INL/EXT-20-59885
Revision 0

Framework for Cyber Risk Management

September 2020

Katya Le Blanc
Shannon Eggers
Robert Youngblood

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory
operated by Batelle Energy Alliance, LLC

INL/EXT-20-59885
Revision 0

Framework for Cyber Risk Management

Katya Le Blanc
Shannon Eggers
Robert Youngblood

September 2020

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Page intentionally left blank

ABSTRACT

The purpose of this report is to describe the activities performed in the risk management area to support characterizing and a managing cyber security risk in the nuclear industry, and to provide a general framework for cybersecurity risk management to inform future research and development for risk analysis methods in the nuclear industry. The purpose of the work in the risk management area is not to replace or supersede existing practice, it is meant to highlight the unique challenges facing the nuclear industry in implementing cybersecurity risk management, and to provide guidance on developing robust risk methods that are consistent with existing methods, policies, and regulation in the US commercial nuclear industry. Specifically, this report constitutes the deliverable for milestone M3CT-20IN1101019 "DOE-NE Cybersecurity for Nuclear Facilities report on Risk Management Framework"

Page intentionally left blank

CONTENTS

FIGURES

TABLES

Page intentionally left blank

ACRONYMS

R&D	Research and Development
NEET	Nuclear Energy Enabling Technologies
CTD	Crosscutting Technology Development
PRA	Probabilistic Risk Assessment
US	United States
NRC	Nuclear Regulatory Commission
DBA	Design Basis Accident
ETA	Event Tree Analysis
FTA	Fault Tree Analysis
HAZOP	Hazard and Operability
NIST	National Institute of Standards and Technology
ICT	Information Communication Technology
NPP	Nuclear Power Plant
I&C	Instrumentation and Control Systems
SSCs	Systems Structures and Components
TTPs	Tactics Techniques and Procedures
ICS	Industrial Control Systems
NEI	Nuclear Energy Institute
CSP	Cyber Security Plan
CDA	Critical Digital Asset
STPA	System-Theoretic Process Analysis
FMEA	Failure Modes and Effects Analysis Failure Modes
FMVEA	Vulnerabilities, and Effects Analysis
IMECA	Intrusion Modes and Effects Criticality Analysis
SCADA	Supervisory Control and Data Acquisition
CVSS	Common Vulnerability Scoring System
CVE	Common Vulnerabilities and Exposures
AT	Attack Trees
VT	Vulnerability Trees
SAG	Security Argument Graph
NESCOR	National Electric Sector Organization Resource
GT	Game Theory
BDMP	Boolean Logic Driven Markov Processes

HHM	Hierarchical Holographic Modeling
CHASSIS	Combined Harm Assessment of Safety and Security for Information Systems
AHP	Analytic Hierarchy Process
SVM	Support Vector Machine
CVE	Common Vulnerabilities and Exposures
NVD	National Vulnerability Database
SW	Software
HW	Hardware
FSAR	Final Safety Analysis Report
NERC	North American Electric Reliability Corporation
CIP	Critical Infrastructure Protection
API SRA	American Petroleum Institute's Security Risk Assessment
TEPA	Top Event Prevention Analysis
DA	Digital Asset
HAZCADS	Hazard and Consequence Analysis for Digital Systems
CCF	Common Cause Failure

Page intentionally left blank

Framework for Cyber Risk Management

1. Introduction

The Nuclear Energy Enabling Technologies Crosscutting Technology Development (NEET CTD) cybersecurity research, development and deployment (RD&D) program is a cross-cutting research program that is intended to bridge critical gaps related to cybersecurity to enable advances in Department of Energy-Nuclear Energy Research and Development (R&D) programs. The NEET CTD program addresses needs and requirements that are common to multiple programs. These needs, in turn, will be translated into actionable R&D activities and coordinated with the other R&D programs. This report addresses the activities in the risk management area of the program.

The purpose of this report is to describe the activities performed in the risk management area to support characterizing and a managing cybersecurity risk in the nuclear industry, and to provide a general framework for cybersecurity risk management to inform future research and development for risk analysis methods in the nuclear industry. There are many existing cybersecurity risk management frameworks (e.g., Ross, 2012; Radack 2011). The purpose of the work in the risk management area is not to replace or supersede existing practice. The area aims to highlight the unique challenges facing the nuclear industry in implementing cybersecurity risk management, and to provide guidance on developing robust risk methods that are consistent with existing methods, policies, and regulation in the US commercial nuclear industry.

Nuclear power plant control systems have historically been largely analog and mostly isolated from networks that attackers could exploit. As plants embark on modernization efforts to increase efficiency and develop cost savings, many digital technologies are likely be incorporated into the control room and in the field for maintenance and field operations. While most utilities may avoid modifying safety systems, the consequences of a cyber attack on upgraded non-safety systems could have severe economic consequences, and risk analysis needs to effectively establish that there are not unidentified interactions that could have consequences to safety systems. Further, new reactors including light water technology small modular reactors, advanced reactors, and microreactors are employing mostly digital control systems, higher levels of automation, and are considering operational concepts such as remote and autonomous operations. Key to the successful deployment of these technologies is a capability to effectively characterize the consequences and likelihood of cybersecurity events including attacks from a determined adversary.

Although cyber risks have received increasing attention in the nuclear industry (Kim, 2014), there is no consensus on how to quantify or even prioritize cyber risks for the industrial control systems in nuclear power plants (Cherdantseva et al. 2016). Because of rapidly changing technologies and rapidly adapting adversary capabilities, understanding possible cyber-attack scenarios is significantly more difficult than understanding the scenarios associated with other system-level failure modes of engineered systems. The methods for generating scenarios for analysis is well developed for a range of system types and failure modes in the nuclear industry but these methods were not developed for cyberattack, and will need to be adapted for use in characterizing cyber risk in the nuclear domain.

This report includes a general introduction to risk management and the simplified risk framework developed on this project. The annexes of this report include detailed documentation of the main activities conducted in the project, which include a review of risk analysis methods for cybersecurity, a description of the development of tools to support top event prevention analysis as part of the risk analysis process, and a brief description of fuzzy analytical hierarchy process (AHP) as a possible tool for quantification in the risk analysis process. The report is structured in this way because the

documentation provided in the annexes are meant to be stand-alone descriptions of the activities described. The relevant details are summarized in the framework.

2. Risk Management

Annex A provides a description of risk management and a comprehensive review of risk analysis methods for cybersecurity. In summary, risk management involves identifying risks, quantifying, or prioritizing those risks against a threshold of risk tolerance, mitigating the risks, and monitoring or evaluating the mitigations over time. The nuclear industry has robust quantitative risk analysis methods for identifying scenario and failure modes in engineered physical systems, namely probabilistic risk assessment (PRA). There are existing frameworks and methodologies for assessing cybersecurity risks. However, as described in Annex A, they are either qualitative in nature, which can sometimes make it challenging to prioritize risks, or are immature and untested.

The first step in risk management is risk analysis. The primary method that commercial NPPs use to evaluate risk is PRA. The PRA technique described by the US Nuclear Regulatory Commission (NRC, 1975) attempts to estimate public risks posed by NPPs by examining the potential paths by which nuclear fuel could melt and release radiation to the environmental. PRAs in the nuclear industry have evolved into a logical framework for identifying the likelihood and consequences of design basis accidents (DBA)—those postulated accidents that could lead to radiation release impacting the health and safety of the public.

PRAs are model-based graphical techniques that use plant assets and design along with historical data (i.e., vendor, plant, and industry data on equipment and events) to determine the likelihood of an event and the frequency of potential consequences. In the nuclear industry, a PRA using event tree analysis (ETA) and fault tree analysis (FTA) results in the development of ‘minimal cutsets’ and estimation of core damage frequency. Minimal cutsets are the sequence of events or failures that must happen for a top event to occur in a fault tree analysis (FTA) model.

2.1 Challenge of Including Cybersecurity in Risk Analysis

As stated above, risk assessment methods for commercial nuclear power plants are mature and commercial power plants use the methods to characterize safety risks to the public. These methods were developed and refined in a time when the technology was largely analog or locally controlled. Consequently, the risks to these facilities were typically the result of design flaws, random failures, or degradation of the equipment over time. While challenging, capturing and modeling these failures is relatively straightforward. With the introduction of digital equipment and networked control systems, the risk to systems depend on the reliability of both the physical systems and information systems in these facilities. These information systems introduce new ways in which systems can fail, especially when it comes to the security of those information systems.

System risk is generally characterized as a set of scenarios, associated frequencies (or likelihoods), and associated consequences. In a physical system, modeling the impact requires a thorough understanding of the components in the system, the interactions among those components, and the likelihood of failure of the components in the system. Information systems add another layer of complexity to modeling risk because they may add previously unknown interactions among components in the physical system; they are also additional components that can fail themselves. Additionally, since threats to information systems include intelligent adversaries who may try to intentionally compromise or sabotage the system (rather than simply considering random failures or degradation), determining what can happen and how likely it is becomes ever more complex.

For a risk assessment method to adequately capture the risk that digital and information systems introduce (also known as cybersecurity risk), the method must capture the new ways in which the system can fail. The challenge is that digital components, which often contain general purpose computing power, can have new and often unexpected impacts to the system. By definition, general purpose computers can be programmed to do anything, so identifying how a digital component can influence a system, includes identifying where there are digital components the system, what those components are designed to do, and what they *can* be made to do (in addition to considering what they are supposed to do).

2.2 Incorporating Cybersecurity into Risk Analysis

A starting point for considering cybersecurity risks is to model the system by carefully considering how digital components can influence system operation, including both how digital systems are *intended* to be operated and how they *can* be operated. This means in addition to modeling the physical components and systems, the risk analyst needs to model the how data flows within the system, what physical actions digital components can influence, and the consequences of compromised control logic. In essence, the analyst needs to consider system consequences of any arbitrary change to inputs, outputs, or logic of any digital device.

There are several methods that can be used to model the interaction between the physical system and the digital components. One of those methods is using the unsafe control actions defined by Levensen (2011) in systems theoretic process analysis to develop cyber-informed faults trees that can be used in PRA (Williams & Clark, 2019). Another method that can capture the potential hazards between digital components and the physical system is hazard and operability (HAZOP) analysis (Dunjó, Fthenakis, Vílchez, & Arnaldos, 2010). Neither of these methods guarantee that all of the scenarios and consequences of digital components or information systems will be captured. The degree to which they capture the full set of possible failures induced by digital components and cybersecurity risks depends on what factors the analyst considers, how well she understands the interactions between digital components, information systems, and physical components, and how creative she is in identifying ways to cause the system to fail. This is not solely a limitation of cyber risk analysis, it is also present in traditional PRA; however, in cyber risk analysis it is exacerbated due to the larger problem space.

Another important consideration for the risk analyst is to adopt an adversarial mindset. This means that the analyst must consider what a determined saboteur would do or try to do if they had access to the system. Failing to adopt an adversarial mindset may limit the scope of the analysis to mundane faults, scenarios, and consequences and may not capture meaningful hazards that exist on the system.

Regardless of the method used to identify potential hazards due to digital components, the output of this analysis will be input into the scenarios developed for the PRA. Scenarios in the traditional PRA describe which initiating events and failures lead to a consequence. The same is true for cybersecurity risk, but the set of things that the analyst needs to considered is much larger. The analyst needs to identify what can happen if information the system needs to perform its functions is intercepted, modified, or interrupted as well as identifying how that information is compromised. Furthermore, the information the system needs to perform its function and the pathway by which that information is compromised should also be considered.

2.3 Prioritizing Assets with Top Event Prevention Analysis

Once the scenarios and consequences have been identified, the risks need to be prioritized. Annex B provides a description of the work conducted to develop tools to perform top event prevention analysis to prioritize the systems, structures, and components that further and detailed cybersecurity analysis need to focus on. In traditional PRA, risk can be written as a relatively straightforward function of consequence likelihood and scenarios. Likelihoods can be relatively straightforwardly developed from component testing, historical and operational data, and the PRA models. When adding cybersecurity into the equation, the scenarios, likelihoods, and consequences are much more complicated functions with many more potential inputs and a dearth of reliable historical data available for quantification.

$$Risk = f(scenarios,\ liklihood,\ consequence)$$

In traditional risk models, the probabilities or likelihoods are considered to occur as a result of objective causes that can be modeled (or estimated from operating history). Cyber risks are beyond the scope of such considerations. In comparison, cybersecurity risk can be described as a function of consequence, threat, and vulnerability. Where threat and vulnerability comprise a variety of complex factors that can be difficult to measure and quantify (see section 2.4 for more discussion of vulnerability and threat).

$$Risk = f(threat,\ vulnerability,\ consequence)$$

The adapted PRA methods described in Williams and Clark (2019) essentially capture the consequence and a subset of the scenario portions of the cyber risk equation. Once the systems, structures, and components are prioritized using a method like prevention analysis, the analyst needs to evaluate the detailed vulnerabilities and threats in order to evaluate the relative likelihood of a cyber attack leading to undesirable consequence. They must also identify risk treatments, including mitigations. Figure 1 shows a hierarchical representation of the factors that influence consequence. The factors highlighted with bold borders show what is considered in a "cyber-informed PRA" generated via a method similar to Clark's (2019).

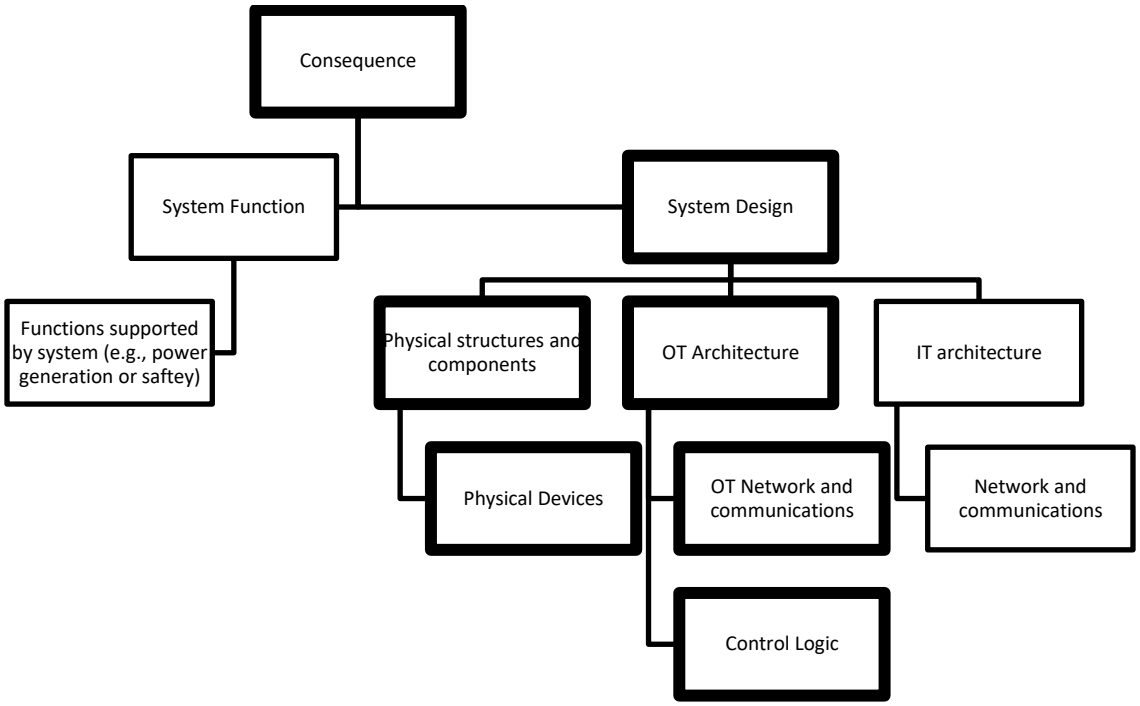

Figure 1. Hierarchical representation of the factors that contribute to consequence.

