A Method to Choose Between Automation and Human Operators for Recovery Actions During a Cyber Attack

Conference Paper · April 2019

CITATIONS
2

READS
106

2 authors:

Douglas Lee Van Bossuyt
Naval Postgraduate School
135 PUBLICATIONS **936** CITATIONS
SEE PROFILE

Bryan O'Halloran
Naval Postgraduate School
64 PUBLICATIONS **476** CITATIONS
SEE PROFILE

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2019) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

17th Annual Conference on Systems Engineering Research (CSER)

A Method to Choose Between Automation and Human Operators for Recovery Actions During a Cyber Attack

Douglas L. Van Bossuyt*[a], Bryan M. O'Halloran[a]

[a]*Department of Systems Engineering, Naval Postgraduate School, Monterey, California 93943, USA*

Abstract

As complex systems such as nuclear power plants, naval ships, critical infrastructure, and other systems become more connected to the internet and digital control interfaces, the chance of a cyber attack causing physical damage to a system and failure of the system increases. In many systems, recovery actions can prevent an incipient failure from causing a system-wide failure. This paper presents a method of determining if a human operator or an automated system is more appropriate to complete a recovery action during a cyber attack. The method is useful during the conceptual phase of system design where architecture changes have minimal impact on the cost and schedule of the system design effort. Practitioners can use the method to make cost and probability-informed decisions. A case study of a spent fuel cooling pool in a nuclear power plant is presented to illustrate the method.

Keywords: System Design; System Architecture; Design Trade-Off Studies; Recovery Actions

1. Introduction and Background

Modern systems are increasingly complex and challenging to design [1]. While the much of the system design effort focuses on adherence to requirements, a significant portion of the of the technical effort is focused on ensuring that the system will not fail and is safe [2]. A significant body of reliability and systems safety methods exist in the literature and in practice where the methods tend to focus on probabilistic failures and hazards combined with corresponding consequences [3, 2, 4, 5, 6, 7, 8]. This approach often identifies failures and hazards that have been seen in previous system, and occasionally will creatively search for and identify new types of failures. The successful conclusion of a reliability and system safety effort is a safe and reliable system.

* Corresponding author. Tel.: +01-831-656-7572.
E-mail address: Douglas.VanBossuyt@nps.edu

Two broad categories of recovery actions can be taken to attempt to stop an incipient fault before it has led to system failure: automated recovery and human recovery actions [9, 10, 11]. Many legacy systems rely on human recovery actions while systems designed in the last three decades are increasingly relying on a mix of human recovery actions and automated recovery actions. While human operators often have a lower probability of success in completing a recovery action, automated systems are often more expensive to implement. As a result, there exists a trade-off between using automated systems and human operators to conduct recovery actions.

Over the last two decades, industrial and defense systems such as power grids, aircraft, nuclear reactors, factories, and other critical systems have become increasingly connected to the internet [12]. With higher connectivity comes greater risk of a malicious and potentially targeted cyber attack [13]. Safeguard such as air-gapped networks have proven to not be a barrier to determined attackers [14]. While a wide range of fault and failure initiators for systems are well understood and quantified [15], predicting the likelihood of a cyber attack and the cyber attack's potential outcomes remains challenging [16, 17].

Some efforts have been made to design systems using early conceptual functional flow block diagrams (FFBDs) [18, 19] to make systems more resilient to and robust against cyber attacks [16, 20]. However, little has been done to understand when a human operator or an automated recovery system is more appropriate to implement a recovery action during an incipient failure event caused by a cyber attack.

1.1. Specific Contributions

This paper specifically contributes a method to aid systems engineers in assessing if a human operator or an automated system is more appropriate to conduct a recovery action as part of an unfolding failure scenario caused by a cyber attack. The method is suitable for the conceptual phase of system design when architectural trade-off studies are being conducted and the cost of making significant changes to system architecture is low. The specific type of cyber attack of interest to this paper is one that occurs during plant operation which directly affects plant operation, such as commanding a valve to close when it should remain open.

2. Methodology

This section presents a methodology to assess if human operators or automated systems are more appropriate for specific recovery actions that a system can take to prevent an incipient fault from causing a system-level failure. In order to use this method, information must be available from generally comparable systems – completely novel system designs using new architecture and new components may not benefit from this method.

