SAND2014-20716C

Assessment of Hazard Analysis Methods for Nuclear Power Cyber Security

Phil Turner, Tim Wheeler, Lon Dawson, Alice Muna
Sandia National Laboratories
PO Box 5800, Albuquerque, NM 87185
pturner@sandia.gov; tawheel@sandia.gov; ladawso@sandia.gov; amuna@sandia.gov

ABSTRACT

Utilities in the electric sector currently identify Critical Digital Assets (CDAs) and apply cyber security controls using a deterministic-based process. Because cyber risk is a relatively new area for most utility control operations, the process for identifying cyber vulnerabilities, threat vectors, and their impacts is not clearly defined. As a result, utilities have adopted strategies that use a shotgun approach of all security controls. These strategies can result in either under or overprotection of systems. This paper presents insights of an EPRI funded study regarding the potential utility of applying several hazard and systems analysis methods for determining cyber risk associated with digital systems.

Although current hazard analysis methods are not ideally suited to assess phenomena important to cyber security (e.g., potential malware insertion and control tactics), this research yields important insights on how to leverage these methods to enable robust cyber security, vulnerability, and consequence assessment of a critical infrastructure facility.

As cyber risk becomes more fully understood, the process by which controls necessary to appropriately protect the system are designed and implemented will be more efficacious. This may allow for an entity to reevaluate their CDAs and provide for a more defensible security plan and better prioritization of risks and resources.

Key Words: hazard analysis, cyber risk, vulnerability, cyber security, critical digital asset

1 INTRODUCTION

Utilities in the electric sector currently identify Critical Digital Assets (CDAs) and apply cyber security controls using a deterministic-based process. Because cyber risk is a relatively new area for most utility control operations, the process for identifying cyber vulnerabilities, threat vectors, and their impacts is not clearly defined. As a result, utilities have adopted strategies that use a shotgun approach of all security controls. These strategies can result in either under or overprotection of systems. This paper presents insights of an EPRI funded study regarding the potential utility of applying several hazard and systems analysis methods for determining cyber risk associated with digital systems.

With regard to operational design, EPRI investigated several methods for hazard and failure analysis of industrial digital instrumentation and control (DI&C) systems. The purpose of the analysis was to evaluate hazard and failure analysis models and methods used in traditional engineering analyses for their comprehensiveness, practicality, and cost-effectiveness regarding the design of industrial DI&C systems for operational purposes. This investigation is documented in *Hazard Analysis Methods for Digital Instrumentation and Control Systems* (EPRI-509) [1]. The methods studied in EPRI-509 are:

1. Failure Modes and Effects Analysis (FMEA),
2. Fault Tree Analysis (FTA),
3. Hazard & Operability Analysis (HAZOP),
4. Systems Theoretic Process Analysis (STPA), and
5. Purpose Graph Analysis (PGA).

The cyber aspect of the DI&C systems was not evaluated in the analysis contained within the original EPRI report. EPRI subsequently has embarked upon a project to determine the feasibility of applying these models for analyzing the cyber security risk of digital systems in the electric sector. The primary focus is to determine the best approach for a cyber-focused risk analysis and then develop a method for cyber informing the hazard analyses with a pilot and tool for implementation. The goal is to determine if the proposed models identified in the EPRI report can be leveraged to increase the efficiency of cyber security programs by avoiding practices that result in under or overprotection of digital assets. A more in-depth discussion on the efficacy of each of these methods can be found in Reference 8.

1.1 Analysis of Hazards Analysis Methods

Insights regarding fundamental challenges, limitations, and possibilities of using traditional hazard and systems analysis approaches (e.g., top-down, bottom-up, bi-directional) are discussed. Each method is presented within the following context (other metrics are also being evaluated but are not represented in this paper):

1. Potential for identification and prioritization of CDAs.
2. Potential for cyber-informing the hazard analysis method or model.
3. Potential for synergistic applications with other methods.

1.1.1 Failure Modes and Effects Analysis (FMEA)

FMEA identifies single equipment failure modes in a design, process, or product, with each failure mode's potential effects on the system. The Functional FMEA (FFMEA) assesses system-level functions and processes from the top-down. The Design FMEA (DFMEA) analyzes failure modes from single components, known as the bottom-up approach [1]. FFMEA identifies effects of systems that fail to perform their design functions in specific ways and DFMEA identifies effects of components that exhibit specific failure modes. Both FMEA methods utilize exhaustive enumeration, meaning each component and each function is examined individually to determine failure modes. The methods then apply inductive reasoning to determine the effects on the system if a failure were to occur.

