# Remote Security Assessment for Cyber-Physical Systems: Adapting Design Patterns for Enhanced Diagnosis

Kazutaka Matsuzaki[1][a], Kenji Sawada[2][b] and Shinich Honiden[3][c]
[1]*Faculty of Global Informatics, Chuo University, Tokyo, Japan*
[2]*The Info-Powered Energy System Center, The University of Electro-Communications, Tokyo, Japan*
[3]*Faculty of Science and Engineering, Waseda University, Tokyo, Japan*

Keywords: Cyber-Physical System, Cyber Security, Communication Robustness Test, Industrial Automation and Control System.

Abstract: This paper presents a novel approach to remote security diagnosis for critical infrastructure systems, focusing on integrating Cyber-Physical Systems (CPS) and cloud-based diagnosis. The proposed method adapts two existing design patterns to address the challenges associated with remote security diagnosis: (1) adapting the "Ambassador Pattern" of the cloud design pattern for virtual extension of the CPS input interface to the cloud computing environment, and (2) adapting the "Data Aggregation Pattern" of the edge computing design pattern for virtual extension of the CPS output to the cloud computing environment. We discuss implementing and evaluating our proposed method in a simulated environment, demonstrating its potential for improving the accuracy and efficiency of remote security assessment. This research contributes to developing secure and reliable CPS by providing insights into effectively adapting existing design patterns for remote security diagnosis.

## 1 INTRODUCTION

As the adoption of cyber-physical systems (CPS) continues to grow, encompassing applications such as distributed solar power plants, cloud-based building management, and smart factory security, the risk of cyberattacks on these systems also rises. Cyberattacks on CPS can result in significant damage to both digital and physical domains, with infrastructure components like pipelines, water treatment facilities, and power grids susceptible to shutdowns and disruptions.

Given the ongoing discovery of new vulnerabilities and the evolving sophistication of cyber attackers, it is crucial to continuously assess the efficacy of cybersecurity measures. Existing security evaluation frameworks include certification tests for control systems and embedded devices, assessing known vulnerabilities and communication robustness. However, conducting such tests on operational systems poses challenges in terms of time, manpower, and cost constraints.

To reduce the burden of on-site testing, this paper explores the potential of remote testing over a network as an alternative approach to enhancing cybersecurity in CPS. We investigate a cloud-based diagnostic method aimed at minimizing on-site testing while efficiently evaluating the security posture of these systems. Our analysis begins with identifying the challenges that can be addressed through remote diagnostic innovations, followed by highlighting the remaining issues that warrant further investigation.

The paper is structured as follows: Section 2 discusses the challenges of CPS security diagnostics; Section 3 introduces a cloud-based security diagnostic architecture; Section 4 evaluates the addressed challenges; Section 5 examines the remaining challenges; and Section 6 presents a conclusion.

[a] https://orcid.org/0000-0003-2337-2686
[b] https://orcid.org/0000-0001-8935-0434
[c] https://orcid.org/0000-0003-1385-3996



Matsuzaki, K., Sawada, K. and Honiden, S.
Remote Security Assessment for Cyber-Physical Systems: Adapting Design Patterns for Enhanced Diagnosis.
DOI: 10.5220/0012132600003555
In *Proceedings of the 20th International Conference on Security and Cryptography (SECRYPT 2023)*, pages 805-812
ISBN: 978-989-758-666-8; ISSN: 2184-7711

# 2 CPS SECURITY ASSESSMENT CHALLENGES

## 2.1 Assumed Situation

To clearly understand the assumed situation, Figure 1 illustrates a simplified model of CPS security diagnostics and the associated challenges. In this model, the diagnostic device (DD) transmits anomalous data to the CPS's input interface (IN) for diagnostic purposes. While the specific data transmitted depends on the protocol supported by the input interface, this paper assumes using TCP/IP.

The DD observes the output (OUT) of the CPS to evaluate its security measures. Potential output interfaces for monitoring include digital, analog, serial (e.g., RS-485), TCP/IP, register (e.g., Modbus), LED, and others.