2.4 Vulnerability and Threat Analysis

Vulnerability and threat analysis are key components in cyber risk analysis. Evaluating vulnerabilities in an information system depends on the specific context of the devices, components, and the environments, in which the digital components are deployed. Exhaustive analysis of all the digital components on a system would be prohibitively expensive and difficult to perform; therefore, it is extremely important to not only prioritize the consequences, but also prioritize the systems, structures, and components leading to that consequence.

The term vulnerability has numerous definitions. The National Institute of Standards and Technology (NIST) alone has 18 different definitions of vulnerability in its standards and documents. One of the definitions often used by NIST defines vulnerability as "*Weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source* (Johnson, Dempsey, Ross, Gupta, & Bailey, 2011). Note that this NIST definition of vulnerability includes threat, which highlights that fact that threat and vulnerability are not independent constructs.

Figure 2 Shows the factors that influence information system vulnerability. Characterizing the vulnerability in the system relies on understanding the assumptions made in the software design, configuration of the system, social and organizational practices for users of the system, and a whole host of other factors that don't easily generalize to other systems. This highlights the importance of focusing on the specific context of interest on the vulnerability assessment phase of any risk analysis.

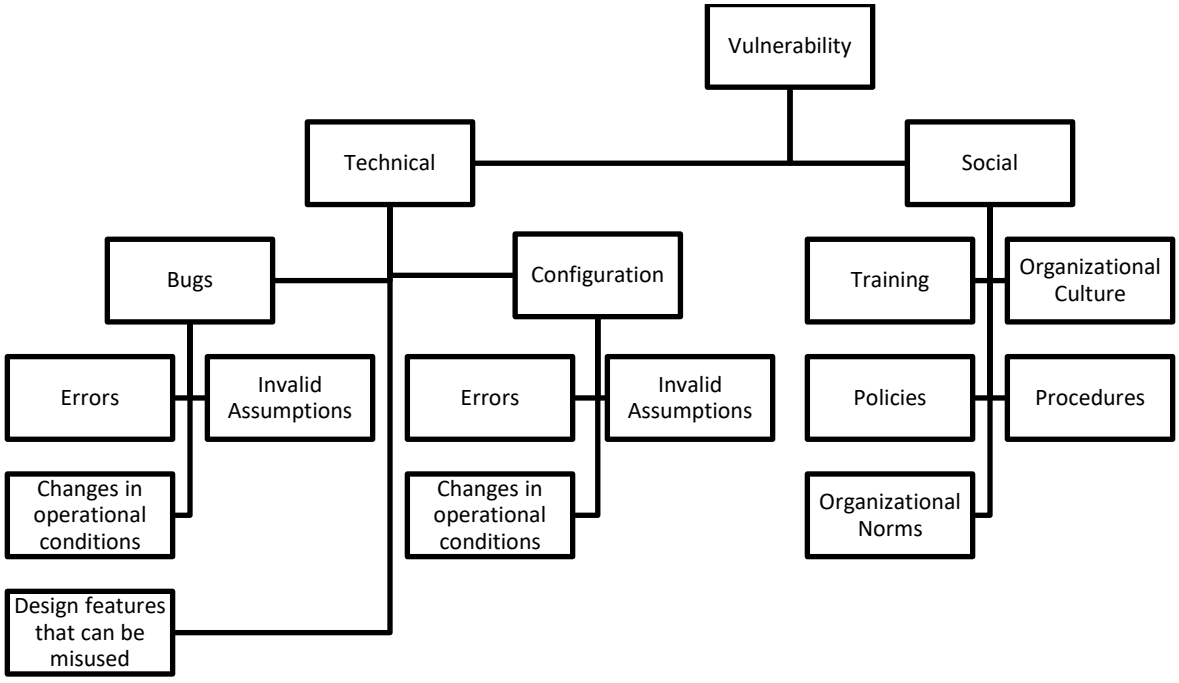

Figure 2. Hierarchical representation of the factors that influence vulnerability.

As indicated by the NIST definition of vulnerability referenced above, threat is tightly coupled with the concept of vulnerability. Understanding what adversaries are likely to target and what capabilities they have is extremely important in determining whether a particular vulnerability could be exploited, and thus how likely it is. This, in turn, is a factor in prioritizing control and mitigation for identified cybersecurity risks. Future research should address a systematic way to incorporate threat analysis into the risk management process.

3. Summary and Remaining Challenges

Figure 3. Simplified Cybersecurity Risk Management Framework.

The framework described in this report defines a systematic process for performing risk analysis that takes cybersecurity issues in digital systems into consideration, and is illustrated in Figure 3. Even if this challenging and complex risk process is perfectly followed, significant challenges in cyber risk analysis remain. The main challenge to cyber risk analysis is quantification. Quantification relies on the ability to both quantify the consequence and quantify the likelihood of those consequences. Because technology changes rapidly, and because vulnerabilities rely on specific context of the system as well as the ability of an intelligent and adaptable adversary, generating reliable probabilities of compromise is extremely challenging. Annex C describes work performed to evaluate using a fuzzy Analytical Hierarchy Process (AHP) to address quantification. The researchers on this project determined that the process as described would not significantly improve quantification; however, there may be a way to adapt it to aid in quantification. Future research should address ways to improve the process described herein by capturing the appropriate data and methodology for cyber risk quantification.

Additionally, as Figure 1 indicates, even with a good system model that is appropriately cyber-informed, the methods described in Williams and Clark (2019) do not capture all of the factors that influence scenario and consequence. Digital components can also be conduits for attacks to propagate or move through the system. For instance, a component itself may be inconsequential to a function or process, yet it may still be used to gain access or adversely impact another component. Future research should evaluate what risks are missing in the analysis and should map the analysis performed in the vulnerability analysis steps of the risk management process back to the PRA models to capture the additional scenarios and consequences.

Another remaining challenge is understanding how effectively and efficiently this process can be followed by industry as part of the standard risk management process. These methods have been developed by researchers and, to date, have only been applied to single systems. For this process to be effective, it needs to capture interactions between systems and systems-of-systems, including the entire plant. Future research needs establish how feasible this process is when applied to entire power plants, including the effectiveness of capturing system-to-system interactions.

Future research should also address how to perform effective vulnerability analysis for the information systems and digital components for the high priority systems structures and components identified in emerging adapted PRA methods like those presented in Williams and Clark (2019). The research should identify the processes and tools that can be used to evaluate how attackers could compromise the digital components and information systems that lead to unsafe control action on the system in ways that compromise system functions. Future research should also characterize how humans interact with the system because human actions are a source of vulnerability that has not been extensively studied in this

space. Finally, future research should develop threat modeling techniques that map to the vulnerabilities identified and the research should develop a framework for using threat and vulnerability as inputs to decision making in the risk analysis process.

Page intentionally left blank

4. REFERENCES

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2016). A review of cyber security risk assessment methods for SCADA systems. *Computers & Security*, 56, 1-27.

Dunjó, J., Fthenakis, V., Vílchez, J. A., & Arnaldos, J. (2010). Hazard and operability (HAZOP) analysis. A literature review. Journal of hazardous materials, 173(1-3), 19-32.

Johnson, A., Dempsey, K., Ross, R., Gupta, S., & Bailey, D. (2011). Guide for security-focused configuration management of information systems. NIST special publication, 800(128), 16-16.

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

Kumar, A., & Ramana, M. V. (2011). The limits of safety analysis: severe nuclear accident possibilities at the PFBR. Economic & Political Weekly,43(46), 44-48.

Leveson, N. (2011). Engineering a safer world: Systems thinking applied to safety. Mit Press.

Radack, S. (2011). Managing information security risk: organization, mission and information system view (No. ITL Bulletin March 2011). National Institute of Standards and Technology.

Ross, R. S. (2012). Guide for conducting risk assessments NIST special publication 800-30 revision 1. US Dept. Commerce, NIST, Gaithersburg, MD, USA, Tech. Rep.

US Nuclear Regulatory Commission. (1975). Reactor safety study: An assessment of accident risks in US commercial nuclear power plants (Vol. 88).

Williams, A. D., & Clark, A. J. (2019). Addressing Cyber Hazards in Nuclear Power Plants with Systems Theoretic-Informed Fault Trees (No. SAND2019-6969C). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).SP 800-30. Revision 1. Guide for conducting risk assessments, 2012.

Williams, A. D., & Clark, A. J. (2019). Using Systems Theoretic Perspectives for Risk-Informed Cyber Hazard Analysis in Nuclear Power Plants (No. SAND2019-2873C). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

SP 800-39. Managing Information Security Risk: Organization, Mission, and Information System View, 2011.

Page intentionally left blank

Annex A
Cyber Risk Review

Survey of Cybersecurity Risk Analysis Approaches for the Nuclear Industry

Shannon Eggers[1] and Katya Le Blanc[2]
Idaho National Laboratory, Idaho Falls, ID 83415, [1]Shannon.Eggers@inl.gov, [2]Katya.LeBlanc@inl.gov

Abstract*: While safety risk management using probabilistic risk analysis methods and software tools at nuclear power plants is well-established, cybersecurity risk analysis is still an immature field with unproven techniques. As the nuclear fleet continues to adopt digital instrumentation and control systems, the ability to more effectively and efficiently evaluate and mitigate cybersecurity risk becomes increasingly more important. The nuclear industry is currently researching methods to improve cybersecurity risk analysis. This paper evaluates the strengths and weaknesses of existing cybersecurity risk analysis methods when applied to the nuclear industry, thereby providing guidance for future research into safety- or consequence-informed cybersecurity risk analysis.*

1. INTRODUCTION

Risk is inherent in all aspects of an organization independent of industry or sector. Risk management is a standard practice by organizations to minimize the adverse effects of loss, such as financial, operational, environmental, political, organizational, and cyber. There are many systematic methods to identify and categorize risk scenarios. In fact, Paul et al. reports that there are over 200 risk management methods and guidelines throughout the world [1].

Despite the large number of risk management methods in use, adequate techniques for analyzing cybersecurity risk are unavailable for most industries. Often, organizations evaluate cybersecurity risk for their business information communication technology (ICT) environments based on a cost-benefit analysis. The financial impacts from a cyber event in ICT may be quantifiable based on lost revenue and equipment replacement costs. Less tangible financial impacts are also possible, such as damage to a company's reputation.

In contrast to ICT environments, impacts from a cyber event on an industrial control system (ICS) at a chemical, petroleum, or manufacturing facility can range from financial damage to loss of life, depending on the severity of the event. Furthermore, while a cyber attack at a chemical plant could cause hundreds of injuries or fatalities within the plant, radiological sabotage at a nuclear power plant (NPP) could affect the health and safety of thousands of individuals both inside and outside the plant.

The existing U.S. nuclear fleet is slowly upgrading plant equipment to digital technology due to aging, obsolescence, and operability concerns. In addition, advanced nuclear plants, such as generation III+ reactors, small modular reactors, and microreactors, plan to use digital instrumentation and control (I&C) systems. While the advancement and installation of digital technology improves the efficiency and reliability of NPPs, this technology introduces new risks due to cybersecurity concerns. As a result, NPPs are required to provide high assurance of protection against these cybersecurity risks. Understanding and defining cybersecurity risks is necessary for developing a risk-informed cybersecurity program. The remainder of this paper provides a survey on existing cybersecurity risk analysis methods, the current gaps in these methods, and recommendations on future research to enable adoption of risk-informed cyber practices within the nuclear industry.

2. BACKGROUND

2.1 Traditional Risk Management

Risk management is the process by which organizations identify all possible risks to assets, evaluate these risks against their risk tolerance, and respond to the risk based upon their tolerance. Risk management is a mature field that has been in existence for almost 40 years. While definitions of risk

vary, Kaplan defines risk as the "possibility of loss or injury" and the "degree of probability of such a loss" [2]. The first step, risk analysis, is traditionally defined as a process that answers three questions [3]:

(i) What can go wrong?
(ii) What is the likelihood it will go wrong?
(iii) What are the consequences if it goes wrong?

Thus, risk is the complete set of triplets including the scenario (or undesired event), likelihood (or probability of the scenario), and consequences (or impact of the scenario).

During risk evaluation, the second step of the risk management process, an organization rates their risk exposure against their risk tolerance to determine the risk significance of an event or events. Although risk evaluation includes prioritizing the risks based on likelihood and consequence, it is important to recognize the risk is not simply the product of probability and consequence but rather a function of probability and consequence:

Risk = f (scenario, likelihood, consequence)

For example, a low-probability, high-consequence event resulting in fatalities will have a much different risk significance to an organization than a high-probability, low-consequence event despite potentially having the same result when multiplying consequence ratings by probability.

The final step in risk management is risk response or risk treatment. After identifying and evaluating risks, an organization typically has four choices—risk acceptance, risk avoidance or elimination, risk transfer, or risk mitigation. Risk mitigation involves reducing the likelihood and/or severity of the consequence by implementing changes or controls in the organization or process. Risk response is often a financial decision based upon the cost of risk mitigation balanced against an organization's risk tolerance.

2.2 Traditional Safety Risk Analysis

Historically, safety concerns in the aerospace, chemical, and nuclear industries drove the development and application of risk analysis techniques. In 1975, the first accepted method for fully quantifying risk was published in WASH-1400 (NUREG 75/014), a Reactor Safety Study sponsored by the U.S. Atomic Energy Commission [4]. The probabilistic risk assessment (PRA) technique in WASH-1400 attempted to estimate public risks posed by NPPs by examining the potential paths by which nuclear fuel could melt and release radiation to the environmental. Since the initial publication of WASH-1400, PRAs in the nuclear industry evolved into a logical framework for identifying the likelihood and consequences of design basis accidents (DBA)—those postulated accidents that could lead to radiation release impacting the health and safety of the public. In an NPP, there are three levels to a PRA—level 1 evaluates the frequency of core damage, level 2 evaluates the probability of specific release of radioactive material, and level 3 evaluates the frequency of adverse public health or environmental occurrences.

PRAs are model-based graphical techniques that use plant assets and design along with historical data (i.e., vendor, plant, and industry data on equipment and events) to determine the likelihood of an event and the frequency of potential consequences. In the nuclear industry, a PRA using event tree analysis (ETA) and fault tree analysis (FTA) results in the development of 'minimum cutsets' and estimation of core damage frequency. Minimum cutsets are the sequence of events or failures that must happen for a top event to occur in a fault tree analysis (FTA) model. The top event in a nuclear PRA is typically a DBA.

2.3 Cybersecurity Risk Analysis

Quantitative safety risk analysis relies heavily on known historical data for functional failure and accident analysis. Safety PRAs address incidents with adverse consequences that are unexpected and unintentional. Nuclear safety PRAs may consider failure of an operator to perform an action, but they do not evaluate deliberate or malicious acts intended to cause damage. Moreover, safety PRAs have difficulty modeling digital systems, structures, and components (SSCs). Analog SSCs in I&C systems fail in expected ways; therefore, modeling the functional failure of an analog SSC is straightforward in a

PRA. Digital SSCs, however, fail in unexpected ways and the set of all failure modes is often unknown. It is also not always clear how failure of a digital SSC affects the required safety functions.