1.1.1.1 Potential for identification and prioritization of CDAs

The DFMEA method neither identifies nor prioritizes which digital devices receive a full-blown analysis. Instead it requires the analysis of all the digital devices before the method can prioritize them. Once the analysis has been completed, a user could prioritize the devices based on the effects of the failures; however this method is not efficient. A combination FFMEA and a DFMEA would be beneficial in this case. The FFMEA can identify undesired system outcomes and trace them to their root causes while a DFMEA can analyze those root causes fully.

1.1.1.2 Potential for cyber-informing the method

FMEA is a useful method for identifying how component failures may occur. It is not suitable for characterizing system behavior following a cyber event since it does not characterize complex interactions or account for the timing of events. Other hazard analysis methods, e.g., PGA or STPA, are more appropriate for analyzing unique complex interactions. If failure modes of digital components are well understood, then the method is able to readily evaluate the failure modes without any changes to the methodology.

1.1.1.3 Potential for synergistic applications with other methods

The FMEA method can be used with other methods. DFMEA in particular can be coupled with a top-down method to limit the scope of the DFMEA. Since the DFMEA analyzes single failure modes better than most other methods, it can be used to analyze specific components the top-down method

identifies as critical. The system consequences derived from both FMEA methods may also be useful in understanding consequences that would feed into a cyber-informed risk analysis.

1.1.2 Fault Tree Analysis (FTA)

Fault trees are deductive logic models referred to as top-down in approach [5]. Development of a Fault Tree Analysis (FTA) model starts by defining the occurrence of a top event that represents an undesirable outcome for a facility or process (e.g., core damage, interruption of electricity generation), or simply a single system or set of several systems (e.g., loss of primary coolant system integrity). FTA is used to identify combinations of failure modes of structures, systems, and components (SSCs) that would lead to failure of systems to perform their intended functions. A fault tree represents a logical structure through which faults, or component failure modes, can be propagated from bottom up through logic AND/OR gates to render a Boolean equation of all combinations of failures that would cause the undesirable top-event to occur. Each term, or combination of faults, is referred to as a "cut set."

1.1.2.1 Potential for identification and prioritization of CDAs

FTA could be used to generate quantitative insights regarding digital vulnerabilities consistent with the ability of FTA to provide quantitative insights in traditional systems analysis. It appears that the potential exists that FTA importance measures could be calculated using cyber related analogs to probability [2] that provide a quantitative measure of a digital asset's susceptibility to cyberattack. Analogous to probabilities, these metrics of cyberattack vulnerability for individual digital assets could be used to generate quantitative FTA results such as importance measures for digital assets. These results in theory could be used to establish relative rankings of digital assets against their potential as cyber-related vulnerabilities to a system or plant.

1.1.2.2 Potential for cyber-informing the method

Initial assessments of the use of FTA for cyber security assessments indicate that there is significant potential to incorporate cyber components directly into FTA. Instrumentation and control (I&C) components traditionally have not been explicitly modeled in system fault trees because industrial I&C systems were largely analog. As the industry evolves from analog to digital I&C, the industry is confronting the challenge of adapting Probabilistic Risk Assessment (PRA) fault tree models for I&C components to address modeling and quantification of DI&C plant features.

The FTA cut set analysis of a circulating water system (CWS) in Reference 6 reveals that the CWS system is vulnerable to eight different double DI&C faults, completely independent of any human or hardware faults. Other cut sets show that certain DI&C faults, when occurring in conjunction with a single human error, would render the CWS inoperable. A potential insight here is that the system is vulnerable to failure through several sets of double digital asset faults without any accompanying hardware or human failures. It is further revealed that the failure of other digital assets would contribute to CWS failure only in conjunction with other non-cyber faults.

1.1.2.3 Potential for synergistic applications with other methods

FTA and PRA historically have benefited from insights gained through other analysis methods to ensure accurate systems modeling, component failure mode identification, subtle system interactions, and system success criteria. With regard to cyber-informing FTA, cyber faults and resultant impacts upon system behavior would be key inputs most probably derived from other hazard assessment methods such as FMEA and HAZOPS.