Monitoring is necessary during two distinct phases:

- Assessing whether essential functions are maintained while transmitting anomalous data for diagnosis.
- Verifying that the system returns to normal operation after sending abnormal data.

The challenge lies in the possibility of obtaining inaccurate diagnostic results due to various factors. These inaccuracies can occur in four different ways:

(a) Anomalous data may not reach the CPS, causing it to be disregarded as an invalid test case. Consequently, potential issues with security measures might be overlooked.

(b) Monitoring data may not be observable from the CPS, leading to determining abnormalities or dismissing an invalid test case.

(c) In cases where monitoring data is not observable from the CPS, anomalies might be missed, causing the test case to proceed undetected.

(d) When verifying whether the system returns to normal operation after transmitting abnormal data, the inability to observe monitoring data from the CPS may result in misjudgements of abnormalities or false negatives.

By addressing these potential inaccuracies, we aim to enhance the effectiveness and reliability of CPS security diagnostics.

## 2.2 Related Research

In recent years, research on network-based security testing (e.g., fuzzing) for control systems has gained prominence at international conferences.

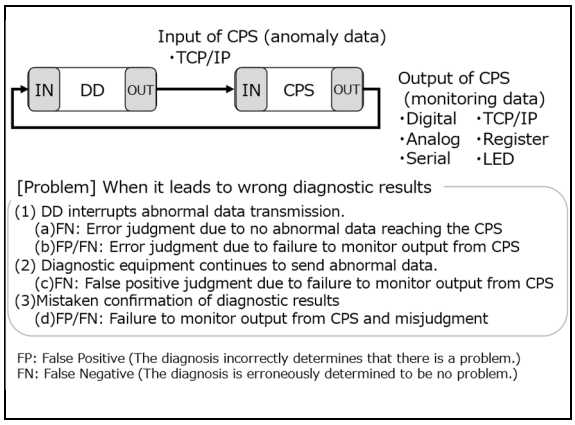



Figure 1: CPS security diagnostic model and current issues.

These studies not only focus on improving test performance but also organizing the requirements for test environments specific to control systems. For instance, it is emphasized that equipment under test should be monitored as comprehensively as possible (Pfrang, 2018). Monitoring targets encompass network interfaces and digital inputs and outputs. Unlike fuzzing tests for software, essential functions are maintained by analyzing communication and hardware signals. Since it is not feasible to access the device's internals using a debugger, inferences about the device must be drawn from externally observable phenomena during testing.

Efforts to enhance the performance of control system fuzzing tools are also underway. For example, Polar (Luo, 2019) employs machine learning to identify the functional code of the control system and conducts fuzzing by progressively modifying anomalous data. This study aims to explore whether such tools can be applied to test targets located at considerable distances from one another. Another study of interest uses fuzzing to evaluate security measures for entire control systems (Chen, 2019).

## 2.3 Security Testing Efforts in Past Demonstrations

Security testing for IEC 61850-based systems has demonstrated that communication robustness tests could be performed using remote access VPNs (L2TPv2) and VPNs between locations (L2TPv3) (Matsuzaki, 2020, Wilkerson, 2022). When employing a VPN between locations, networks can communicate as if they were part of the same LAN.

In one instance, a communication robustness test of the IEC 61580-MMS protocol was conducted via a VPN using DTLS. In another case, a compact portable terminal (e.g., Raspberry Pi) was brought

near the test target and operated as an LAC/LNS[4] for L2TPv3, creating a virtual network to verify the feasibility of communication robustness testing.

Figure 2 illustrates the response status to the issue using naïve methods, including those demonstrated in past years. The following three mechanisms were added as naïve solutions:

**Remote Access VPN:** The diagnostic equipment and CPS were connected through a remote access VPN. Specifically, a VPN client device using DTLS was placed on the CPS side, and a controller was installed on the diagnostic equipment side.

**DAQ (Data Acquisition):** A digital signal was branched from the terminal block contact in the CPS and used as an input to DAQ.

**DAQ Client:** The DAQ client accesses DAQ via TCP/IP, periodically obtains data values held by DAQ, and presents them to the user via a GUI.

