Uncover the Veil: The First Lesson We Learn from a Real-world Bulk Power System

Xi Qin, Neil Ortiz, Alvaro A. Cardenas @ University of California, Santa Cruz
Kelvin Mai @ University of Texas at Dallas
{xqin9, nortizsi, alacarde}@ucsc.edu; Kelvin.Mai@utdallas.edu

Baskin Engineering UC SANTA CRUZ

ABSTRACT

In the last two decades, the communication technologies used for supervision and control of critical infrastructures such as the power grid, have been migrating from serial links to Internet-compatible network protocols. Despite this trend, the research community has not explored or measured the unique characteristics of these industrial systems. In this research we perform the first study of a real Supervisory Control And Data Acquisition (SCADA) network in the power grid. We develop a new protocol parser that can be used to analyze packets not conforming to standards, find attributes to profile the SCADA network, and identify several anomalies which underscore the difficulties in designing an anomaly detection system for a federated network, where different devices are under the control of different power companies. Our work provides valuable insights for network monitoring and system state estimation of such systems.

RECENT WORK OF POWER GRID SECURITY

Systems	Protocols	System Type
Power distribution substations [2][5]	DNP3	Real-world distribution substations
Power grid [6]	DNP3	Emulated testbed
Power grid [7][8]	IEC 104	Emulated Testbed

CONTRIBUTIONS

- To our best knowledge, we are the first to characterize the SCADA network of a real bulk power grid, propose and testify three hypotheses that break the research community's presumptions about the SCADA network.
- We study the IEC 60870-5-104 (IEC 104) protocol and investigate the federation behaviors in this multi-administrative network for the first time.
- We show that the SCADA network in this work has unconventional behaviors than other SCADA networks, which helps the normal behavior baseline settings in the anomaly detection system.
- We also discover the trace of the operators incompletely upgrading the legacy protocols to new TCP/IP compatible standards. The protocol non-compliance issue leads to malformed IEC 104 packets.
- We have made our IEC 104 parser available [3] to the research community that can parse the above malformed packets.

BACKGROUND

Problem description:

- SCADA networks of Industrial Control Systems (ICS) are fragile to various malicious attacks and the network characterization is essential to build the anomaly detection system.
- IEC 60870-5-104 is one of the ICS protocols under attack during the 2016 Ukraine power outage [1].
- For security researchers, the real-world operational systems are hard to access.
- The closest previous work analyzes a much smaller power **distribution** system [2]. **Distribution systems** focus on smaller geographical areas such as cities.
- **Generation and transmission** represent **the bulk** of the power grid. The bulk system is much larger and more sophisticated with a federated network covering multiple power companies under different administrative domains.