1.1.3 Hazard and Operability Analysis (HAZOP)

The Hazard and Operability Analysis (HAZOP) method identifies the potential hazards in a system and the potential operational disturbances or deviations that lead a system to deviate from expected behaviors. The HAZOP method uses exhaustive enumeration, meaning every identified hazard,

operational disturbance, or deviation must be examined individually for potential causes of failure. Once this has been completed, inductive reasoning is used to determine the effects of each hazard on the system. The HAZOP method uses “guide words” to provide structure to the analysis and push the analyst toward analytical completeness. The idea is to use each guide word in the context of the potential hazard, operational disturbance, or deviation to determine if the affected process deviates from intended design.

1.1.3.1 Potential for identification and prioritization of CDAs

The HAZOP method looks at possible causes that would alter a process from its intended function. These causes could be failures of digital assets. CDAs could be identified in this way and prioritized based on the effects of the process. The only CDAs that are analyzed are for the classes of hazards, operational disturbances, or disruptions. The combined effects of multiple failures are not analyzed; therefore the CDAs that could contribute to this combined failure are neither identified nor prioritized. The HAZOP method also analyzes the system at a certain level; vulnerabilities for individual digital devices can affect larger digital components that are not analyzed in the HAZOP method.

1.1.3.2 Potential for cyber-informing the method

A HAZOP analysis is not suitable for characterizing system behavior for a complex cyber event because it does not identify unique interactions or account for the timing of events. A HAZOP analysis is suitable for analyzing system parts individually. For highly interconnected systems like computer networks, a deviation at one part of a system may have a cause elsewhere. Other hazard analysis methods, e.g., PGA or STPA, are more appropriate for identifying and analyzing unique, complex interactions and consequences.

1.1.3.3 Potential for synergistic applications with other methods

The HAZOP method can be used easily with both top-down and bottom-up methods. A top-down method, like FTA or FFMEA can limit which processes are analyzed in the HAZOP. The HAZOP method can be used to limit the scope of a bottom-up method like DFMEA by identifying components of interest for the DFMEA to analyze on a deeper level.

1.1.4 Systems Theoretic Process Analysis (STPA)

Systems Theoretic Process Analysis (STPA) was developed to identify accident scenarios that encompass the entire accident process, not just the electromechanical components. STPA analysis models, for example, may include design errors, software flaws, component interaction accidents, and cognitively complex human decision-making errors [1]. The STPA method is intended to capture unique failure modes that may have been missed or misunderstood in traditional hazard analyses. The term “control action” used in the STPA method describes the effect that a controller (human, machine, or both) has on an actuator and, ultimately, on the controlled process. One of the goals during analysis is to determine the context in which each control action can be hazardous. The two main steps in STPA include: (1) Identifying the potential for inadequate control of the system which could lead to a hazardous state and (2) determining how each potentially hazardous control action identified could occur.

1.1.4.1 Potential for identification and prioritization of CDAs

The STPA method is not intended to identify all CDAs. It identifies them through their impact to the control processes [7]. In this way, identified digital assets already are prioritized based on their impact to a control action. One of the problems with this methodology is that the process of analyzing each control action is tedious when working with a complex system. A complete analysis, however, would identify digital assets that are critical for complex system states.

1.1.4.2 Potential for cyber-informing the method

Similar to HAZOP, states are postulated for each control action with resulting behaviors. For each control action, all the possible states of the controller feedback signals are documented along with the combination of the various states that result in a hazardous control action. The pathways for possible cyber security incidents become evident when the documentation is generated. By including all possible states, the analysis is thorough, capturing unique hazardous behaviors that may not have been captured in traditional hazard analysis methods.

The STPA analysis is not intended to analyze how component failures might occur; it could be used to describe the system behavior following a cyber event or attack (e.g., cyber common cause failures). The strength of STPA is that it analyzes all the interactions between a controller and control process and how multiple interactions can affect a system. A cyber impact from an event, such as the loss of a network or loss of data integrity, could be translated easily into system effects with a properly mapped out STPA analysis. The STPA method also analyzes human control over a process as a feedback loop, not as a chain of directly related events. This analysis of potential human responses can be used to analyze behaviors during or after a cyber event.