Although most items were partially addressed, several challenges remained:

a) The team was able to run numerous test cases on the TCP/IP protocol stack. Still, issues persisted with Layer 2 (e.g., Ethernet) and the reachability of anomalous data using broadcast.

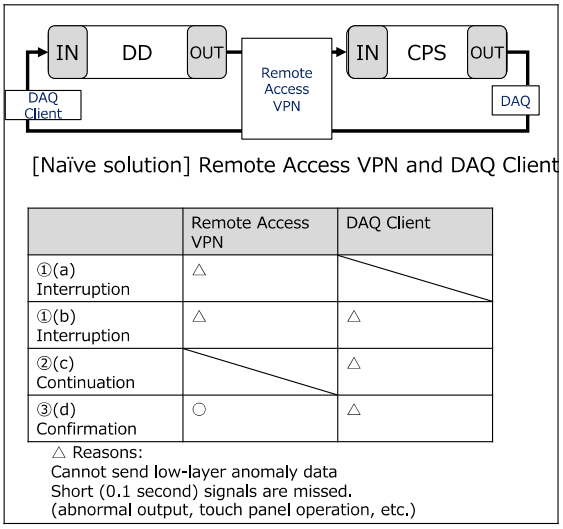
b) The DAQ client was used via VPN to acquire current values at high frequency continuously. However, there remained the issue of being unable to detect pulse data for short periods (about 0.1 second) due to the access frequency of the DAQ client and network latency.

c) Network latency occasionally made it challenging to perform essential function monitoring. The problem remained that essential function outages could be overlooked, which were short but longer than the test requirements.

d) The issue persisted that after sending abnormal data, when monitoring whether the system had returned to normal operation status, it might not be able to detect the condition that allowed the user to operate the touch panel.

# 3 ADAPTING DESIGN PATTERNS FOR REMOTE

Figure 3 shows the proposed input/output extension design pattern and the improvement status of the




| | Remote Access VPN | DAQ Client |
|---|---|---|
| ①(a) Interruption | △ | |
| ①(b) Interruption | △ | △ |
| ②(c) Continuation | | △ |
| ③(d) Confirmation | ○ | △ |

△ Reasons:
Cannot send low-layer anomaly data
Short (0.1 second) signals are missed.
(abnormal output, touch panel operation, etc.)

Figure 2: Status of addressing challenges with naïve Solutions.

issue. The proposed method addresses the security diagnosis issue by simultaneously adapting two existing design patterns.

**Adapting the Ambassador Pattern for Virtual Extension of CPS Input Interface to the Cloud**

This approach adapts the "Ambassador Pattern" from the cloud design pattern[5] to the security diagnosis. It uses a site-to-site VPN instead of the previous section's remote access VPN. The *OUT* of the diagnostic equipment is virtually on the same local network segment as the *IN* of the *CPS*. Specifically, LAC/LNS equipment is brought in, and an L2TPv3 VPN is temporarily established between the diagnostic device (*DD*) and the *CPS*. In the ambassador pattern, a proxy for remote access is deployed for communication with external services. Therefore, the proposed method can be described as a "dispatched" ambassador pattern, which is realized by bringing in equipment on the CPS side.

**Adapting the Data Aggregation Pattern for Virtual Extension of CPS Output to the Cloud**

In this approach, the "Data Aggregation Pattern" of the edge computing design pattern[6] is adapted for the security diagnosis. Instead of checking the output of the *CPS* using the pull-type method by the DAQ client as described in the previous section, the data

[4] L2TP Access Controller / L2TP Network Server

[5] Microsoft, Ambassador pattern, https://docs.microsoft.com/en-us/azure/architecture/patterns/ambassador

[6] Amazon Web Services, Industrial edge design considerations, https://docs.aws.amazon.com/whitepapers/latest/industrial-iot-architecture-patterns/industrial-edge-design-considerations.html

aggregated on the DAQ is delivered to the *DD* side using the push-type method. Using digital output as an example, data is aggregated and conditioned on the DAQ, relays are turned on and off, and the data is transmitted to the *DD* side as digital output using the remote I/O mechanism. This method makes it possible to convey short pulse data that would otherwise be overwritten on the DAQ side to the *DD* side. In the data aggregation pattern, data is temporarily aggregated by an edge computer in an environment with unstable network communication and periodically summarized and sent to the center. In the proposed method, data is aggregated and stored so that it is not overwritten and reliably transmitted to the *DD* side through push-type distribution.

