolution protocol Address Resolution Protocol (ARP) spoofing, Bridge Data Protocol Unit (BDPU) attacks, etc. [124]. In SDN, all network flows and backup paths are specifically defined in the controller, eliminating the need for MAC tables or RSTP [124]. Additionally, SDN does not rely on "cast types", because it uses traffic engineering to process forwarding behavior [124]. SDN uses flow match rules to whitelist network flows, significantly diminishing adversarial lateral movement within a network level [124] [121]. Lateral movement is one of the key steps identified in the ICS Cyber Kill Chain [4] that adversaries use in their attack campaigns against ICS networks. By preventing lateral movement and detecting unauthorized movement attempts, SDN can significantly empower the nuclear industry in their network defense and recovery in the event of a cyber-attack. Lastly, SDN, provides asset owners the capability to easily create zones within levels without negatively impacting operability or functionality [124]. Zones add an additional layer of isolation and simplify the administration, communication, and application of computer security measures [124] [64]. Using zones in addition to vertical network levels would align plants closer to the guidance and best practices described in the IAEA's NST047 [64], as opposed to NEI 08-09 [91] and NRC Regulatory Guide 5.71 [99].

5.5 Future Research

The next phase of this research effort will include experiments to support the definition of "inherently secure". The experiments will include conducting various network-based cyber-attacks attacks against an example ICS. The attacks will include a variety of tactics, including those in the top 25 items on the Common Weakness Enumeration [83] list. The experiment will conduct the attacks against two network architectures. The first, a network architecture that utilizes traditional networking equipment and the second using SDN equipment. The effectiveness of the cyber-attacks against the two architectures will be compared and contrasted. The ultimate goal of the experiments will be to provide empirical evidence that SDN can be used as an effective mitigation of the most common network-based cyber-attacks against ICSs. Additional experiments will be designed to demonstrate how to evaluate whether other types of equipment are "inherently secure". This may include assessing any digital equipment against the most commonly known vulnerabilities, such as those in the top 25 items on the Common Weakness Enumeration list. Because evaluating the effectiveness of cybersecurity controls is a regulatory requirement listed in NEI 08-09 [91], this future research may be useful for nuclear plants looking for guidance on how to

conduct those kinds of assessments.

5.6 Conclusion

The goal of achieving high assurance of the security of nuclear plant networks and systems will require more than traditional IT controls, such as those listed in NIST 800-53 [96], NIST 800-82 [127], or appendix D of NEI 08-09 [91]. High assurance will be easier to achieve, if the industry moves towards implementing equipment that promotes an "inherently secure" environment. Equipment can be said to be "inherently secure", if it is immune to many of the most commonly known vulnerabilities, such as those in the top 25 items on the Common Weakness Enumeration list, as well as be resilient enough to resist zero-day attacks. Using SDN cannot by itself, guarantee high assurance of the security of nuclear plant networks. However, SDN is better for OT environments than traditional networking, because by design, it mitigates many of the common tactics and pathways used by advanced adversaries. Ultimately, high security assurance will require a combination of a cross-disciplinary team of cybersecurity and engineering professionals, security-conscious practices and procedures, overlapping layers of protection and detection, in addition to using "inherently secure" equipment, such as SDN.

Chapter 6: Earlier Detection of the Active Insider With Face Recognition

We discovered in our research that insider threats are not viewed as seriously as external threats, like a cyberattack. But when companies had an insider threat, in general, they were much more costly than external incidents. This was largely because the insider that is smart has the skills to hide the crime, for months, for years, sometimes forever. -Dr. Larry Ponemon

6.1 Introduction

In many disciplines, including nuclear security, the insider threat is often cited as the most serious security problem (Probst). Even though nuclear security has been an established discipline for decades, the risk associated with a malicious insider is still considered the most difficult to mitigate. The biggest factor as to why this problem remains so tough to solve is the information a malicious insider has on the internal workings of equipment, processes, and people. Additionally, a malicious insider is significantly more likely to have unsupervised physical access to vital systems, equipment, or material. Nuclear power plants operate at a high level of safety to reduce the likelihood of the HCEs associated with nuclear energy. However, the potential for high consequence events such as a radiological release or power disruption to electric grid exist. These consequences, however, are not risks that can simply be accepted. The nuclear industry has implemented several programs, processes, and technology to address the risk associated with these types of consequences. A malicious insider armed with cyber capabilities poses a serious threat to nuclear power operations and presents a severe risk that must be reduced to an acceptable level.

6.2 Background

6.2 Risk Management

Risk can be calculated as a function of the impact or consequence of an undesired event, multiplied by the likelihood of the event's occurrence [67] [34] [18]. Consequences can be categorized a variety of ways, but in the context of this chapter and nuclear power operations, the author will categorize consequences as either catastrophic, unacceptable, undesirable, acceptable, and desirable. Likelihood is composed of three factors, the attractiveness of the target to the adversary, the capabilities of the adversary, and a vulnerability of the target [65]. In any risk mitigation strategy, risks are identified, quantified, and prioritized [18]. The quantification of each risk allows asset owners to prioritize risks and

their management based on their values [87]. There a several options for managing risks once they are identified, quantified, and prioritized. Risk can be reduced, mitigated, transferred, avoided, or accepted [87] [18]. Risk transference is when the risk is transferred from one party to another, such as an insurance policy [18] [87]. Risk avoidance is when an entity chooses to stop a course of action in order to completely avoid an identified risk [18] [87]. Risk acceptance is when a risk is identified and quantified, but control measures are not applied, this risk is simply accepted [87] [18]. Most risk management strategies focus on reducing or mitigating risk [18]. Reducing and mitigating risk requires the application of control measures [18] [87]. Reducing risk and mitigating risk are the same approach but with different outcomes [87][18]. Mitigation implies that the control measures have been put in place to completely prevent a risk from occurring [87]. Risk reduction is when control measures are applied that diminish the consequence and or likelihood of the event, but not completely [34] [18].

6.2 Cyber Risk Management

Cyber risk refers to a subset of risk that exists due to digital interconnectivity [18]. Residual risk refers to the amount of risk remaining after risks have been reduced by controls or mitigations [87]. In the context of cybersecurity and nuclear power reactors, residual cyber risk is the amount of cyber risk remaining, after the licensees have completed the full implementation of their cybersecurity programs, as described in Chapter 2.

6.2 Threat

In nuclear security, cyber is not considered a new threat. It is considered a new capability of existing threat actors [63]. "The Threat" refers to a threat actor with motivation, intention, and capability to commit intentional unauthorized acts directed at nuclear material [65]. A threat actor is an individual, group, organization, or government that conducts or has the intent to conduct detrimental activities [65]. In nuclear security, classifying threat actors is done for three reasons, to identify those threats that can impact nuclear operations, to understand actual threat characteristics, and distinguish between lower and higher consequence threat impacts [65]. There are three attributes to threat, motivation or intent, capabilities, and opportunity [65] [63] [33]. Motivation and intent are the reason and the goal the adversary seeks to achieve [65]. Traditional threat actor TTPs and motivations continue to evolve [33] [32]. Threat capabilities refer to the threat actor's ability and tools to successfully achieve their intent; TTPs [65]. Opportunity, in the context of threat, is the knowledge of vulnerabilities as well as the ability to leverage and exploit them to breach a system [65] [63]. Attack opportunities are created when threat capabilities exceed complexity to exploit vulnerability [65] [63]. Threat analysis is an essential

component of risk analysis [18]. Risk analysis includes the study of the susceptibility of targets to threat TTPs resulting in system accessibility and consequence of unauthorized access [18] [33]. Studying the threat is a critical component to building a robust cybersecurity program [91] [9] [33]. Threat analysis allows a security program to identify and evaluate threats specific to their site [33]. Threat motivation, intent, and capabilities can identify the attack opportunities and potential consequences [65] [33]. The likelihood of those opportunities and the consequences are used to calculate risk [18]. A risk management program is incomplete without threat analysis [18] [33].