1.1.4.3 Potential for synergistic applications with other methods

STPA could be blended with other hazard analysis methods to create a more efficient and more complete analysis. STPA can be paired with top-down methods, such as FTA. This could limit the scope of the STPA analysis, thus making the workload more manageable. STPA could identify processes or components on which to perform a single failure analysis. Like other top-down methods, STPA does not analyze single failures unless each controller feedback signal is considered in isolation; this goes against the intended purpose of STPA. DFMEA, for example, could be used in conjunction with STPA to analyze these types of failures for a system.

1.1.5 Purpose Graph Analysis (PGA)

A Purpose Graph is a figure that illustrates system features in terms of observables, states, goals and processes. The State Graph combines measured characteristics (observables) into states. The process graph combines system processes into goals. The state and process graph together form the purpose graph that facilitates complex system analysis for 1) state redundancy, interdependence, and attribute diversity, 2) goal interactions (direct and indirect) and 3) process redundancy, interactions, interdependence, etc. [1].

1.1.5.1 Potential for identification and prioritization of CDAs

PGA is excellent for complex system analysis. It would help in designing systems to be more resilient to any failure, including those from cyber events. Given a set of cyber events, PGA could help determine which digital assets were more important to the system impact, which could aid in the prioritization of digital assets.

1.1.5.2 Potential for cyber-informing the method

PGA by itself does not identify how component failures might occur. It could be readily enhanced to describe system behaviors following cyber events (e.g., cyber common cause failures). This is because thorough analysis of hazards from PGA makes understanding complex interactions between system observables, processes, states, and goals easier and more intuitive. Once a hazard is known, PGA helps to understand and mitigate the associated system effects. The state graph shows the relationship between measured physical characteristics (observables) and system states. The state graph analysis could be extended beyond an analysis of only the physical characteristics to all aspects of the associated sensor information – how the sensor data is verified accurate, transported, and stored. An observable sensor has a digital lifecycle in software design and maintenance that may extend to other similar sensors. The lifecycle of the sensor-related information needs to be considered, including common cause failures from

signal processing, network transport, system control, data storage, etc. Because the state graph couples observables with systems states, the cyber impacts from an event such as the loss of a network or loss of data integrity can be easily translated into system effects. Similarly, there will be a process graph that includes the communication and control processes and be subject to the systematic analyses of PGA.

1.1.5.3 Potential for synergistic applications with other methods

Both PGA and FMEA begin with function/process mapping, so coupling of these methods is somewhat intuitive and efficient. As described in EPRI-509, if the PGA method is applied first, the process would yield hazards along with the relevant components and processes. FMEA can focus exclusively on the devices or components that could cause or contribute to hazards, then determine the failure modes or failure mechanisms that could lead to such hazards.

1.2 Cyber Informed Risk Analysis

There is evidence that adversaries have been performing exploits and compromising systems important to national infrastructure [3]; additional interest seems highly likely. Since this type of adversary would be a high-tier actor, the adversary is considered motivated and sufficiently capable [4]. The systems would need to be protected from various levels of adversaries. Some controls will prevent low-level adversaries from being able to compromise systems while not addressing mid and high-level adversaries. When evaluating the cyber risks associated with the system, understanding the difficulty of carrying out attacks is important to contextualize that risk to make a determination of the exploitability of the system. The goal in understanding and mitigating the cyber risk elements of the system is to raise the requirements for successful attack or exploit. In the case of critical infrastructure, the goal is to raise those requirements significantly, requiring much more difficult and complex attacks to achieve desired adversarial consequences. The difficulty for the defender is in knowing what level of resources to allocate to the problem and where to place them.