A case study based on a generic nuclear power plant spent fuel cooling pool is presented to illustrate the methodology. Spent fuel cooling pools are used to maintain fuel temperature below the threshold for damaging the cladding used to contain enriched uranium fuel elements after the fuel has come out of a reactor core. The fuel requires a period of water immersion cool-down of several years before it is transferred into long-term storage [21].

2.1. Preparatory Step

Prior to using the method, a variety of models and data products of the system must be assembled. These include: 1) a FFBD system model [18], 2) a database of similar system designs with function-to-component mapping that includes failure and prognostics and health management (PHM) [22] data [23, 24, 25, 26], 3) and an analysis of the system using the PHM system design method based on functional modeling of a system presented by L'Her et. al. [25] and including the function failure identification and propagation (FFIP) method of functional risk analysis [27, 28]. The family of FFIP functional failure analysis methods uses functional models and component-to-function databases of failure information to determine the likelihood of a system failing due to a varity of initiating events both internal and external to the system of interest. Recent FFIP developments include an expansion of the FFIP methodology into systems of systems [29]. In L'Her et. al.'s method, a functional system architecture is first analyzed for failure likelihood using techniques broadly similar to FFIP but with specific augmentation for failure propagation timing. The failure information is then used to determine where PHM sensors might be installed to detect an incipient failure while there is still time for a recovery action to be taken to get the system to a safe system state. It is important to

Failure Scenario Chain of Events	Probability of Occurrence (Per Year)
Heat Exchanger Motor Valve #1 Fails Closed, Heat Exchanger Motor Valve #2 Fails to Open, Water Boils Off From Cooling Pool	2.7E-4/yr
Pump #1 Fails to Operate, Pump Valve #3 Fails Closed, Water Boils Off From Cooling Pool	5.3E-5/yr
Pump #2 Fails to Operate, Pump Valve #1 Fails Closed, Water Boils Off From Cooling Pool	5.3E-5/yr
Heat Exchanger #1 Clogs, Heat Exchanger #2 Clogs, Emergency Cooling Water Supply Exhausted, Water Boils Off From Cooling Pool	1.8E-6/yr
Drain Valve Fails Open, Emergency Cooling Water Valve Fails to Open, Water Boils Off From Cooling Pool	2.7E-7/yr

Table 1. Top five most likely failure scenarios determined Using L'Her et. al.'s method and the FFIP method of functional risk analysis

note that L'Her et. al.'s method does not include analysis of malicious cyber attacks and further does not examine if a human operator or an automated system is more appropriate to perform the recovery action.

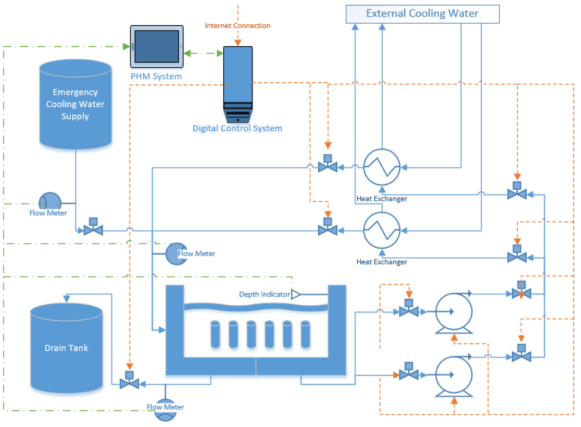

Fig. 1. Spent fuel pool P&ID with proposed PHM system. Three flow meters and a depth gauge have been added to the spent fuel pool design to monitor water moving through the system and detect potential incipient faults. The solid lines represent piping while the dashed lines represent data and control wiring. Dashed and dotted lines indicate data flow from the PHM sensors to the PHM system. External cooling water is supplied to the system from a source outside of the analysis conducted in the case study. The digital control system is connected to the internet.

Due to space constraints, a FFBD of the spent fuel cooling pool system is not shown here. Similar FFBDs can be seen in [25, 30]. Additionally, a database of similar systems with requisite function-to-component mapping including failure and PHM data is not shown here due to space constraints. Similar databases of function-to-component mapping with specific data (i.e.: reliability, mass, sustainability, failure probability, etc.) can be seen in [23, 25, 31, 32]. Analysis to design a PHM system for the spent fuel pool cooling system was completed including using L'Her et. al.'s method to place PHM components, and a resulting P&ID is shown in Figure 1. A table of the top five most likely failure scenarios as analyzed using the FFIP method of functional risk analysis is shown in Table 1. It is useful to note that the scenarios shown in Table 1 are displayed in a manner similar to cutsets found in probabilistic risk assessment (PRA) [33].