The insider threat is subset of threat [103] [65]. For example, if a threat actor such as a political activist or nation state had motivation to attack a facility, they may look to utilize an insider to accomplish some or all of their goals. The insider may be just one component of an adversary's overall attack campaign [98] [141]. However, the insider would have or could gain valuable knowledge of the facility, its equipment, processes, procedures, and personnel [98] [103] [65] [141]. The insider also is more likely to have more or better opportunities to achieve their goals [98] [141]. An insider with cyber capabilities may circumvent control measures in place to defend against cyber-attacks originating from external sources, such as rogue wireless access points or bridging network air-gaps with portable media [36] [112] [32] [103] [98]. A cyber-enabled insider poses a serious threat to nuclear power operations and is a component of their risk management program [103] [102] [101] [98].

6.2 Current Insider Mitigation at Nuclear Power Plants

Nuclear security has been concerned about the insider threat long before cyber was threat capability [98]. 10 CFR. 73.55(b)(7) and (b)(9), provides the necessary flexibility for nuclear power plant licensees to address the complexities of an insider threat [103]. The NRC staff has, nonetheless, established the minimum criteria required to meet the DBT goal of mitigating the active insider, active violent insider, or passive insider in Regulatory Guide 5.77, Insider Mitigation Program. The minimum criteria include a "critical group" for individuals performing certain job functions, an initial security determination of employees, ongoing drug and alcohol testing, periodic psychological assessments, annual reviews by immediate supervisors, and periodic re-investigations [98].

6.2 Face Recognition Technology

Face recognition technology refers to technology capable of identifying a person from a digital images or videos [74]. While it is only in the last few years that face recognition capabilities have been taken seriously, it has roots as early back as the 1960s [74]. There are many differing variations and techniques for face recognition [74]. However, generally, it works by comparing selected facial features from a given image

with faces, known or unknown, from other images [74]. Face recognition has many use cases including thwarting criminals and terrorists, locating missing persons, medical diagnosis and assistance, purchase validation, advertising, and access control [74]. The author posits that face recognition technology can be implemented at nuclear power plants to reduce the cyber risk posed from active insiders.

6.3 Solution

6.3 Representative Physical Protection System

Modern PPS rely on numerous and varied types of digital components to provide protection, detection, and delay functions for the security of nuclear fuel and plant operation [61] [120] [75]. CDAs within PPS include edge devices, Field Distribution Boxes (FDBs), and the head end system [28] [61]. Edge devices include both interior and exterior sensors, access control mechanisms, and cameras. FDBs consist of PLCs including local processors, input/output panels, and multiplexing units [28]. The purpose of FDBs is to facilitate communication and control from the head end system to the edge devices [28]. This communication is usually specified to be redundant and independently routed [28]. The head end system, also called Access Control and Detection (AC&D) system, includes servers and workstations, such as the Local Alarm Station (LAS), Central Alarm Station (CAS), and Network Video Recorder (NVR). See Figure 6.1 for an overview of common digital assets within a PPS. Figure 6.2 illustrates a generic data flow of how PPS manage alarms, camera feeds, and other important security data points. Typically, an access authorization database containing biometric data of authorized personnel is used by the AC&D system to ensure physical access to various plant areas is controlled and managed [91] [28] [61]. See Figure 6.3 for a simplified network diagram of a typical PPS.

System Design

As discussed in Sections 6.2, nuclear power plants are concerned about the insider threat and have many processes and programs in place to lower the risk associated with an active insider. One of these processes is called the “two person rule”, which requires essential tasks and actions be taken by multiple people. The concept originates from the military and was designed to prevent accidental or malicious launch of nuclear weapons by a single individual [135]. The purpose of the proposed system is to use face recognition technology to reduce the cyber risk posed from active insiders by passively validating the two-person rule and detecting active insiders earlier in the Cyber Kill Chain [60]. At nuclear power plants, engineering modifications, also known as design changes, undergo an exhaustive and bureaucratic process that can take upwards of 18 months to complete [52] [55] [91] [41]. This novel system design