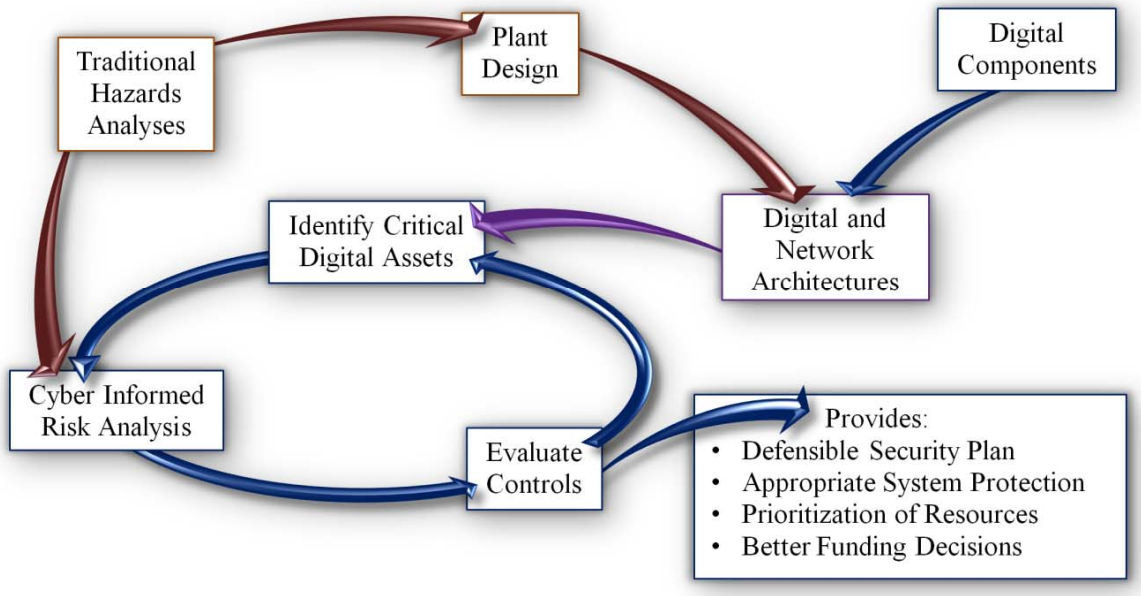

Figure 1. Process Flow for Cyber Analysis.

Figure 1 shows a visual representation of the process at a high level for performing a cyber risk analysis of the system which would allow for better vetting of CDAs and the protections placed in and around the system. The system can be evaluated at the component level and at the system level. A cyber

informed risk analysis can be conducted with inputs from the various hazard analysis methods, as appropriate.

The hazards analysis methods evaluated each provide unique elements that can be leveraged to help understand the cyber risk posed to the system. Traditionally, these methods have been used during plant design or upgrades with results applied to improve the plant. Relatively recent years have seen the addition of digital components, which adds complexity when considering supply chain and interconnectivity of these systems. This additional complexity adds unique vulnerabilities and attack vectors which has not previously existed.

The risk analysis is mainly focused on vulnerabilities, the difficulty in exploiting those vulnerabilities, and system consequences. The system is evaluated with two main focus areas, risks arising from interconnectivity of systems and components and risks inherently associated with individual components, as shown in Figure 2. At the component level, the analysis includes the supply chain associated with hardware and software, internal processes such as the process for updating device firmware, and networks that support those operations, among others.

