SAND2018-6150C

# An Engineering Systems Framework for Analyzing Nuclear Security

Adam D. Williams[1]

[1]*Sandia National Laboratories*, Albuquerque, NM, USA, adwilli@sandia.gov*

Current approaches to nuclear security are best described by the International Atomic Energy Agency's guidance to develop 'risk-based physical protection systems and measures' intended to deter, detect, delay, respond and (if necessary) mitigate malicious acts regarding nuclear materials. These approaches can produce elegantly designed physical protections systems that may be limited by untenable assumptions (for mathematical tractability) or well stated descriptions of desired behaviors that rely on vague, imprecise notions of security-improving characteristics. More to the point, noted nuclear security culture expert Dr. Igor Khripunov noted a lack of guidance on "assessing the human factor in detection, delay and response."

But no one has yet figured out a way to understand specifically how organizational and human factors might influence PPS effectiveness. This conference paper summarizes recent research exploring and developing a framework that evaluates system-level interactions between the technical nuclear security systems and human/organizational behaviors to determine overall security performance. Technical systems encompasses both the PPS and physical infrastructure on which it sits and is described by the traditional system effectiveness measure. Similarly, human/ organizational behaviors include formal (e.g., official roles and responsibilities) and informal aspects (e.g., networks of information flow and internal power dynamics) that are manifested in security procedures and concepts of operations (CONOPs).

The Systems-Theoretic Framework for Security (STFS) uses security task completion to explain how human behavior is required to enact the technical system—and the technical system is necessary to guide human behavior—to achieve desired levels of security performance. This interaction is based on the logic that the adequate completion of security tasks, defined as performance specifications based on the PPS design, is required to achieve desired levels of security performance. STFS, then, argues that desired security performance is achieved when the PPS, human/organizational behaviors and their interactions support the validity of such performance requirements to enable adequate security task completion. Further, STFS aids in identifying where organizational influences on security task completion may be varying enough from what the technical system designers expected to undermine the assumptions on which they based their estimates of system performance.

## INTRODUCTION[1]

Take into account this intriguing description of a physical protection system (PPS) upgrade closeout visit to a Russian nuclear facility in the 1990s, related by a former nuclear security manager. After arriving in Russia, the U.S. team of security subject matter experts (SMEs) was informed that the date of the inspection fell on a newly established national holiday and that no one would be at the facility to host them. The inspection team did not consider this a problem, since they assumed that security personnel would be on site to protect the special nuclear material (SNM). Upon arriving at the site, the team found neither guards nor central alarm station (CAS) operators present, and the

[1] This conference paper is a summary of Chapter 1 in [1].

***SAND2018-XXXX.** Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525

PPS was turned off. The team eventually learned that the PPS was always turned off on weekends and holidays in order to meet power allocation goals from the local community. According to this former nuclear security manager, their Russian counterparts felt this was acceptable security behavior—ultimately putting the SNM at risk in order to meet an electricity use quota. Despite having the *right technology* in place, the interaction of technology with human operators resulted in reduced security at this facility.

Current approaches to nuclear security[2] are best described by the International Atomic Energy Agency's (IAEA) guidance to develop "risk-based physical protection systems and measures" intended to deter, detect, delay, respond and mitigate malicious acts on nuclear materials. To hinder any success of potential adversary actions, these approaches rely on *defense-in-depth* strategies. In addition, recent initiatives on nuclear security culture focus on staff vigilance and motivation to improve security. These approaches can produce elegantly designed technical solutions for physical protection that leave our human factors or descriptions of desired human and organizational behaviors that rely on vague, imprecise notions of how these behaviors connect to the use of technological systems and actual security performance.