Contrary to the occurrence of random events in safety risk analysis, security risk includes intentional attacks by intelligent and adaptive adversaries. Due to the challenges with applying traditional PRA methodologies to these deliberate, dynamic acts against digital SSCs, significant efforts are underway to develop methodologies for identifying and evaluating cybersecurity risks. Many security risk approaches define cybersecurity risk as:

Cybersecurity Risk = Threat x Vulnerability x Consequence

where threat is a hostile action, vulnerabilities are exploitable weaknesses, and consequence is the impact of the attack. This formula fails to recognize that risk is not multiplicative but a function of attributes. Therefore, the following equation is more accurate:

Cybersecurity Risk = f (threat, vulnerability, consequence)

Additionally, as described in [5], security risk assessments are "more an art than science". Cybersecurity risk analysis is often subjective, relying on expert opinion to quantify likelihood of occurrence and expected damage. The threats and vulnerabilities also continuously change as adversaries become smarter, threat vectors change, and technology advances. Therefore, it is not only difficult to determine the current state, it is nearly impossible to predict the future state. As Oppliger succinctly writes, "we must admit that we've reached a dead end and that our nice mathematical formula for quantifying risks hardly works in practice and is therefore useless." [6]

2.3.1 What can go wrong?

With over 200 risk management methods and guidelines around the world [1], there have been many attempts by researchers to solve the problem of accurately determining cybersecurity risk. Reviews of cybersecurity risk standards and cybersecurity risk analysis methods for ICS are provided by [5, 7-12]. As with traditional risk analysis, qualitative or quantitative techniques are also used in cybersecurity risk analysis to answer Kaplan's questions. When answering the question 'what can go wrong', researchers define the digital asset(s) and the possible malicious threats against those assets to create scenarios or sets of scenarios. The probability of a threat includes adversary knowledge, motivation, intent, characteristics, and capabilities including those tactics, techniques, and procedures (TTPs) an adversary will use to compromise an asset. Threat analysis is often the domain of intelligence analysts.

2.3.2 What is the likelihood?

What is the likelihood, or probability, a threat will occur? This question can be broken down into evaluating the likelihood of an attack occurring and the likelihood of an attack succeeding. The answer to this question is often a function of the threat, asset vulnerabilities, and security controls. As stated, vulnerabilities are weaknesses or flaws in an asset that an adversary can exploit. Security controls are technical, physical, and administrative countermeasures put in place to mitigate cyber threats and vulnerabilities. For instance, an adversary may attempt to attack a control system from a plant's outward-facing internet—this attack may have a high probability of occurring based upon accessibility to the internet but a low probability of succeeding based upon security controls used to implement secure architecture, such as network segregation, firewalls, and data diodes. Likelihood of the attack occurring is also dependent on the 'attractiveness' of the compromise to an adversary. The attractiveness of a compromise, however, is different for each adversary as it depends on the adversary's motives, intent, and skill. Knowledge regarding likelihood of an attack occurring and succeeding is often the domain of plant and security specialists.

2.3.3 What are the consequences?

In cybersecurity, the answer to Kaplan's final question, "what are the consequences if an attack occurs?", is often discussed in terms of the C-I-A triad (confidentiality—integrity—availability). Loss of confidentiality is often considered the least important consequence for ICS. These reconnaissance or data gathering attacks, however, may be used to plan future, more damaging attacks. Further, loss of company

or facility data may be financially damaging or otherwise detrimental to the company. Conversely, integrity and availability attacks may result in safety-related (i.e., radiological sabotage, loss of life, injury), financial-related (i.e., lost generation, equipment damage), or reputation-related consequences. Integrity attacks impact the truthfulness of a system—data, logic, or command modification may adversely affect system operation. Availability attacks impact data and communication flow in a system. For instance, a denial of service attack may prevent control system communication, resulting in adverse system operation. Determining the consequences or impacts of a successful attack is often the domain of plant specialists.

2.4 ICS Cybersecurity Risk Assessment Standards and Guidelines

Voronca surveyed worldwide standards used for risk assessment at energy companies [13]. While this review was not specific to cybersecurity risk assessment, the report concluded that many of the standards and guidelines provide generalist approaches that do not capture the specificities of critical energy infrastructures. European risk assessment methods are further behind U.S. standards, in part due to the fragmented infrastructure and differences in security culture [12]. Knowles et al. also performed a detailed review of standards and guidelines for ICS. They concluded that guidance for managing cybersecurity risks in control-system-specific publications is both too high-level and scarce [9].

The problem of inadequate risk assessment standards in critical infrastructure is magnified when addressing cybersecurity risk analysis in ICS environments, including the nuclear industry. Several cybersecurity risk assessment standards exist for information communication technology (ICT) environments; however, the challenges ICT environments face from cyber threats are much different than those faced in control systems. An inventory of methods and tools available for network and information security risk management is provided by ENISA [14].

Standards incorporating information security risk management include NIST SP 800-30 Rev 1 [15], NIST SP 800-39 [16], and IEC 27001:2013 [17]. Standards incorporating ICS security risk management include NIST SP 800-82 Rev 2 [18] and IEC 62443-3-2 (in draft) [19]. While these standards provide high-level information on cybersecurity risk assessments, they lack the implementation details necessary to appropriately evaluate and capture risk to a facility due to a cyber threat. The chemical industry, on the other hand, provides a systematic approach for qualitative or quantitative security risk assessment in ANSI/API standard 780 [20] and their white paper titled "*Security Vulnerability Assessment Methodology*" [21].

In early 2011, the U.S. Nuclear Regulatory Commission (NRC) commissioned a task force to develop more comprehensive and holistic risk-informed, performance-based regulatory approaches to ensure the safe and secure use of nuclear material [22]. The NRC's risk-informed approach to regulatory decision making considers insights from PRAs in conjunction with other engineering insights to complement the agency's deterministic approach and defense-in-depth philosophy. The International Nuclear Safety Group at the International Atomic Energy Agency (IAEA) also developed a framework for an integrated risk-informed decision process in 2011 to provide guidance on incorporating deterministic considerations with probabilistic analyses [23].

Although the NRC is transitioning to risk-informed approaches, this transition is not complete with regard to cybersecurity regulations. In fact, there are currently no actionable cybersecurity risk management tools from the NRC or Nuclear Energy Institute (NEI) for NPP use. Currently, the U.S. nuclear industry follows Regulatory Guide (RG) 5.71 [24] and NEI 08-09 [25] for implementation of an NPP's cybersecurity plan (CSP). As part of a CSP, digital assets are classified as critical digital assets if they are associated with safety-related, important-to-safety, security, or emergency preparedness functions or if they are supporting equipment which, if compromised, adversely impacts those functions [26]. While this CDA classification and the subsequent CDA consequence assessments outlined in NEI 13-10 [27] attempt to narrow the scope of the program, it falls short of a risk-informed process. Furthermore, even though section C.13 in RG 5.71 (and corresponding section E-12 in NEI 08-09) is titled 'Evaluate and Manage Cyber Risk,' the section is primarily focused on vulnerability scanning tools and does not provide specific risk analysis guidance.

3. METHODOLOGY

3.1 Survey Format

Tools are publicly available for stepping through the cybersecurity risk management process, including the risk analysis phase. However, since none of these tools meet all criteria for establishing the level of cybersecurity risk for ICT and ICS environments [28], research and development into improved analysis methods is ongoing. In order to evaluate these techniques and their potential applicability to the nuclear industry, we studied risk analysis methods in various application domains. This survey only considers the risk analysis phase of the risk management process. Table 1 provides a list of attributes compared in the survey. The number of citations for an article was excluded from this survey as a measure of industry adoption or acceptance since this number will be under-represented for newly developed methods.

Table 1. Attributes compared in the survey.

Attribute	Description
Method	Approach used for risk analysis, including (a) model-based, graphical, (b) model-based, non-graphical, (c) formula-based, and (d) combination methods. Each method is described in the section 3.2.
Type	Technique used to calculate or determine a risk value for a plant, system, or component, including quantitative, semi-quantitative, and qualitative.
Domain	Application domain in which the risk analysis method was demonstrated, including include energy (i.e., utility, smart grid, nuclear), aerospace, railway, maritime, oil and gas, chemical, enterprise-level IS/ICT, and generic control systems.
Goal	Purpose for the risk analysis technique, including financial, security control prioritization, or other objectives.
Rigor	Level of effort needed to implement the method established as low, medium, or high. Methods that are very time and resource intensive (i.e., if modeling of every digital component in a plant is required) were rated as high rigor. Methods that incorporate an implementation tool or database to assist with the process were scored with a lower level of rigor.
Source Data	Source data is any data used as input into the analysis method, including expert opinion, threat databases, vulnerability databases, scenario databases, and established attack classification tools.
Maturity	Maturity is the readiness level for use in the nuclear industry established as low, medium, or high for a given risk analysis technique. A method that is still theoretical without practical examples for application has a low maturity level while a method that is currently in use has a high maturity level. Methods that are not easily adaptable for use in the nuclear industry are rated as low maturity.
Example	Examples included in the reference papers.
Gaps	Identified gaps or weaknesses that potentially impact the method's adoption in the nuclear industry.
Starting Basis	The starting point for the analysis, including asset, impact, threat, or vulnerability. (Note: The starting basis for the risk analysis methods are excluded in Table 2.)

3.2 Method Attribute Description

Method defines the approach used for risk analysis, including (a) model-based, graphical, (b) model-based, non-graphical, (c) formula-based, and (d) combination methods. As shown in Figure 4, risk

analysis methods can be classified as model-based or formula-based. If model-based, the models can be further subclassified into those that use graphical or non-graphical tools. Model-based, graphical methods use visual techniques, such as fault tree analysis (FTA), event tree analysis (ETA), attack tree analysis, vulnerability tree analysis, and system-theoretic process analysis (STPA) to represent systems. Graphical models are logic techniques that systematically describe pathways within a system to identify and categorize deviations.

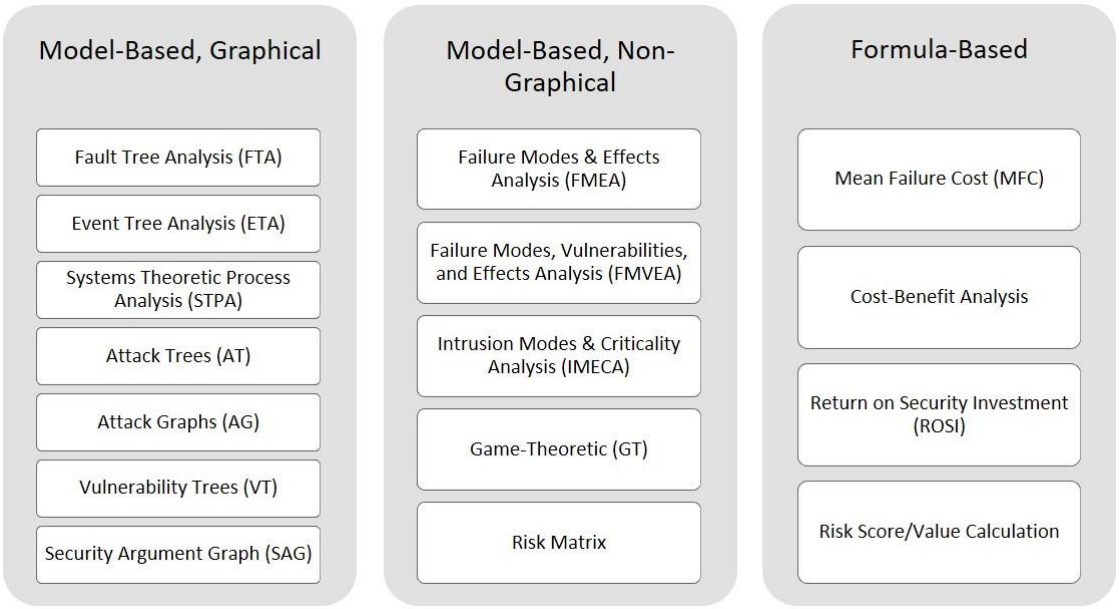

Figure 4. Categories of risk analysis methods.

Model-based, non-graphical risk analysis methods may use game-theoretic frameworks or matrix-based tools, such as failure modes and effects analysis (FMEA), failure modes, vulnerabilities, and effects analysis (FMVEA), and intrusion modes and effects criticality analysis (IMECA), to identify risk. Formula-based risk analysis techniques use a calculation to quantify risk, such as mean failure cost or return on security investment. Combination methods use more than one of the approaches to analyze risk.

4. RESULTS

A summary of attribute findings for the surveyed risk analysis methods are listed in Table 2.

Table 2. Review of cybersecurity risk analysis methods.

Method[1]	Type	Domain(s)	Goal	Rigor	Source Data	Maturity	Example?	Gaps for Nuclear	Ref
Formula-Based									
MFC	QT or SQ	Smart Grid, Utility	Cost-benefit analysis	H	Experts	M	X	No safety focus	[29-31]
ROSI	QT	ICT	Cost-benefit analysis	H	Equipment costs, Experts	M	X	No safety focus	[32]
Risk Score	SQ	SCADA, Maritime, Petroleum, Chemical	Identify need for controls	M/H	Experts, CVSS/CVE	L/M	X	No relationship between systems; Subjective data	[33-38]
Model-Based, Non-graphical									
Risk Matrix	SQ	Railway, Chemical, ICT	Prioritize controls	M/H	Experts, Historical threat data, IEC 62443-3-2 rankings	M/H		Requires analysis of every asset	[15, 20, 39, 40]
GT	SQ	Utility, Smart Grid	Identify optimal strategy	H	ENISA threat landscape, MARGERIT, CVE, SECCRIT, Experts	L	X	Cost focus (no safety focus)	[41, 42]
Risk Matrix-GT	SQ	Chemical	Identify optimal strategy	UNK	API SRA results as input	M/H		Lacks detail	[43]
Model-Based, Graphical									
SAG	SQ	Smart Grid	Calculate probability of scenario success	H	NESCOR failure scenarios	M	X	Uses predefined scenarios	[44]
PN	SQ	Chemical	Identify need for controls	H	Experts	L	X	Focused on overall facility security	[45]
Model-Based, Graphical Combinations									
Attack Tree (AT) Combinations									
AT-VT	SQ	SCADA, DCS	Prioritize controls	H	Experts	L		Threat index is cost focused not safety focused; no final risk determination	[46, 47]
AT-CD	QL	Air Traffic Control	Identify need for controls	H	EBIOS database	L	X	Incomplete set of threats for NPP	[1]
AT-AHP	SQ	ICT	Prioritized linear cost function	H	CAPEC, CVE	L	X	Cost focus (no safety focus)	[48]
FMEA/FMEVA Combinations									
FMEA-AT	QL	Railway	Identify need for controls	M	CAPEC, CWE	L		Lacks detail	[49]
FMEA-HHM AT	SQ	SCADA	Identify need for controls	H	Experts	L	X	Requires analysis of every asset	[50]
FMEVA-STPA	SQ	Machinery	Cyber-inform safety analysis	H	STRIDE	L	X	Requires analysis of every asset	[51]
FMEVA-CHASSIS	QL	Automotive	Cyber-inform safety analysis	M	Experts	L	X	Lacks detail Not ongoing	[52]

Method[1]	Type	Domain(s)	Goal	Rigor	Source Data	Maturity	Example?	Gaps for Nuclear	Ref
FTA Combinations									
FTA-ETA-AT	SQ	ICS, Chemical	Cyber-inform safety analysis	H	Experts	M	X	Requires analysis of every asset	[53, 54]
FTA-BDMP	QT	ICS	Cyber-inform safety analysis	H	Experts	L	X	Mean times to failure or success are arbitrarily chosen	[55]
FTA-STPA	QL	NPP	Cyber-inform safety analysis	H	Experts	M		Intensive to describe all STPA control actions in NPP; Impact to current PRA	[56]
FTA-STPA-AG	SQ	NPP	Cyber-inform PRA	H	CVE, CCE, NVD, CVSS, Experts	M	X	Intensive to describe all STPA control actions in NPP; Impact to current PRA	[57]
Other combinations									
IMECA & SVM	SQ	NPP	Prioritize controls	M	CVE, NVD, Experts	L	X	Only focuses on known vulnerabilities	[58]
BN & ETA	SQ	NPP	Cyber-inform PRA	H	Experts (for cyber)	L	X	Limited model of attack vectors and controls	[59]

[1] Acronyms are defined in the following section.