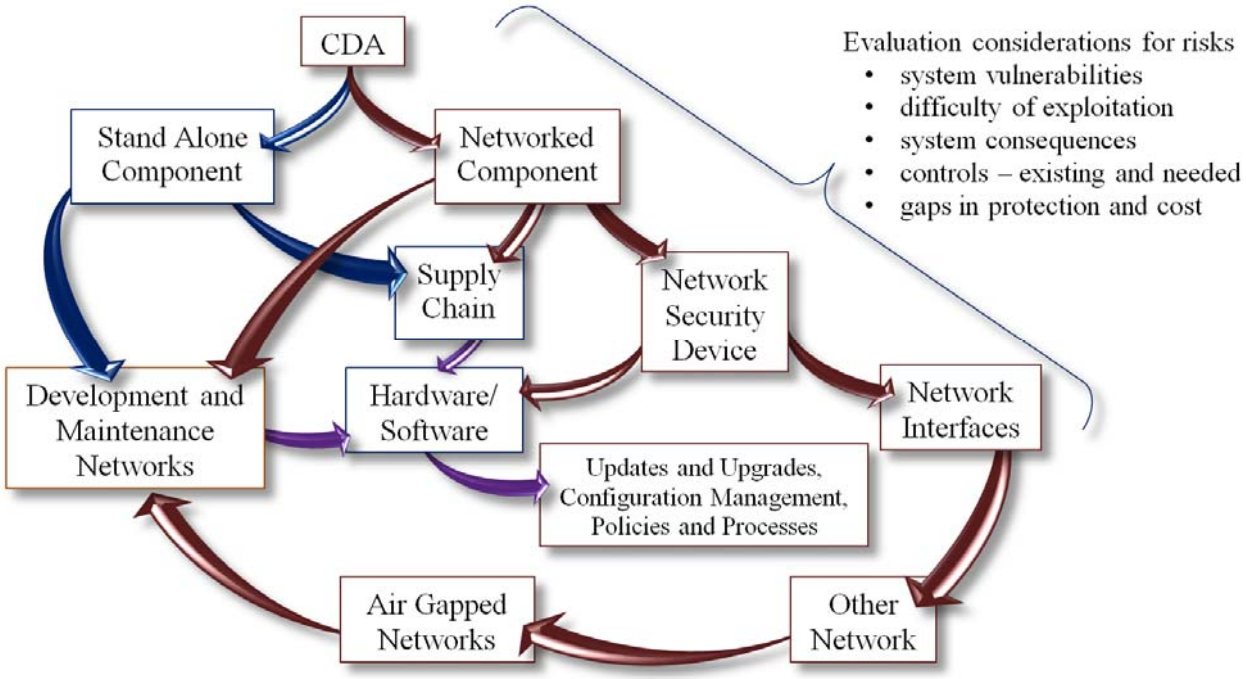

Figure 2. System Considerations for Cyber Analysis.

For interconnected components, the analysis considers cyber common cause failures such as shared components, common attack paths, communication or network protection devices and technologies, and vulnerabilities resulting from the combinations of those elements. The original hazard analyses might show where system redundancy is needed from a physical reliability perspective. This redundancy could add additional hardware that includes common digital controllers, which could be interconnected through the same communications network. This connectivity is an example of a cyber common cause failure. In this case, redundancy would have reduced the resiliency of the system to survive in the case of a successful a cyberattack against those components.

2 CONCLUSIONS

Cyber risk management is a relatively new field and processes for identifying vulnerabilities, threat vectors, and the impacts are not clearly defined. Although current hazard methods are not ideally suited to

assess phenomena important to cyber security (e.g., potential malware insertion and control tactics), this research yields important insights on how to best leverage these methods to enable robust cyber security vulnerability and consequence assessment of a critical infrastructure facility. As the risk is more fully understood, it will enable better understanding of the controls necessary to appropriately protect the system. This may allow for an entity to reevaluate the CDAs that have been identified and provide for a more defensible security plan.

Research is continuing to understand more fully how to incorporate cyber elements into existing hazard analysis methods and to understand the most effective and efficient ways of using the outputs from those methods. A pilot process will then be developed with a potential tool, or tools, identified that could aid in the analysis process.

3 ACKNOWLEDGEMENTS

This research into applying hazards analysis models and methods to better enable understanding of cyber risk in the electric utility sector has been funded by the Electric Power Research Institute (EPRI). This paper reflects this ongoing research to develop a methodology and tools for cyber risk analysis [8].

4 REFERENCES

1. *Hazard Analysis Methods for Digital Instrumentation and Control Systems*, 3002000509, EPRI, Palo Alto, CA (2013).
2. Peter Mell and Karen Scarfone, National Institute of Standards and Technology, Sasha Romanosky, Carnegie Mellon University, *CVSS: A Complete Guide to the Common Vulnerability Scoring System, Version 2.0*. NIST (2007).
3. *APT1: Exposing One of China's Cyber Espionage Units*. Mandiant (2013).
4. Defense Science Board (DSB), *Task Force Report: Resilient Military Systems and the Advanced Cyber Threat*. Office of the Undersecretary of Defense for Acquisition, Technology and Logistics (2013).
5. U.S. Nuclear Regulatory Commission, *Fault Tree Handbook*. NUREG-0492, USNRC (1981).
6. *Modeling of Digital Instrumentation and Control in Nuclear Power Plant Probabilistic Risk Assessments*, 1025278, EPRI, Palo Alto, CA (2011).
7. *An STPA Primer*, Version 1, Massachusetts Institute of Technology, http://sunnyday.mit.edu/STPA-Primer-v0.pdf (2013).
8. *Interim Analysis of Hazard Models for Cyber Security*, 3002003248, EPRI, Palo Alto, CA (2014).