2.2. *Step 1: Hazard/Threat Analysis*

The first step of the method is a hazard and threat analysis of the system specifically with respect to cyber attack. While there are many potential ways to conduct such an analysis, the authors suggest using the method O'Halloran et. al. developed to assess cyber physical attacks on systems [16]. O'Halloran et. at.'s method develops the worst case scenario cyber physical attacks via a genetic algorithm approach where the worst attacks are sought out through successive generations of potential attacks.

Practitioners are advised to carefully inspect the list of potential hazards and down-select to the most probable or most concerning hazards. While the O'Halloran et. al. method does produce the worst cyber physical hazards possible for a system, it is still necessary to review the list of hazards and validate that they are possible. Further, it is important to validate that the "worst" hazards output by the O'Halloran et. al. method are in fact the worst for the system based

Scenario Label	Hazard Scenario
Scenario 1	Cyber Attack Disables Pumps #1 and #2, Emergency Cooling Water Tank Depleted, Water Boils Off From Cooling Pool
Scenario 2	Cyber Attack Opens Drain Valve, Emergency Cooling Water Tank Depleted, Water Boils Off From Cooling Pool
Scenario 3	Cyber Attack Closes Pump Valves #1 and #3, Emergency Cooling Water Tank Depleted, Water Boils Off From Cooling Pool
Scenario 4	Cyber Attack Opens Drain Valve, Cyber Attack Prevents Emergency Cooling Water Valve From Opening
Scenario 5	Cyber Attack Closes Heat Exchanger Valves #1 and #3, Emergency Cooling Water Tank Depleted, Water Boils Off From Cooling Pool
~~Scenario 6~~	~~Cyber Attack Disables Pumps #1 and #2, Cyber Attack Opens Drain Valve, Cyber Attack Prevents Emergency Cooling Water Valve From Opening~~
~~Scenario 7~~	~~Cyber Attack Closes All Valves, Water Boils Off From Cooling Pool~~

Table 2. Spent fuel cooling pool hazard list as identified by O'Halloran et. al.'s method. A horizontal line through a hazard indicates that the hazard has been removed from further consideration due to being ruled as invalid for further analysis.

on an analysis of the core functionality of the system [1]. The resulting list of cyber attack hazards can then be used in the subsequent steps of this methodology.

The top seven hazards identified by O'Halloran et. al.'s method for the spent fuel pool case study are shown in Table 2. Two of the hazards have been struck from the list after review by a system expert as being invalid due to the digital control system segmenting valve operations in such a manner that a cyber attack is unlikely to affect multiple valve groupings at the same time.

2.3. Step 2: Analyze Recovery Actions

The second step of the methodology analyzes potential recovery actions that can be taken to stop an incipient fault and either place the system in a safe state or a nominal operating state. While some systems such as nuclear reactors are designed and operated in such a way that a safe shutdown state is the objective of a recovery action [10, 3], other systems use recovery actions to bring a system back to a nominal operating state [34, 35]. The recovery actions developed from L'Her et. al.'s work [25] as part of the preparatory step of the method presented here is recommended as the starting point to determine what recovery actions can be performed either by an automated system or by humans. However, new recovery action development may be justified due to potentially new cyber attack failure initiating events being identified in Step 1 of the method that were not assessed previously. Many methods exist to develop appropriate recovery actions for systems such as [10, 9].

Table 4 shows how the recovery actions identified by L'her et. al.'s method map to the hazards identified from O'Halloran et. al.'s method. Note that there are no recovery actions identified on two of the hazards from L'her et. al.'s method due to the method truncating scenarios that were deemed to be of low probability or specific initiating events that were not included in the analysis due to L'her et. al.'s method not assessing cyber attack initiating events. Recovery actions shown in an italics font are those that were developed independent of L'Her et. al.'s method after the need for a recovery action to a cyber attack was identified by O'Halloran et. al.'s method.