➢ We got dataset from **a real Bulk Power Grid!!**

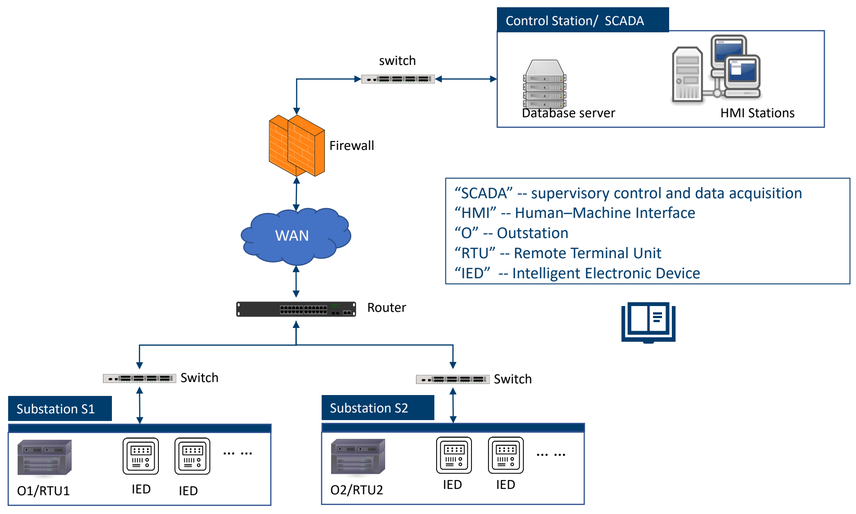


Figure 1. Schematic representation of a SCADA network in the power grid

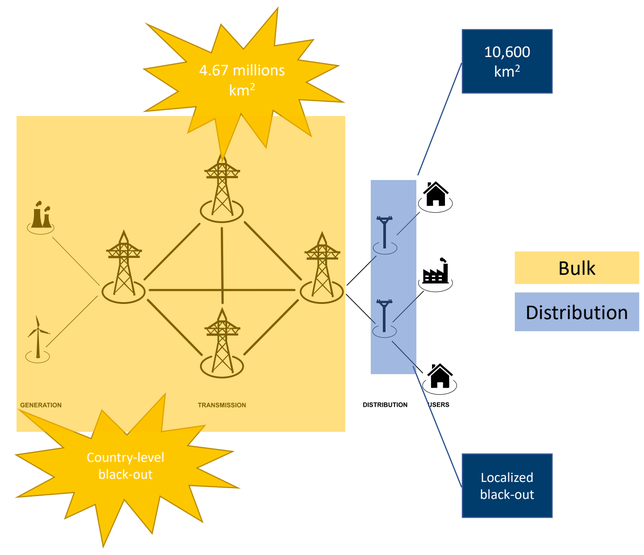


Figure 2. A bulk v.s. a distribution system: Differences in geographic and power black-out impacts

CONVENTIONAL INTERPRETATIONS OF SCADA V.S. OUR HYPOTHESES

The research community have understood the SCADA network as such:

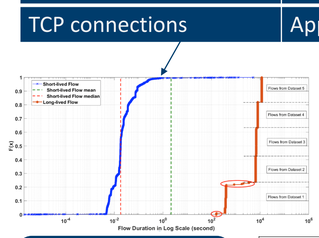
1. SCADA network is isolated from the Internet and can stay stable and predictable in decades with rare configuration changes. Therefore, the security study and network monitoring are not urgent.
2. The SCADA network has its own set of industrial control protocols. It's assumed that all the endpoint devices follow the communication standards to make the transmission contents readable. The network traffic should be easily parsed.

We have the following hypotheses based on our observations:

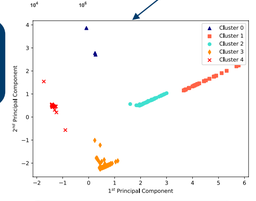
I. SCADA network has migrated to TCP/IP compatible network. It can have relatively frequent reconfiguration and upgrades within two years with significant topology changes. Many TCP connections between endpoints are instable and short-lived.
II. We expect to extract the communication patterns using the traditional deep packet inspection with the help of unsupervised machine learning methods. We can learn the underlying physical behaviors for process-based anomaly detection by examine the contents of process variables.
III. Not all the endpoint devices meet the protocol compliance. It leads to malformed packets in the network monitoring if not handled properly.

METHODOLOGY OF NETWORK CHARACTERIZATION [4]

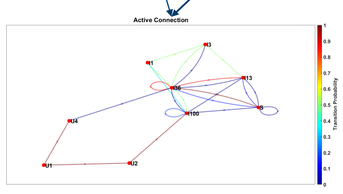
TCP Header	TCP Payloads	
TCP connections	Application Protocol Control Level	Application Protocol Data Level



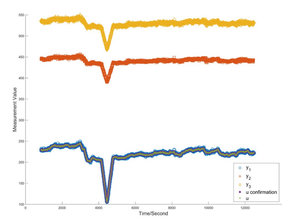
TCP connection duration analysis: CDF plot



