INL/CON-23-75326-Revision-0

Developing an AI-Powered Zero-Trust Cybersecurity Framework for Malware Prevention in Nuclear Power Plants

December 2023

Sajedul Talukder, Syed Alam, Palash Kumar Bhowmik

Changing the World's Energy Future

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/CON-23-75326-Revision-0

Developing an AI-Powered Zero-Trust Cybersecurity Framework for Malware Prevention in Nuclear Power Plants

Sajedul Talukder, Syed Alam, Palash Kumar Bhowmik

December 2023

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Developing an AI-Powered Zero-Trust Cybersecurity Framework for Malware Prevention in Nuclear Power Plants

Sajedul Talukder,[*] Palash Kumar Bhowmik,[∩] Piyush Sabharwall,[∩] Syed Bahauddin Alam [†▽]

[*] *Computer Science, Southern Illinois University, Carbondale, IL, USA, sajedul.talukder@siu.edu*
[∩] *Irradiation Experiment & Thermal Hydraulics Analysis, Idaho National Laboratory*
[†] *Nuclear, Plasma & Radiological Engineering, University of Illinois at Urbana-Champaign, IL, USA*
[▽] *National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Champaign, IL, USA*

INTRODUCTION

This paper presents the development of an AI-powered Zero-Trust cybersecurity framework for malware prevention in nuclear power plants. The framework aims to enhance the security of critical systems within nuclear power plants by adopting the principles of Zero-Trust and leveraging artificial intelligence (AI) technologies. By assuming no implicit trust in any user or device and continuously authenticating and authorizing access, the framework ensures a robust defense against malware attacks. The integration of AI allows for the detection and prevention of malware through behavioral analytics, endpoint protection, network segmentation, and continuous monitoring. The paper discusses the key considerations, steps, and technologies involved in developing this framework, emphasizing the importance of regular updates, training, compliance, and auditing. The proposed framework serves as a comprehensive approach to safeguarding nuclear power plants from sophisticated malware threats and protecting the integrity and safety of critical infrastructure.

With the increasing digitization drive in nuclear power plants, cybersecurity has become a paramount concern. The vulnerability of computer network architectures, particularly within the operational zone, calls for the development of robust cyber-defense systems. This study proposes an integrated multi-layer cyber-defense system with a zero-trust-infused malware prevention framework to combat cyber threats specifically targeting nuclear power plants.

The objectives of this study are twofold. Firstly, it aims to identify and understand the flaws and vulnerabilities in the existing computer network architectures within nuclear power plants, particularly within the operational zone. The recent digitization drive has introduced potential weaknesses that can be exploited by intelligent and strategic cyber threats. By comprehensively analyzing the existing network infrastructure, this study aims to uncover these vulnerabilities. Secondly, the study addresses the challenge of protecting safety-critical digital and non-digital assets within the network from intelligent cyber threats. To achieve this, a robust intrusion prevention framework will be developed, leveraging artificial intelligence (AI) technologies. By infusing the principles of zero-trust into the framework, the study aims to create a resilient defense against cyber threats.

The specific focus of this study is on developing a robust intranet architecture within the operational zone of nuclear power plants. By implementing a zero-trust strategy, which assumes no implicit trust in any user or device, the study aims to fortify the network against potential breaches and unauthorized access. The integration of AI in the intrusion prevention framework is crucial to detect and prevent malware attacks. AI algorithms can analyze network traffic patterns, identify anomalous behavior, and detect known malware signatures or indicators of compromise. This intelligent approach enhances the overall security posture within the operational zone, ensuring the protection of safety-critical assets. By developing a robust intranet architecture and implementing a zero-trust strategy, this study seeks to address the vulnerabilities introduced by the digitization drive in nuclear power plants. The proposed AI-based intrusion prevention framework, fortified with the principles of zero-trust, aims to safeguard the operational zone from intelligent and strategic cyber threats.

In the following sections, the study will delve into the identification of flaws and vulnerabilities within the existing network architectures, the development of the intrusion prevention framework, the implementation of the zero-trust strategy, and the integration of AI technologies. Through this comprehensive approach, the study aims to contribute to the enhancement of cybersecurity in nuclear power plants, ensuring the safety and integrity of critical assets within the operational zone.

ZERO-TRUST FRAMEWORK

This study proposes a Zero Trust Model (ZTM) as an autonomous cyber defense system for nuclear power plants. The ZTM utilizes artificial intelligence (AI) and machine learning (ML) safeguards to block known vulnerabilities, detect and mitigate endpoint vulnerabilities, and prevent zero-day exploits. The goal is to establish secure and dynamic connections while maintaining a high-security posture at the edge operations of nuclear power plants. The proposed architecture incorporates threat intelligence feeds, behavior analysis, whitelisting, blacklisting, and AI/ML-enabled safeguards to provide comprehensive protection against intelligent malware. By taking a holistic approach, combining secured network architecture, advanced AI methods, dedicated hardware security modules, and human-developed rules, the ZTM aims to prevent the entry of intelligent malware in real-time. This proactive approach ensures the safety of millions of lives and protects critical security assets from compromise.