Because these approaches struggle to account for the connection between technological systems and human and organizational behaviors, they do not reflect the real daily security performance at operating nuclear facilities. As indicated in the opening anecdote, security at nuclear facilities, like nuclear safety, is predicated on how well individual actions, global expectations and operational assumptions coexist with local cultural norms.[2] Ignoring these interdependencies can result in insufficient performance and suggests a need to better incorporate them into nuclear security analysis. More to the point, according to one nuclear security culture expert:

> While the International Atomic Energy Agency [IAEA] has released methodologies on evaluating vulnerabilities and physical protection, it has not yet introduced guidelines on *assessing the human factor in detection, delay, and response* the three main pillars of security. [3, p. 39-40] (Emphasis added)

In response, the Systems-Theoretic Framework for Security (STFS) uses security task completion to explain how human behavior is required to enact the technical system—and the technical system is necessary to guide human behavior—to achieve desired levels of security performance.

## CHALLENGES TO SECURITY AT NUCLEAR FACILITIES

Several publicized nuclear security incident—such as the 2007 armed attack at the Pelindaba Nuclear Research Center in South Africa [4] or the 2018 armed raid on the housing complex for workers of Brazil's Angra Nuclear Power Plant [5]—further demonstrate the importance of better understanding the role of human or organizational behaviors in meeting security performance requirements.

For example, in the summer of 2012 the Y-12 National Security Complex (Y-12)—commonly referred to as the *Fort Knox of Nuclear Security*—was breached by three elderly protestors. Despite the fact that no SNM was accessed—let alone stolen— the significance of this security breach is

[2] Consistent with recent trends, in this study *physical protection* refers to the collection of technologies arranged to meet performance requirements and *nuclear security* refers to all aspects on protecting nuclear materials and facilities.

multifaceted, prompting numerous calls for sweeping changes to the security protocols, structures and organizations for nuclear sites within the U.S. Various post-event reporting interpreted the events in different ways. Some argued that the security system *worked* since no nuclear material had been compromised, while others indicated that systemic and organizational issues made the site vulnerable.[6][7][8] This event illustrates how human and organizational behaviors can influence security performance at nuclear facilities—and challenge the efficacy and effectiveness of current approaches to nuclear security.

Such incidents illustrate two important trends. First, the challenges to securing nuclear materials and facilities are varied, continuing and evolving. Second, human and organizational behaviors (and their interactions with security technologies) are important for adequate security performance to mitigate these challenges. Despite the common understanding within the nuclear security profession that "good security is 20 percent equipment and 80 percent culture"[3] [9, p. 10], current approaches still struggle to capture human, social and organizational factors in nuclear security analysis.

## INCLUDING HUMAN/ORGANIZATIONAL BEHAVIORS IN NUCLEAR SECURITY

Traditional approaches to nuclear security systems make assumptions about how the PPS will be used in operation that ignore organizational behaviors and interactions with the PPS itself. PPS designers also tend to make decisions based on an envisioned set of individual work actions necessary to meet PPS performance goals. The Design Evaluation Process Outline (DEPO) is a world renowned security analysis technique that combines technical measures of detection, delay and response into a stochastic measure of overall security performance, specifically the security effectiveness measure.[10] DEPO describes security performance in terms of probabilistic influences—e.g., the probability that a particular sensor alarms when an adversary enters a prohibited area or the probability that the response force is able to intercept the adversary—on competing timelines of required adversary action to achieve a malicious act and the response force actions necessary to protect nuclear facilities and materials. Such common technology-based solutions assume that PPS operations will both align with the individual security actions envisioned by designer and remain the same over time. This analytical logic supports a view that security incidents result from independent causes (e.g., broken cameras) and can be resolved with increasing technological reliability that often discounts the role of human behaviors.

In response, *nuclear security culture*-based approaches try to provide insight into how the human and organizational behaviors influence security performance. By leveraging the common concept of organizational culture[4], the IAEA offers descriptive characteristics of individual behavior to "foster more effective nuclear security" and management systems that "prioritize security." Describing eight leadership behaviors (e.g., effective communication), five personnel behaviors (e.g., vigilance) and 17 management system attributes (e.g., training or self-assessment) also assumes that once these are established, the organizational influences will subsequently support adequate security performance. While not fundamentally inaccurate, this approach does not link these desired behaviors to actual change in security performance—almost assuming that improvement can occur independently of PPS. Neither of these current approaches accounts for the socio-technical interactions that influence individual security actions as suggested by recent events.