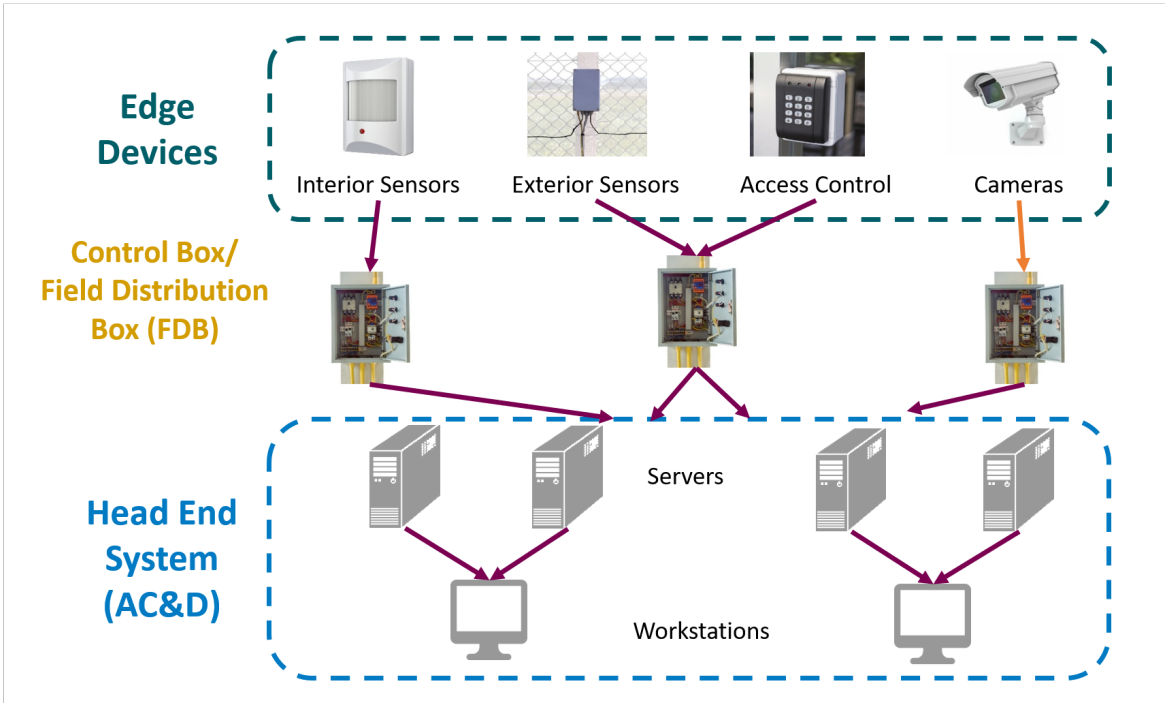

Figure 6.1: Digital Overview of a PPS

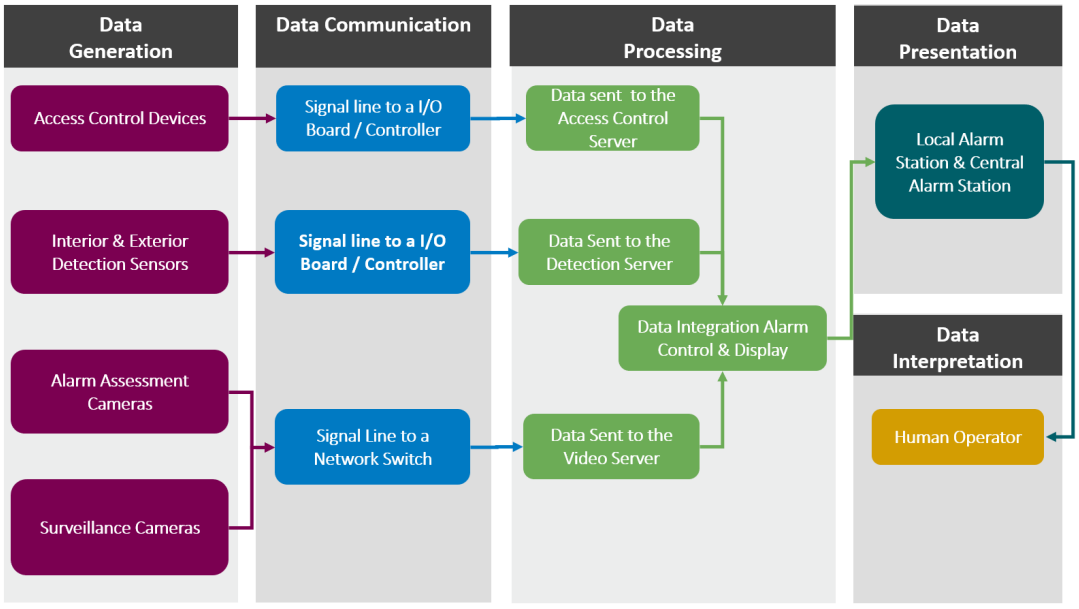

Figure 6.2: Data Management in a PPS

is focused on providing the desired functionality, while minimizing the design change effort and impact on network or system resources. In order to meet these design considerations, the system will integrate open-source face recognition libraries with the existing physical security infrastructure and equipment. Integrating into the existing system offers operational benefits in addition to the significant cost savings from not requiring new cameras, and workstations with expensive Graphics Processing Units (GPUs), etc. Operational benefits to this novel design include the ability to turn any camera at the facility into a face recognition camera, a central database for managing personnel access, and the ability to inject alerts into CAS and LAS.

The system will run on a virtual machine on one of the servers in the PPS network. It utilize an open-source software called Face_Recognition, that was designed for recognizing and manipulating faces using Python or the command line [50] . This software library is built using Dlib's state-of-the-art face recognition built with deep learning. DLib's model has an accuracy of 99.38% on the Labeled Faces in the Wild benchmark [50]. The system will leverage the OpenCV library to connect to the existing cameras on the PPS network. OpenCV, is a library of programming functions mainly aimed at real-time computer vision [109]. The system will use a simple web interface to enable a security officer to strategically select cameras for face detection and recognition. The officer can link the access authorization employee badge database and work authorization database to allow the system validate the two-person rule and detect active insiders earlier in the cyber-attack kill chain. [60]. When a face is detected, the system tries to recognize it. If recognized, then the system validates that the person has both access to that area and a reason, such as an authorized work order, to be in that area. If the face is not recognized, the system will simply try again on the next frame. This is because the system only checks one frame at a time and the person may not be directly facing the camera in every frame. The system will continue to try and recognize the face until the face leaves the frame. If the system sees a face, but cannot recognize it before it leaves the view of the camera, the system will log the anomaly and prompt security for a 'random' security round. Physical security programs usually require security personnel to do rounds of an area at both scheduled and random intervals [61] [103] [102] [98]. This proposed system can function as an initiator of random security rounds. A generic program flow for the proposed system is illustrated in Figure 6.4.

This initial proposed design can be quickly implemented it for four or less strategically chosen cameras, such as those at choke points or doors that lead to vital areas. Adding more than four cameras to the system as depicted in Figure 6.3 and Figure 6.4 may add undesirable stress to the PPS network resources. However, with a few modifications to the design, a facility could implement this system on most, if not, all of its cameras, allowing for additional operational benefits such as anomaly detection by fingerprinting

routes and times. The general program flow and design would be the same except a 'sensor' would be deployed for every four cameras to distribute the computation reduce network stress. The 'sensor' is a small computer such as a Raspberry Pi that monitors four or less cameras and passes relevant information back to the central node running on the virtual machine. The 'sensor' taps into the live feeds of the cameras using OpenCV and Python. It scans the feeds for faces. If a face is detected, then the 'sensor' generates a face encoding for that face. The encoding is an image of the face represented as a NumPy [106] array of 128 float numbers. A string representation of that encoding and the timestamp are stored locally on the 'sensor' in a database. The central node can choose a person, generate their face encoding, and then poll the 'sensors' to see all of the cameras that identified that face and when. This polling can effectively track a persons movements across the facility over a period of time. See Figure 6.5 for an illustrated diagram of the scaled system implementation. Additionally, this system can also be useful in emergency situations. Security could quickly and easily calculate the number of personnel in a given area, as well as their identity.

6.4 Conclusion

The nuclear industry has implemented several measures to reduce the HCE, cyber-attacks, and the insider threat. However, a malicious insider, armed with cyber capabilities continues to pose a serious threat to nuclear power operations and illustrates a severe risk that must be reduced to a more acceptable level. Face recognition technology has grown exponentially in recent years. It has spread from small, niche use cases like casinos to broad use by the general public, such as providing access control to personal mobile phones. Face recognition has many use cases including biometric access control and as shown in this chapter, can be integrated into the insider mitigation program to help address risks associated with the insider threat. The proposed system could passively audit and validate access to vital areas and the two-person rule. Further research for this work may include the expansion of this system to provide an early detection of malicious insiders through fingerprinting of their movements and access times. High risk scenarios often include the insider threat, due to their specialized knowledge and their frequent attack opportunities. Nuclear security must continue to evolve their control measures to address the risk posed by the insider, especially the cyber-enabled malicious insider.