5. DISCUSSION

5.1 Attribute Comparison

5.1.1 Method

Model-Based, Graphical Risk Analysis Methods

Graphical risk models are very effective on smaller scales. For large systems such as industrial control systems, there are often cognitive scalability issues. The systems are too complex to render and comprehend within a graphical model.

FTA is a graphical model developed in the 1960's that represents, in symbolic logic model, the cause and effect relationships between combinations of events leading to an identified top undesired event. FTA contains only those activities that contribute to the top event and may be created using quantitative or qualitative techniques. Quantitative FTA, as used in PRA, identifies event probability at each step – the probability is propagated up to the top event to calculate an overall probability of occurrence. The sequence of events (or group of initiators) that, if all occur and cause a top event, is termed a cut set. Although a very resource-intensive process, FTA is useful for identifying single points of failure as well as vulnerabilities and potential mitigations.

Attack trees (AT) are a variation of FTA in which an attack is the top event instead of an overall system fault or DBA. In an attack tree, analysts predict the path adversaries will follow based on known TTPs and evaluate scenario likelihood instead of failure rate or probability. Attack trees are useful for identifying weaknesses, however, they are difficult to use on large systems or plants because of their complexity. Attack graphs are similar to attack trees but use a different visual format to indicate entry points, exit points, nodes and attack pathways.

Like FTA, vulnerability trees (VT) are top-down approaches that decompose the relationship between a top vulnerability and the sequence of vulnerabilities an adversary must exploit in order to reach the top. Vulnerability trees help inform attack scenarios an adversary may follow in order to exploit an SSC. Like attack trees, however, vulnerability trees are complex and difficult to use on large systems.

While FTA is a top-down approach, ETA is a bottom-up approach. ETA is also a symbolic logic model which, starting with an initiating event, identifies the sequence of propagating events leading to a final undesired event or loss. ETA may use qualitative or quantitative techniques, has a clear order from beginning to end, and can account for mitigations. ETA, however, is complex, resource intensive, and requires a new tree for each initiating event.

STPA models systems into hierarchical control structures which are then used to identify unsafe control actions for which mitigation measures can be used [60]. STPA first identifies top level accidents or hazards to avoid then identifies the control actions leading to the top event. STPA-Sec modifies STPA to incorporate both safety and security. While STPA moves beyond identifying system failures to find complex causal chains of events in control structures, analyzing all interactions between controllers and system level components, including human interactions, is very resource intensive in a complex environment.

Security Argument Graph (SAG) is tool developed for the smart grid using failure scenarios defined by the U.S. National Electric Sector Organization Resource (NESCOR). The SAG tool provides a graphical representation that connects mal-activity processes with system components and threat agents to evaluate the probability of a failure scenario occurring. Currently the tool only applies to the NESCOR scenarios [44]. In addition, a semi-quantitative method using weighted fuzzy petri-nets (PN) was developed by Zhou, et al. to evaluate an overall risk value for a facility [45]. While this PN method may provide valuable information for overall facility risk, it does not provide enough detail for use in ICS.

Model-Based, Non-Graphical Risk Analysis Methods

Risk analysis methods integrating game theory model intelligent interactions between adversaries and defenders [41, 43]. Certain game theory (GT) techniques model the adversary's strategy to cause as much damage as possible in order to evaluate security control implementations to optimize cybersecurity

spending. While GT may improve protections from cyber events, the models require quantitative data based on inexact assumptions [43]. In addition, full games for large systems are thought to be too complex [61].

FMEA is a systematic, non-sequential bottom-up method that identifies known or potential failures, problems, or errors based on historical or inferential data at the component level. In cybersecurity risk analysis, FMVEA is often used since it incorporates a systematic review of component level vulnerabilities and how these vulnerabilities are susceptible or can be targeted during a cyber event. FMVEA is a resource-intensive process if performed on every digital SSC in a plant.

IMECA is a further modification of FMEA that examines the effects of intrusions during system operation [58]. IMECA is a bottom-up approach that identifies the vulnerabilities for each component and its criticality to system operation. IMECA is also a resource-intensive process.

Other non-graphical model techniques use traditional heat-map risk matrices to determine a risk score or value based upon parameters such as likelihood and impact. The parameters may be calculated via a formula or applied using a ranking (i.e., High, Medium, Low). The two parameters are then combined on a matrix to identify the resulting risk score. Often, the risk matrix technique is used to determine a risk score with and without security controls to evaluate and prioritize control implementations [15, 20, 39].

Formula-Based Risk Analysis Methods

Formula-based techniques are typically present in non-safety environments such as enterprise ICT. Econometric methods calculate mean failure cost, cost-benefit of risk mitigation, or return on security investment using quantitative data on asset values and security control implementation costs [29-32]. These calculations are often used to prioritize control implementations to align with an organization's risk tolerance.

Other formula-based methods calculate risk values using semi-quantitative and quantitative data. Various formulas have been studied to 'quantify' risk to determine both relative risk and security control prioritization. The challenge with these formulas is that the source data is often ranked and based on expert opinion. The results, therefore, are often subjective.

Combination Risk Analysis Methods

With combination methods, techniques are combined to present a more complete representation of the risk from cyber events. Bow-tie methods in safety PRA combine top-down FTA with bottom-up ETA. Researchers, such as in [53] and [54], combine bow-tie analysis with attack tree or threat analysis to cyber-inform a safety PRA. Still others integrate safety and cybersecurity by combining FTA with Boolean logic Driven Markov Processes (BDMP) [55], STPA [56], and attack graphs (AG) [57]. Many researchers also combine FMEA/FMVEA with attack trees [49], Hierarchical Holographic Modeling (HHM) attack trees [50], STPA [51], or Combined Harm Assessment of Safety and Security for Information Systems (CHASSIS) [52].

Attack trees have been combined with vulnerability trees in [46] and [47] to evaluate threat-impact (TI) and cyber-vulnerability (CV) indices in SCADA applications. While a risk value is not determined, the TI and CV values are evaluated with and without controls applied to determine if the impact from a cyber-event is reduced. Attack trees were also combined with chain diagrams (CD) to improve cognitive scalability of large systems, such as found in air traffic control systems. Argyropoulos et al. recently evaluated combining attack trees in the Secure Tropos security-by-design model with likelihood metrics determined by the Analytic Hierarchy Process (AHP) used in software engineering [48]. This AT-AHP approach expresses the level of threat mitigation as a linear cost function for cost-benefit analysis in applying security controls.

ETA was combined with Bayesian Networks (BN) by Shin et al. to numerically evaluate a cyber PRA for an NPP reactor protection system (RPS) [59]. The method relies on experts to determine and rank threats and mitigations as input into a BN model which informs a cybersecurity risk index (CSRI) calculation. The CSRI value is used as an input into the ETA of a safety PRA to cyber-inform the PRA.

IMECA and Support Vector Machine (SVM) tools were combined in [58]. SVM was used to develop a vulnerability classifier tool derived from common vulnerabilities and exposures (CVE) and national vulnerability database (NVD) data to define vulnerability probabilities and severities. This data

was used in an IMECA model to calculate system risk based on probability and damage related to the vulnerabilities in each SW and HW component within the system.

5.1.2 Type

Risk analyses either use quantitative, semi-quantitative, or qualitative techniques to derive the level of risk for a plant, system, or component. Although many authors describe their method as quantitative, fully quantitative methods for determining cybersecurity risk are potentially flawed. In safety risk analysis, historical data on equipment failure, adverse events, and environmental factors are available to calculate probabilities and uncertainties in quantitative methods such as PRA. Even if complete histories on cyber attacks or compromises were available, the continuous changes associated with threat vectors, adversarial skills, and technological advances would make the use of this data irrelevant.

Many of the self-described quantitative methods are, in fact, semi-quantitative methods that use numerical values based on expert opinion. For instance, an author may describe a technique that applies numerical rankings to qualitative values or ranges (i.e., 1, 2, 3 for High, Medium, Low impact) as quantitative. Since no true numerical probabilities or uncertainties are applied to these qualitative observations, methods that use these devices are more properly classified as semi-quantitative. Our literature review found that there were no fully quantitative methods since expert opinion is required for at least a portion of the analysis and expert opinion is highly subjective and often unrepeatable. However, methods that use econometric data to calculate costs were considered quantitative in this paper.

5.1.3 Application Domain

Traditionally, formula-based, econometric risk analysis methods are used to prioritize security control implementations in ICT environments. These methods provide IT managers the knowledge required to determine financially optimized cybersecurity risk treatments. In contrast, the focus in OT environments is to ensure continuous, safe operation. Therefore, model-based or combination risk analysis approaches are typically used in OT environments to prioritize security controls that will be most effective in eliminating or reducing the impact of a cyber-event on safety and production.

5.1.4 Goal

While the end goal of risk analysis is to identify and inform a company on the risks they face, the aim of the techniques evaluated for cybersecurity risk analysis vary on how this risk is presented. In some cases, the goal is a cost-benefit analysis that calculates financial values, such as mean failure cost [29-31], return on security investment (ROSI), or prioritized linear cost functions [48], in order to decide where best to allocate cybersecurity mitigation funds [32]. Other techniques are focused on cyber-informing safety analyses [51-57, 59], identifying components, functions, or pathways requiring security controls based on risk level [1, 33-35, 37, 38, 45, 49, 50, 58], or prioritizing implementation of security controls [20, 39, 40, 46, 47]. Still other decision-support methodologies use game theory to identify optimal attack strategies and the resultant optimal defense strategies [41-43] or are concerned with how risk changes during an attack [62].

5.1.5 Rigor

A challenge with some of the methods reviewed is the extensive time and resources that are required to model every digital component in a plant. Methods that have high levels of rigor may not be readily adopted in the nuclear industry because they require too much time and expertise to complete. Since the nuclear industry is currently focused on methods that streamline processes to improve NPP efficiencies to enhance economic competitiveness, expensive analysis may prove counterproductive. None of the methods reviewed in this survey were determined to be low rigor. Techniques with high rigor may lead to more robust risk determination, however, the extensive resources required to use these methods may make them challenging for the nuclear industry to accept and implement.

5.1.6 Source Data

An ideal ICS cybersecurity risk analysis is a repeatable process which results in the same risk determination regardless of who performs the analysis. Unfortunately, much of the analysis with cybersecurity risk relies upon expert opinion—expert opinion on threats and adversaries, asset vulnerabilities, and impacts or consequences. While there are well-established approaches for expert elicitation, these processes are time intensive and may not result in the same outcome if repeated.

To provide insight into potentially useful data sources that may improve repeatability, we documented the sources, if any, for each method. These data sources varied based upon the type of risk analysis performed. Vulnerability data and scores based upon common vulnerabilities and exposures (CVE), common weakness enumeration (CWE), common vulnerability scoring system (CVSS), and the national vulnerability database (NVD) are used by some researchers [41, 48, 49, 57, 58, 62-64].

Threat databases from ENISA, MAGERIT, and SECCRIT are integrated into the tool developed by Gouglidis et al. [41]. Failure scenarios from NESCOR are used in [29] and [44]. Jillepalli et al. use threat categories from NIST SP 800-82 [31], Hutle et al. use threat level tables from the HMG IS1 UK standard (now withdrawn) [40], and Paul and Vignon-Davillier use the threat database from EBIOS [1]. The API standard 780 and Ralston et al. suggest using threat data based on facility, national, and global histories [20, 46].

Aside from threat and vulnerability databases, the common attack pattern enumeration and classification (CAPEC) tool from Mitre is used by [48] and [49] to identify common attack patterns and applicable countermeasures. Researchers use costs associated with current equipment and countermeasures for cost-based analyses. System designs and facility drawings are commonly used for asset identification and pathway evaluations.

5.1.7 Maturity

The majority of methods reviewed have a low maturity level. Low maturity level indicates that the method may include a risk analysis framework or theoretical discussion but lacks sufficient steps or methodology for immediate use. Methods with high maturity level, such as the ANSI/API STANDARD 780, include detailed steps for risk analysis [20].

5.1.8 Gaps

In theory, cybersecurity risk management frameworks and risk analysis methods seem like they would be straightforward to implement. Additionally, it may seem appropriate to simply add “cyber” considerations to mature methods like PRA. In practice, however, cybersecurity risk analysis is very difficult, especially in ICS environments. And, as indicated by the amount of ongoing research in this field, there is not yet an industry-accepted risk analysis methodology. Cybersecurity risk analysis is difficult because human adversaries are intelligent, unpredictable, persistent, and adaptable. It is impossible to map all possible attack scenarios that might lead to core damage at an NPP. It is also impossible to remove or fully mitigate all cybersecurity risk unless a facility is built completely without digital components. Thus, the ideal cybersecurity risk analysis provides a repeatable method to identify when risks exceed an NPP’s risk tolerance such that mitigation strategies or security controls can be implemented to reduce the risk to an acceptable level. Since most of the techniques surveyed in this paper rely on expert opinions or analysis, repeatability is often challenging.

Table 2 identifies specific gaps for each method. Overall, this survey found that there is not yet a tool that provides repeatable, actionable risk analysis methods with the appropriate level of rigor for use in an NPP. While there is ongoing research to integrate safety and security into a cyber-informed PRA, this approach may be undesirable to the industry as it potentially adds unnecessary complexity to existing PRAs. In addition, techniques that rely on analyzing the pathway or control logic for every digital asset in an NPP may also be untenable as NPPs have thousands of digital assets. The resources required to perform such cybersecurity risk analyses are potentially prohibitive.

5.1.9 Starting Basis

Techniques differ based upon the starting point for the analysis. The analysis may start with the asset, impact, threat, or vulnerability. Asset-initiated techniques start with an asset inventory and apply risk analysis to the assets. Many financially driven, quantitative risk analysis techniques, such as those calculating mean failure cost, start with the asset. Impact-initiated techniques consider the consequences and losses associated first, and then develop risk analyses based upon those impacts. Game-theoretic techniques, combined security and safety techniques, and top event analysis techniques (i.e., FTA) generally start with impact analysis. Threat-initiated techniques start with threat identification followed by risk analysis based upon those threats. Methods starting with threat determination often evaluate the adversaries' TTPs and determine the likelihood of attack upon a given component or attack tree. Finally, vulnerability-initiated techniques start with identifying the vulnerabilities of systems and components to evaluate their associated security risk and determine how best to mitigate those vulnerabilities. Risk analysis methods that combine techniques, such as bow-tie analysis, often have more than one starting point. While this attribute was omitted from Table 2, the starting basis and availability of data is key to successful risk analysis.