Novel System Achitecture

This robust intranet architecture will also include an intrusion prevention software system to prevent intelligent malware from being installed onto any digital systems. The proposed project will also leverage an existing testing facility featuring a network of equipment, devices, and computers on which we will implement and evaluate our proposed framework in a realistic setting. Existing network architectures

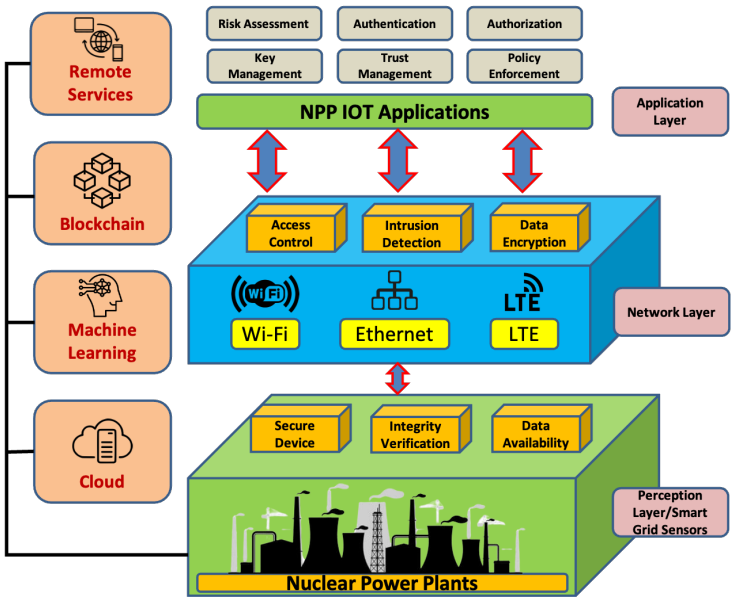

Fig. 1. AI-Powered Zero-Trust Cybersecurity Framework

inside nuclear power plants—especially inside the operational zone—typically implement a private intranet that connects multiple computers in a peer-to-peer (P2P) network architecture. These computers are considered almost impenetrable because (1) the intranet is completely isolated from the outside world (i.e., the Internet), and (2) data diodes regulate the flow of data in only one direction (i.e., from the operational zone to user zones such as the front office).

Review of Previous Malware Attacks

Along with physical security, anti-virus software, verification of supply chain, etc., it seems that employing an "air gap" [1] to isolate the operational zone will very reliably protect the vulnerable P2P intranet from complex cyber threats. However, this is a start misconception, as demonstrated by the incident involving the Stuxnet Trojan Worm [2] which was first discovered in Iran's Natanz nuclear facility. It was the world's first malware that could thrive and replicate itself inside a completely secured intranet isolated from the Internet. It completed all its intended malicious tasks without anyone noticing its presence. Following Stuxnet, several variants of it have emerged, becoming ever stealthier and more dangerous (e.g., Duqu [2011], Flame [2012], Havex [2013], Triton [2017], and Unnamed [2018]). With the rapid advancement of artificial intelligence (AI), adversarial state and non-state actors are developing increasingly powerful autonomous malware [3] that can thrive undetected inside a secured intranet for years and inflict serious damage to the facility.

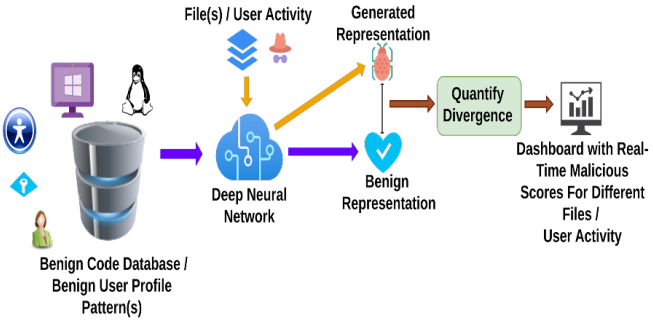

Fig. 2. Detailed look inside the detection mechanism

Malware Prevention Strategies

Development of an in-house architecture that is dedicated to the country's Nuclear power plants and will be able to detect such next-generation intelligent malware that intentionally targets the most secured zone is therefore becoming of utmost importance. Our proposed architecture takes a holistic view of this problem. It leverages secured network architecture, state-of-the-art AI methods, dedicated hardware security modules, and human-developed rules in order to nip this problem in the bud (i.e., by preventing, in real-time, the entry of intelligent malware (not just detecting it after it installs itself), potentially safeguarding millions of lives and protecting the country's critical security assets from being compromised.

TECHNICAL APPROACH TO THE SOLUTION

Individual machines are the starting point in our proposed system architecture, in which we aim to incorporate our intel-

ligent, AI-driven Host Intrusion Prevention and Monitoring System (HIPMoS). The primary task is to develop a novel, state-of-the-art intrusion prevention system that is operating-system-independent and trained on byte-level code (binary) rather than on source code (human-readable text) so as to make it more adaptable to different computing environments.