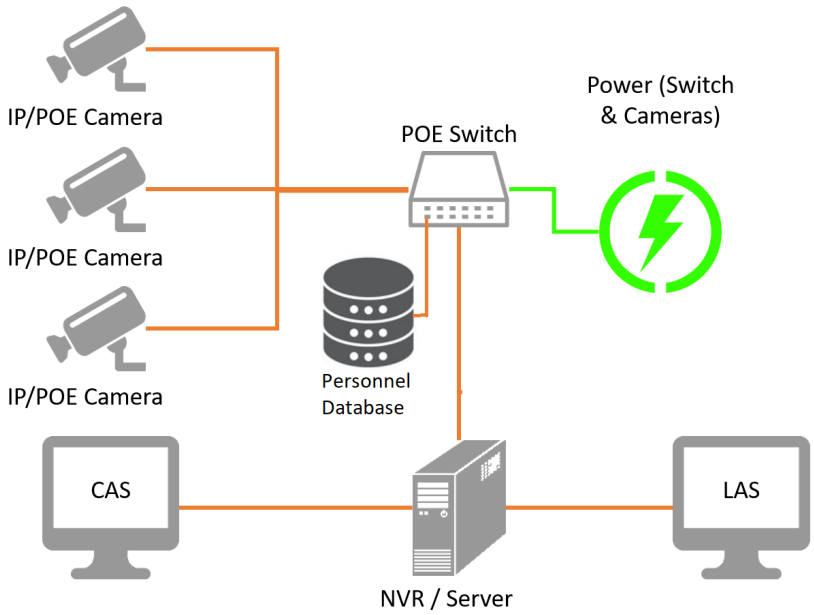

Figure 6.3: Simplified PPS Network

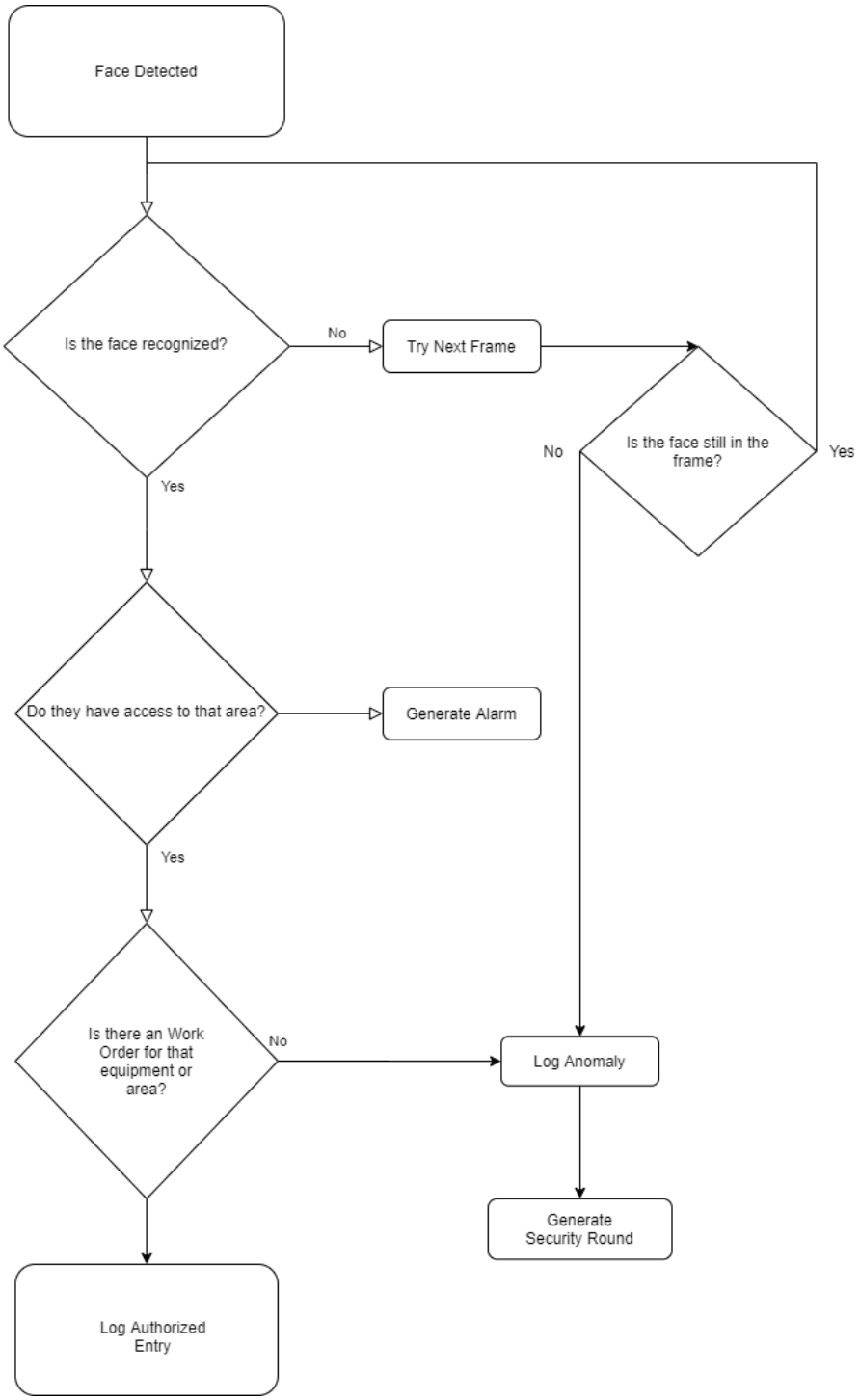

Figure 6.4: Passive Monitoring Program Flow

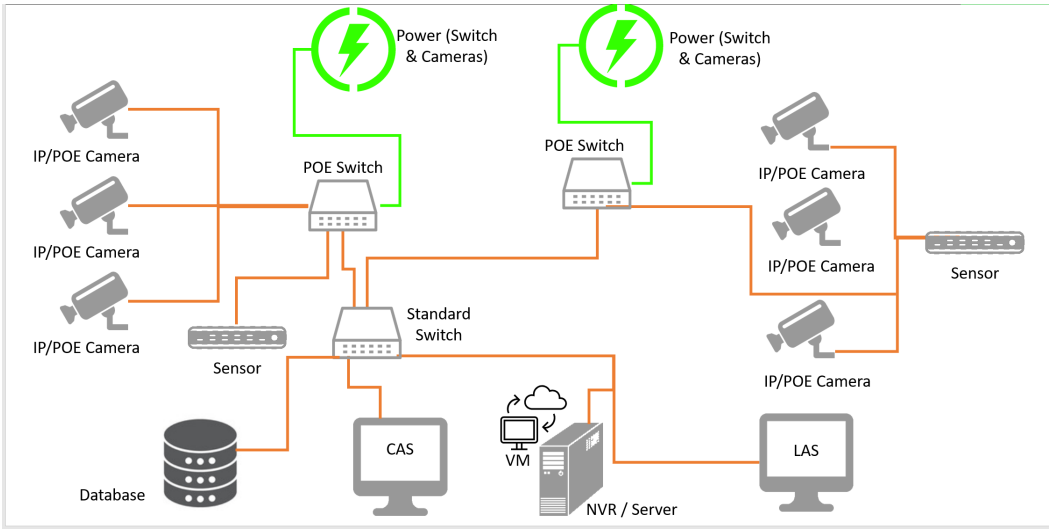

Figure 6.5: Scaled System Implementation

Chapter 7: Summary and Conclusions

There are 99 active nuclear power plants generating nearly 20% of the United States' electricity. Sophisticated adversaries, with destructive cyber capabilities, pose a threat to both the operation of these reactors and the health and safety of the public, because of their potential to cause HCEs using cyber-attacks. Nuclear power plants use DBT as the key input for designing systems that protect against acts of radiological sabotage, or prevent the theft of special nuclear material. DBT is a profile of the type, composition, and capabilities of an adversary. However, due to significant challenges, the United States chose to add cybersecurity requirements to their DBT, instead of using traditional DBT analysis for cyber. Unfortunately, as documented in Chapter 2, these cybersecurity requirements, even if implemented correctly, may not be sufficient to defend against a persistent adversary with advanced cyber capabilities. The use of a compliance-based approach left nuclear power plants unable to quantitatively measure their ability to defend against adversaries with cyber capabilities. This dissertation identified residual cyber risk at nuclear power plants, proposed a novel approach to developing cyber DBTs, and proposed two solutions for significantly increasing the cybersecurity posture of nuclear power plants. Quantifying cyber capabilities and their impacts to specific plant functions bounds the threat from cyber adversaries so that protections, detections, and responses can be scoped and scaled appropriately. Appendix A includes research ideas from this dissertation that the author recommends be expanded upon to continue the maturity of cybersecurity in the nuclear industry.

References

[1] Robert Anderson, Jacob Benjamin, Virginia Wright, Luis Quinones, and Jonathan Paz. Cyber-informed engineering. 2019.

[2] Robert Anderson and Joseph Price. Cyber-informed engineering: The need for a new risk informed and design methodology. 2015.

[3] Anastasios Arampatzis. What is nei 08-09? https://www.tripwire.com/state-of-security/ics-security/what-is-nei-08-09/, 2019.

[4] Michael J. Assante and Robert M. Lee. The industrial control system cyber kill chain. https://www.sans.org/reading-room/whitepapers/ICS/paper/36297, 2015.

[5] James Bamford. Nsa snooping was only the beginning. meet the spy chief leading us into cyberwar. https://www.vanityfair.com/news/2011/03/stuxnet-201104, 2013.

[6] BBC. Estonia fines man for 'cyber war'. http://news.bbc.co.uk/2/hi/technology/7208511.stm, 2008.

[7] BBC. Shamoon virus targets energy sector infrastructure. https://www.bbc.com/news/technology-19293797, 2012.

[8] BBC. Hack attack causes 'massive damage' at steel works. https://www.bbc.com/news/technology-30575104, 2014.

[9] James Beardsley. Nrc cyber security regulatory overview. https://www.nrc.gov/docs/ML1727/ML17278A744.pdf, 2017.

[10] Jacob Benjamin. Digital steganography: The effectiveness of current detection software. https://www.utica.edu/academic/library/Benjamin_J_2012.pdf, 2012.

[11] David Bianco. The pyramid of pain. https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html, 2013.

[12] B. E. Binde and et al. Assessing outbound traffic to uncover advanced persistent threat. https://www.sans.edu/student-files/projects/JWP-Binde-McRee-OConnor.pdf, 2011.