5.2 Use of Methods in Nuclear Facilities

The power reactor cybersecurity program in the U.S. is largely programmatic and compliance-based without the inclusion of risk analysis techniques to risk-inform the processes. In addition, NPP licensees commented in open forums during the 2019 NRC Power Reactor Cyber Security Program Assessment that the program is overly conservative and does not appropriately focus on protecting NPPs against radiological sabotage as prescribed in 10 CFR 73.54 [65]. Specifically, this cybersecurity rule requires "high assurance that digital computer and communication systems and networks are adequately protected against cyber attacks, up to and including the design basis threat (DBT) [26]." Radiological sabotage is a key tenet of the DBT as described in 10 CFR 73.1 [66]. 10 CFR 73.2 defines radiological sabotage as "any deliberate act directed against a plant… which could directly or indirectly endanger the public health and safety by exposure to radiation." [67]

The NRC and licensees have expressed concerns that the number of digital assets scoped in as CDAs is far larger than originally expected at the start of the program. Although the security control assessment guidance in NEI 13-10 reduced the number of controls evaluated for some digital assets by introducing direct and indirect CDAs, the programmatic requirements to address the controls is still cumbersome, expensive, and often may not provide the desired cyber-protections against radiological sabotage. Risk-informing CSP processes (i.e. CDA classification, security control assessments) would incorporate safety significance or relative risk along with other engineering insights to deliver a holistic, rational, and cost-effective approach to provide high assurance of protection against cyber attacks that could result in radiological sabotage.

Risk-informing processes or programs in the nuclear industry is often defined as using an NPPs PRA in combination with deterministic evaluations (i.e., engineering analysis, expert judgement, experience) to guide decision making. While PRA is one technique for identifying risk, it is better suited to safety analysis; the quantitative requirements of PRA do not effectively pivot to cybersecurity risk analysis. Several of the risk analysis approaches surveyed combine a safety PRA with cybersecurity risk analysis or cyber-inform a safety analysis. In theory, this is a logical progression towards developing a holistic risk analysis. In practice, however, incorporating CDAs or their control logic, connections, and/or pathways into an existing plant PRA is problematic. For instance, incorporating cyber aspects into a safety risk analysis greatly increases the scope of a PRA, which increases rather than decreases the burden of the existing cybersecurity programs. Furthermore, PRAs are quantitative techniques that use historical equipment data and known events to evaluate probabilities of unexpected, unintentional safety incidents—this type of data is largely absent for cybersecurity events. In addition, digital SSCs not only fail in unexpected ways, but modeling deliberate, intentional attacks is challenging in a PRA. Since the U.S. nuclear industry has indicated a desire to streamline CSP processes to improve efficiencies while

maintaining or improving cyber-protection against radiological sabotage, adding greater complexity via cyber-informed PRAs or safety risk analyses may be in direct opposition to this improvement pathway.

Cost-based cybersecurity risk analysis that ignore safety or ICS concerns are also inappropriate solutions for the nuclear industry. These risk analysis techniques provide useful cost-benefit analysis information for ICT environments; however, pure financial analyses that ignore system interactions and production or safety impacts are unsuitable for ICS environments. In addition, pure quantitative cybersecurity risk analysis is unattainable as threats and vulnerabilities are constantly changing and there is no hard data available to quantify current or future scenario likelihoods.

Returning to the traditional set of triplets in cybersecurity risk analysis—threat, vulnerability, and consequence—an NPPs Final Safety Analysis Report (FSAR) can be used to identify the design basis accidents (DBA) that must be protected against to minimize radiological sabotage. Therefore, safety-informing cybersecurity risk analyses may offer more insight into protecting a nuclear facility against a DBT than cyber-informing safety risk analyses. Safety-informed cybersecurity risk analysis may also potentially be simpler to implement and maintain than a cyber-informed PRA.

While protecting a nuclear facility against a DBT is regulated by the cybersecurity rule, owners also want to protect their plants from non-DBT cyber-events that might cause economic losses from plant shutdown, equipment damage, or intangible effects, such as reputation. The risk evaluation and risk treatment for those digital assets that cannot impact radiological release, or the health and safety of the public could fall outside of the NRC regulatory guidance and, therefore, be subject primarily to the facility's decision-making process. Then again, a facility may still be subject to regulatory guidance under the North American Electric Reliability Corporation (NERC) Critical Infrastructure Protection (CIP) regulation. If an NPP generates less than 1500 MWe, however, it is designated as low impact and has limited NERC-CIP requirements, such as cybersecurity awareness, physical access control, electronic access control, and incident response. Consequence-informing cybersecurity risk analysis, with inclusion of both radiological exposure, safety impacts and non-radiological, plant impacts, may enable the industry to 'right size' their cybersecurity program by using regulatory guidance to apply the highest security controls to those CDAs impacting safety and more business-driven, cost-effective processes to prioritize security control implementation to the remaining digital assets.

Although determining likelihood is challenging, cybersecurity risk analysis techniques that develop qualitative or semi-quantitative risk scores to identify those SSCs that must be protected to prevent a DBA may have the most promise for nuclear industry adoption. The ability to evaluate an SSC's cybersecurity risk before and after security control implementation is also a highly important feature for the tool. Several techniques reviewed expand upon the traditional cybersecurity risk analysis set of triplets to include more detailed analysis of threats and vulnerabilities. The American Petroleum Institute's Security Risk Assessment (API SRA) methodology defines the likelihood of attack based upon attractiveness of an asset to a given threat and likelihood of the success based upon the vulnerability and attack attempt [20]. Extensions of the API SRA method are presented in [43, 45] and [37].

Tam and Jones developed a visual risk value using system vulnerability, ease of exploit, and attacker rewarder as axes [34]. Vulnerability is a function of attack vector, asset vulnerability, and consequence. Ease of exploit is a function of attacker profile, asset type, attacker resources, and implemented security controls. Attacker reward is a function of attacker profile, asset type, attacker's goal, and consequence. Although this system was designed as a visual identification for maritime cybersecurity risk, the technique could potentially be adopted within the nuclear industry to develop a risk score.

Risk values are also calculated in [33] and [62]. The technique developed by Kure et al. derives semi-qualitative scores for asset criticality, vulnerability impact, and likelihood but it uses an arbitrary weighting factor for each asset that would potentially be challenging to define for all digital assets in an NPP [33]. Wu et al. calculates risk as a function of attack severity, attack success probability, and attack consequence where attack severity is a function of frequency, intensity, and stealth of attack; success probability is a function of a vulnerability's ease of exploit, number of authentication times required, and exploitation level location; and consequence is a function of economic loss, casualties, environmental damage, and repair cost.

These semi-quantitative risk calculations show promise for determining safety-informed cybersecurity risk in the nuclear domain especially if an NPP's DBA is used to inform the consequence, impact, or severity values. The DBA could also be used to inform the asset type, asset criticality, or asset attractiveness values. Further research is necessary to determine if these techniques, or variations of these techniques incorporating DBA-informed data, provide a useful methodology for risk-informing the CDA determination and security control assessment processes in an NPP's CSP.

6. CONCLUSIONS AND FUTURE WORK

Cybersecurity risk assessment remains a challenge in all industries. The inherent unknowns associated with current and future cyber threats, vulnerabilities, and adversaries prohibit the use of meaningful quantitative risk assessments. Thus, we are forced to evaluate cybersecurity risk qualitatively or semi-quantitatively. While it is important to use meaningful and repeatable analyses to ensure accurate and maintainable risk-informed security control implementation decisions, these cybersecurity risk analysis solutions do not yet exist.

This paper reviews existing cybersecurity risk analysis techniques and evaluates their strengths and weakness for use in the nuclear industry. Methodologies that increase the complexity of the cybersecurity program without providing an increase in protection level (i.e., cyber-informed safety risk analysis) will be challenging to implement in the nuclear industry. Conversely, cost-based approaches do not include the requisite level of safety focus for industry adoption. Safety-informed or consequence-informed cybersecurity risk analysis methods that qualitatively or semi-qualitatively determine a risk value show promise for use in the nuclear industry. Future research will be performed to determine if these methods, or variations of these methods as informed by an NPP's DBA and FSAR, will benefit the industry to drive cyber-informed decision making to minimize radiological sabotage.

7. ACKNOWLEDGEMENTS

This research was funded by the U.S. Department of Energy Office of Nuclear Energy under DOE Idaho Operations Office, Contract DE-AC07-05ID14517.

8. REFERENCES

[1] Paul, S. and R. Vignon-Davillier, "Unifying traditional risk assessment approaches with attack trees," *Journal of Information Security and Applications,* vol. 19, no. 3, pp. 165-181, 2014/07/01/ 2014.

[2] Kaplan, S., "The words of risk analysis," *Risk analysis,* vol. 17, no. 4, pp. 407-417, 1997.

[3] Kaplan, S. and B.J. Garrick, "On the quantitative definition of risk," *Risk analysis,* vol. 1, no. 1, pp. 11-27, 1981.

[4] United States Nuclear Regulatory Commission, W.D.C., "Reactor safety study An assessment of accident risks in US commercial nuclear power plants Executive summary," International Atomic Energy Agency (IAEA)1975, Available: http://inis.iaea.org/search/search.aspx?orig_q=RN:35053391.

[5] Tweneboah-Koduah, S. and W.J. Buchanan, "Security risk assessment of critical infrastructure systems: a comparative study," *The Computer Journal,* vol. 61, no. 9, pp. 1389-1406, 2018.

[6] Oppliger, R., "Quantitative Risk Analysis in Information Security Management: A Modern Fairy Tale," *IEEE Security & Privacy,* vol. 13, no. 6, pp. 18-21, 2015.

[7] Cherdantseva, Y. *et al.*, "A review of cyber security risk assessment methods for SCADA systems," *Computers & Security,* vol. 56, pp. 1-27, 2016.

[8] Chockalingam, S., D. Hadžiosmanović, W. Pieters, A. Teixeira, and P. van Gelder, "Integrated safety and security risk assessment methods: a survey of key characteristics and applications," in

International Conference on Critical Information Infrastructures Security, 2016, pp. 50-62: Springer.

[9] Knowles, W., D. Prince, D. Hutchison, J.F.P. Disso, and K. Jones, "A survey of cyber security management in industrial control systems," *International Journal of Critical Infrastructure Protection,* vol. 9, pp. 52-80, 2015.

[10] Voronca, S. and S. Voronca, "Survey of existing risk assessment and management standards applied worldwide, for power companies," in *6th International Conference on Modern Power Systems*, Cluj-Napoca, Romania, 2015, pp. 369-373.

[11] Kriaa, S., L. Pietre-Cambacedes, M. Bouissou, and Y. Halgand, "A survey of approaches combining safety and security for industrial control systems," *Reliability engineering & system safety,* vol. 139, pp. 156-178, 2015.

[12] Giannopoulos, G., R. Filippini, and M. Schimmer, "Risk assessment methodologies for Critical Infrastructure Protection. Part I: A state of the art," *JRC Technical Notes,* 2012.

[13] Voronca, S.L., "Analising some of the existing risk assessment and management standards applied worldwide, for energy companies," *Journal of Sustainable Energy,* vol. III, no. 2, pp. 77-84, 2012.

[14] (ENISA), E.U.A.f.N.a.I.S. (May 29, 2019). *Inventory of Risk Management / Risk Assessment Methods and Tools*. Available: https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory

[15] *SP 800-30. Revision 1. Guide for conducting risk assessments*, 2012.

[16] *SP 800-39. Managing Information Security Risk: Organization, Mission, and Information System View*, 2011.

[17] *IEC 27001:2013, Information technology - security techniques - information security management systems - requirements*, 2013.

[18] Stouffer, K., V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "SP 800-82. Revision 2. Guide to industrial control systems (ICS) security," *National Institute of Standards and Technology,* 2015.

[19] *IEC 62443-3-2, Security Risk Assessment and System Design, Draft.* , 2017.

[20] *ANSI/API STD 780 Security Risk Assessment Methodology for the Petroleum and Petrochemical Industries*, 2013.

[21] Institute, A.P. and N.P.R. Association, "Security vulnerability assessment methodology for the petroleum and petrochemical industries," 2003: API.

[22] NRC. (2019, July 9, 2019). *Risk-Informed Activities*. Available: https://www.nrc.gov/about-nrc/regulatory/risk-informed/rpp.html

[23] Group, I.N.S., *A Framework for an Integrated Risk Informed Decision Making Process*. Vienna: International Atomic Energy Agency, 2011.

[24] "Cyber Security Programs for Nuclear Facilities," U.S. Nuclear Regulatory Commission, January 2010.

[25] "NEI 08-09, Cyber Security Plan for Nuclear Power Reactors, Rev 6," Nuclear Energy Institute, April 2010.

[26] *10 CFR 73.54, Protection of Digital Computer and Communication Systems and Networks,* U.S. Nuclear Regulatory Commission, 2009.

[27] "NEI 13-10 Cyber Security Control Assessments, Rev 5," Nuclear Energy Institute, February 2017.

[28] Baybutt, P., "Issues for security risk assessment in the process industries," *Journal of Loss Prevention in the Process Industries,* vol. 49, pp. 509-518, 2017.

[29] Abercrombie, R.K., F.T. Sheldon, K.R. Hauser, M.W. Lantz, and A. Mili, "Risk Assessment Methodology Based on the NISTIR 7628 Guidelines," in *2013 46th Hawaii International Conference on System Sciences*, Hawaii, 2013: IEEE, 2013.

[30] Chen, Q., R.K. Abercrombie, and F.T. Sheldon, "Risk assessment for industrial control systems quantifying availability using mean failure cost (MFC)," *Journal of Artificial Intelligence and Soft Computing Research,* vol. 5, no. 3, pp. 205-220, 2015.
[31] Jillepalli, A.A., F.T. Sheldon, D.C. de Leon, M. Haney, and R.K. Abercrombie, "Security management of cyber physical control systems using NIST SP 800-82r2," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2017, pp. 1864-1870.
[32] Bojanc, R. and B. Jerman-Blažič, "An economic modelling approach to information security risk management," *International Journal of Information Management,* vol. 28, no. 5, pp. 413-422, 2008.
[33] Kure, H., S. Islam, and M. Razzaque, "An Integrated Cyber Security Risk Management Approach for a Cyber-Physical System," *Applied Sciences,* vol. 8, no. 6, p. 898, 2018.
[34] Tam, K. and K. Jones, "MaCRA: A Model-Based Framework for Maritime Cyber-Risk Assessment," 2019.
[35] Papa, S.M., W.D. Casper, and S. Nair, "Availability based risk analysis for SCADA embedded computer systems," in *Proceedings of the International Conference on Security and Management (SAM)*, 2011, p. 1: The Steering Committee of The World Congress in Computer Science, Computer ….
[36] Moore, D.A., "Security risk assessment methodology for the petroleum and petrochemical industries," *Journal of Loss Prevention in the Process Industries,* vol. 26, no. 6, pp. 1685-1689, 2013.
[37] Landucci, G., F. Argenti, V. Cozzani, and G. Reniers, "Assessment of attack likelihood to support security risk assessment studies for chemical facilities," *Process Safety and Environmental Protection,* vol. 110, pp. 102-114, 2017.
[38] Caralli, R.A., J.F. Stevens, L.R. Young, and W.R. Wilson, "Introducing octave allegro: Improving the information security risk assessment process," Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst, 2007.
[39] Braband, J., "Towards an IT Security Risk Assessment Framework for Railway Automation," *arXiv preprint arXiv:1704.01175,* 2017.
[40] Hutle, M., G. Hansch, and W. Fitzgerald, "D2. 2 Threat and Risk Assessment Methodology," *Tunneling and Underground Space Technology,* vol. 24, no. 3, pp. 269-277, 2015.
[41] Gouglidis, A., J. Busby, D. Hutchison, S.N. Shirazi, S. König, and A.Z. Galbis, "Deliverable 2.3. Software tools for hybrid risk management in SCADA networks," in "Hybrid risk management for utility networks," 2017.
[42] Schauer, S., S. König, M. Latzenhofer, and S. Rass, "Identifying and managing risks in interconnected utility networks," presented at the SECUREWARE 2017: The Eleventh International Conference on Emerging Security Information, Systems and Technologies, Rome, Italy, 2017.
[43] Zhang, L., G. Reniers, B. Chen, and X. Qiu, "Integrating the API SRA methodology and game theory for improving chemical plant protection," *Journal of Loss Prevention in the Process Industries,* vol. 51, pp. 8-16, 2018.
[44] Jauhar, S. *et al.*, "Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios," pp. 319-324, 2015.
[45] Zhou, J., G. Reniers, and L. Zhang, "A weighted fuzzy Petri-net based approach for security risk assessment in the chemical industry," *Chemical Engineering Science,* vol. 174, pp. 136-145, 2017.
[46] Ralston, P.A.S., J.H. Graham, and J.L. Hieb, "Cyber security risk assessment for SCADA and DCS networks," *ISA Transactions,* vol. 46, no. 4, pp. 583-594, 2007.

