

# On the Application of Causal Inference to Incident Response in Nuclear Facilities

The digitalization of Instrumentation and Control (I&C) systems in nuclear facilities introduces the potential for cyber-attacks to result in operational effects on critical systems. There have been several high-profile incidents that have demonstrated this potential, including the Stuxnet virus in 2010 [1], which targeted the nuclear sector, and the cyber-attacks to electrical distribution networks in the Ukraine in 2015 [2]. In these cases, and in others, a cyber-attack resulted in significant consequences for the physical process under control.

Threats of this type implement a so-called kill chain [3] –a series of activities that an adversary must complete to achieve their goal –that can be realized over several months and involves the compromise of numerous computer systems. It is possible that such an attack will be blended with a physical compromise of a target facility, e.g., via an insider. An adversary will attempt to remain stealthy (undetectable) until the point the attack on the physical system is executed. Therefore, indicators that a facility has been compromised via a cyber-attack could be weak and not readily attributable to its observable consequences. Consequently, a cyber-attack may exhibit the characteristics of a fault or another non-malicious root cause. This can make the execution of appropriate and targeted computer security incident response activities difficult.

To help detect the effects of threats that target the physical domain, there are several approaches that aim to identify anomalous behaviour that is being exhibited by a process [4]. In many cases, these approaches involve learning the normal behaviour of the process and detecting deviations from the learned norm. Abnormalities could be caused by a cyber-attack or other non-malicious challenges. To determine whether a detected anomaly has been caused by a cyber-attack (or otherwise), other systems that indicate the root cause of an anomaly can be used. For example, intrusion detection and anti-virus systems can be used to support a hypothesis that the root cause of a detected anomaly in a process behaviour is a cyber-attack.

For cyber security, combining this information is typically achieved by correlating indicators; however, this may yield misleading insights as correlated events are not necessarily causally related. An approach to addressing this issue is to use causal models [5][6] –which can capture expert knowledge –that describe the relationships between indicators of anomalous behaviour and the likelihood they have a certain root cause. Using such models, an operator can infer the likely root cause, e.g., expressed in terms of a system state, associated with detected anomalous behaviour, and use these inferences to guide an incident response.

In this paper, we present a novel application of an approach to causal inference to support incident response in nuclear facilities. This approach builds on anomaly detection algorithms that aim to detect deviations from the normal operation of nuclear processes and systems within a facility. The aim is to indicate and evaluate the utility of using causal models to infer the root causes of anomalous behaviour to support computer security incident response. This is done using scenarios that have been developed in the IAEA Coordinated Research Project (CRP) J02008 on incident response in nuclear facilities. In the project, a hypothetical Nuclear Power Plant (NPP), called Asherah, has been developed, which includes a simulation model of a Pressurized Water Reactor (PWR). Experimental results using the Asherah simulator, which has been coupled with representative computer systems, indicate how causal models can be used to determine the root cause of anomalous behaviour. We comment on how this capability can expedite incident response activities in nuclear facilities.

## REFERENCES

- [1] R. Langner, “To Kill a Centrifuge A Technical Analysis of What Stuxnet’s Creators Tried to Achieve,” November 2013, Available online: <https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf>
- [2] R. M. Lee, M. J. Assante, T. Conway, “Analysis of the Cyber Attack on the Ukrainian Power Grid,” March 2016, Available online: [https://ics.sans.org/media/E-ISAC\\_SANS\\_Ukraine\\_DUC\\_5.pdf](https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf)
- [3] M. J. Assante, R. M. Lee, “The Industrial Control System Cyber Kill Chain,” October 2015, Available online: <https://www.sans.org/reading-room/whitepapers/ICS/paper/36297>
- [4] A. L. Buczak, E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, October 2015. doi: 10.1109/COMST.2015.2494502
- [5] J. Pearl. “The Seven Tools of Causal Inference, with Reflections on Machine Learning,” Communications of the ACM vol. 62, no. 3, pp. 54-60, February 2019. doi: <https://doi.org/10.1145/3241036>
- [6] I. Friedberg, X. Hong, K. McLaughlin, P. Smith, P. C. Miller: “Evidential Network Modeling for Cyber-Physical System State Inference,” IEEE Access, vol. 5, pp. 17149-17164, 2017.

## **Gender**

## **State**

Austria

**Authors:** Dr PIATKOWSKA, Ewa (AIT Austrian Institute of Technology); Mr ALLISON, David (AIT Austrian Institute of Technology); SMITH, Paul (AIT Austrian Institute of Technology); HEWES, Mitchell (IAEA)

**Presenters:** Dr PIATKOWSKA, Ewa (AIT Austrian Institute of Technology); SMITH, Paul (AIT Austrian Institute of Technology)

**Track Classification:** CC: Information and computer security considerations for nuclear security