---

[3] This quote is attributed to former Department of Energy *security czar* and former commander of U.S. strategic forces General Eugene Habiger.

[4] For more, please see [11].

By ignoring these interactions and discounting organizational behaviors, security assessment is limited in its ability to reconcile daily security performance with facility operational requirements (e.g., the deleterious effect of turning off facility power at the Russian nuclear facility on security performance in the opening anecdote). And yet at the same time, facility security forces are expected to perform at high levels of vigilance even as nuclear facility organizations forget about the undermining effects of a "chilled work environment" characterized by retribution for identifying shortcomings [12] or the perception that security expenditures are drains on limited funds and lower priority than meeting operational needs [13]. As one expert puts it "every dollar that a facility manager spends on protection is a dollar *not* spent on revenue-generating production" [14]. More pointedly, a visit by Dr. Matthew Bunn[5] to a Russian nuclear institute in the mid-2000s illustrates the necessity for incorporating organizational influences on performance into security analysis, noting:

> … that inside the hallway leading to the vault where a substantial amount of weapons-grade nuclear material was stored, there were two portal monitors that personnel had to pass through, one after the other, an American machine and a Russian machine. When asked why, the site official conducting the tour said that the building next door made medical isotopes, and on Thursdays, when the chemical separations were done to get the desired isotopes from the remainder, so much radiation went up the stack that it set off the American-made portal monitor. So on Thursdays, they turned off the American-made monitor and relied on the less sensitive Russian one. Of course, every insider was aware of this practice, and would know to plan an attempted theft for a Thursday, making the existence of the American portal monitor largely pointless. [14, p.11]

In other words, there is still a need to better understand the relationship between human/organizational behaviors (e.g., not wanting high levels of false alarms from medical isotope production on Thursdays), PPS technology operations (e.g., turning off the American portal monitor) and security performance (e.g., reduced detection capability and increased opportunity for a malicious act by an insider adversary).

## A NEW APPROACH: THE SYSTEMS-THEORETIC FRAMEWORK FOR SECURTY

The objective of [1] is to improve the understanding of security performance at nuclear facilities, in terms of the quality with which detection, delay and response functions are achieved. The primary hypothesis is based on the idea that both technical (e.g., PPS) and non-technical (e.g., the organization with security authority and responsibility) affect security performance. This introduces the need to understand the dynamics by which these two components influence security performance. First—and not often accounted for in traditional approaches—real security operations suggest that individual actions during daily security work practices affect both the PPS and the security organization (e.g., interaction). Second, the current level of performance influences both the PPS and the security organization (e.g., feedback). This suggests that such daily work practices can help explain these feedback and interaction dynamics. In theory, such individual security actions are affected by security procedures. Stated another way, security procedures outline

[5] Former nuclear security advisor to the Office of Science & Technology Policy; current Professor of Practice at Harvard University's Kennedy School of Government and co-Principal Investigator for the Project on Managing the Atom.

expected security-related behaviors of individuals. Yet, individual security actions are also influenced by the operational environment, which includes both the security technology available and organizational factors that affect implementation of security responsibilities. The technological availability is related to whether or not the PPS makes it easy or difficult for individuals to complete their related security actions. Security performance then results from how well individual security actions achieve security functions with a given PPS design, influenced by the security organization and current levels of performance.