[47] Patel, S.C., J.H. Graham, and P.A. Ralston, "Quantitatively assessing the vulnerability of critical information systems: A new method for evaluating security enhancements," *International Journal of Information Management,* vol. 28, no. 6, pp. 483-491, 2008.
[48] Argyropoulos, N., K. Angelopoulos, H. Mouratidis, and A. Fish, "Risk-aware decision support with constrained goal models," *Information & Computer Security,* vol. 26, no. 4, pp. 472-490, 2018.
[49] Birr, P., M. Hetzer, and S. Petretti, "IT security risk analysis and threat mitigation for railway applications," presented at the Fast abstracts at International Conference on Computer Safety, Reliability, and Security (SAFECOMP), Trondheim, Norway, 2016, 2016. Available: https://hal.laas.fr/hal-01370249
[50] Henry, M.H. and Y.Y. Haimes, "A comprehensive network security risk model for process control networks," *Risk Analysis: An International Journal,* vol. 29, no. 2, pp. 223-248, 2009.
[51] Kivelä, T., M. Golder, and K. Furmans, "Towards an approach for assuring machinery safety in the IIoT-age," *Logistics Journal: Proceedings,* vol. 2018, no. 01, 2018.
[52] Schmittner, C., Z. Ma, E. Schoitsch, and T. Gruber, "A case study of FMVEA and CHASSIS as safety and security co-analysis method for automotive cyber-physical systems," in *Proceedings of the 1st ACM Workshop on Cyber-Physical System Security*, 2015, pp. 69-80: ACM.
[53] Abdo, H., M. Kaouk, J.M. Flaus, and F. Masse, "A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie – combining new version of attack tree with bowtie analysis," *Computers & Security,* vol. 72, pp. 175-195, 2018.
[54] Mohr, R., "Evaluating Cyber Risk in Engineering Environments: A Proposed Framework and Methodology," in "Information Security Reading Room," SANS Institute, 2016, Available: https://www.sans.org/reading-room/whitepapers/ICS/evaluating-cyber-risk-engineering-environments-proposed-framework-methodology-37017.
[55] Piètre-Cambacédès, L. and M. Bouissou, "Modeling safety and security interdependencies with BDMP (Boolean logic Driven Markov Processes)," in *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2010, pp. 2852-2861: IEEE.
[56] Clark, A.J., A.D. Williams, A. Muna, and M. Gibson, "Hazard and Consequence Analysis for Digital Systems–A New Approach to Risk Analysis in the Digital Era for Nuclear Power Plants," 2018.
[57] (EPRI), E.P.R.I., "Program on Technology Innovation: Cyber Hazards Analysis Risk Methodology Phase II: A Risk Informed Approach," 2015.
[58] Zelinko, I., V. Kharchenko, and K. Leontiev, "Cyber Security Assessment of Component Off-the-Shelf Based NPP I&C System Using IMECA Technique," in *2017 25th International Conference on Nuclear Engineering*, 2017, pp. V009T15A034-V009T15A034: American Society of Mechanical Engineers.
[59] Shin, J., H. Son, and G. Heo, "Cyber Security Risk Evaluation of a Nuclear I&C Using BN and ET," *Nuclear Engineering and Technology,* vol. 49, no. 3, pp. 517-524, 2017.
[60] Leveson, N., *Engineering a safer world: Systems thinking applied to safety*. MIT press, 2011.
[61] Fielder, A., E. Panaousis, P. Malacaria, C. Hankin, and F. Smeraldi, "Decision support approaches for cyber security investment," *Decision Support Systems,* vol. 86, pp. 13-23, 2016.
[62] Wu, W., R. Kang, and Z. Li, "Risk assessment method for cyber security of cyber physical systems," in *2015 First International Conference on Reliability Systems Engineering (ICRSE)*, 2015, pp. 1-5: IEEE.
[63] McCrory, F.M., "Cyber Informed Risk Analysis (CIRA) for Nuclear Power Cyber Security," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2015.
[64] Wheeler, T.A., "Cyber Informed Risk Analysis (CIRA) for Nuclear Power Cyber Security," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2015.
[65] "Power Reactor Cyber Security Program Assessment," U.S. Nuclear Regulatory Commission, ML19175A211, July 12, 2019.

[66] *10 CFR 73.1, Purpose and Scope, 10 CFR 73 Physical Protection of Plants and Materials,* U.S. Nuclear Regulatory Commission, 2008.
[67] *10 CFR 73.2, Definitions (for 10 CFR 73 Physical Protection of Plants and Materials)*, 2017.

Page intentionally left blank

Annex B
Application of Top Event Prevention Analysis (TEPA) to Assessment and Mitigation of Cyber Vulnerabilities

Application of Top Event Prevention Analysis (TEPA) to Assessment and Mitigation of Cyber Vulnerabilities

Robert Youngblood

1. Top Event Prevention Analysis (TEPA)

1.1 What TEPA does: what question it answers

Starting from a fault-tree / event-tree model of plant safety systems, TEPA answers the following question: What systems, structures, and components (SSCs) do we need to "protect (invest to prevent the failure of)" in order to meet a given safety criterion? [Youngblood and Oliveira 1989; Youngblood and Worrell, 1995; Worrell and Blanchard 1995] By "protecting" an item, we do not mean "guarantee absolutely that the subject item will never fail;" we mean "apply engineering resources sufficient to achieve good (not necessarily perfect) item reliability (or, for the digital asset (DA) problem, resistance to attack)." How TEPA fits into reactor safety analysis is summarized below under "Background." An important issue addressed in reactor safety analysis is precisely that of deciding what SSCs to invest in, and from a certain point of view, that question resembles the question of deciding what DAs need protection. Correspondingly, at least some aspects of the DA problem are arguably addressable using TEPA, given good models of plant systems.

TEPA starts with results from a risk model, and generates insights from those results. It does not by itself do much to compensate for weaknesses in the model, but turns out to be a very interesting way to examine the model.

The calculations done by TEPA are arduous, but the results have simple meanings, *and can be checked independently of TEPA software.*

1.2 Background

For present purposes, we oversimplify early reactor safety analysis as follows:

> We need to show by analysis that the plant can cope with a carefully formulated set of challenges to plant safety: loss of coolant accidents, loss of offsite power, etc., with boundary conditions and analysis assumptions carefully specified. The acceptance criterion for the plant is "protection of the core, assuming the initiating event (the "challenge"), plus failure of the most limiting active component in the mitigating systems, plus all the other analytical conservatisms required by regulatory guidance." The requirement that protection be able to succeed despite a single failure is called the "single-failure criterion." If we believed that we could achieve perfect reliability in plant systems, we would not apply this criterion; but we don't believe that, and we do need redundancy, diversity, margin, etc.
>
> If you can show, for every challenge on your list, that the plant systems satisfy the single-failure criterion and protect the core, then the plant design is deemed to provide adequate protection. In general, in order to satisfy the single-failure criterion, the analysis needs to "take credit for" two or more success paths, so that when one path is assumed to be failed by the single failure, the equipment still operating is sufficient to satisfy the core protection requirement. The phrase

"success path" has a detailed technical meaning in fault-tree / event-tree modeling, but operationally, a success path is a complement of equipment (and perhaps operator actions) whose successful operation yields "success." A *minimal* success path is a complement of equipment that is sufficient to succeed, but is no longer sufficient if any one element is removed from that complement.

Within the above construct, if sufficient failures occur, none of the success paths operate, and the systems fail to protect the core. A combination of failure events or conditions that fails all the success paths is a "cut set," and a "minimal cut set" is a cut set that is minimal in the sense that if one or more elements are removed from the set, it is no longer sufficient to fail all of the success paths.

Back to the question: What systems, structures, and components do we need to "protect (invest in order to prevent the failure of)" in order to meet a given safety criterion?

To borrow a term from TEPA, a correct answer to the above question – a list of components whose protection satisfies a given criterion – is a "prevention set." For the classical reactor safety analysis described above, a prevention set is required to include enough equipment that failure of any single element leaves at least one success path still functional. For simple systems, prevention sets satisfying the single-failure criterion can sometimes be derived from plant walkdowns and inspection of system diagrams. If a system includes precisely the equipment needed to satisfy the single-failure criterion, and no more, then the prevention set is simply "everything there is." But for complex and highly redundant systems, and combinations of systems that share dependencies on support systems (electrical power, component cooling, operator actions, etc.), manual identification of prevention sets is not reliable; and if the safety requirement is more ambitious than the single-failure criterion, then unless the systems are exceptionally simple, we absolutely need something better than manual inspection, because the selection problem (choosing a collection of things that work together) is too complicated to be done manually. Even without going beyond the single-failure criterion, given the broad set of safety challenges that we are nowadays concerned with, something better is needed.

TEPA is "something better."

1.3 Mechanics of TEPA

1.3.1 Inputs to TEPA:

The minimal cut sets from a fault-tree / event-tree model, and

a prevention criterion.

Examples of prevention criteria are:

- Prevent at least two elements of every cut set.
- Prevent at least three elements of every cut set.
- Prevent at least N elements of every cut set.

- Prevent enough elements of every minimal cut set to drive the joint failure probability of prevented events below a specified probability threshold (assuming that the events that are not prevented all fail).

In extant versions of TEPA, admissible prevention criteria operate at the cut set level, for a reason that will be made clear below. For this reason, TEPA is not a true global optimizer, but its prevention sets nevertheless perform very well at the system level.

1.3.2 Processing in TEPA:

In essence, the processing is as follows:

1. For each cut set, a logic expression is written, describing all of the ways of satisfying the given prevention criterion for that cut set. For example, Level 2 prevention calls for protecting two things in every cut set. By this criterion, protection for cut set A*B*C is achieved by preventing A*B OR A*C OR B*C. If we are working with a probability threshold, calling for prevention of enough things to drive cut set probability below that threshold, then we need to identify the combinations of components whose failure probabilities combine to satisfy that criterion.
2. Since we must prevent all of the cut sets in order to prevent the top event, we form the logical "AND" of all the cut-set-prevention expressions, expand this typically enormous Boolean expression, and reduce it. Every term in the resulting expression is a minimal prevention set: a complement of things that satisfies the prevention criterion (or criteria) for every cut set.

1.3.3 Output of TEPA:

Every term in the output expression derived in "processing" is a minimal prevention set: a prevention scheme that satisfies the criterion (or criteria) from which it was derived. Notionally, one picks the "best" one, and implements the implied prevention measures.

It turns out that for simple TEPA based on literal count, the prevention sets are unions of complete success paths. This leads to an intuitively appealing narrative, which is not only useful in explaining why TEPA works, but may also be useful in explaining why some prevention sets are better than others, even at the same nominal level of prevention.

It should be apparent from the above discussion that imposing cut-set-level criteria, as TEPA does, is not the same thing as globally optimizing system performance. But it turns out that most prevention sets perform well globally.

1.4 Assessing "Performance" of Prevention Sets for Comparison Purposes

Even for a given prevention criterion, not all prevention sets perform equally well (they do not all provide the same reliability at the system level). In order to choose a good prevention set for implementation, we need a practical way to compare them. In order to assess the performance of a given prevention set, we can re-run the logic model with all events NOT in that prevention set assumed to be failed. This is a drastic assumption, but a meaningful case to consider; after all, if we implement that prevention set, we are not protecting against those events that are not in the prevention set.

We can also estimate the cost of the implied prevention measures. A simple estimate for screening purposes is the following: impute a cost to prevention of each event in the model, and estimate the cost of a prevention set as the sum of the costs imputed to its elements. Before a final decision is made about which prevention set to implement, one would want to improve on this approximation, but it is arguably useful as a preliminary screening.

For every prevention set, then, there is a reliability metric and an associated cost, so we can draw a scatter plot of prevention set performance versus cost. Moreover, we can include, on the same plot. prevention sets from a spectrum of prevention criteria, in order to see what performance levels turn up. We are likely to see a noninferior surface emerging from this plot, which is arguably of interest to decision-makers.

A plot showing the noninferior surface is, in principle, very interesting, and a worthwhile goal for the present development. Let us briefly describe the effort associated with the calculations.

- Just getting the prevention sets can be an arduous process in real problems. That said, Blanchard has been able to do very large problems, after a lot of experimentation and some software development. [Blanchard 2009, among many, many others] Our present capability suffices to do at least "meaningful" problems. (Blanchard has a software advantage that we are not positioned to overcome.)
- Re-running the logic model for each prevention set is straightforward in principle, and computationally feasible, but setting up the inputs manually for a large number of prevention sets is completely impractical for real (not toy) problems; the process really needs to be automated. In many of the problems done by Blanchard, the client is interested in minimizing the number of components in a given category that need to be protected, such as motor-operated valves (MOVs). For purposes of illustration, Blanchard may simply pick the prevention set having the fewest MOVs, and show that its performance is pretty good, by executing the re-run of the logic model for just that prevention set. This typically shows that its reliability performance is satisfactory, even if all components not in the prevention set are assumed to be failed.

Blanchard also has software providing an interesting visual perspective on prevention sets, useful for helping clients intelligently choose from among the options. This capability will not be discussed further here, but it is a very interesting and worthwhile capability. Within the present effort, independent development of such a capability is not being contemplated.

Some of the extensions accomplished or underway are aimed at improving this process, especially with a view to addressing the DA problem.

2. TEPA Application to Deciding What Subset of Digital Assets (DAs) to Protect

2.1 Special Things about the DA Problem

We do not know the probability of hack-induced failure of a DA. We do not know the probability of an attempted hack, and whatever we mean by "protection" of a DA is completely notional. In essence, the present strategy is to decide right now what DAs to focus on, and decide later on the specifics of protection when we know what we are focusing on, recognizing that if we decide that it is impractical to protect certain DAs in a given prevention set, we will go back and choose a different

prevention set that does not contain those DAs; or, if no such prevention set exists, we will have concluded that the present design is vulnerable.

All that said: how can we assess the efficacy of a candidate prevention scheme for the DA problem, where we don't really know the probabilities?

2.2 Application of TEPA within HAZCADS

An approach to assessment of a prevention set (including DAs) based on sensitivity analysis has been illustrated by Blanchard in an EPRI report on HAZCADS. [Gibson 1989] It is inappropriate to quote here in detail from that proprietary EPRI report, but it is fair to recapitulate certain key ideas, which have been around for a long time.