Communication pattern method I: Clustering analysis



Communication pattern method II: Markovian Chain analysis



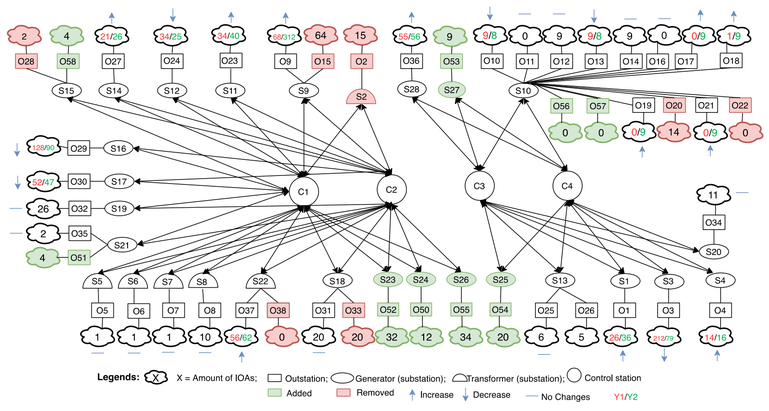
Physical behavior of process variables: Automatic Generation Control

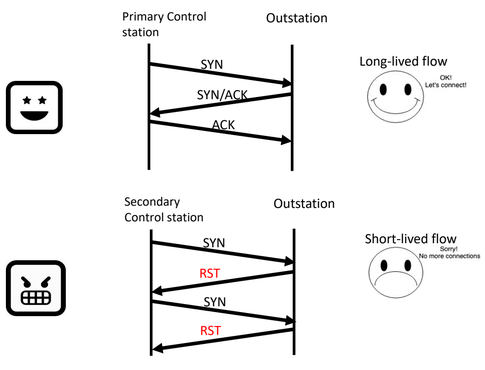
HYPHOTHESIS I TESTIMONY

Over two years, 75% of the power outstations didn't retain the connections.

Over 90% of the TCP connections last less than one second in the secondary connections. The flow white-listing in network monitoring should consider the network redundancy.

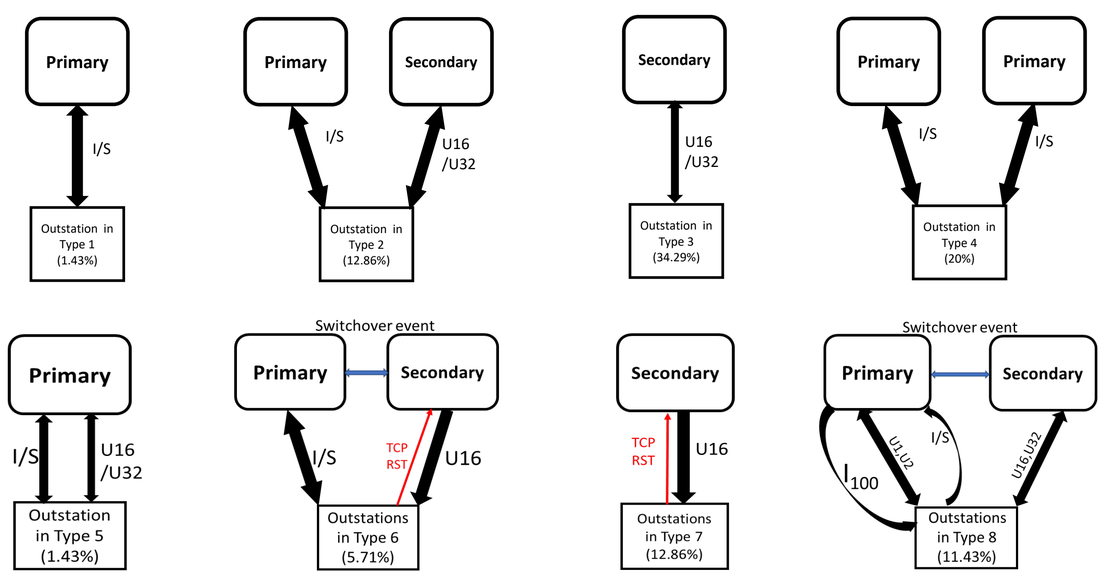
The changes of the SCADA network over two years with a one-year gap





