ASSESSING TRENDS IN NUCLEAR SECURITY THAT IMPACT THE PREVAILING SITUATION: NON-NUCLEAR EMERGING TECHNOLOGIES – CYBER AND ARTIFICIAL INTELLIGENCE

Jor-Shan Choi[*]

Berkeley Nuclear Research Center (affiliate)
Lawrence Livermore National Laboratory (retired)

February 2020

I INTRODUCTION

Critical safety, security, and emergency-preparedness systems (SSEP) at nuclear facilities, including nuclear power plants (NPPs), are susceptible to cyberattacks if the nuclear facilities use digital systems to obtain and store vital information, control and account for nuclear materials, or monitor and operate safety equipment. The damage from cyberattacks, initiated externally or aided by insiders, can range from loss of confidential data and sensitive information to theft of nuclear or radioactive materials to a radiological release.

The entry into force of the amended Convention on the Physical Protection of Nuclear Materials (CPPNM/A) in May 2016 ushered in a new nuclear security regime in which states parties to the CPPNM/A are obligated to maintain a physical protection regime that protects all nuclear materials from theft and nuclear facilities from sabotage.

The threats against which states parties must protect their nuclear materials and facilities, however, have evolved significantly since 2005, when the CPPNM/A was originally adopted. In 2021, when states parties will convene to review the implementation and adequacy of the convention "in the light of the then prevailing situation," they will need to address how emerging technology has impacted nuclear security and how it may impact it in the future. Now, more than ever before, stakeholders within each state party, such as its government, regulator, and operators of nuclear facilities, must recognize that for nuclear security to continue to be effective, they must apply a broader definition of physical protection of nuclear facilities and materials that includes cyber protection. This recognition must occur before a cyber-mediated theft of nuclear materials or sabotage of a nuclear facility leads to catastrophic results.

---

[*] Dr. Jor-Shan Choi, an affiliate of Berkeley Nuclear Research Center at UC Berkeley, retired from Lawrence Livermore National Laboratory (LLNL) in 2008 after 21 years of service. He was a project professor at the University of Tokyo from 2008 to 2011 and a visiting professor at Tokyo Institute of Technology in 2012. Dr. Choi worked at the International Atomic Energy Agency in Vienna from 1998-2001 and was a Science Fellow at the Center for International Security and Cooperation at Stanford in 1995. Prior to joining LLNL in 1987, he had 13 years of industrial experience working for Bechtel Corporation. Dr. Choi has a BS in Electrical Engineering and Computer Science, a MS and a PhD in Nuclear Engineering, all from UC Berkeley. He is also a Professional Engineer in the State of California. Opinion expressed here are those of the author, not necessary of their affiliations.

The need for protection of computer-based systems (including instrumentation and control (I&C) systems) is established in the International Atomic Energy Agency's (IAEA) Nuclear Security Recommendations on Physical Protection of Nuclear Material and Nuclear Facilities (INFCIRC/225/Revision 5), paras. 4.10 & 5.19, which both state that: "Computer-based systems used for physical protection, nuclear safety and nuclear material accountancy and control should be protected against compromise (e.g., cyberattack, manipulation, or falsification) consistent with the threat assessment or design basis threat."

The IAEA has identified three categories of cyber threats that could pose serious problems for civilian nuclear facilities, including NPPs: (1) compromise of safeguards or security systems resulting in unauthorized removal of nuclear material; (2) sabotage of a nuclear facility resulting in physical damage and/or radiological release; and (3) espionage, resulting in the exfiltration and exploitation of sensitive nuclear information.[1] All three cyber threats are feasible by external hackers, although the impacts of these threats, especially for (1) and (3), would be much more devastating if insiders were involved.