Specific information about potential recovery actions must be collected including: 1) specific recovery action (e.g.: open a valve, connect a cable, etc.), 2) amount of time available to take the recovery action, 3) information relevant to conducting human reliability analysis (HRA) [10, 36, 37] on the recovery action, 4) information relevant to developing an automated system to conduct the recovery action. Next, a HRA and an analysis of an automated recovery system for each specific recovery action can be conducted to determine the following information: 1) probability of success or failure of the recovery action, 2) cost of implementing the recovery action, and 3) cost of a cyber attack succeeding and causing system failure.

2.4. Step 3: Trade-off Study Between Automated and Human Recovery Actions

The third step of the method develops a trade-off study between automated and human recovery actions. In order to make decisions more easily, the authors advocate placing all decision-making criteria into the cost domain. The decision-making criteria are based on several factors including the following:

2.4.1. Probability of Successful Recovery

This factor addresses the chance that a recovery will be successful. Three broad categories of input are considered to achieve the final probability. Specifically, the probability of a recovery action being successful, P_{RS}, is the mathematical product of the probability of detecting that a malicious act has occurred, P_{DMA}, the probability of correctly identifying the recovery action to be taken, P_{IRA}, and the probability that the recovery action is successful, P_{RAS}, as shown in Equation 1. These values are assigned independently for each malicious act and separately for human operators and automated systems. HRA methods follow the basic outline described above to determine human operator probability of successfully completing a recovery action. Cheung et al., presents research on the automated detection of systems, which can be used to inform the probability of detecting that a malicious act has occurred [38].

$$P_{RS} = P_{DMA} * P_{IRA} * P_{RAS} \tag{1}$$

Similar techniques are used to determine the probability of an automated system completing a recovery action although with focus on understanding the ability of a system to use its PHM system to detect the malicious act and determine the correct recovery action [39, 20].

2.4.2. Recovery Action Cost

The cost to implement a human operator recovery action versus an autonomous system recovery is a critical factor in determining which is optimal. This factor estimates the cost of each independently. Recovery action cost values should estimated on an equivalent timescale basis for accurate comparison (e.g.: yearly, weekly, etc.). In this paper, the authors use an annual cost basis.

Practitioners are advised to determine recovery action costs that are appropriate for their specific systems. Nuclear power plants often have higher recovery action costs than water treatment plants, for instance. Regardless of the system being investigated, it is important to capture the entire cost of a recovery action including but not limited to the following factors: 1) fully burdened salary of operators, maintenance personnel, and other people involved with the recovery action or with maintaining the automated recovery system ($COST_{Humans}$), 2) cost of all equipment and hardware both for automated recovery systems and for human-operated recovery systems ($COST_{Equip}$), 3) maintenance costs not already captured ($COST_{Maint}$), 4) in specific circumstances, residual costs associated with a recovery action ($COST_{Res}$) such as in the case of emergency neutron poisoning in a nuclear reactor to quench the fissile reaction [40], and 5) any other cost factors identified by the practitioner ($COST_{O}ther$). Equation 2 demonstrates how the cost factors are combined into a recovery action cost ($COST_{Recovery}$).

$$COST_{Recovery} = COST_{Humans} + COST_{Equip} + COST_{Maint} + COST_{Res} + COST_{Other} \tag{2}$$

2.4.3. Successful Attack Costs

The cost of a successful cyber attack is important to the evaluation of the trade-off between human operators and autonomous systems being used to conduct recovery actions. If the cost of a successful cyber attack is very high, this may indicate that a more expensive but more likely to succeed recovery action is justified. However, if the cost of a successful cyber attack is low, a less reliable recovery action may be more cost effective.

Measuring the cost of a successful cyber attack can consider but is not limited to: 1) the cost of repairing the system ($COST_{Repair}$, 2) the cost of system downtime ($COST_{Downtime}$), 3) the cost of remediation of any negative consequences of the system failure such as toxic chemical release, radiological release, damage to nearby infrastructure, etc.($COST_{Remediation}$), and 4) any other identifiable costs($COST_{OtherSucc}$). The authors recommend tracking each of these costs separately but using a total cost of successful attack ($COST_{SuccAtt}$) in the next step. Equation 3 shows how to calculate the cost of a successful attack. Table 3 shows the results of the analysis of the recovery actions for both human operators and automated systems.

$$COST_{SuccAtt} = COST_{Repair} + COST_{Downtime} + COST_{Remediation} + COST_{OtherSucc} \tag{3}$$