Observation:

The assumption is made that any particular DA affects a particular identifiable set of physical components (pumps, valves, …). Functionally, a successful hack of a given DA will simply alter the behavior of its associated components. Notionally, then, we can model hacks by adding the implied component behaviors to the affected component models in our existing fault trees. If failure of a particular valve to open can be caused by a hack, then in addition to "Valve 10 Fails to Open for Mechanical Reasons," the fault tree will also contain "Valve 10 Fails to Open Because of Hack of DA X." So if we can do TEPA on the original model, we can notionally[†] do it for the model augmented by consideration of hacks. We can find prevention sets that tell us what sets of {hardware to protect from failure and DAs to protect from hacks}.

If we already know what *hardware* we are going to end up protecting, we can condition our TEPA on that, and streamline the analysis considerably. In fact, notionally, if we already know what success paths we are protecting, we could begin by considering the merits of just protecting the DAs associated with those paths. We may find that it is sufficient to protect a subset of those DAs.

However, since we don't know the probability of hack, or the probability of hack success, the evaluation of prevention set efficacy is more subtle than it seemed to be for previous TEPA applications where we only needed component reliability information.

One way to think about the merits of a particular prevention set is to conduct the following sensitivity study. Assume that its DAs can be protected to a level where the probability of a successful hack of a DA is reduced to a level comparable to, or less than, that of the probability of failure of the associated components *without* hack. One can additionally assume that component failures induced by hack of unprotected DAs occur with unit probability. Based on these sets of assumptions, we can quantify the top event, and compare that result with the result of the original model without consideration of any DAs. Since we have added failure modes to the original model, and assumed that some of them actually occur with unit probability, the new answer will be worse than the original answer; but *if we are protecting enough DAs*, it may not be much worse. A "good" result from such a study can result if enough success paths are being protected that even if some components are taken down, enough of the system survives to get by.

The Blachard application in the HAZCADS report is based on some of the above ideas.

Caveat: the above assumptions sound reasonable, but let us consider an example of a situation that could violate them. Suppose part of the system of interest is located in a fire zone where the sprinkler system can be activated by a hack, and suppose further that the sprinklers may damage electronics in the system of interest. Is the sprinkler system part of the logic model of the system of interest? If so,

[†] We say "notionally" because adding failure modes to an already-large problem may make it significantly more difficult.

then we will naturally address its hackability when we address the hackability of all the system's components. If not, we may need to broaden the scope of our system's model. Anything co-located with our system, but not part of our system, may have the potential to influence our system; that sort of modeling issue is beyond the scope of the present discussion, but may turn out to be very important, just as fire scenarios, flood scenarios, tornados, hurricanes, and seismically-initiated scenarios have turned out to be important.

2.3 TEPA-Related Capabilities that Would Be Useful in Dealing Convincingly with the DA Problem

As implied above, TEPA can provide MANY solutions (prevention sets), and a good way is needed to identify a preferred prevention set. A general approach was outlined above, but the mechanics of doing it have not yet been completely implemented. Several analogous processes can be considered, as described below.

Call this the "Classic Approach:"

- Given: a collection of prevention sets
- For each prevention set:
- Quantify the cost of the prevention set.
  - For example, impute costs to each element, and compute the cost of the prevention set as the sum of those imputed costs
- Quantify the top event metric.
  - Create a value block (a probability table) in which all basic events in the model are set to the level implied by the prevention set, and create a SETS user program that re-runs the model with unprotected events set to "fail."
- Given the cost and the performance figure of merit for each prevention set, choose one that performs well and is affordable.

A potentially more interesting approach would be the following. Call it the "Multi-State Approach." [Youngblood 2020]

> Instead of considering events that are either prevented or not, generate prevention sets that allow for different levels of protection of a given element. Instead of either preventing event A, or not, consider driving A's probability down to .1, or .01, or .001, or .05, or … . A prevention set will then specify each component's level of protection. Some prevention sets will rely on protecting a few components to extremely high levels (at extremely high cost); some will take credit for protecting everything a little bit; some will achieve a balance in between. Clearly, there will be many more prevention sets, corresponding to variations on how prevention resources are allocated among elements of each prevention set. This approach was executed some years ago for seismic hardening; the same could be done for human error, and it may prove to be interesting to consider different levels of hardening for DAs.

Yet another potentially worthwhile approach is the following. Call it the "Success-Path-Based" approach.

Determine the success paths in the original model.

For each prevention set,

1. Determine what success paths it comprises.
2. Derive the cut sets implied by those success paths. Quantify them. Note that this step has tacitly set all unprotected events to "fail."

Now examine these cut sets, manually or using software. By construction, all of the events in those cut sets are "prevented;" but are there cut sets whose particular elements have significant potential for common-cause failure (CCF) due to hacking? If so, we have to revisit the tacit independence assumption in the assignment of hacking probabilities in our model. The usual practice is to model CCF explicitly in the original logic model, by applying widely known parametric models to groups of similar components, but even if we did that, a cut set containing a conjunction of hack-induced events cries out to be scrutinized more closely for CCF potential, even if the component types are diverse and would not traditionally be modeled as vulnerable to a common cause. Such a conjunction appearing in multiple cut sets should be examined very closely.

The machinations described above do not, by themselves, add value relative to the classical approach, which also needs the CCF scrutiny. The real potential value of this approach lies in the improved understanding of why a given prevention set is as good as it is, or why it is no better than it is, and in the potential to improve upon the minimal prevention sets by supplementing them with additional success paths.

The emphasis on success paths also implicitly shows how to begin to implement the TEPA thought process if we know how our system works (we know our success paths) but don't have a logic model for it.

If prevention set performance is good (reliability is high and cost is low), and the CCF potential is minimal or has been addressed somehow, then the prevention set is a good candidate for implementation.

It would be interesting to try to combine the multi-state idea with the success-path-based idea.

3. Status and Path Forward

3.1 Discussion

In order to clarify the discussion of details of the current status, the original process of TEPA is illustrated below in slightly more operational detail.

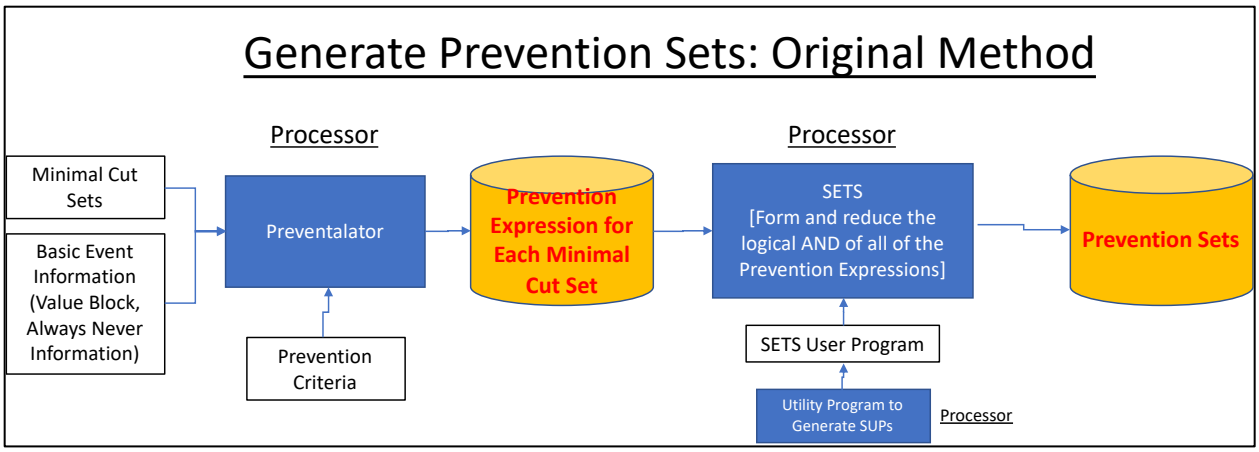

Figure 5 : Original Method for Generating Prevention Sets

The process starts on the left. Starting with the minimal cut sets, basic event information like component type and failure probability, and prevention criteria (such as the single-failure criterion), the "Preventalator" generates a logical prevention expression for each cut set. In order to get system-level prevention sets, these must be ANDed together, and the resulting expression must be reduced. When Prevention Analysis started out (Youngblood 1989), the only practical way to do this was to use SETS [Worrell 1985], a flexible general-purpose tool for manipulating logic expressions. SETS was developed by Richard Worrell at Sandia, beginning in the 1960's; it was among the first logic-model tools to be developed, and for many years, SETS was regarded as the gold standard in that category. Today, other tools claim to be faster-running or capable of running larger risk models, but that claim is based in part on their having been optimized for the characteristics of problems of interest to their market. SETS is a more general-purpose tool. It has to be told what to do through a "SETS user program," which some would call a command script for SETS.‡ For meaningful prevention analysis problem sizes, that SETS user program has to be generated by computer.

Given all of this, one obtains Prevention Sets on the right of Figure 5. For future reference, note the number and kind of interfaces between processes, and between the user and the processes.

An integrated TEPA capability has been developed by Blanchard and Worrell, and has been applied commercially in numerous studies. Since it is a commercial tool, the notional illustration below is simply based on the present author's understanding:

‡ This is both good news and bad news; it means that SETS is usable only by people who know what they are doing, but it also means that such people have the flexibility needed to solve unusual problems that would crash logic programs that tried to implement cookbook solution processes. Nowadays, this is less of a problem than it used to be.

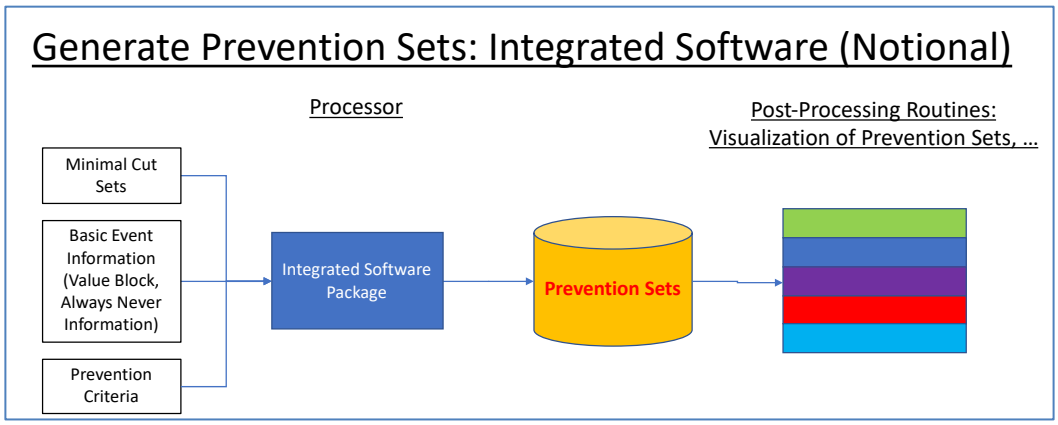

Figure 6. Cartoon of Blanchard / Worrell Commercial Software

This tool is also based on SETS, but not the standalone version used in the present work; the integrated tool comprises some of the steps called out in the "original" Figure *6* (associated with the Preventalator and the SETS user program), and on the right adds some post-processing capability. Using the "Original" capability requires enough competence with SETS to generate the SETS user program; in the integrated commercial tool, the SETS user program is generated automatically, and never sees the light of day. In that regard, the commercial tool is far more turnkey.

Following are key points to be made regarding the comparison:

There is a fair amount of user involvement in making the original version work, and for large problems, a fair amount of extra effort. In part, this reflects the work needed to interface distinct processors (SETS and the Preventalator).

For reasons that are only implicit in the diagrams, the commercial tool actually has a significant advantage in execution of problems of typical size; the SETS script needed to execute the problem is generated automatically under the hood, a few commands at a time. This has advantages whose details are beyond the scope of the present discussion.

However, a sophisticated user lacking access to the source code of the commercial tool may find it much easier to conduct methodological experiments with the original version, which was itself a methodological experiment.

Figure 7 below illustrates an example of a methodological experiment based on the "success-path-based" approach.

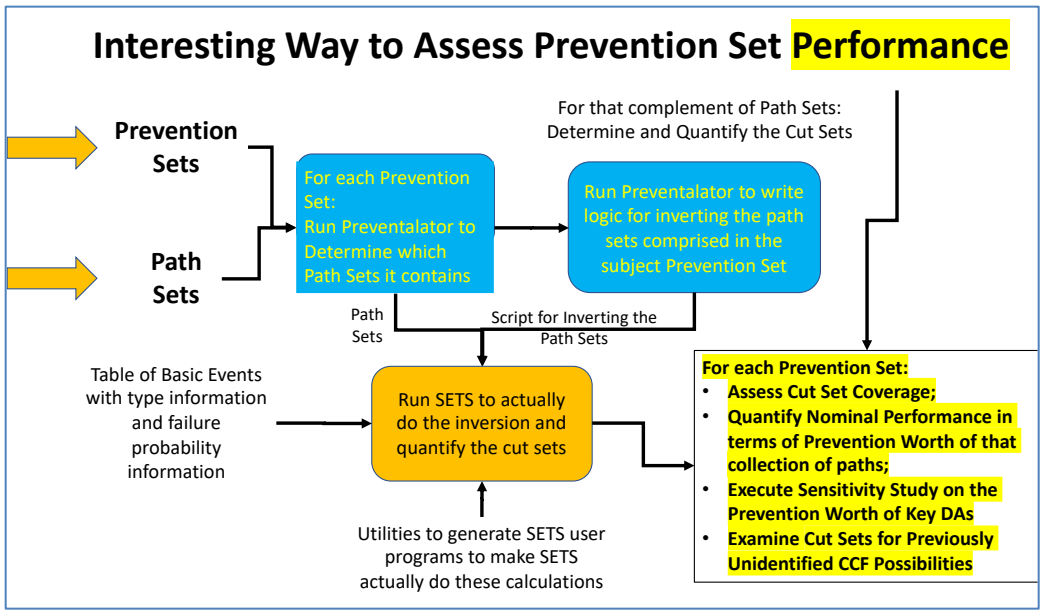

Figure 7: A Process for Assessing the Performance of Prevention Sets for Purposes of Comparison

This process starts at the left, with the prevention sets and path sets already generated. For each prevention set, we determine which path sets it contains, determine and quantify the cut sets implied by that specific collection of paths, assess the coverage of those cut sets by the current prevention set, and apply various sensitivity studies (including importance measures) for assessing the balance and robustness of the prevention set.

The possible interest in this experiment derives in part from the belief that it is useful to assess prevention sets based on an explicit understanding of what success paths each prevention set contains, and the degree to which of those prevention sets overlap each other. We can come at this from the other direction by re-reducing the cut sets with all events that are NOT in the prevention set assumed to occur; this allows us to quantify the top event metric associated with that prevention set, but may furnish less insight into what paths are included, and why the prevention set is as good or as bad as it is. This is a research topic.

Is it useful to evaluate the "Prevention Worth" importance measure [Youngblood 2001] for DAs in given Prevention Sets? Arguably yes. This would tell us which DAs are critical to the most valuable success paths. Details of that discussion are beyond the scope of this brief summary, but it is noted that the Prevention Worth calculations are a trivial add-on to the "success-path-based" approach.

3.2 Status

Table 3 provides some indication of what capabilities actually exist currently, and which are still works in progress.

Prevention set generation exists in "beta," meaning that it works for the test problems that have been run, but only a few have been run. If we are getting our inputs from other modelers, those inputs may be formatted in odd ways.

The processes associated with assessment of prevention sets are partially developed but are still being debugged ("gamma" or "delta").

Note:

Beta: The implementation exists, and generally works, but is designated "beta" because we have not yet run cases with enough different problems to be sure that we have adequately covered variations in problem characteristics.

Gamma: So far, the scripts (the SETS user programs) are highly problem-specific, so generating them can be 90% automatic but not 100% automatic. You have to know what you are doing.