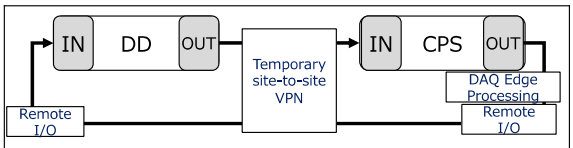
The proposed method can be used in a stable network environment.

a) The reachability of Layer 2 (e.g., Ethernet) and broadcast anomalous data is also addressed.

b) Depending on the performance of the DAQ, changes as short as 0.1 seconds can be monitored and communicated to the *DD* so that short pulse data (about 0.1 seconds) can be detected.

c) As with the above, depending on the performance of the DAQ, monitoring of essential functions can be performed in milliseconds. It is also possible to supplement the suspension of essential functions for short periods, which cannot be overlooked as a test.

d) When monitoring whether the system returns to normal operation status after sending abnormal data, we confirmed that the system can detect a 0.1-second operation, even when the user's touch panel operation is enabled.




[Proposed Method]

| | Temporary site-to-site VPN | DAQ Edge Processing with Remote I/O |
|---|---|---|
| ①(a) Interruption | ○ | |
| ①(b) Interruption | ○ | ○ |
| ②(c) Continuation | | ○ |
| ③(d) Confirmation | ○ | ○ |

Inadequate Reasons:
When latency is high, jitter error countermeasures are necessary.
Limitations of existing SW/HW when there are many objects to be monitored

Figure 3: Proposed Input/Output Extension Patterns and Status of Issue Improvement.

# 4 IMPLEMENTATION AND EVALUATION

To evaluate the ability of the proposed method to meet the challenges under certain conditions in a simulated environment, we conducted three different tests:

1. Comparison of the number of automated tests by testing on the local network and testing with the proposed method.

   For comparison, we ran a test suite on TCP/IP on the test equipment. The number of test cases within the test suite automatically executed by local testing environment was compared with the number of test cases when performed in a cloud-based environment.

2. Comparison of the number of automated tests in industrial protocols.

   From the tests in the first point, a similar comparison was made only for the test suite for IEC 61850 (MMS), which is increasingly being introduced in the power sector. The IEC 61850 software was changed to another software, and the digital output of the equipment under test was also subject to monitoring.

3. Pulse Signal Monitoring Test.

   We conducted a monitoring test of the proposed method in a cloud-based environment. The test was performed by changing the analog output for the test in a steady time series and sending a pulse output to the diagnostic device for 100 milliseconds when the condition was met. When this pulse is detected, the diagnostic device records it as "abnormal" in the test case.

The diagnostic device used was the Achilles Test Platform (ATP), a commercial product with a test suite for each protocol. The ATP does this by terminating the test case in progress and moving on to the next test case.

## 4.1 Evaluation Experiment 1: Number of Test Cases (General)

A comparison of the number of cloud-based automated tests was conducted by testing on the local network and by using the proposed method.

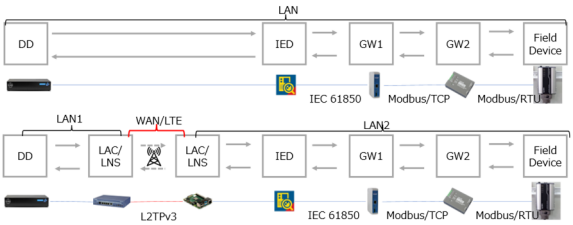



Figure 4: Test environment for the experiment 1.

The test suite used includes the following:

- Major grammar exams for Ethernet, IP, TCP
- Ethernet Load Testing
- IEC 61850 (MMS) testing

The test subject was Omicron IEDScout[7] as an Intelligent Electronic Device (hereinafter IED) simulator. The test input was limited to port TCP 102, and the monitoring covered network communication and the availability of port TCP 102.