To that end, the System-Theoretic Framework for Security (the STFS) for evaluating these system-level interactions between PPS and human/organizational behaviors to describe overall security performance (discussed in more detail in Chapter 5). the STFS is consistent with the tenets of engineering systems [15] and other research exploring security as an emergent property of complex systems [16][17][18][19]. Figure 1 illustrates security performance as driven by individual security actions—in terms of PPS design, security task completion expectations and the security organization interactions—using a set of eleven links (the numbered arrows) and four feedback loops (denoted with Roman numerals). These dynamics (and feedback loops) help explain how continuous improvement can create a virtuous cycle of vigilance for strong security performance or how repetitive small, negative shifts in security behaviors can lead to increasingly large drifts from desired states of performance, e.g., [20][21][22].[6] Ultimately, the STFS uses a security task completion construct to explain how the interactions between PPS and human/organizational behaviors result in security performance as an emergent property of nuclear facilities.

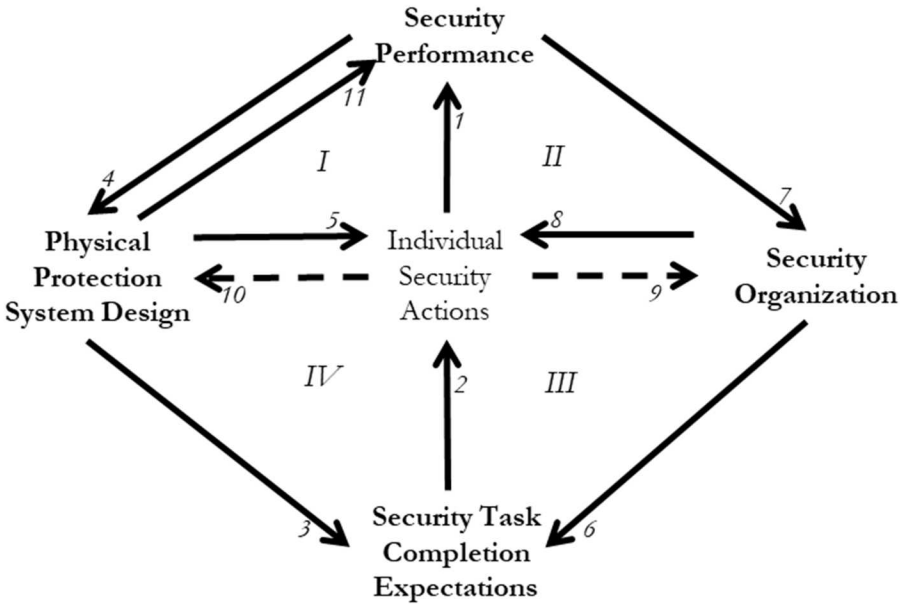


**Figure 1. Overview of the Systems-Theoretic Framework for Security**
***Solid arrows [⟶] indicate existing influences and dashed arrows [- - -⟶] indicate influences unrecognized by current state-of-the-art approacnes.***

[6] The conceptual similarities between security and safety as emergent properties [16] suggest this concept of "drift from safety" is also relevant to security. It is also important to note the major distinction between safety and security—namely that the former is founded on protecting against accidents that happen by random occurrences and the latter is founded on the need to protect against the conscious, deliberate actions of an intelligent adversary looking for ways to defeat your defense measures.

This framework is based on the logic that the high-quality completion of security tasks envisioned in the PPS design is necessary—though in most cases not always sufficient—to accomplish high level security functions. Achieving desired levels of security performance requires understanding the influences and dynamics affecting security task completion. *Security task* is conceptually similar to other constructs for describing interactions influencing desired outcomes. For example, control actions guide systems away from hazardous states in System-Theoretic Accident Model and Process (STAMP)-based approaches and hedging/shaping actions maintain the validity of expectations of future operational contexts in assumption-based planning.[23] More specifically, the security task completion construct is predicated on the validity of three behavioral performance requirements: (1) *the required task is identified and assigned*, (2) *the standard for task completion is met*, and (3) *meeting these standards of task completion is sufficient to achieve primary PPS security functions*. In other words, the influences that invalidate any of these behavioral performance requirements can result in inadequate security performance. The STFS argues that influences, dynamics and interactions that reinforce these three behavioral performance requirements contribute to security performance.