Development of the Intrusion Prevention Module

Most intrusion prevention systems work at the network packet level, meaning that, following packet analysis, certain packets may be discarded if deemed suspicious, preventing any host computer from getting infected [4, 5]. However, our proposed software module can analyze malware not only from network packets (in this case, packets received over the intranet), but more importantly, scrutinize any files before they are installed onto the machine(s), irrespective of the underlying operating system. These files can be software update patches, new software applications, etc., and may originate from any external device such as a laptop or USB.

Development of the User-Behavior Monitoring Module

Our proposed HIPMoS also has a user-behavior monitoring tool. Using deep representation learning, this AI-driven tool is first trained on regular employee behavior patterns as the employees follow standard operating procedure during regular operations as per Nuclear-Regulatory-Commission-approved Standard Technical Specifications. And athough every employee, especially inside the operational zone, is authorized and even monitored 24/7, our module will track and analyze each and every movement on an assigned machine the employees' digital signatures as applications are opened, mice are clicked on different applications, confidential files are accessed, keystroke are made, etc. [6, 7]. The proposed tool can raise an alert if any authorized employee's digital behavior appears suspicious. Fig. 3 shows the entire HIPMoS training and deployment cycle.

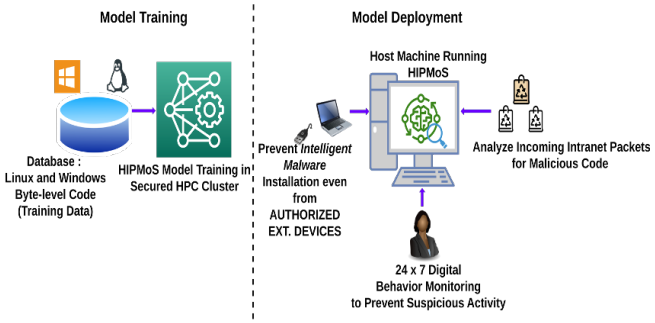

Fig. 3. The entire HIPMoS deployment cycle

CONCLUSION

In conclusion, the proposed Zero Trust Model (ZTM) presents a comprehensive and advanced approach to network security in nuclear power plants. By integrating artificial intelligence (AI) and machine learning (ML) safeguards, the model aims to provide autonomous cyber defense, effectively blocking known vulnerabilities, detecting and mitigating endpoint vulnerabilities, and preventing zero-day exploits. One of the key strengths of the ZTM is its emphasis on ease of use and versatility to meet the demands of power plant managers. With secure and dynamic connections, the model enables quick interpretation of alerts, efficient sharing of misconfigurations, and the provision of necessary security controls at machine speed. This ensures a high-security posture at the edge operations of nuclear power plants, where critical systems and assets are located. The ZTM's autonomous cyber defense capabilities are of particular importance in the context of nuclear power plants. By blocking known vulnerabilities through threat intelligence feeds and employing behavior analysis, whitelisting, blacklisting, and AI/ML-enabled safeguards, the model significantly enhances the security infrastructure. Additionally, the ability to detect and mitigate endpoint vulnerabilities reduces the risk of potential breaches and compromises. Furthermore, the ZTM addresses the challenge of zero-day exploits posed by "Intelligent Malware." With a holistic approach that combines secured network architecture, advanced AI methods, dedicated hardware security modules, and human-developed rules, the model aims to prevent the entry of intelligent malware in real-time. This proactive defense strategy not only protects critical security assets but also has the potential to save lives and prevent potentially catastrophic incidents. By adopting the proposed ZTM, nuclear power plants can enhance their cybersecurity posture and protect the integrity and safety of their operations. The model's integration of AI and ML technologies, along with its versatile and user-friendly design, makes it a valuable framework for combating emerging cyber threats. As the digital landscape evolves, the ZTM serves as a robust defense mechanism, providing a foundation for the ongoing protection of nuclear power plant infrastructure and the preservation of national security.

REFERENCES

1. E. BYRES, "The air gap: SCADA's enduring security myth," *Communications of the ACM*, **56**, *8*, 29–31 (2013).
2. R. LANGNER, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, **9**, *3*, 49–51 (2011).
3. B. GUEMBE, A. AZETA, S. MISRA, V. C. OSAMOR, L. FERNANDEZ-SANZ, and V. POSPELOVA, "The emerging threat of ai-driven cyber attacks: A Review," *Applied Artificial Intelligence*, **36**, *1*, 2037254 (2022).
4. C.-M. CHEN, Y.-L. CHEN, and H.-C. LIN, "An efficient network intrusion detection," *Computer communications*, **33**, *4*, 477–484 (2010).
5. V. PAXSON, R. SOMMER, and N. WEAVER, "An architecture for exploiting multi-core processors to parallelize network intrusion prevention," in "2007 IEEE Sarnoff Symposium," IEEE (2007), pp. 1–7.
6. S. TALUKDER, I. I. SAKIB, F. HOSSEN, Z. R. TALUKDER, and S. HOSSAIN, "Attacks and defenses in mobile ip: Modeling with stochastic game petri net," in "2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)," IEEE (2017), pp. 18–23.
7. S. TALUKDER and Z. TALUKDER, "A survey on mal-

ware detection and analysis tools," *International Journal of Network Security & Its Applications (IJNSA) Vol*, **12** (2020).