Kai Goebel and Shantanu Rane
PARC/SRI International, USA

AI in Industrial IoT Cybersecurity

AI's role within the domain of Industrial IoT (IIoT) is somewhat obscured by other limelight-stealing feats of AI, such as the creation of convincing deep fakes featuring politicians or celebrities. However, the critical importance of IIoT in domains like consumer goods manufacturing, healthcare, power generation, and transportation mandates a closer examination of the new capabilities and pitfalls spawned by AI technologies.

In recent years, well-funded teams of expert hackers have exploited Supervisory Control and Data Acquisition (SCADA) systems as well as Programmable Logic Controller (PLC) systems, orchestrating acts of sabotage across a variety of industrial sectors spanning water utilities, petrochemical facilities, nuclear facilities and power grids. Given this context, it is extremely important to ensure that AI-enhanced Industrial IoT does not inadvertently provide a conduit for unauthorized system access and misuse.

Like most technological advances applied to cybersecurity, AI assumes a dual role: In the hands of security practitioners, it functions as a guardian, fortifying defenses against everyday threats. In the hands of attackers, it enables novel attack mechanisms, posing risks to an IIoT system's safety and integrity. There is a third aspect in which AI is relatively unique: Even as the impetus to deploy AI technology grows, there remain aspects of AI that the scientific community does not yet fully understand. This creates a possibility of adverse outcomes resulting from unexplained weaknesses in key AI components, such as deep neural networks. We divide our reflection on the role of AI in IIoT security into 4 main areas.

I. Device Security

Modern AI approaches are being developed to help security experts analyze software running on IIoT devices, identify and chain vulnerabilities that might breach security or privacy [3]. These insights can drive the development of patches, and innovative defenses based on software compartmentalization that can limit the influence of an adversary who has already gained access to the system [2]. By combining vulnerability information with analysis of historical usage patterns and device telemetry, modern AI tools can enhance predictive maintenance capabilities of an IIoT system.

Unfortunately, these same approaches can be used to build malware that is adept at bypassing conventional security measures by employing tactics like constant mutation and evading signature-based detection. They will learn to automatically identify and select high-value device targets, exploit vulnerabilities in the device software and exfiltrate sensitive data with remarkable efficiency. Attackers thus have at their disposal new ways to develop powerful, continuously evolving, advanced persistent threats (APTs).

II. Connectivity and Network Security

A key aspect of IIoT security relates to the exploitation of the network connectivity and network traffic. AI approaches have leveraged machine learning and natural language processing to obtain superior threat intelligence from security alerts and threat feeds in real time. They also enable more comprehensive and automatic scanning of networks for known vulnerabilities.

This has resulted in improved intrusion detection in IIoT networks. Traditionally, intrusions have been detected by comparing the incoming network traffic against recognized patterns or signatures. New approaches have been developed to detect intrusions based on more complex and adaptive network behaviors, rather than comparing against pre-packaged attack signatures. Once intrusive behavior is discovered, it can be swiftly blocked to provide improved network protection.

Generative AI can support organizations in bolstering their defenses. This includes tasks like reviewing device and network code for efficiency, identifying potential security vulnerabilities, and exploring novel tactics that malicious actors may employ.

Digital Object Identifier 10.1109/MCI.2023.3332547
Date of current version: 8 January 2024

Corresponding author: Kai Goebel (e-mail: kai.goebel@sri.com).

1556-603X

Unfortunately, advances in AI have also facilitated the creation of more sophisticated network attacks. In particular, AI primitives can be leveraged to craft, automate and amplify Distributed Denial-of-Service (DDoS) attacks. AI-driven rapid identification of known vulnerabilities makes it possible for adversaries to determine and attack high-value targets while evading detection as long as possible.

III. Configuration Security

AI techniques – both the modern ML-driven approaches and classical symbolic approaches – play an extremely useful role in IIoT configuration. This involves improving the security posture of an IIoT system by manipulating its software components, networking components and hardware settings on devices. There are novel mechanisms that automatically consult vulnerability databases, alter composed system configurations and provide human-readable insights about the security trade-offs being made in order to achieve specific functionalities [4]. This is significant, given that two-thirds of the security-specific downtime of critical infrastructure is attributable to misconfiguration [5].

IV. Alignment Challenges

Recently, it has been recognized that the functionality of AI algorithms may not automatically align with the properties desired by its human designers. This is especially true for AI approaches that learn from data, e.g., deep neural networks and large language models. Collectively, these challenges constitute the AI alignment problem, and they must be resolved before AI algorithms are widely deployed. Unfortunately, experience suggests that AI deployment will precede careful analysis, and we will continue to discover potentially serious security-related alignment problems in AI-based IIoT systems.

Alignment challenges have been attributed to incomplete or statistically biased training data, incorrect optimization functions within the machine learning core, lack of consideration of risk attitudes in automated decision making, and other reasons. A particularly concerning example is in the arena of using generative AI, e.g., Large Language Models (LLMs), to write code. Software written by AI – or with the help of an AI assistant – has been found to be less secure than that written by human beings [1]. Furthermore, the illusion of correctness communicated by AI assistants such as ChatGPT creates an insidious situation: Security flaws in AI-generated code will propagate as engineers become more complacent about systematic testing of the software. In the ever-evolving landscape of IIoT cybersecurity, AI brings promise as well as peril. It serves as a double-edged sword, capable of enhancing defenses while simultaneously enabling innovative attacks and introducing complex challenges. In this landscape, the never-ending cat-and-mouse game in cybersecurity is poised to intensify with the rapid adoption of AI techniques by friendly and hostile actors.

References

[1] A. Clark, "New AI wave will find uses and abuses in cybersecurity," Axios, 2023. Accessed: Oct. 14, 2023. [Online]. Available: https://www.axios.com/2023/02/17/cybersecurity-ai-tech-chatgpt-bing

[2] Compartmentalization and Permissions Management (CPM), "Defense Advanced Research Projects Agency (DARPA) broad agency announcement," Funding Opportunity Number HR001123S0028, Apr. 2023. [Online]. Available: https://sam.gov/opp/d624b5ffc9d24e38b966a714cbbb4f7d/view

[3] Intelligent Generation of Tools of Security (INGOTS), "Defense Advanced Research Projects Agency (DARPA) broad agency announcement," Funding Opportunity Number HR001120S0049, Jun. 2023. [Online]. Available: https://sam.gov/opp/7afef7eed5db4ff490971d0667cbaa48/view

[4] H. Soroush et al., "SCIBORG: Secure configurations for the IoT based on optimization and reasoning on graphs," in *Proc. IEEE Conf. Netw. Secur.*, 2020, pp. 1–10.

[5] M. Nunnikhoven, "The top worry in cloud security for 2021," TrendMicro Rep., Jan. 2021. [Online]. Available: https://www.trendmicro.com/en_vn/research/21/a/the-top-worry-in-cloud-security-for-2021.html