This approach aids in identifying where organizational influences on security task completion may vary enough from what PPS designers expected to undermine the assumptions on which they based their estimates of system performance. Based on this security task completion construct, I argue that the STFS offers a framework for assessing indicators of organizational influences likely to affect the validity of these behavioral performance requirements (and assumptions of human behavior made in DEPO) and, therefore, influence the resultant security performance at nuclear facilities. For example, the STFS argues that there is a difference in security performance between Facility A where internal assessments of security performance meet requirements and a strong PPS preventive maintenance program exists versus Facility B where internal assessments report requirements are met, but almost daily maintenance is required for portions of the PPS to be operational. There is a higher expectation for the PPS to sufficiently protect the nuclear facility in the former than in the latter, despite the similarly reported internal assessments of adequate PPS performance. As a more concrete example, in the anecdote that opened this chapter, the pressure to appease local government authorities challenged the assumption that *the security task is identified and assigned* for powering and operating the PPS 24 hours a day/seven days a week.

Supporting empirical studies in [1] helped identify a set of ten (10) such organizational influences and eight (8) quality indicators related to nuclear security performance. Such organizational influences includes *providing communication and feedback channels* and *supporting an attitude of self-responsibility*, while the quality indicators include *frequency of silencing alarms as registered in central alarm station (CAS) logs*. Security assessors can then use data-derived influences and indicators to probe the validity of the STFS' behavioral performance requirements—thereby identifying whether some of the performance assumptions that informed the estimate of PPS effectiveness may be unjustified. If so, either expected security performance needs to be readjusted or the identified destructive organizational weaknesses need to be addressed.

For example, consider the need to keep *closed* an emergency exit in a building storing SNM during non-emergency operations. By design, this supports both the probability of detection and delay associated with the PPS achieving its expected level of performance. Yet, at a given nuclear facility this door is routinely propped open by security officers for quick access to smoke breaks[7]—thus directly subverting the estimated ability of the PPS to detect and delay adversary action. The

[7] This is a scenario routinely used by the IAEA during security workshops and training programs.

*frequency of silencing alarms as registered in central alarm station (CAS) logs* quality indicator provides evidence to the security assessor of the degree to which this emergency exit door is appropriately kept closed.[8] Similarly, the *providing communication and feedback channels* and *supporting an attitude of self-responsibility* organizational influences help point the security assessor to dynamics challenging the validity of the behavioral performance requirements that resulted in undesired security performance. Moreover, the presence of a high *frequency of silencing alarms as registered in central alarm station (CAS) logs* would more likely have informed security management of this ongoing practice that introduced a new challenge to security performance at the facility—leveraging the organizational influences to support, rather than challenge, desired levels of security performance.

This framework can also help designers and assessors of nuclear security performance identify areas of improvement not captured by traditional approaches to nuclear security system design and assessment. From an evaluation perspective, the STFS provides assessors with a framework to investigate organizational influences affecting inadequate security task completion—based upon reports of undesired security performance—from either longer-term processes (e.g., formal oversight) or shorter-term observations (e.g., daily security practices). Likewise, the STFS offers a framework for incorporating quality indicators and organizational influences into both formal monitoring/inspection protocols and informal behavioral observation regimes. The underlying logic of the STFS also suggests it can help PPS analysts challenge assumptions on human behavior at the PPS design stage. This could provide an opportunity for a PPS design robustness evaluation by varying the level of security function accomplished (e.g., detection) based on indicators of organizational influences. The STFS could also aid in identifying potential interventions for improving security performance by redirecting oppositional organizational influences to support security task completion.