When considering the development of a design basis threat, due attention must be paid to insider threats. In the case of cyber sabotage, external hacker(s) with or without the aid of overt insider(s) could attack the digital I&C systems in a facility, physically damaging hardware devices (e.g., valves, pumps, generators, safety-interlocks, etc.) which could result in fuel or core damage and radiological release. In the case of cyber compromise or espionage, covert insider(s) could manipulate the material accountancy data at a nuclear material processing plant with the aim of diverting or stealing nuclear materials or exfiltrate sensitive information, such as blueprints of facilities or physical protection plans, to aid in sabotage or theft of nuclear materials. Insider(s) can take advantage of their access privileges, complemented by their authority and knowledge, to bypass physical protection elements.

With the advance in digital, cyber, and artificial intelligence (AI, or machine learning (ML)) technologies, I&C systems at many nuclear facilities have been upgraded to digital systems. Even the physical protection systems for nuclear materials and facilities have evolved from traditional "guns, guards, and gates" as the primary form of detection, delay, and response to include digital equivalents. Replacing legacy analog I&C systems, including those controlling the physical protection of the facility, with digital systems helps facility operators overcome obsolescence issues and enhance operational efficiency, availability, and performance. However, the application of digital technologies in protection and control systems has made them vulnerable to cyberattacks. Similarly, more and more information is being digitized and stored and transmitted on digital platforms. Doing so reduces costs and is more efficient, but increases the risk of loss of sensitive security information. For instance, the IAEA stores much of its safeguards inspection data digitally, transmits surveillance data via virtual private network,

[1] V. Boulanin and T. Ogilvie-White, "Cyber Threats and Nuclear Dangers," APLN-CNND Policy Brief No 17 (November 2014), https://cnnd.crawford.anu.edu.au/sites/default/files/publication/cnnd_crawford_anu_edu_au/2014-11/policy_brief_no_17_-_cyber_threats_and_nuclear_dangers.pdf

and scans and stores old confidential data. Such sensitive data about nuclear facilities would be valuable to would-be attackers if cyber attackers could access it.

The growing use of digital technologies and information operations will increasingly invite additional cyber intrusions and threats. As high-end cyber threat activity continues to become more sophisticated and AI/ML tools become easier to use, the level of expertise required by hackers is decreasing. These emerging trends in cyber technology impact the "prevailing situation" in which states must implement and review their CPPNM/A obligations. Given the rapid pace of change in the cyber age, it is difficult for states to predict how future advances in cyber will impact nuclear security, both positively and negatively, and therefore impossible for them not to regularly engage in dialogue with one another on how to adjust their nuclear security practices to new realities, including the implementation of the CPPNM/A.

II CYBER SABOTAGE

Threat

Over 20 cyber incidents, some accidental and some deliberate, have occurred at nuclear facilities, including NPPs, around the world since 1990.[2] The most recent theft of data from an administration network occurred in India's largest NPP in November 2019.[3] These incidents demonstrate that even NPPs are vulnerable to untargeted malware and targeted cyberattack. Despite the nuclear industry's warning that cyberattacks could cause massive physical damage and loss of life, only two cyberattacks are known to have significantly disrupted nuclear facility operations. These are the SLAMMER worm, which disabled the control room safety parameter display system at Davis Besse NPP in 2003, and which blocked plant operators' access to reactor core information,[4] and the STUXNET attack on an Iranian fuel enrichment plant in Natanz in 2009/2010, which physically destroyed around 1,000 centrifuges.[5]

The U.S. nuclear industry has experienced several cyber anomalies severe enough to cause plant emergencies and reactor shutdowns. One occurred at the Browns Ferry NPP in 2006 and the other at the Hatch NPP in 2008.[6] At Browns Ferry, both the plant's condensate demineralizers and recirculation pumps have digital equipment and embedded microprocessors that communicate data over the Ethernet Local Area Network. The Browns Ferry control network produced more traffic than the digital

---

[2] A. Van Dine, M. Assante, P. Stoutland, "Outpacing Cyber Threats – Priorities for Cyber Security at Nuclear Facilities," NTI, December 2016.

[3] J. M. Porup, "How a Nuclear Plant Got Hacked – Plugging Nuclear Plants into the Internet Makes them Vulnerable Targets for Nation-State Attacks," IT World, December 9, 2019.