Recovery Action	Human / Automation	Probability of Recovery Success	Recovery Action Cost	Successful Attack Cost
Recovery 1	Human Operator	0.85	$0.5M/yr	$10M
	Automated System	0.78	$1.2M/yr	$10M
Recovery 2	Human Operator	0.52	$2.1M/yr	$12M
	Automated System	0.85	$2.7M/yr	$12M
Recovery 3	Human Operator	0.92	$0.75M/yr	$15M
	Automated System	0.95	$0.5M/yr	$15M
Recovery 4	Human Operator	0.7	$1.2M	$9M
	Automated System	0.82	$1.4M	$9M
Recovery 5	Human Operator	0.87	$0.6M	$17M
	Automated System	0.75	$0.3M	$17M

Table 3. Results of analysis of recovery actions performed by either human operators or automated systems. Costs are on a yearly basis.

2.5. Step 4: Recovery Decision-Making

With the information developed in the previous three steps, decisions can now be made between a human operator or an automated system being used to complete a recovery action. As was stated in Step 3, using a cost basis to make decisions can help to make decisions more easily. A risk number, where risk is defined as the probability of an event occurring multiplied by the outcome of the event, can be developed to represent both the case where a system is successfully recovered from a cyber attack ($R_{N-Success}$), shown in Equation 4, and the case where a system fails due to a cyber attack ($R_{N-Failure}$), shown in Equation 5 . It should be noted that the method presented in this paper does not assign a probability to cyber attacks either individually or as a category of initiating events. This is intentional and represents the current state of knowledge over the likelihood of cyber attacks in the future. Predicting zero-day exploits is notoriously difficult [41].

$$R_{N-Success} = COST_{Recovery} * P_{RS} \tag{4}$$

$$R_{N-Failure} = (COST_{Recovery} + COST_{SuccAtt}) * (1 - P_{RS}) \tag{5}$$

The risk associated with both outcomes ($R_{N-Rec-Outs}$) can then be added together to develop a risk of the choice of either having a human operator or an automated system attempt the recovery action, shown in Equation 6.

$$R_{N-Rec-Outs} = R_{N-Success} + R_{N-Failure} \tag{6}$$

A decision-maker can now analyze the risk numbers for each recovery action to determine if a human operator or an automated system is more appropriate. In the case where a the architecture of a system is still in flux, it may be useful to broadly evaluate if a system is better served by having recovery actions performed entirely by automated systems or human operators. To make this determination, the risk numbers of all human operator recovery actions are added together and the risk numbers of all automated system recovery actions are separately added together ($R_{N-Humans}$ and $R_{N-Automation}$), then the two summed risk numbers are compared to see which is lower – the lower risk number indicates that it is the better choice, as shown in Equations 7 and 8.

$$\sum R_{N-Rec-Outs_{Human}} = R_{N-Human} \tag{7}$$

$$\sum R_{N-Rec-Outs_{Automation}} = R_{N-Automation} \tag{8}$$

Table 5 shows the risk numbers for each recovery action. Decisions for each recovery action where the choice is either a human operator or an automated system are listed. The bottom of the table shows the summed risk numbers for all recovery actions being taken by human operators versus all recovery actions being taken by automated systems. If the architecture of the case study system were still in flux, the values of $R_{N-Humans}$ and $R_{N-Automation}$ indicate that using automation across all recovery actions is a lower overall risk.

Recovery Action	Hazard Scenario	Recovery Action Description
Recovery 1	Scenario 1	Pump Control Override to Restart Pumps
Recovery 2	Scenario 2	Install and Turn On Pump Between Drain Tank and Cooling Pool to Recycle Water
Recovery 3	Scenario 3	Install and Open Backup Pump Valves #1 and #3
Recovery 4	Scenario 4	Install and Turn On Pump Between External Water Source and Emergency Cooling Water Tank
Recovery 5	Scenario 5	Install and Turn On Fire Water Source to Replenish Cooling Pool

Table 4. Specific recovery actions that can be taken to stop incipient failures caused by cyber attacks. Recovery actions in italics font indicate that they were not previously identified by L'her et. al.'s method.

3. Discussion and Future Work

While other methods such as L'Her et. al. [25] can aid in designing a PHM system, the method presented here is specifically useful in determining if a human operator or an automated system should be used to perform a recovery action during a cyber attack-initiated incipient failure. As noted elsewhere in this paper, the probability of a cyber attack is not part of the assessment of the method due to the inability of a cyber attack to be assigned a probability, especially over significant lengths of time. This limitation may be corrected in the future if a satisfactory means of predicting cyber attack frequency is developed.