As a condition for determining an abnormality, we set a loose condition that the monitored object should recover within 30 seconds after it is no longer captured so that the test case can be implemented as much as possible.

Under these conditions, the study was conducted three times with direct connection in a local network environment and via the cellular network (LTE) for comparison.

The results were almost identical in all cases, with the last test case conducted. There is no statistically significant difference in the number of test cases performed (Figure 5, Figure 6).

Figure 6 shows the number of test cases with anomalies. On average, more anomalies are detected when the TCP port is added to the monitoring conditions. This can be seen because of the difficulty in determining test results using the same port to which data is sent.

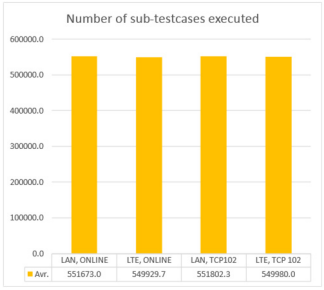



Figure 5:Number of sub-testcases executed.

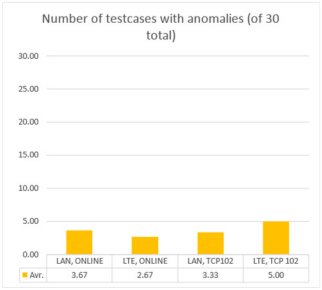



Figure 6: Number of test cases with anomalies.

## 4.2 Evaluation Experiment 2: Number of Test Cases (IEC 61850)

A comparison of the number of automated tests in industrial protocols was performed. As a change from Experiment 1, the software implementing IEC 61850 was changed to a different type, and the output interface of the device under test was included in the monitoring conditions (Figure 8).

The test set of ATP used includes the following:

- IEC 61850 (MMS) testing

The libiec61850[8] sample program was used as the test subject. The test input was limited to port TCP 102, and the monitoring was network sparsity, availability of port TCP 102, and digital output of IEDs (should always hold High). Other conditions were the same as in Evaluation Experiment 1.

Figure 8 shows the number of tests and the percentage containing anomalies. All of the tests eventually resulted in an abnormality. In fact, the software also ended abnormally. Although the number of tests conducted was the same each time, the number of anomalies found differed slightly (35, 30, 33 times). Still, as in Experiment 1, it is thought to contain some judgment errors.

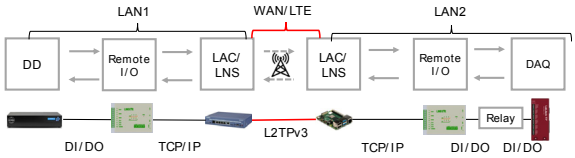



Figure 7: Test environment for the experiment 3.

[7] OMICRON, IEDScout, https://www.omicronenergy.com/en/products/iedscou/

[8] https://github.com/mz-automation/libiec61850

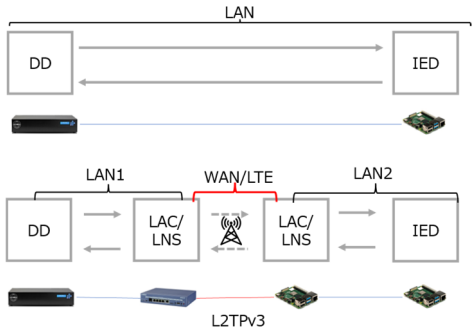



Figure 8: Test environment for the experiment 2.

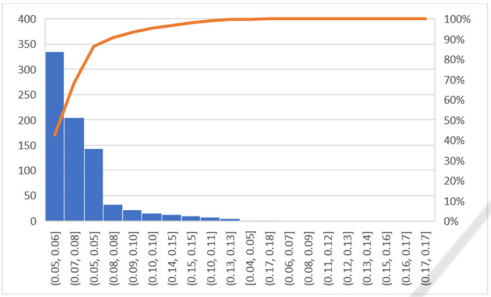


Figure 9: Latency in the environment 3.
The leftmost bar graph shows that there were about 340 ICMP exchanges between 0.05-0.06 seconds (about 45% of the total).

