DATASETS GENERATION WITH VIRTUAL CYBER PHYSICAL SYSTEM DESIGN AND CYBER-ATTACK SIMULATION ON NUCLEAR FACILITIES

Yue Xiao	**Feiyan Dong**	**Shi Chen**	**Kazuyuki Demachi**
The University of Tokyo	The University of Tokyo	The University of Tokyo	The University of Tokyo

ABSTRACT

With the increasing adoption of digital instrument and control systems in nuclear facilities, cyber-attacks pose serious threats and bring a new issue to nuclear security. In response, authorities have published relevant criteria for preventing cyber-attacks in security culture and recommended to adopt Defense in Depth (DiD) strategy to cyber-security of nuclear facilities. To this end, we introduce deep learning-based time series analysis to detect cyber-attacks on nuclear facilities. Given the difficulty of deep learning models training due to the lack of cyber-attacks data on nuclear facilities, we design a virtual cyber physical system (CPS) to simulate cyber-attacks and generate cyber-attack datasets. The virtual CPS deploys the human machine interface (HMI), programmable logic controller (PLC) and controlled devices connected via Modbus/TCP protocol to simulate the control processes in nuclear facilities. The simulation of cyber-attacks consists of two steps carried out separately on two independent Local Area Networks (LANs), i.e., site LAN #1, site LAN #2. In cyber-attacks, hackers obtain access to office computers in NPPs through site LAN #1 by buffer overflow attack and collect network traffic data on the attacked host. Then, we simulate hackers using the attacked host to attack the PLC through site LAN #2 to cause confusions collect network traffic data on the attacked PLC device to generate the first part of the dataset. This simulation imitates a real cyber-attack activity that could happen in nuclear facilities. And we also collect the data of the device status simultaneously to monitor the physical layer condition. To demonstrate the practicality of the generated datasets, verification experiments are performed on the proposed deep learning model. Moreover, the design of virtual CPS allows flexible generation of extensive cyber-attacks datasets, which has significant benefits for further evaluation of the design of cyber security systems for nuclear facilities and the implementation of DiD.

INTRODUCTION

As the reliance on digital instrumentation and control (I&C) devices in nuclear power plants (NPPs) grows [1], the risks of cyber-attacks on these facilities increase as well. The consequences of such attacks can be catastrophic, leading to widespread damage and loss of life. Therefore, it is crucial to implement effective cyber security measures to protect NPPs from cyber threats. Several

cyber-attacks on NPPs have been reported in recent years. For example, the Stuxnet worm was introduced into Iran's nuclear facility in 2010 through a USB drive by insiders, and it destroyed the centrifuges used in uranium enrichment. In 2014, a phishing email was used to download malware onto the network of a South Korean nuclear power plant, allowing attackers to steal information including designs, manuals, and personal documents. These incidents demonstrate the vulnerability of NPPs to cyber-attacks and highlight the need for better cyber security measures.

Digital I&C devices to be applied in NPPs should be designed to meet the licensing requirements in terms of security, and cyber security is a most important issue [2, 3]. The Cyber-Physical Systems (CPS) in NPPs can be divided into four parts, as shown in Fig.1a: the Corporate Wide Area Network (WAN), the Site Local Area Network (LAN), the Data Acquisition Systems, and the Control and Safety Systems.