Recovery Action	$R_{N-Rec-Outs_{Human}}$	$R_{N-Rec-Outs_{Automation}}$	Decision
Recovery 1	2	3.4	Human
Recovery 2	7.86	4.5	Automation
Recovery 3	1.95	1.25	Automation
Recovery 4	3.9	3.02	Automation
Recovery 5	2.81	4.55	Human
			Summation
		$R_{N-Automation}$	16.72
		$R_{N-Human}$	18.52

Table 5. Summary table of risk numbers. Units are $probability * Cost(\$)/1E6$.

In some cases, such as when loss of life may be involved or other very bad outcomes of a successful cyber attack are possible, it may be useful to use a weighted method of decision-making. Van Bossuyt et. al. [42], Van Bossuyt et. al. [43], and Van Bossuyt et. al. [44] provide several related methods for weighting significant outcomes using utility theory and other techniques. This is especially important in situations where the entire cost of a successful cyber attack may not be fully calculable.

Uncertainty in the data sources may cause decisions between human and automated recovery actions to not be as clear-cut as the above case study shows. Uncertainty may come from probability of recovery success, recovery action cost, and successful attack cost. The authors suggest practitioners attempt to decrease uncertainty through refining the variables with the largest impact on uncertainty in the model until a clear decision has been identify.

While cyber attack risks to operating plans has received a great deal of attention and are the focus of this work, the entire systems engineering process is vulnerable. For instance, malicious code could be inserted into critical software repositories, configuration datasets could be changed, electronic records and logbooks could be altered, a plant's design baseline could be doctored, inventory control systems could be changed, regulatory compliance documents could be altered, and even system reference models during the design process could be tampered with. Although the method and case study presented below intentionally only addresses a small subset of potential cyber attacks during system operation, the authors caution practitioners to remember the entire lifecycle of the system may be attacked.

The case study presented above is of a small subsystem within a larger system. Analysis of a large system or a system of systems using the method presented in this paper may result in interesting conclusions such as human operators being preferred at a subsystem level but automation being preferred at a system or system of systems level. The authors recommend evaluating at the highest level practical.

One potential outcome of using this method may be to redesign a system to have a longer period of time available to complete a recovery action. This often will lead to an increase in the probability of a human operator successfully

completing recovery actions which may help to lower costs in cases where human operators are less expensive than automation systems. This may also allow for sufficient time for human operators to double check automation system performance to prevent a cyber attack from succeeding at causing the system to fail.

In the future, an extension of this method may be made to more closely analyze staffing levels within a system to provide a more nuanced cost basis for human operators. Currently the method assumes that one operator will only perform one recovery action. However, a single human operator may be in charge of multiple recovery actions or several human operators might be needed to accomplish one recovery action.

The method may be extensible to represent more complex recovery efforts where human operators and automated recovery systems work together. This may result in a wide range of potential choices between entirely human operator and entirely automated system-driven recovery. Other strategies such as human operators and automation systems backstopping each other may also provide fruitful for future research.

Another future expansion of the work is to develop an ability to better understand how cyber attacks may interfere with recovery actions. For instance, a cyber attack may disable digital display panels or may even spoof the digital display panels [14], thus making a human operator's job of successfully identifying an incipient failure and diagnosing the failure less likely. A cyber attack may simultaneously target both the control system and the automated recovery system which may both initiate an incipient failure and allow the incipient failure to result in a system-level failure due to the automated recovery system being rendered ineffective.

4. Conclusion

The method presented in this paper provides a means for assessing if a human operator or an automated system is more appropriate as the actor in a recovery action in a cyber attack scenario on a physical system. By being positioned for use during the conceptual phase of systems design when architectural trade-off studies are being conducted and the cost of making significant changes to a system are low, this method allows systems engineers to make informed recovery action decisions. Practitioners can use this method to make cost and probability-informed decisions that otherwise would not be possible during conceptual design.

Acknowledgements

This research is partially supported by the Naval Postgraduate School (NPS). Any opinions or findings of this work are the responsibility of the authors, and do not necessarily reflect the views of the sponsors or collaborators.