[13] Broadcom. The shamoon attacks. https://community.broadcom.com/symantecenterprise/communities/community-home/librarydocuments/

viewdocument?DocumentKey=281521ea-2d18-4bf9-9e88-8b1dc41cfdb6&CommunityKey=1ecf5f55-9545-44d6-b0f4-4e4a7f5f5e68&tab=librarydocuments, 2012.

[14] Sergio Caltagirone and Amy Bejtlich. Operationalizing threat intelligence in ics. https://www.sans.org/webcasts/disc-ics-virtual-conference-114285, 2020.

[15] CEA. The cost of malicious cyber activity to the u.s. economy. https://www.whitehouse.gov/wp-content/uploads/2018/03/The-Cost-of-Malicious-Cyber-Activity-to-the-U.S.-Economy.pdf, 2018.

[16] Mike Chapple. Malware jumps the air gap. http://www.gocertify.com/articles/security-matters-malware-jumps-the-air-gap.html, 2015.

[17] Anton Cherepanov and Robert Lipovsky. Industroyer: Biggest threat to industrial control systems since stuxnet. https://www.welivesecurity.com/2017/06/12/industroyer-biggest-threat-industrial-control-systems-since-stuxnet/, 2017.

[18] Jason Christopher. Threat management and nerc cip: Own it before you get owned. https://dragos.com/blog/industry-news/threat-management-and-nerc-cip-own-it-before-you-get-owned/, 2020.

[19] CISA. Russian government cyber activity targeting energy and other critical infrastructure. https://www.us-cert.gov/ncas/alerts/TA18-074A, 2018.

[20] CISA. Alerts — cisa. https://www.us-cert.gov/ics, 2020.

[21] CISA. Nuclear reators, materials, and waste. https://www.cisa.gov/nuclear-reactors-materials-and-waste-sector, 2020.

[22] CISA. Ransomware impacting pipeline operations. https://www.us-cert.gov/ncas/alerts/aa20-049a, 2020.

[23] CobaltStrike. Advanced threat tactics for penetration testers. http://www.cobaltstrike.com, 2020.

[24] C.P.Sumathi, T.Santanam, and G.Umamaheswari. A study of various steganographic techniques used for information hiding. *International Journal of Computer Science & Engineering Survey (IJCSES)*, 4(6), 2013.

[25] Richard Dahl. Fixing rg 5.71 / nei 08-09. http://cmplid.com/fixing-rg-5-71nei-08-09/, 2016.

[26] Jana Dittman and et al. Steganography and steganalysis in voice-over ip scenarios: operational aspects and first experiences with a new steganalysis tool set. *SPIE*, 5681, 2005.

[27] DNI. Cyber threat framework. https://www.dni.gov/index.php/cyber-threat-framework, 2020.

[28] DOE. Physical security systems assessment guide. https://www.energy.gov/sites/prod/files/2017/02/f34/PhysicalSecuritySystemsAssessmentGuide_Dec2016.pdf, 2016.

[29] DOE. 5 fast facts about nuclear energy. https://www.energy.gov/ne/articles/5-fast-facts-about-nuclear-energy, 2018.

[30] Dragos. Trisis analysis of safety system targeted malware. https://dragos.com/wp-content/uploads/TRISIS-01.pdf, 2017.

[31] Dragos. 2019 year in review: Ics vulerabiltiites. https://dragos.com/wp-content/, 2019.

[32] Dragos. 2019 year in review: Lessons learned from the front lines ics cybersecurity. https://dragos.com/wp-content/, 2019.

[33] Dragos. 2019 year in review: The ics landscape and threat activity groups. https://dragos.com/wp-content/, 2019.

[34] Dragos. Improving ot defense and response with consequence-driven ics cybersecurity scoping. https://dragos.com/blog/industry-news/combating-cyber-attacks-with-consequence-driven-ics-cybersecurity/, 2019.

[35] Dragos. Dragos world view. https://dragos.com/dragos-threat-intelligence/, 2020.

[36] Dragos. Top 5 ics cybersecurity myths. https://dragos.com/blog/industry-news/top-5-ics-cybersecurity-myths-whats-preventing-your-organization-from-reducing-risk/, 2020.

[37] Dragos. Xenotime. https://dragos.com/resource/xenotime/, 2020.