## CONTRIBUTIONS OF THE STFS TO IMPROVING NUCLEAR SECURITY

This dissertation supplements existing approaches to analysis of nuclear security. Traditional engineering design approaches nor nuclear security such as DEPO assume that human performance is fixed and not affected by the organizational context in which the PPS is expected to operate. Another set of approaches focus on security culture and emphasize steps to strengthen the organization's focus on security, but do little to identify how that strengthened focus might connect to improved security performance. The System-Theoretic Framework for Security (STFS) helps designers, analysts, and inspectors to understand how the technology and organization responsible for nuclear security interact—and how those interactions contribute to stronger or weaker security performance. In particular, this approach focuses on *security task completion* to explain how security technologies and organizations interact as a complex system. This provides the opportunity to explore how these interactions affect whether the human elements of this complex system in fact accomplish the security tasks with the level of quality envisioned by PPS designers. This approach offers an additional set of questions (or categories of questions) that security system designers, analysts, and inspectors can explore to deepen their understanding and generate more effective interventions.

[8] International best practice guides for security recommend alarming such emergency exits as an additional measure of detection. As pointed out by Dr. Bunn, this *assumption* is not always true in implementation in nuclear facilities around the world.

These empirical studies in [1] reveal a number of broad categories of issues that are often not considered in traditional DEPO assessments. Some of these issues (though not all) are considered in current nuclear security culture approaches, but without clear connections to security performance. This framework neither completely articulate nor is comprehensive in addressing *every* aspect of achieving high security performance. Rather, the STFS describes how the technology of the security system and the characteristics of the security organization interact to affect individual—and collective—security behaviors, which in turn contribute to security performance at nuclear facilities.

Similar to the arguments of the National Academy of Science's 2011 Committee on Risk-Based Approaches for Securing the DOE Nuclear Weapons Complex, observations captured in this paper describe how traditional nuclear security approaches may be limited by untenable design assumptions or vague, imprecise notions of security culture characteristics. In response, this NAS Committee noted that:

> The NNSA should adopt what the committee terms a 'total systems approach' to characterize the interactions and dependencies of security countermeasures at its facilities. Such an approach is commonly used as an initial step in assessing the risks associated with a complex technological system [24, p.1]

The STFS could provide a starting point for exploring analytical options to meet this total *systems approach* standard that both incorporates the traditional perspectives on security and overcomes their respective limiting factors. The STFS's description of security as resulting from patterns of practice and feedback loop dynamics aid in identifying potential interventions for improving security performance by redirecting oppositional organizational influences—and undesirable socio-technical interactions—to support security task completion. Taken together, the conclusions of this thesis position the STFS as a viable option for improving nuclear security by describing how these interactions affect whether the human elements of this complex system in fact accomplish the security tasks with the level of quality envisioned by PPS designers. By invoking systems theory, organization science and an engineering systems approach, this framework redefines security as a complex system property resulting from the alignment of technologies, policies, expectations, design requirements, daily work practices and their interactions. While this framework does not address *every* aspect of achieving high security performance, the STFS offers a structured thought process for how security technologies and organizations interact to affect security behaviors that contribute to security at nuclear facilities.

## REFERENCES


[1] Williams, A. (2018). Beyond Gates, Guards & Guns: The Systems-Theoretic Framework for Security at Nuclear Facilities. *PhD Dissertation.* Massachussets Institute of Technology.

[2] Meshkati, N. (1995). Cultural Context of the Safety Culture: A Conceptual Model and Experimental Study. *Conference Proceedings of the International Topical Meeting on Safety Culture in Nuclear Installations (INIS-MF-14747)*, (pp. 261-270). Vienna.

[3] Khripunov, I. (2014). *The Human Dimension of Security for Radioactive Sources: From Awareness to Culture.* Athens, GA: Center for International Trade & Security, the University of Georgia.

[4] Birch, D., & Smith, R. (2015). Nuclear Waste: The Assault on Pelindaba. *The Center for Public Integrity*. The Center for Public Integrity.