References

[1] E. Crawley, B. Cameron, D. Selva, System architecture: strategy and product development for complex systems, Prentice Hall Press, 2015.
[2] U. DoD, Mil-std-882e, department of defense standard practice system safety, US Department of Defense.
[3] J. Zamanali, Probabilistic-risk-assessment applications in the nuclear-power industry, IEEE transactions on reliability 47 (3) (1998) SP361–SP364.
[4] C. A. Ericson, Event tree analysis, Hazard Analysis Techniques for System Safety (2005) 223–234.
[5] E. J. Henley, H. Kumamoto, Probabilistic risk assessment: reliability engineering, design, and analysis, IEEE, 1981.
[6] U. N. R. Commission, et al., Regulatory Guide 1.174: An Approach for Using Probabilistic Risk Assessment in Risk-informed Decisions on Plant-specific Changes to the Licensing Basis, US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, 2002.
[7] M. Stamatelatos, H. Dezfuli, G. Apostolakis, C. Everline, S. Guarro, D. Mathias, A. Mosleh, T. Paulos, D. Riha, C. Smith, et al., Probabilistic risk assessment procedures guide for nasa managers and practitioners.
[8] Army, Failure modes, effects and criticality analysis (fmeca) for command, control, communications, computer, intelligence, surveillance, and reconnaissance (c4isr) facilities, United States Government (2006).
[9] L. M. Weston, D. W. Whitehead, N. L. Graves, Recovery actions in pra [probabilistic risk assessment] for the risk methods integration and evaluation program (rmiep): Volume 1, development of the data-based method, Tech. rep., Sandia National Labs. (1987).
[10] D. Gertman, H. Blackman, J. Marble, J. Byers, C. Smith, et al., The spar-h human reliability analysis method, US Nuclear Regulatory Commission.
[11] C. L. Smith, Calculating conditional core damage probabilities for nuclear power plant operations, Reliability Engineering & System Safety 59 (3) (1998) 299–307.
[12] S. Sridhar, A. Hahn, M. Govindarasu, et al., Cyber-physical system security for the electric power grid., Proceedings of the IEEE 100 (1) (2012) 210–224.