[38] DVICE. The 7 worst cyberattacks in history (that we know about). https://web.archive.org/web/20141112155600/http:/www.dvice.com/archives/2010/09/7_of_the_most_d.php, 2010.

[39] The Economist. Estonia fines man for 'cyber war'. https://www.economist.com/international/2007/05/24/newly-nasty, 2007.

[40] EICAR. Download anti malware testfile. https://www.eicar.org/?page_id=3950.

[41] EPRI. Cyber security technical assessment methodology. Technical Report 3002008023, Electric Power Research Institute, Palo Alto, CA, 2016.

[42] ESET. Vulnerability cve-2017-0144 in smb exploited by wannacryptor ransomware to spread over lan. http://support.eset.com/ca6443/?locale=en_US&viewlocale=en_US, 2017.

[43] Sean Bodmer et al. *Reverse Deception: Organized Cyber Threat Counter-Exploitation*. McGraw-Hill Osborne Media, 2012.

[44] Nicolas Falliere. Stuxnet infection of step 7 projects. https://www.symantec.com/connect/blogs/stuxnet-infection-step-7-projects, 2010.

[45] FireEye. Cyber threat intelligence. https://www.fireeye.com/solutions/cyber-threat-intelligence.html, 2020.

[46] Ben Flanagan. Former cia chief speaks out on iran stuxnet attack. https://www.whitehouse.gov/wp-content/uploads/2018/03/The-Cost-of-Malicious-Cyber-Activity-to-the-U.S.-Economy.pdf, 2011.

[47] J. D. Foley and A. Van Dam. *Fundamentals of Interactive Computer Graphics*. Addison-Wesley, Reading, MA, 1982.

[48] Thomas Fox-Brewster. An nsa cyber weapon might be behind a massive global ransomware outbreak. https://www.forbes.com/sites/thomasbrewster/2017/05/12/nsa-exploit-used-by-wannacry-ransomware-in-global-explosion/#2876effe599b, 2017.

[49] S. Freeman, C St Michel, R. Smith, and M. Assante. Consequence-driven cyber-informed engineering (cce). Technical Report INL/EXT-16-39212, Idaho National Laboratory, Idaho Falls, ID, 2016.

[50] Adam Geitgey. Ageitgey face_recognition. github.com/ageitgey/face_recognition, 2019.

[51] Ryan Gibson. Stego-lsb. https://pypi.org/project/stego-lsb/#description, 2019.

[52] Bill Gross. Cyber security and nuclear power. https://www.slcatlanta.org/KY2016/presentations/EE_William_Gross.pdf, 2016.

[53] Bill Gross. Cybersecurity threats are always evolving, so should nuclear cybersecurity regulation. https://www.nei.org/news/2019/cybersecurity-threats-evolving-nuclear-regulation, 2019.

[54] Michael Gross. A declaration of cyber-war. https://www.vanityfair.com/news/2011/03/stuxnet-201104, 2011.

[55] D. R. Harp and B. Gregory-Brown. It/ot convergence: Bridging the divide. https://ics.sans.org/media/IT-OT-Convergence-NexDefense-Whitepaper.pdf, 2019.

[56] Anna-Mari Heikkila. *Inherent safety in process plant design : an index-based approach*. PhD thesis, 1999-05-08.

[57] D. Hesse, J. Dittmann, and A. Lang. Network based intrusion detection to detect steganographic communication channels - on the example of images. IEEE, 2004.

[58] Robert Hirsch. *Exploring Colour Photography: A Complete Guide*. Laurence King Publishing, 2004.

[59] John D. Howard and Thomas A. Longstaff. A common language for computer security incidents. Technical Report SAND98-8667, Sandia National Laboratories, Albuquerque, NM, 1998.

[60] Eric M. Huchins, Michael J. Cloppert, and Rohan M. Amin. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. https://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf, 2011.

[61] IAEA. Guidance and considerations for the implementation of infcirc/225/rev.4, the physical protection of nuclear material and nuclear facilities. https://www-pub.iaea.org/MTCD/publications/PDF/te_967rev1_prn.pdf, 2000.

[62] IAEA. Development, use and maintenance of the design basis threat. https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1386_web.pdf, 2009.

[63] IAEA. Nuclear security threat assessment, design basis threats and representative threat statements. https://www.iaea.org/sites/default/files/18/08/nst058-dpp.pdf, 2016.

[64] IAEA. Computer security techniques for nuclear facilities. https://www-ns.iaea.org/downloads/security/security-series-drafts/tech-guidance/nst047.pdf, 2017.

[65] IAEA. Design basis threat. https://www.iaea.org/topics/security-of-nuclear-and-other-radioactive-material/design-basis-threat, 2019.

[66] IAEA. Power reactor information system. https://pris.iaea.org/pris/, 2020.

[67] Koen Van Impe. Simplifying risk managmenet. https://securityintelligence.com/simplifying-risk-management/, 2019.

[68] FireEye Intelligence. Triton attribution: Russian government-owned lab most likely built custom intrusion tools for triton attackers. https://www.fireeye.com/blog/threat-research/2018/10/triton-attribution-russian-government-owned-lab-most-likely-built-tools.html, 2018.

[69] James Judge. Steganography: Past, present, future. http://www.sans.org/reading_room/whitepapers/stenganography/steganography-past-present-future_552, 2001.

[70] James Judge. Steganography & digital watermarking tools. http://www.jjtc.com/Steganography/tools.html, 2011.

[71] Stanley Kaplan and B. John Garrick. On the quantitative definition of risk. *Risk Analysis*, 1(1):11–27, 1981.

[72] Gary Kessler. Steganography: Hiding data within data. http://www.garykessler.net/library/steganography.html, 2001.

[73] Dale Klein. Risk management and security. https://www.nrc.gov/docs/ML1006/ML100640508.pdf, 2010.

[74] Yaroslav Kuflinski. How ethical is facial recognition technology? https://towardsdatascience.com/how-ethical-is-facial-recognition-technology-8104db2cb81b, 2019.

[75] Alan J. Kuperman and Lara Kirkham. Protecting u.s. nuclear facilities from terrorist attack: Re-assessing the current "design basis threat" approach. In *INMM 54th Annual Meeting*, 2013.

[76] Ralph Langner and Perry Pederson. Why cyber security risk cannot simply be "managed" away. https://www.brookings.edu/wp-content/uploads/2016/06/cybersecurity_langner_pederson_0225.pdf, 2013.

[77] Robert M. Lee. Crashoverride: Analysis of the threat to electric grid operations. https://dragos.com/wp-content/uploads/CrashOverride-01.pdf, 2017.

[78] Max Londberg. Russia infiltrated kansas nuclear plant's business network, fbi and dhs say. https://www.kansascity.com/news/local/article205581509.html, 2018.

[79] Heather Mackenzie. Shamoon malware and scada security: What are the impacts? https://www.wired.com/2015/02/nsa-acknowledges-feared-iran-learns-us-cyberattacks/, 2012.

[80] Malwarebytes. Malwarebytes can't detect eicar test virus. https://forums.malwarebytes.com/topic/9994-malwarebytes-cant-detect-eicar-test-virus/, 2009.

[81] N. Meghanathan and L. Nayak. Steganalysis algorithms for detecting the hidden information in image, audio, and video cover media. *International Journal of Network Security & Its Application*, 2(1), 2010.