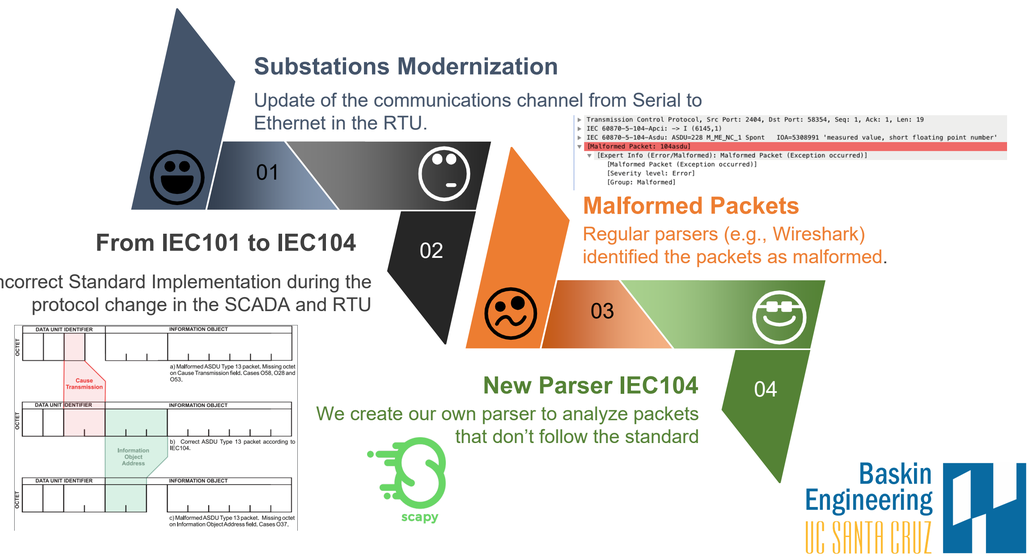
HYPHOTHESIS II TESTIMONY

Communication patterns extracted from Markovian Chains and clustering are potential in establishing the white rules for network monitoring of outstations.



HYPHOTHESIS III TESTIMONY

Industrial systems may retain certain non-standard legacy characteristics from previous technologies.



REFERENCES

[1] Greenberg, A. (2020). Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers. Anchor.

[2] Irvene, C., Shekari, T., Formby, D., & Beyah, R. (2019, December). If I Knew Then What I Know Now: On Reevaluating DNP3 Security using Power Substation Traffic. In Proceedings of the Fifth Annual Industrial Control System Security (ICSS) Workshop (pp. 48-59).

[3] Neil Ortiz. 2020. https://github.com/Cyphysecurity/IEC104_Parser.git

[4] Mai, K., Qin, X., Ortiz, N., Molina, J., & Cardenas, A. A. (2020, October). Uncharted Networks: A First Measurement Study of the Bulk Power System. In Proceedings of the ACM Internet Measurement Conference (pp. 201-213).

[5] Formby, D., Srinivasan, P., Leonard, A. M., Rogers, J. D., & Beyah, R. A. (2016, February). Who's in Control of Your Control System? Device Fingerprinting for Cyber-Physical Systems. In NDSS.

[6] Lin, H., Zhuang, J., Hu, Y. C., & Zhou, H. (2020, January). DefRec: Establishing Physical FunctionVirtualization to Disrupt Reconnaissance of PowerGrids' Cyber-Physical Infrastructures. In The Proceedings of 2020 Network and Distributed System Security Symposium (NDSS).

[7] Liu, Y., Ning, P., & Reiter, M. K. (2011). False data injection attacks against state estimation in electric power grids. ACM Transactions on Information and System Security (TISSEC), 14(1), 1-33.

[8] Lin, C. Y., & Nadjm-Tehrani, S. (2018, May). Understanding iec-60870-5-104 traffic patterns in scada networks. In Proceedings of the 4th ACM Workshop on Cyber-Physical System Security (pp. 51-60).

Acknowledgement: This material is based on research sponsored by Air Force Research Laboratory under agreement number FA8750-19-2-0010, and by NSF CNS 1929406.