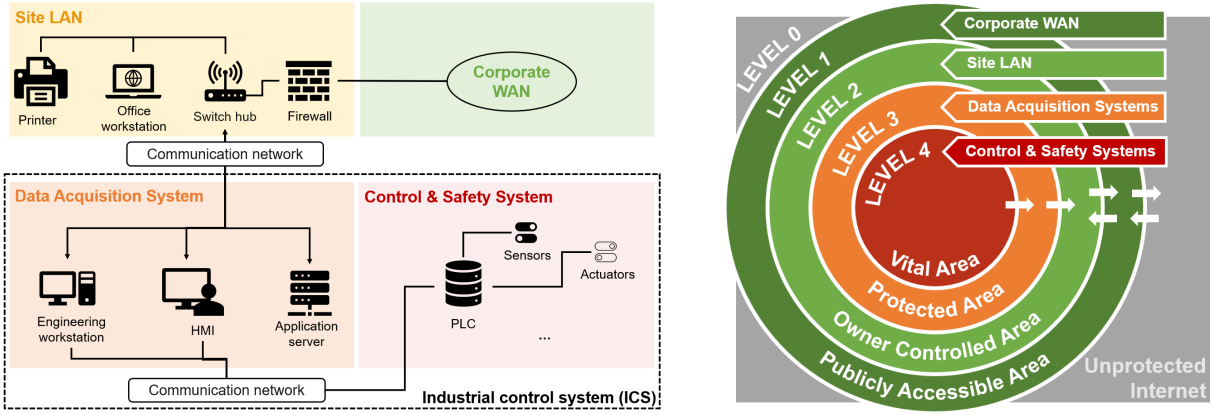

Figure 1. (a) Cyber-physical systems in nuclear facilities.
(b) Defense in Depth for cyber security.

The U.S. Nuclear Regulatory Commission (NRC) recommends adopting Defense in Depth (DiD) for cyber security in NPPs. DiD divides the CPS into four levels, corresponding to the cyber and physical layers. Levels 1-3 are related to the cyber layer, while level 4 involves the physical layer. The cyber and physical layers are both essential components of cyber security in NPPs. In addition, the Physical Protection System (PPS) is widely used to protect NPPs from threats. The PPS consists of four steps: deterrence, detection, delay, and response, with detection being a crucial step for the proper functioning of the system. If detection fails, the delay and response steps will not be activated.

The implementation of effective cybersecurity measures in nuclear power plants (NPPs) faces a major challenge in the form of limited availability of data on cyber-attacks specific to NPPs. This lack of sufficient data can adversely affect the training of machine learning models used for cybersecurity purposes. In order to overcome this challenge, we propose the use of digital twins to generate training data with diverse distributions. A digital twin is a virtual replica of a physical system that can simulate cyber-attacks on a virtual cyber-physical system (CPS) in an NPP.

To achieve better detection efficiency, it is crucial to design a digital twin that can monitor both the physical and cyber layers simultaneously. Such a digital twin can provide a more comprehensive view of the system and enable a more accurate detection of cyber-attacks. This approach can improve the effectiveness of the cybersecurity measures in NPPs and provide a reliable means of training machine learning models for cybersecurity purposes. In conclusion, digital twins offer a promising solution to the challenge of limited data availability in NPP cybersecurity, and their use can lead to more robust and effective cybersecurity measures.

METHOD

To design the virtual CPS, four representative components of a real CPS structure in nuclear facilities were considered. The framework shows in figure 2 included circuit elements to simulate basic devices, three computers to simulate PLC, HMI, and Workstation, and a hacker from the outside facility to simulate attacks. As an example, the simulation of a pump was demonstrated using an LED to represent the pump's working status. Connect the LED to the Raspberry Pi according to the circuit. After programming with *openplc*® [5] on Raspberry Pi could imitate PLC function. As for the HMI simulation, filling in the Raspberry IP address and the corresponding port number will enable the monitoring and operation of the led status through the HMI panel in another computer.

The simulation of cyber-attacks was carried out separately on two independent Local Area Networks (LANs), i.e., site LAN #1 and site LAN #2. Hackers obtained access to office computers in nuclear power plants through site LAN #1. Then, hackers used the attacked host to attack the PLC through site LAN #2 to cause confusion. Collected digital data on the attacked PLC device to generate the first part of the dataset. This simulation imitates a real cyber-attack activity that could happen in nuclear facilities. Data on the device status was simultaneously collected to monitor the physical layer condition.