[82] Microsoft. Vsecurity update for windows xp sp3 (kb4012598). https://www.microsoft.com/en-us/download/details.aspx?id=55245, 2017.

[83] MITRE. 2019 cwe top 25 most dangerous software errors. https://cwe.mitre.org/top25/archive/2019/2019_cwe_top25.html, 2019.

[84] MITRE. Att&ck ics. https://collaborate.mitre.org/attackics/index.php/Main_Page, 2020.

[85] MITRE. Mitre att&ck. https://attack.mitre.org/, 2020.

[86] MITRE. Xenotime ics att&ck. https://collaborate.mitre.org/attackics/index.php/Group/G0001, 2020.

[87] Gregory Monahan. *Enterprise Risk Management: A Methodology for Achieving Strategic Objectives*. John Wiley & Sons, 2008.

[88] Ellen Nakashima. Us launched cyber attack on iranian rockets and missiles. https://www.washingtonpost.com/world/national-security/with-trumps-approval-pentagon-launched-cyber-strikes-against-iran/2019/06/22/250d3740-950d-11e9-b570-6416efdc0803_story.html, 2019.

[89] Ryan Naraine. Stuxnet attackers used 4 windows zero-day exploits. https://www.zdnet.com/article/stuxnet-attackers-used-4-windows-zero-day-exploits, 2010.

[90] NATO. Cyberwar - does it exist? https://www.nato.int/docu/review/2013/Cyber/Cyberwar-does-it-exist/EN/index.htm, 2013.

[91] NEI. Cyber security plan for nuclear power reactors. https://www.nrc.gov/docs/ML1011/ML101180437.pdf, 2010.

[92] NEI. Nei 10-04, revision 2, identifying systems and assets subject to the cyber security rule. https://www.nrc.gov/docs/ML1218/ML12180A081.pdf, 2012.

[93] NEI. Cyber security control assessments. https://www.nrc.gov/docs/ML1427/ML14276A144.pdf, 2014.

[94] Annalee Newitz. The bizarre evolution of the word 'cyber'. https://io9.gizmodo.com/today-cyber-means-war-but-back-in-the-1990s-it-mean-1325671487, 2013.

[95] NIJ. Digital evidence analysis: Steganography detection. http://nij.gov/topics/forensics/evidence/digital/analysis/steganography.htm, 2012.

[96] NIST. Recommended security controls for federal information systems and organizations. https://nvd.nist.gov/800-53, 2017.

[97] NRC. Nrc glossary. https://www.nrc.gov/reading-rm/basic-ref/glossary/risk-informed-decisionmaking.html.

[98] NRC. Insider mitigation program. https://www.nrc.gov/docs/ML1521/ML15219A609.pdf, 2008.

[99] NRC. Cyber security programs for nuclear facilities. https://www.nrc.gov/docs/ML0903/ML090340159.pdf, 2010.

[100] NRC. Guidance on cyber security plan implementation schedule. https://adamswebsearch2.nrc.gov/webSearch2/main.jsp?AccessionNumber=ML110600218, 2010.

[101] NRC. 73.54 protection of digital computer and communication systems and networks. https://www.nrc.gov/reading-rm/doc-collections/cfr/part073/part073-0054.html, 2016.

[102] NRC. Nrc regulations part 73.1 purpose and scope. https://www.nrc.gov/reading-rm/doc-collections/cfr/part073/part073-0001.html, 2017.

[103] NRC. Requirements for physical protection of licensed activities in nuclear power reactors against radiological sabotage. https://www.iaea.org/sites/default/files/18/08/nst058-dpp.pdf, 2017.

[104] NRC. Nrc: Strategic plan: Nureg-1614. https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1614/, 2018.

[105] NTI. References for cyber incidents at nuclear facilities. https://www.nti.org/analysis/tools/table/133/, 2016.

[106] NumPy. Numpy. https:www.numpy.org, 2020.

[107] Lindsey O'Donnell. The oceanlotus apt is using two new loaders which use steganography to read their encrypted payloads. https://threatpost.com/oceanlotus-apt-uses-steganography-to-shroud-payloads/143373/, 2019.

[108] ONF. Open networking foundation: Software-defined networking. https://www.opennetworking.org/sdn-definition/, 2019.

[109] OpenCV. Opencv. opencv.org/, 2019.

[110] OPSWAT. How to block malicious file uploads. https://info.opswat.com/how-to-block-malicious-file-uploads-opswat-apis-mx.

[111] OPSWAT. Webinar recap: Metadefender kiosk and nrc milestone 8 inspections. https://www.opswat.com/blog/webinar-recap-metadefender-kiosk-and-nrc-milestone-8-inspections, 2017.

[112] OPSWAT. Nuclear — opswat. https://www.opswat.com/blog/tag/nuclear, 2019.

[113] Perry Pederson. Aurora revisited: by its original project lead. https://www.langner.com/2014/07/aurora-revisited-by-its-original-project-lead/, 2014.

[114] Nicole Periroth. In cyberattack on saudi firm, u.s. sees iran firing back. https://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html, 2012.

[115] Nicole Periroth and David E. Sanger. Cyberattacks put russian fingers on the switch at power plans, u.s. says. https://www.nytimes.com/2018/03/15/us/politics/russia-cyberattacks.html, 2018.

[116] Nicole Perlroth and Scott Shane. In baltimore and beyond, a stolen n.s.a. tool wreaks havoc. https://www.nytimes.com/2019/05/25/us/nsa-hacking-tool-baltimore.html, 2019.

[117] John Peterson, Michael Haney, and R.A. Borrelli. An overview of methodologies for cybersecurity vulnerability assessments conducted in nuclear power plants. *Nuclear Engineering and Design*, 346:75–84, 2019.

[118] F.A. Petitcolas and et al. Information hiding: A survey. *IEEE*, 1999.

[119] Associated Press. Us launched cyber attack on iranian rockets and missiles. https://www.theguardian.com/world/2019/jun/23/us-launched-cyber-attack-on-iranian-rockets-and-missiles-reports, 2019.

[120] Tudor Radulescu. Correlated analysis of physical protection and cyber security measures for nuclear sites. In *Proc. International Scientific Conference Strategies XXI*, volume 2 of *The Complex and Dynamic Nature of the Security Environment*, pages 242–251, Bucharest, Romania, 2015. University of Bucharest.

[121] Rene Rietz, Radoslaw Cwalinski, Hartmut Konig, and Andreas Brinner. An sdn-based approach to ward off lan attacks. Hindawi, 2018.

[122] Neil J. Rubenking. Is your antivirus working? https://www.pcmag.com/news/is-your-antivirus-working, 2013.

[123] Scythe. Adversary emulation with scythe. http://www.scythe.io, 2020.

[124] SEL. Software-defined networking. https://selinc.com/solutions/p/software-defined-network/, 2019.

[125] Mike Sieffert, Rodney Forbes, Charles Green, Leonard Popyack, and Thomas Blake. Assured information security: Stego intrusion detection system. DFRC, 2004.

[126] Georg Steinhauser, Alexander Brandl, and Thomas E. Johnson. Comparision of the chernobyl and fukushima nuclear accidents: A review of the environmental impacts. volume 487, pages 800–817. ELSEVIER, 2013.