[4] C. Baylon, R. Brunt, & D. Livingstone, " Cyber Security at Civil Nuclear Facilities – Understanding the Risks," Chatham House, September 2015.

[5] R. Langner, "To Kill a Centrifuge," 2013.https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf

[6] J. R. Thomson, "Nuclear Power Plant Cybersecurity Incidents," 2012. *https://*www.safetyinengineering.com/FileUploads/Nuclearcybersecurityincidents_1349551766_2.pdf

equipment could handle (or the equipment malfunctioned and flooded the Ethernet with spurious traffic). This disabled the variable frequency drive controllers and caused the unit 3 reactor to shut down. At Hatch, an engineer updated software for a business-network computer to synchronize diagnostic data collected from the process control network. While rebooting the computer, the synchronization program reset the data on the process control network, which interpreted the change as a sudden drop in the reactor's cooling-water reservoirs and initiated a reactor shutdown.

Although these events are not believed to have been deliberate attacks on the digital systems supporting critical NPP operations, they illustrate some of the different types of disruptive effects that could be deliberately engineered by malicious actors. These events inadvertently reinforced the U.S. nuclear industry's false confidence that although cyberattacks at NPPs could disrupt power generation, they could not cause devastating core damage or radiological releases because safety mechanisms would shut down the reactor first. However, from a cyber security perspective, a deliberate denial-of-service attack against Browns Ferry could have had serious safety consequences if it was part of a coordinated campaign that included other attacks that prevented an automatic reactor shutdown. Similarly, malicious software deliberately embedded in network systems at Hatch could have compromised its safe operation if the plant operators did not understand the interdependence of the network configurations or recognize the safety implications of a software update to plant equipment.

In each of these cases, examining the specific information technology (IT) systems involved reveals vulnerabilities created by reliance on digital technologies without adequate measures to prevent or mitigate cyberattacks. This suggests that vulnerability assessments for nuclear facilities' protection and control systems should incorporate steps taken to strengthen each digital and cyber component against cyberattacks[7]

Defense

All digital and microprocessor systems are potentially vulnerable to cyberattack. Whether or not those vulnerabilities could be leveraged to disrupt operations via a specific attack scenario depends on whether appropriate defensive measures have been taken. The next two examples provide some lessons about available defenses.

In the event at the Davis-Besse NPP in 2003, the SLAMMER worm infected 75,000 computer servers worldwide. The staff at Davis Besse had not addressed the vulnerability that SLAMMER worm exploited because they didn't know about the patch that Microsoft had released six months earlier.[8] The SLAMMER worm traveled from a consultant's computer to the corporate network by a privilege access bridging over the firewall. It then traveled to the plant process control network. The traffic generated by the worm clogged the corporate and control networks and crashed a plant process computer. Plant personnel could not access the safety parameter display system for 4 hours and 50 minutes. Losing the safety parameter display system could have been very serious because operators depended on it to

[7] J. Choi, N. Gallagher, C. Harry, "An effect-centric approach to assessing the risks of cyberattacks against the digital instrumentation and control systems at nuclear power plants," paper presented at the ICONS2020 in Vienna, February 2020.

[8] K. Poulsen, "SLAMMER worm crashed Ohio nuke plant network," Security Focus, 2003-08-19.

actively adjust plant operations. Luckily, there was an analog backup readout printer providing the safety parameters of the plant at the time.