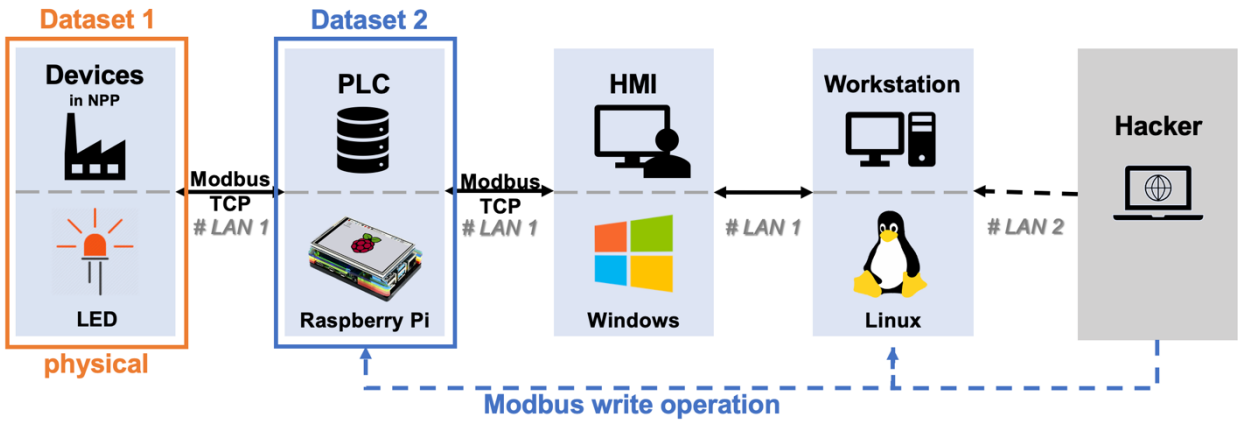

Figure 2. Simulation of attacking the virtual cyber-physical systems (CPS).

During the collection, data was generated for an hour, including normal behavior (pressing button, and HMI operations) and malicious behavior of Modbus attacks by workstations to turn the LED on and off, which simulate the working and stopping of pumps in real facilities. The

attacks were programmed to occur randomly and less than or equal to six times in an hour to mimic real attacks. Data was constantly collected at the Raspberry Pi through *Wireshark*® [6] during the entire experiment.

RESULTS & DISCUSSIONS

The article presents two datasets: physical data and digital data. The physical data set contains the feature of time, LED condition (on/off), and actions like normal behavior of press button and HMI operation, also abnormal behavior of cyber-attacks. The LED condition during the whole time is shown in Figure 3, and Table 1 is extracted from the dataset that shows the important time of actions. The collection lasts for one hour from 19:35 to 20:35. The LED is humanly controlled 2 times by pressing the button, and HMI operated 6 times. Attacks happened 4 times.

Table 1. Action schedule

Time	Action
19:35:36	Start
19:35:40	Turn on
19:36:06	HMI_turn off
19:42:16	HMI_turn on
19:43:36	Attack1_turn off
19:48:06	Attack2_turn on
19:48:17	Turn off
19:48:41	HMI_turn on
20:01:36	Attack3_turn off
20:05:26	Attack4_turn on
20:07:19	HMI_turn off
20:13:52	HMI_turn on
20:13:54	HMI_turn off
20:35:27	Stop

Table 2. Statistical results of dataset 2

Action	Attack	Normal	
		HMI	None
Number	8	12	7,148
Total	7,168		

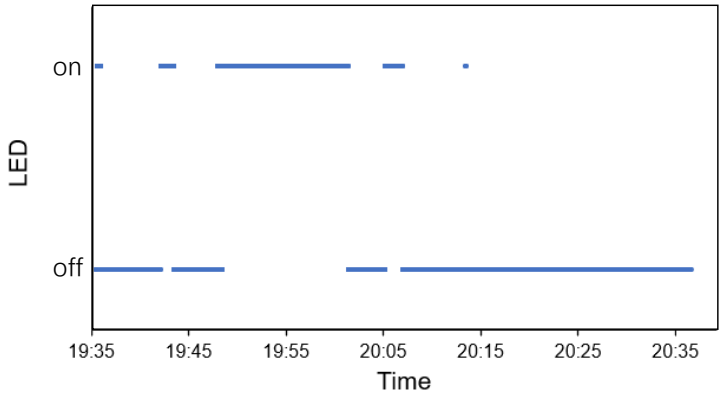

Figure 3. LED condition

The digital data set collected a total of more than 7,000 messages shown in Table 2. It contains 8 messages of 4 attacks and 12 messages for 6 times HMI operations. The features of the dataset contain several information, where the function codes as well as the IP addresses are noteworthy.

Figure 4 analyses the IP addresses and shows that there are two addresses representing the HMI and the workstation. Under normal operation, no workstation address would appear, but the IP address of the workstation shows that the workstation helped the hacker read and write data. The table shows some details. Note that the hacker_write should only be labeled as attacks. However, we also can see that the read action always appears before the write action. This might affect the detection model.

Figure 5 shows the message structure in Modbus TCP. It contains many different bits with different information. Here, the function code needs attention because it contains many different commands including reading and writing. Various codes might be used based on different activities. But in malicious attacks, hackers should use code with write function to attack, like

"05". Figure 6 analyses the function codes, and the whole process consists of only 1 and 5. The table shows that the hacker has used function code 5 as expected. However, we can see it is also used in HMI operation. So we cannot distinguish the attack based on function codes alone.