[127] Keith Stouffer, Suzanne Lightman, Victoria Pillitteri, Marshall Abrams, and Adam Hahn. Guide to industrial control systems (ics) security. https://csrc.nist.gov/publications/detail/sp/800-82/rev-2/final, 2015.

[128] John Thomas, Francisco Luiz de Lemos, and Nancy Leveson. Evaluating the safety of digital instrumentation and control systems in nuclear power plants. Technical Report NRC-HQ-11-6-04-0060, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2012.

[129] Iain Thomson. Exhibitionist shamoon virus blows pcs's minds. https://www.theregister.co.uk/2012/08/17/shamoon_malware_energy/, 2012.

[130] Nathan Thornburgh. The invasion of the chinese cyberspies (and the man who tried to stop them). https://courses.cs.washington.edu/courses/csep590/05au/readings/titan.rain.htm, 2005.

[131] Ellen Nakashima-Craig Timberg. Nsa officials worried about the day its potent hacking tool would get loose. then it did. https://www.washingtonpost.com/business/technology/nsa-officials-worried-about-the-day-its-potent-hacking-tool-would-get-loose-then-it-did/2017/05/16/50670b16-3978-11e7-a058-ddbb23c75d82_story.html, 2018.

[132] Ian Traynor. Russia accused of unleashing cyberwar to disable estonia. https://www.theguardian.com/world/2007/may/17/topstories3.russia, 2007.

[133] Tresys. Tresys: Removable media threat protection. https://www.tresys.com/solutions/network-perimeter-defense/removable-media-threat-protection, 2019.

[134] Phillip L. Turner, Timothy A. Wheeler, and Matt Gibson. Risk informed cyber security for nuclear power plants. In *10th International Topical Meeting On Nuclear Plant Instrumentation, Control, and Human Machine Interface Technologies*, 2017.

[135] USAF. Nuclear surety tamper control and detection programs. https://fas.org/irp/doddir/usaf/afi91-104.pdf, 2013.

[136] Danielle Veluz. Stuxnet malware targets scada systems. https://www.trendmicro.com/vinfo/us/threat-encyclopedia/web-attack/54/stuxnet-malware-targets-scada-systems, 2010.

[137] Verizon. 2019 data breach investigations report. https://enterprise.verizon.com/resources/executivebriefs/2019-dbir-executive-brief.pdf, 2019.

[138] VirusTotal. Virus total. https://www.virustotal.com/gui/home/upload, 2020.

[139] Kristy Westphal. Ssteganography revealed. http://www.symantec.com/connect/articles/steganography-revealed-data-in-data, 2010.

[140] Davey Winder. Trump declares national emergency as foreign hackers threaten u.s. power grid. https://www.forbes.com/sites/daveywinder/2020/05/02/trump-declares-national-emergency-as-foreign-hackers-threaten-us-power-grid/#e7d38523497f, 2020.

[141] James E. Wingate, Glenn D. Watt, Marc Kurtz, Chad W. Davis, and Robert Lipscomb. Defending against insider use of digital steganography. *Annual ADFSL Conference on Digital Forensics, Security and Law*, 2007.

[142] Zack Wittaker. Steganography & digital watermarking tools. https://techcrunch.com/2018/12/17/malware-commands-code-twitter-hidden-memes/, 2018.

[143] WSC. Nuclear power plant simulation. https://www.ws-corp.com/default.asp?PageID=2&PageNavigation=Nuclear-Power-Plant-Simulation, 2020.

[144] Richard Wyman. Consider the consequences: A powerful approach for reducing ics cyber risk. In *Cyber Security: A Peer-Reviewed Journal*. Henry Stewart Publications, 2017.

[145] Kim Zetter. An unprecedented look at stuxnet, the world's first digital weapon. an unprecedented look at stuxnet, the world's first digital weapon. www.wired.com/2014/11/countdown-to-zero-day-stuxnet/, 2014.

[146] Kim Zetter. The nsa acknowledges what we all feared: Iran learns from us cyberattacks. https://www.wired.com/2015/02/nsa-acknowledges-feared-iran-learns-us-cyberattacks/, 2015.

[147] Kim Zetter. The ukrainian power grid was hacked again. https://www.vice.com/en_us/article/bmvkn4/ukrainian-power-station-hacking-december-2016-report, 2017.

Appendix A: Ideas For Future Research

This appendix consists research topics intended to stimulate conversation about future work in the fields of computer science, nuclear engineering, and where they converge in the cybersecurity of nuclear power plants.

A.1 Further Validation of Cyber DBT Approach

In Chapter 4, the validation of the cyber DBT was limited to a hypothetical nuclear facility. The author acknowledges that a true validation would require a small scale feasibility study be conducted at an real operational nuclear facility in order to evaluate the feasibility, duration, cost, etc. of a full-scale implementation. The author believes the best facility for such a study would be Barakah Nuclear Power Plant in the United Arab Emirates, because of its increased digital footprint in comparison to older plants and its complex threat environment.

A.2 Automation of Cyber DBT Development

The proposed approach for developing cyber DBTs in Chapter 4, was a manual process that included identifying targets, refining threat intelligence, and reviewing past events. The next step in evolution for this research effort would be to automate the process through the development of a software tool. This tool would increase the efficiency and scalability of the solution, enabling easier adoption of the approach by the nuclear industry.

A.3 Steganography Detection in Network Monitoring

Steganography detection on host machines already has many viable options [10]. The residual risk identified in Chapter 3 could be mitigated by adding one of those detection programs into the PMMD kiosk software. While some research has been done on the topic of detecting steganography using network monitoring [57] [125], fifteen years later, a commercially available network intrusion detection system that detects steganography is still not available.

A.4 Face Recognition Feasibility Study

Chapter 6 provided a high level system design for implementing open source face recognition technology on an existing PPS at a nuclear facility. The next step for this research effort is a small scale feasibility study be conducted at on PPS equipment configured similarly to those in operational nuclear facilities. This follow-on study would evaluate the feasibility, duration, cost, etc. of a full-scale implementation.

The author believes the best candidates for this study would a facility used for physical security training, such as Tech Area V at Sandia National Laboratories in the United States and the Center of Excellence on Nuclear Security in Beijing, China.

A.5 Validation of Inherent Security

The next phase of the research documented in Chapter 5, should include experiments to support the definition of "inherently secure". The experiments should include conducting various network-based cyber-attacks attacks against an example ICS. Ideally, the attacks would include a variety of tactics, including those in the top 25 items on the Common Weakness Enumeration [83] list. The experiment could conduct the attacks against two network architectures. The first being a network architecture that utilizes traditional networking equipment and the second using SDN equipment. The resulting data could be compared and contrasted to determine the effectiveness of SDN against typical cyber-attacks.

This validation of inherent security could also be applied to other types of equipment. The experiments could be expanded to include assessing any digital equipment against the most commonly known vulnerabilities, such as those in the top 25 items on the Common Weakness Enumeration list. Because, evaluating the effectiveness of cybersecurity controls is a regulatory requirement listed in NEI 08-09 [91], this research may be useful for nuclear plants looking for guidance on how to perform those kinds of assessments.