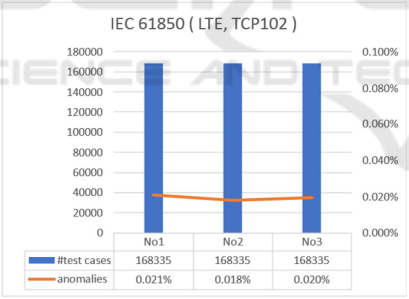



Figure 10: Number of tests resulted in anomalies.

### 4.3 Evaluation Experiment 3: Pulse Signal Monitoring

In evaluation experiment 3, a pulse signal monitoring test was carried out with the aim of verifying the system's capability to accurately capture short-duration communications and signals, such as touch panel operations. Figure 7 shows the test environment which is specialized for monitoring for evaluation of a pattern of virtual extension of CPS output to the cloud. A direct path from the DAQ to the diagnostic device is also set up for the comparison experiment. The diagnostic device was operated in manual test mode to assess the functionality of the monitoring system during the test. The output was transmitted to the diagnostic device through the cellular operator's network, and it was confirmed that the diagnostic device accurately detected the abnormality.

The average latency observed was 0.07 seconds, with the maximum latency reaching 0.18 seconds (occurring only once in a 10-hour period). The distribution of latencies is depicted in Figure 10. There were no instances where the pulse outputs were incompatible, and 100% of the pulses were successfully reported.

## 5 DISCUSSION

### 5.1 Relationship Between Evaluation Experiments and Challenges

Table 1 presents the results of the evaluation experiments conducted in Section 4 and summarizes the relationship between the experiments carried out in Section 4 and the challenges identified in Section 2.

(a) Based on experiments 1 and 2, it can be concluded that the aborted transmission of abnormal data due to undelivered transmitted data has been addressed. However, the network load test was not included in this evaluation, so additional evaluation on load reproducibility is necessary.

(b) Experiments 2 and 3 suggest that the system can handle events interrupting the transmission of abnormal data due to monitoring malfunctions of the output being diagnosed. Nonetheless, further evaluation is needed when dealing with more complex conditions, such as increasing the number of contacts to be monitored or monitoring the TCP port on the transmitting side.

(c) The issue where the transmission of abnormal data is interrupted due to a malfunction in monitoring the output to be diagnosed is also handled similarly to (b).

(d) The misjudgment of diagnostic results due to a monitoring malfunction of the output to be diagnosed does not fall under the risk category of misjudgement when the operator's operation is normal. Other confirmation methods are also considered applicable if they are fundamentally the same as (b) and (c), i.e., normal operation confirmation on transmitting and monitoring.

Table 1: Relationship between evaluation experiments and tasks.

| | Ex 1 | Ex 2 | Ex 3 | |
|---|---|---|---|---|
| (a)Interruption | * | * | | (a-1) The same tests were conducted as those in the local area.<br>(a-2) The same number of tests were conducted in the local area. |
| (b)Interruption | | * | * | (b-2) Equipment output was properly monitored.<br>(b-3) Anomalies as small as 0.1 second were also supplemented. |
| (c)Continuation | | * | * | (c-2) Equipment output was properly monitored.<br>(c-3) Anomalies as small as 0.1 second were also captured. |
| (d)Confirmation | | | * | (d-3) Operations of about 0.1 second were also captured. |

## 5.2 Applicability of This Method

The requirement for monitoring in fuzzing studies and certification schemes is "to cover all outputs." The scope is limited even for monitoring tests in local networks that are not cloud-based. When applying the virtual extension pattern to the cloud in the proposed method, it is possible to reproduce the same monitoring as for testing in a local area network by increasing the number of devices with remote I/O and DAQ. However, even test devices that are certified tools in the certification system have GUI and physical limitations on the number of monitored points. For example, a digital output is physically limited to four contacts. If many contacts are to be monitored to improve diagnostic accuracy, some kind of encoding and expansion of the monitoring GUI may be necessary.