[13] O. Kosut, L. Jia, R. J. Thomas, L. Tong, Malicious data attacks on the smart grid, IEEE Transactions on Smart Grid 2 (4) (2011) 645–658.
[14] R. Langner, Stuxnet: Dissecting a cyberwarfare weapon, IEEE Security & Privacy 9 (3) (2011) 49–51.
[15] J. Polosky, D. Marksberry, C. Atwood, W. Galyean, Nureg/cr-5750 rates of initiating events of us nuclear power plants, US Nuclear Regulatory Commission, Washington DC (USA).
[16] B. M. O'Halloran, N. Papakonstantinou, D. L. Van Bossuyt, Assessing the consequence of cyber and physical malicious attacks in complex, cyber-physical systems during early system design, in: Proceedings of the IEEE International Conference on Industrial Informatics (INDIN), IEEE, 2018.
[17] C. Layne, Cyber attacks against critical infrastructure, Ph.D. thesis, Utica College (2017).
[18] J. E. Long, Relationships between common graphical representations used in system engineering, INSIGHT 21 (1) (2018) 8–11.
[19] R. B. Stone, K. L. Wood, Development of a functional basis for design, Journal of Mechanical design 122 (4) (2000) 359–370.
[20] E. Cohen, J. Bortman, R. Klein, Cyber defense of rotating machinery using an integrated 'fuse' bearing, in: Annual Conference of the Prognostics and Health Management Society, The Prognostics and Health Management Society, 2015.
[21] D. Wang, I. C. Gauld, G. L. Yoder, L. J. Ott, G. F. Flanagan, M. W. Francis, E. L. Popov, J. J. Carbajo, P. K. Jain, J. C. Wagner, et al., Study of fukushima daiichi nuclear power station unit 4 spent-fuel pool, Nuclear technology 180 (2) (2012) 205–215.
[22] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, D. Siegel, Prognostics and health management design for rotary machinery systemsreviews, methodology and applications, Mechanical systems and signal processing 42 (1-2) (2014) 314–334.
[23] M. R. Bohm, R. B. Stone, S. Szykman, Enhancing virtual product representations for advanced design repository systems, Journal of Computing and Information Science in Engineering 5 (4) (2005) 360–372.
[24] S. Szykman, R. D. Sriram, C. Bochenek, J. Racz, The nist design repository project, in: Advances in Soft Computing, Springer, 1999, pp. 5–19.
[25] G. LHer, D. L. Van Bossuyt, B. M. OHalloran, Prognostic systems representation in a function-based bayesian model during engineering design, International Journal of Prognostics and Health Management 8 (2) (2017) 23.
[26] T. Kurtoglu, M. I. Campbell, Automated synthesis of electromechanical design configurations from empirical analysis of function to form mapping, Journal of Engineering Design 20 (1) (2009) 83–104.
[27] T. Kurtoglu, I. Y. Tumer, A graph-based fault identification and propagation framework for functional design of complex systems, Journal of Mechanical Design 130 (5) (2008) 051401.
[28] T. Kurtoglu, I. Y. Tumer, D. C. Jensen, A functional failure reasoning methodology for evaluation of conceptual system architectures, Research in Engineering Design 21 (4) (2010) 209–234.
[29] D. L. Van Bossuyt, B. M. OHalloran, R. M. Arlitt, Irrational system behavior in a system of systems, in: 2018 13th Annual Conference on System of Systems Engineering (SoSE), IEEE, 2018, pp. 343–349.
[30] B. M. O'Halloran, N. Papakonstantinou, D. L. Van Bossuyt, Modeling of function failure propagation across uncoupled systems, in: Reliability and Maintainability Symposium (RAMS), 2015 Annual, IEEE, 2015, pp. 1–6.
[31] B. M. OHalloran, B. Haley, D. C. Jensen, R. Arlitt, I. Y. Tumer, R. B. Stone, The early implementation of failure modes into existing component model libraries, Research in Engineering Design 25 (3) (2014) 203–221.
[32] R. Arlitt, D. L. Van Bossuyt, R. B. Stone, I. Y. Tumer, The function-based design for sustainability method, Journal of Mechanical Design 139 (4) (2017) 041102.
[33] H. Kumamoto, E. J. Henley, Probablistic risk assessment and management for engineers and scientists, Wiley-IEEE, 2000.
[34] Y. Ting, S. Tosunoglu, D. Tesar, A control structure for fault-tolerant operation of robotic manipulators, in: Robotics and Automation, 1993. Proceedings., 1993 IEEE International Conference on, IEEE, 1993, pp. 684–690.
[35] K. L. Butler, N. Sarma, C. Whitcomb, H. Do Carmo, H. Zhang, Shipboard systems deploy automated protection, IEEE Computer Applications in Power 11 (2) (1998) 31–36.
[36] K. M. Groth, A. Mosleh, Deriving causal bayesian networks from human reliability analysis data: A methodology and example model, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 226 (4) (2012) 361–379.
[37] O. Svenson, Time pressure and stress in human judgment and decision making, Springer Science & Business Media, 1993.
[38] S. Cheung, U. Lindqvist, M. W. Fong, Modeling multistep cyber attacks for scenario recognition, in: DARPA information survivability conference and exposition, 2003. Proceedings, Vol. 1, IEEE, 2003, pp. 284–292.
[39] A. Chhokra, N. Mahadevan, A. Dubey, D. Balasubramanian, G. Karsai, Towards diagnosing cascading outages in cyber physical energy systems using temporal causal models, in: Annual Conference of the Prognostics and Health Management Society, The Prognostics and Health Management Society, 2017.
[40] J. O'Brien, Doe fundamentals handbook, nuclear physics and reactor theory, Tech. Rep. DOE-HDBK-1019/2-93, Department of Energy (1993).
[41] M. Bozorgi, L. K. Saul, S. Savage, G. M. Voelker, Beyond heuristics: learning to classify vulnerabilities and predict exploits, in: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2010, pp. 105–114.
[42] D. Van Bossuyt, C. Hoyle, I. Y. Tumer, A. Dong, Risk attitudes in risk-based design: considering risk attitude using utility theory in risk-based design, AI EDAM 26 (4) (2012) 393–406.
[43] D. L. Van Bossuyt, A. Dong, I. Y. Tumer, L. Carvalho, On measuring engineering risk attitudes, Journal of Mechanical Design 135 (12) (2013) 121001.
[44] D. L. Van Bossuyt, I. Y. Tumer, S. D. Wall, A case for trading risk in complex conceptual design trade studies, Research in Engineering Design 24 (3) (2013) 259–275.