In the attack on the Natanz plant, STUXNET had targeted the Siemens Step-7 programmable logic controllers, which controlled cascades of centrifuges, in two separate attacks. During the second attack in late 2009, the hackers took over the centrifuge speed controls and repeatedly ramped the speeds of some centrifuges rapidly from 0.2% to 130% of normal speed. They also altered the speed control readings in the control room display such that the attacked centrifuges' speed appeared to be normal. Over a six-month period, STUXNET destroyed some 1,000 centrifuges in Natanz.[9] The postmortem analysis by some process experts shows that the Natanz plant could have protected against this type of attack by installing the centrifuge rotors with a motor-over-speed-trip or physically hardening the rotors with more advanced materials.[10]

Patches, air gaps, and other IT-based cybersecurity techniques can make it harder for outsider(s) to gain access to critical digital protection and control systems at nuclear facilities, but they cannot protect against insider threats or certain other types of cyberattacks. In addition to the IT methods, a security process hazard analysis should be performed to identify potential vulnerabilities created by the digital systems and find appropriate defense mechanisms. At least four non-IT methods can be used to increase defensive robustness.[11]

- Provide robust administrative controls that protect against cyberattacks. This may be the weakest protection because it depends on people faithfully following the administrative requirements, and people are prone to make mistakes.

- Replace the problematic digital systems or components with analog devices, or provide redundant analog systems for the same function.

- Insert mechanical systems in place of certain digital components, or limit the range over which the digital system can control the problematic function.[12]

- Design or change the process or equipment such that the system's physics prevents hazardous consequences. This may be the strongest protection against cyberattack, but it may also be the most difficult to implement, especially for existing plants.

[9] D. Albright, P. Brannan, & C. Walrond, "Stuxnet malware and Natanz: Update of ISIS December 22, 2010 report," February 2011.
[10] E. Marszal and J. McGlone, "Security PHA Review – for Consequence-based Cybersecurity," International Society of Automation (ISA), 2019.
[11] G. Johnson, "Cyber Robust Systems – The vulnerability of the current approach to cyber security," June 2019.
[12] G. Falco and H. Lin, et. al., "Cyber Risk Research Impeded by Disciplinary Barriers," Science, Vol.366, p.1066-1069, November 29, 2019.

III CYBER COMPROMISE AND ESPIONAGE

Threat

Nuclear security experts are also concerned about cyber compromises and espionage. The threat posed by insiders is of particular concern in many of these scenarios.

A hypothetical scenario might involve an insider deliberately manipulating inventory data to facilitate an undetected diversion and theft of nuclear material from a nuclear material processing facility (e.g., a MOX fuel fabrication plant). The insider could accomplish this by altering the accountancy record and inventory data to hide the removal of nuclear material, then manipulating the criticality safety control limit set for a workstation to move the diverted item into a solid waste station, placing the item in a contact waste container and moving the container to a holding area outside the vital processing area, and finally, smuggling the diverted nuclear material outside the plant. The diversion and theft would remain undetected for weeks due to the cyber manipulation of records until plant operators conduct an inventory check for the entire mass balance area.

Cyber espionage is now commonplace in a variety of sectors and international agencies have already been targeted. It was revealed recently in an internal confidential document from the United Nations (UN) that sophisticated hackers infiltrated UN offices in Geneva and Vienna in 2019 in an apparent espionage operation.[13] The document showed that the hackers had deployed malware to machines that were linked to specific purposes for the hackers. The document also showed that among the accounts known to have been hacked were those of domain administrators who by default had master access to all user accounts in their purview. Reports noted that a flaw in Microsoft's SharePoint software was exploited by the hackers to infiltrate the networks, but that the type of malware used was not known and technicians had not identified the command and control servers on the internet used to exfiltrate information. The UN indicated that the attack resulted in a compromise of core infrastructure components and was determined to be serious.

It is unclear how the attack on the UN was perpetrated, but the possibility that insiders might participate in cyber intrusion–intentionally (e.g., espionage) or inadvertently–vastly complicates an already very challenging cyber threat to organizations like the IAEA that hold confidential data about nuclear facilities that, if stolen, could prove useful to would-be attackers.

In the context of cyber, an insider, as defined by the Carnegie Mellon University Software Engineering Institute Community Emergency Response Team Program,[14] is a current or former employee, contractor, or business partner who:

- Has or had authorized access to an organization's network, system, or data,

[13] News article, "Leaked report shows United Nations suffered 'sophisticated' hack," Associated Press January 30, 2020

[14] The CERT Insider Threat Center, Common Sense Guide to Mitigating Insider Threats, Fifth Edition, Software Engineering Institute, Carnegie Mellon University, December 2016.