In this test environment, latency was relatively small, and jitter issues did not affect the results. Further empirical research is needed to test the response to jitter errors in areas with weak radio waves and poor network conditions.

When considering actual use cases of diagnostics, it is also necessary to consider the increased burden on testers who temporarily set up DAQ and other equipment in the local system. As for the testers, they will have to carry in the monitoring equipment, install it, set it up, and witness the implementation of the test, which will be costly in proportion to the scale of the target to be monitored. When determining the parameters that define the monitoring conditions, if the parameters are optimized for the local environment, the burden on local workers will increase, which will not lead to a reduction in the cost of the test. Establishing a methodology to reduce the total cost, including the burden on workers, is also considered necessary.

Developing strategies to minimize the workload on local testers and optimize the overall testing process will be crucial in ensuring the success of the proposed method. This could involve creating comprehensive guidelines and automating certain aspects of the installation and monitoring processes to streamline the workflow. Additionally, exploring innovative ways to expand the monitoring GUI and improve diagnostic accuracy without overwhelming the system or personnel will be essential.

# 6 CONCLUSION

In this paper, we presented a novel method for remote security assessment of critical infrastructure systems using virtual extension design patterns. The proposed method addresses the challenges and limitations faced by traditional remote security assessment approaches, including network stability, reachability of Layer 2 data, and the monitoring of essential functions in real time.

Through the implementation and evaluation of the proposed method, we demonstrated its ability to effectively address these challenges in a simulated environment. The evaluation experiments showed promising results, with the proposed method successfully handling aborted transmission of abnormal data, ensuring proper output monitoring to be diagnosed, and maintaining consistent performance in various testing scenarios.

However, further research and development are needed to optimize the proposed method for real-world applications. This includes expanding the scope of monitoring, addressing latency and jitter issues, and minimizing the burden on local workers during the testing process. Developing comprehensive guidelines and automating certain aspects of the installation and monitoring processes will be essential in streamlining the workflow and reducing overall costs.

In conclusion, the proposed method offers a promising solution for the remote security assessment of critical infrastructure systems. By addressing the existing challenges and limitations, this method can significantly enhance cybersecurity and protect critical infrastructures in the face of ever-evolving threats.

# REFERENCES

Boehme, M., Cadar, C., & Roychoudhury, A. (2021). Fuzzing: Challenges and Reflections. In IEEE Software, 38(3), pp. 79-86. doi: 10.1109/MS.2020.3016773

Chen, Y., Poskitt, C. M., Sun, J., Adepu, S., & Zhang, F. (2019). Learning-guided network fuzzing for testing cyber-physical system defences. In Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (ASE '19). IEEE Press, 962–973. DOI: https://doi.org/10.1109/ASE.2019.00093

Luo, Z., Zuo, F., Jiang, Y., Gao, J., Jiao, X., & Sun, J. (2019). Polar: Function Code Aware Fuzz Testing of ICS Protocol. ACM Trans. Embed. Comput. Syst., 18(5s), Article 93, 22 pages. https://doi.org/10.1145/3358227

Matsuzaki, K., Sawabe, N., Maeda, R., Suzuki, D., Matsuura, T., & Hamada, H. (2020). Cybersecurity Evaluation Methodology for Distributed Energy Resources: Industrial Demonstration. In IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society. Singapore, 2020, pp. 2169-2174, doi: 10.1109/IECON43393.2020.9254422.

Pfrang, S., Meier, D., Friedrich, M., & Beyerer, J. (2018). Advancing Protocol Fuzzing for Industrial Automation and Control Systems. In Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018). DOI: 10.5220/0006755305700580

Serpanos, D., & Katsigiannis, K. (2021). Fuzzing: Cyberphysical System Testing for Security and Dependability. In Computer, 54(9), pp. 86-89. doi: 10.1109/MC.2021.3092479

Wilkerson, C., & Hariri, M. E. (2022). IEC 61850-Based Renewable Energy Systems: A Survey on Cybersecurity Aspects. In 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe). Prague, Czech Republic, 2022, pp. 1-6, doi: 10.1109/EEEIC/ICPSEurope54979.2022.9854539.