Modbus Application Protocol (MBAP) Header				Protocol Data Unit(PDU)	
Transcation Identifier	Protocol Identifier	Length Field	Unit ID	Func Code	Func Data

Figure 4. ADU Message Structure of MODBUS TCP

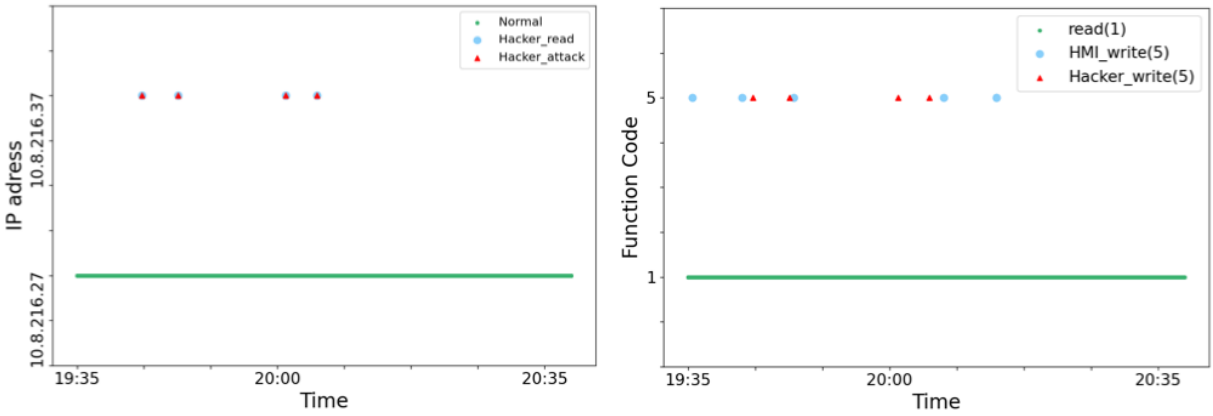

Figure 5. Statistical results of the IP address Figure 5. Statistical results of the IP address

Therefore, the data of "37 IP address along with 5 function code" is labeled as 1, indicating malicious behavior of attacks. Others are labeled as 0. Calling back to the dataset with a label of 1, we find the information shows the hacker's actions in physical reality turning the LED on and off. In the next step, the labeled dataset can be used for detection.

Table 3. Label method of the dataset 2

Action		IP	F-code	Label
Normal		10.8.216.27	1	0
HMI operation		10.8.216.27	5	0
Hacker	**Read**	10.8.216.37	1	0
	Write	10.8.216.37	5	1

CONCLUSION

In this paper, we presented two datasets that were generated simultaneously from the cyber and physical layers of a virtual cyber-physical system. The datasets were linked by timestamp, which allowed for a more efficient detection of anomalies. Our design of the system was flexible and

can be easily adjusted or expanded to generate infinite datasets. Overall, these datasets will be useful for developing new anomaly detection models.

Moving forward, our future work aims to develop an anomaly detection model for multiple layer datasets. The first step will be to verify the datasets and then prototype the development of the model. Ultimately, we hope to develop a robust anomaly detection model that can be applied to various cyber-physical systems.

In conclusion, this study highlights the importance of considering both the cyber and physical layers of a system when generating datasets for anomaly detection. Our approach can be applied to other systems and can lead to more efficient and accurate anomaly detection.

REFERENCES

[1] Kim, S., Heo, G., Zio, E., Shin, J., & Song, J. G. (2020). Cyber attack taxonomy for digital environment in nuclear power plants. Nuclear Engineering and Technology, 52(5), 995-1001.

[2] P.A. Khand, “Attack Tree Based Cyber Security Analysis of Nuclear Digital Instrumentation and Control Systems” the Nucleus, vol. 46, 2009, pp. 415e428, 4.

[3] Kim, D. Y. (2014). Cyber security issues imposed on nuclear power plants. Annals of Nuclear Energy, 65, 141-143.

[4] U.S. Nuclear Regulatory Commission, "Regulatory Guide 1.180: Guidelines for Digital Instrumentation and Control Systems at Nuclear Power Plants," NUREG-0800, 2018

[5] https://openplcproject.com/

[6] https://www.wireshark.org/