[5] Phillips, D. (2018). Armed raid on nuclear workers' housing raises fears over Brazil's two reactors. *The Guardian (U.S. Edition)*. London, United Kingdom. <https://www.theguardian.com/world/2018/jan/12/brazil-nuclear-reactor-armed>
[6] Schlosser, E. (2015). Break-In At Y-12. *The New Yorker*, 46-69.
[7] Zak, D. (2013). The Prophets of Oak Ridge. *The Washington Post*.
[8] Zak, D. (2014). 'The Prophets of Oak Ridge' and Facility Guard Adjust to Fallout of Nuclear Break-In. *The Washington Post*.
[9] Bunn, M., & Sagan, S. (2014). *A worst Practices Guide to Insider Threats: Lessons from Past Mistakes.* Cambrdge, MA: American Academy of Arts & Sciences.
[10] Garcia, M.L. (2008). *The Design and Evaluation of Physical Protection Systems (2nd Ed.).* Butterworth-Heineman.
[11] Schein, E. (1990). "Organizational Culture." *American Psychologist* 45 (2): 109-119.
[12] Zipp, Y. (2014). 'Chilled work environment' found in Palisades Nuclear Plant's security department, NRC says. *Kalamazoo Gazette (via MLIve Media Group).* Kalamazoo, Michigan. <http://www.mlive.com/news/kalamazoo/index.ssf/2014/03/nuclear_regulatory_commission_5.html>
[13] Stockton, P. (2002). *Nuclear Power Plant Security: Voices from Inside the Fence.* Project on Government Oversight. Retrieved March 24, 2016, from http://www.pogo.org/our-work/reports/2002/nss-npp-20020912.html
[14] Bunn, M. (2005). Incentives for Nuclear Security. *Proceedings of the 46th Annual Meeting of the Institute of Nuclear Materials Managment (INMM).* Phoenix, AZ: INMM.
[15] de Weck, O., Roos, D., & Magee, C. (2011). *Engineering Systems: Meeting Needs in a Complex Technical World.* Cambridge, MA: The MIT Press.
[16] Young, W. E. (2015). A System-Theoretic Security Analysis Methodology for Assuring Complex Operations Aginst Cyber Disruptions. *PhD Dissertation*. Massachusetts Institute of Technology.
[17] Williams, A. D. (2015). Beyond a Series of Security Nets: Applying STAMP & STPA to Port Security. *Journal of Transportation Security, 8*(3-4), 139-157.
[18] Williams, A. D., Osborn, D., Homan, R., Jones, K. A., Kalinina, E. A., Parks, M. J., . . . Cohn, B. (2016). A New Look at Transportation Security: A Complex Risk Mitigation Framework for the Security of International Spent Nuclear Fuel Transportation. *International Conference on Nuclear Security: Commitments and Actions.* Vienna, Austria: International Atomic Energy Agency.
[19] Electric Power Research Institute. (2015). *Program on Technology Innovation: Cyber Hazards Analysis Risk Methodology (Phase II: A Risk Informed Approach | 3002004997).* Palo Alto, CA: EPRI.
[20] Dekker, S. (2011). *Drift into Failure: From Hunting Broken Components to Understanding Complex Systems.* Farnham, England: Ashgate Publishing Limited.
[21] Leveson, N. (2012). *Engineering a Safer World: Systems Thinking Applied to Safety.* Cambridge, MA: MIT Press.
[22] Rasmussen, J. (1997). Risk Management in a Dynamic Society: A Modelling Problem. *Safety Science, 27*(2/3), 183-213.
[23] Dewar, J. (2002). *Assumption-Based Planning: A Tool for Reducing Avoidable Surprise.* Cambridge University Press.
[24] Committee on Risk-Based Approaches for Securing the DOE Nuclear Weapons Complex. (2011). Understanding and Managing Risk in Security Systems for the DOE Nuclear Weapons Complex (Abbreviated Version). Washington, DC: The National Academies Press.