- Has intentionally exceeded or used that access in a manner that negatively affected the confidentiality, integrity, availability, or physical well-being of the organization's information or information systems or workforce,

- Has colluded with outsiders, including organized crime groups, and foreign organizations or governments.

Most cyberattacks involve some kind of privileged access, including stolen privileged access. For instance, attackers that gain entry to an enterprise (e.g., computer systems, networks, and/or control systems) can compromise the identity of an employee and then use that employee's permissions to plan vectors of attack. This is what so many cyberattacks involve: a bad person using a good person's identity, or a good person turning into a bad person. Both cases result in the same consequences.

Focusing on insider threats is important because it has been documented extensively that many people leave companies and try to take company intellectual property with them. Employees can be bribed, have pressure put on their families, or become disgruntled because they feel the company hasn't treated them well. There are many different reasons why someone who was trustworthy when hired could become a risk to the company.

Defense

While stopping insider threats completely is virtually impossible, a proactive defense using AI or ML could provide an optimal and cost-effective answer to the challenges of insider compromise or espionage. ML, relying on big data for an identity analytic and a behavior analytic, could find the unknowns involved in the risks of identity access (either legitimate or fake), as well as the threat and concealment of intent based on behavior changes.[15] The ML software algorithms observe system access and usage, and estimate in real time whether there is an insider threat. Ironically, while the technology was intended to help defenders more rapidly identify and fix vulnerable systems, it is equally effective for adversarial use in finding and exploiting systems. Sophisticated malware is already using ML to detect when it is being monitored and to alter its behavior to escape detection.

A ML algorithm must be trained on large training data sets to be effective, and its effectiveness would depend on the quantity and quality of the data. Large training data sets may be available, but are often incomplete because people and organizations are influenced by liability and reputational concerns and withhold data about embarrassing cyber events. The relevance and integrity of the data set are additional factors affecting the quality of the data. While simulated data sets are convenient to generate, they may not properly encapsulate reality and the human dimension of adversarial actions.

[15] S. Nayyar, "Borderless Behavior Analytics – Who's Inside? What're They Doing?" P. 148, ISBN: 134:978-1986763332, 2018.

IV. CONCLUSIONS

Advances in digital, cyber, and artificial intelligence technologies have resulted in changes in nuclear facilities that could impact nuclear security. Digital upgrades of protection and control systems have many benefits, but also significant drawbacks when it comes to introducing new cyber vulnerabilities at nuclear facilities. These added vulnerabilities mean that the "prevailing situation" in which states parties must implement and review their CPPNM/A obligations has changed over the years, and will likely continue to change in the future as new technologies, threats, vulnerabilities and defenses emerge.

Emerging defense tactics against cyber threats indicate that some risks associated with cyberattacks on the digital protection and control systems at nuclear facilities, including NPPs, can be mitigated by technical, physical, or administrative measures, but nuclear facility operators will be unlikely to eliminate all cyber vulnerabilities due to constraints in costs, resources, and capabilities. To set priorities for protection against cyber threats, stakeholders such as governments, regulators, and operators will need a more systematic way to assess the consequences of cyber disruption scenarios involving IT systems that support important facility functions. It will also be beneficial for states, particularly at a regional level, to cooperate and pull resources and capabilities together in countering the emerging cyber threats.

The scale and complexity of the unknown attack space by malicious insiders are expected to grow to an extent far beyond human capability for detection. A proactive defense with AI or ML analytics and algorithms is needed to provide an optimal and cost-effective answer to the challenges of insider threats, cyber compromise, and cyber espionage. As states parties gather in 2021 to review implementation and adequacy of the CPPNM/A "in the light of then prevailing situation," they must acknowledge that developments in both the cyber threat and in the measures to mitigate them require in-depth discussions of how these developments impact their implementation of the convention and their thinking about the need to hold future review conferences.