Delta: This seems to work, but is at the debugging stage. Moreover, some variations on the test problems don't always yield prevention sets that are in fact unions of complete paths, so we are still ironing out both the process and the narrative. (Prevention sets derived from simple event-count prevention analysis DO seem to be unions of success paths, but we are still analyzing cases where the prevention sets were developed based on cut set probabilities, and interpreting the multi-state case from a success-path point of view remains to be done.)

Table 3. Current Status of Processes and Capabilities Mentioned Above

Function	Status	Comments
Generate Cut Set Prevention Expressions for count-based or probability-based or multi-state-probability-based cases	Beta	The 1988 version did count-based, applying different prevention reasoning to different basic event types. The 1990 version did probabilities, after a fashion. The 1998 version did multi-state prevention analysis. The current version incorporates all of these things in a much more transparent way.
Generate analogous logic needed to obtain Path Sets	Beta	In the current version, this is done by default. In the 1988 version, it would have required a separate step.
Write path set inversion equations	Beta	This is just a special case of Cut Set Prevention equations.
Script Generation for Interfacing with SETS	Easy enough, but Gamma	This is problem-specific. Current method is to look at problem characteristics and try to optimize the scripts accordingly.
Carry out simple coverage assessment for a given prevention set	Beta	Compare every cut set to the prevention set to determine how many cut set elements are "covered" and how many are not
Determine, for each prevention set, which path sets it contains	Delta	Being debugged

3.2.1 Open Issues

Which variations on classical prevention analysis yield prevention sets that are, in fact, unions of path sets?

Note: For cases where prevention sets are unions of path sets, it is useful to view the prevention set concept as a generalization of the path set concept. But we have shown that when we force the prevention sets to include certain events, because we know *a priori* that we are going to protect them regardless of the current analysis, the prevention set structure is not a simple union of path sets. What does this observation tell us about whether we ought to be doing that?

Does the path set interpretation still work in the multi-state case?

Should we be using Prevention Worth? Prevention Worth was mentioned earlier. The concept was illustrated many years ago; it works in success space, and can very naturally provide perspective on the roles of particular DAs in particular Prevention Sets.

Can we establish a clear correspondence between old-fashioned Seismic Margins analysis [Budnitz *et al.* 1985] and TEPA? If so, does this help us with the DA problem? Does it help us to explain TEPA to a broader audience?

3.3 Path Forward

Observations:

- This is not a software development project. It is a methodology research project. When we have a firm handle on what's useful and what's feasible, we can consider generating a software requirements document.
- TEPA has a long history, but we are still determining what is feasible and what is useful for the DA problem specifically.
- Test problems from different sources have posed different challenges to the Preventalator. It is believed that the current version of the Preventalator is fairly robust, but more test problems from different sources are needed for testing the robustness of the implementation.
- Most of the calculations mentioned here have been implemented. A few have not, but they seem straightforward, and will either fail early for presently unforeseen reasons, or be successfully completed with little difficulty.

We have a meaningful test problem that includes basic event information (probabilities) but does not address DAs, and we have another meaningful test problem that addresses DAs but does not contain meaningful basic event information. We need a meaningful test problem that both considers DA performance *and* provides meaningful basic event information. Then, In addition to challenging the computational robustness of the processes, we can carry out exercises like the following.

- Assuming that the model contains operator actions, we can experiment with a multi-state version of DA behavior and operator behavior, and establish the usefulness (or otherwise) of the multi-state approach. (The expectation here is that it will be very useful).
- We need to establish the value (or otherwise) of the "Prevention Worth" idea for this class of problems.
- For fundamental reasons, we need to establish the conditions under which the success-path narrative is straightforward, and when it is not, whether that matters.
- We need to establish this point for the multi-state case in particular.

Page intentionally left blank

Annex C

Using Fuzzy AHP to Evaluate Cyber Risk

Using Fuzzy AHP to Evaluate Cyber Risk – Progress Report 3/18/2020

Analytical Hierarchy Process (AHP) is a multi-criteria decision-making process that uses fuzzy arithmetic operations. AHP typically has three steps: (1) structuring the problem into a hierarchy of a goal, criteria, and sub-criteria; (2) establishing pairwise comparisons between elements at each level; and (3) establishing weighted priority. Adding fuzzy-based techniques to AHP translates linguistic variables (i.e., high, moderate, low) to fuzzy numbers to allow for uncertainties in the decision-making process.

The following steps were used to develop cyber risk fuzzy AHP for a nuclear plant.

1. Identifying the goal, criteria, and sub-criteria. The following figure is a work-in-progress. The analysis described in the following steps uses only adversarial sub-criteria and eliminates a couple others.

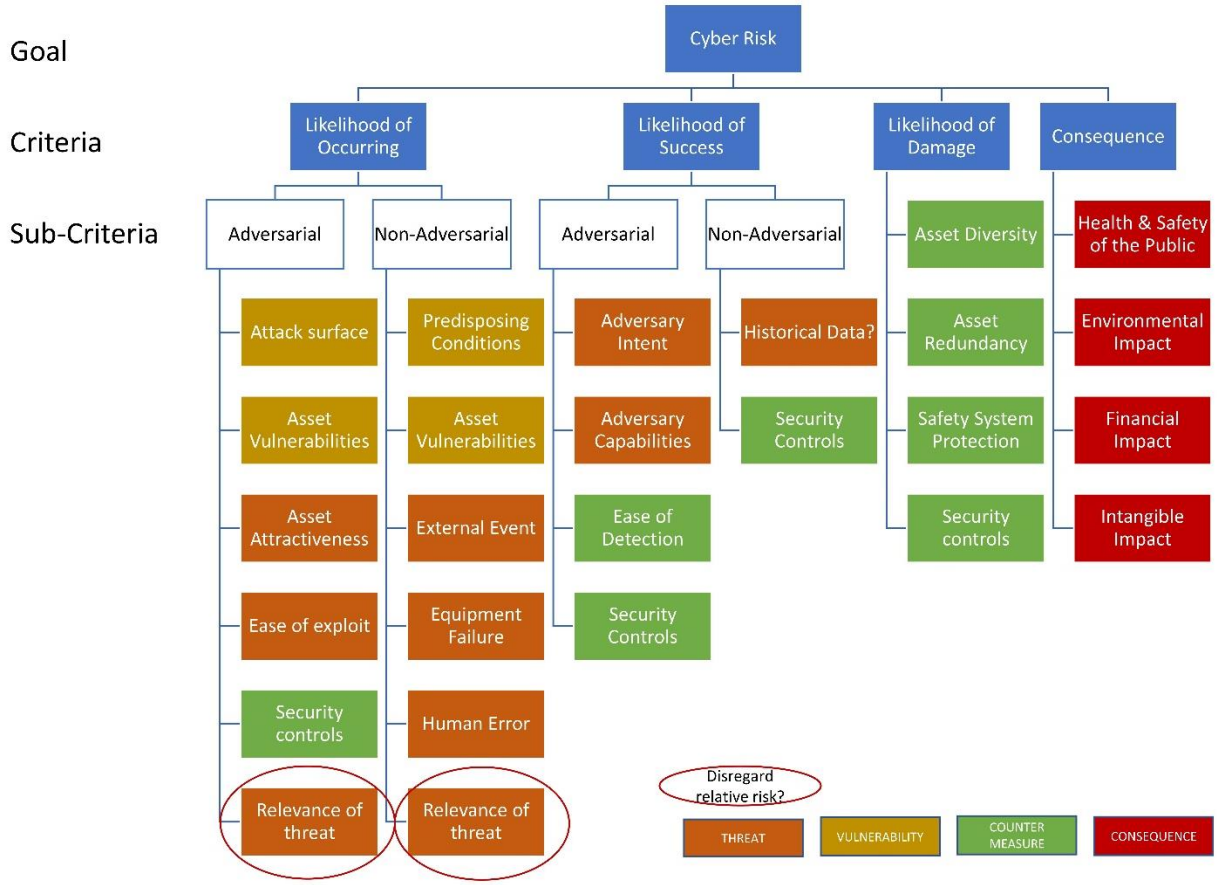

Figure 8. Hierarchy of cyber risk at a nuclear plant.

The criteria in the initial analysis includes the following:

Adversarial Criteria		
Num	**Positive Effect**	**Description**
C1		Likelihood of Attack Occurring
C2		Likelihood of Attack Success
C3	1	Likelihood of Damage
C4		Consequence

C1_1		Attack Surface
C1_2		Asset Vulnerability
C1_3		Asset Attractiveness
C1_4		Ease of Exploit
C1_5	1	Security Controls
C2_1		Adversary Intent
C2_2		Adversary Capabilities
C2_3	1	Security Controls
C3_1	1	Asset Diversity
C3_2	1	Asset Redundancy
C3_3	1	Safety System Protection
C3_4	1	Security Controls
C4_1		Health and Safety Impact
C4_2		Environmental Impact
C4_3		Financial Impact
C4_4		Intangibles Impact

Where C1, C2, C3, and C4 are the main criteria with each of the sub-criteria numbered based upon the main criteria (based on Figure 8).

2. Perform a pairwise comparison between the criteria and each sub-criteria where the pairwise comparison scale is based upon the following:

Pairwise Comparison Scale			
Number shorthand	**Fuzzy Number**	**Definition**	**Explanation**
1	(1,1,1)	Equally preferred	Activities contribute equally
2	(1,2,3)		
3	(2,3,4)	Weakly preferred	One activity slightly favored
4	(3,4,5)		
5	(4,5,6)	Strongly preferred	One activity is strongly favored
6	(5,6,7)		
7	(6,7,8)	Very strongly preferred	One activity is strongly favored and demonstrated in practice
8	(7,8,9)		
9	(8,9,9)	Absolutely preferred	Evidence favoring one activity is of highest possible order of affirmation
1/x	(1/x+1, 1/x, 1/x-1)		
1/9	(1/9,1/9,1/8)		

The pairwise comparison looks at the preference of the row over column.

Pairwise Comparisons					
Criteria	**C1**	**C2**	**C3**	**C4**	
C1	1	1/3	1/5	1/7	
C2	3	1	1/5	1/7	
C3	5	5	1	1/7	
C4	7	7	7	1	
Sub-Criteria	**C1-1**	**C1-2**	**C1-3**	**C1-4**	**C1-5**
C1-1	1	1/4	4	1/5	1/5
C1-2	4	1	4	1/4	1/5
C1-3	1/4	1/4	1	1/5	1/6
C1-4	5	4	5	1	1/4
C1-5	5	5	6	4	1
Sub-Criteria	**C2-1**	**C2-2**	**C2-3**		
C2-1	1	1/5	1/5		
C2-2	5	1	1/5		
C2-3	5	5	1		
Sub-Criteria	**C3-1**	**C3-2**	**C3-3**	**C3-4**	
C3-1	1	1/2	1/4	1/3	
C3-2	2	1	1/4	1/3	
C3-3	4	4	1	1/3	
C3-4	3	3	3	1	
Sub-Criteria	**C4-1**	**C4-2**	**C4-3**	**C4-4**	
C4-1	1	5	5	5	
C4-2	1/5	1	1/4	1/2	
C4-3	1/5	4	1	4	
C4-4	1/5	2	1/4	1	

3. The numbers in the table are converted to their fuzzy numbers and run the AHP calculation to derive weights for each pairwise comparison. These weights are then normalized.

Adversarial Criteria				
Num	**Positive Effect**	**Description**	**Group Weights**	**Overall Weights**
C1		Likelihood of Attack Occurring	0.042409	
C2		Likelihood of Attack Success	0.115623	
C3	1	Likelihood of Damage	0.287342	
C4		Consequence	0.554626	
			1	

C1		Likelihood of Attack Occurring		
C1_1		Attack Surface	0.106663	0.005
C1_2		Asset Vulnerability	0.179846	0.008
C1_3		Asset Attractiveness	0.033918	0.001
C1_4		Ease of Exploit	0.286859	0.012
C1_5	1	Security Controls	0.3927141	0.017
C2		Likelihood of Attack Success		
C2_1		Adversary Intent	0.071588	0.008
C2_2		Adversary Capabilities	0.333333	0.039
C2_3	1	Security Controls	0.5950781	0.069
C3	1	Likelihood of Damage		
C3_1	1	Asset Diversity	0.08635	0.025
C3_2	1	Asset Redundancy	0.147732	0.042
C3_3	1	Safety System Protection	0.363286	0.104
C3_4	1	Security Controls	0.4026321	0.116
C4		Consequence		
C4_1		Health and Safety Impact (rad sabotage)	0.513723	0.285
C4_2		Environmental Impact	0.066952	0.037
C4_3		Financial Impact	0.300662	0.167
C4_4		Intangibles Impact	0.1186621	0.0661

4. The sub-criteria are then evaluated based upon the following scale (the criteria that are indicated as a positive effect are inverted).

Criteria Scale

Num	Fuzzy Num	Description	Membership Function Type
1	(0,0,0.1,0.2)	Very High,Very Large, Very Easy	Trapezoidal
2	(0.1,0.25,0.25,0.4)	Large, High Easy	Triangular
3	(0.3,0.5,0.5,0.7)	Moderate	Triangular
4	(0.6,0.75,0.75,0.9)	Low, Small, Difficult	Triangular
5	(0.8,0.9,1,1)	Very Low, Very Small, Very Difficult	Trapezoidal

Criteria Scores

Criteria	Score
C1	4
C2	1
C3	2

C4	3
C1_1	4
C1_2	5
C1_3	**3**
C1_4	2
C1_5	1
C2_1	4
C2_2	2
C2_3	5
C3_1	3
C3_2	3
C3_3	4
C3_4	2
C4_1	4
C4_2	4
C4_3	5
C4_4	3

5. These scores are converted to fuzzy numbers to calculate the weighted score for each sub-criteria and an overall relative risk score.

Adversarial Criteria						
Num	**Positive Effect**	**Description**	**Group Weights**	**Overall Weights**	**Fuzzy Score**	**Weighted Score**
C1		Likelihood of Attack Occurring	0.042409			
C2		Likelihood of Attack Success	0.115623			
C3	1	Likelihood of Damage	0.287342			
C4		Consequence	0.554626 1			
C1		Likelihood of Attack Occurring				
C1_1		Attack Surface	0.106663	0.005	0.750	0.003
C1_2		Asset Vulnerability	0.179846	0.008	0.933	0.007
C1_3		Asset Attractiveness	0.033918	0.001	0.500	0.001
C1_4		Ease of Exploit	0.286859	0.012	0.250	0.003
C1_5	1	Security Controls	0.392714 1	0.017	0.067	0.001
C2		Likelihood of Attack Success				
C2_1		Adversary Intent	0.071588	0.008	0.750	0.006
C2_2		Adversary Capabilities	0.333333	0.039	0.250	0.010
C2_3	1	Security Controls	0.595078 1	0.069	0.933	0.064

C3	1	Likelihood of Damage					
C3_1	1	Asset Diversity	0.08635	0.025	0.500	0.012	
C3_2	1	Asset Redundancy	0.147732	0.042	0.500	0.021	
C3_3	1	Safety System Protection	0.363286	0.104	0.750	0.078	
C3_4	1	Security Controls	0.402632	0.116	0.250	0.029	
			1				
C4		Consequence					
C4_1		Health and Safety Impact (rad sabotage)	0.513723	0.285	0.750	0.214	
C4_2		Environmental Impact	0.066952	0.037	0.750	0.028	
C4_3		Financial Impact	0.300662	0.167	0.933	0.156	
C4_4		Intangibles Impact	0.118662	0.066	0.500	0.033	
							Total Relative Risk Score
			1	1		0.666	