Cybersecurity Anomaly Detection in SCADA-Assisted OT Networks Using Ensemble-Based State Prediction Model

Venkateswara Reddy Motakatla,[1] Jiazi Zhang,[1] Chen-Ching Liu,[2] Clifton Black,[3] Hongming Zhang,[4] and Seong Choi[1]

1 National Renewable Energy Laboratory
2 Virginia Polytechnic Institute & State University
3 Southern Company
4 Utilicast

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-84582
August 2023

Cybersecurity Anomaly Detection in SCADA-Assisted OT Networks Using Ensemble-Based State Prediction Model

Venkateswara Reddy Motakatla,[1] Jiazi Zhang,[1] Chen-Ching Liu,[2] Clifton Black,[3] Hongming Zhang,[4] and Seong Choi[1]

1 National Renewable Energy Laboratory
2 Virginia Polytechnic Institute & State University
3 Southern Company
4 Utilicast

Suggested Citation
Motakatla, Venkateswara Reddy, Jiazi Zhang, Chen-Ching Liu, Clifton Black, Hongming Zhang, and Seong Choi. 2023. *Cybersecurity Anomaly Detection in SCADA-Assisted OT Networks Using Ensemble-Based State Prediction Model*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-84582. https://www.nrel.gov/docs/fy23osti/84582.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-84582
August 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Electricity under the Cyber Research and Development Program. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Executive Summary

Cybersecurity threats to power systems are gradually increasing because of the increased sophisticated interactions between information technology (IT) and operational technology (OT) networks. A false data injection attack that aims to compromise the supervisory control and data acquisition (SCADA) measurements and disturb system operation is one such cybersecurity threat. Such attacks can potentially lead to significant operational issues at control centers and substations and hence result in severe physical consequences. To avoid catastrophic failures across the power grid resulting from these attacks, it is essential to arm the OT network with real-time vulnerability assessment tools. To this end, this paper outlines various drawbacks of the Purdue Enterprise Reference Architecture model to defend against cyberattacks in OT networks. Further, a novel ensemble-based state prediction model is proposed to detect cybersecurity anomalies in SCADA-assisted OT networks. The proposed model uses control center-level generation and load forecasts, scheduled, and forced outages, power flow solutions, and substation-level historical data. The hypothesis of the proposed scheme relies on the fact that additional control center and substation data can hardly be accessed and compromised by attackers. One vital feature of the proposed scheme is the future hour-ahead prediction of the operational feasibility of the SCADA measurement range at the control center and substation in real time that helps to detect anomalies in measurements across both the substation and the control center.

Table of Contents

List of Figures

1 Introduction

The growth of utility-scale renewable generation, distributed energy resources (DERs), and transportation electrification has increased uncertainty and cybersecurity risks in power grids. Existing power systems are monitored and controlled via the interaction of information technology (IT) and operational technology (OT) networks [1]. Toward this end, the utility industry has adopted the Purdue Enterprise Reference Architecture model-based Information Technology (IT) Cybersecurity to defend against cyberattacks in OT networks [2]. The Purdue model, which was first introduced in 1992, was initially for IT staff to enhance the understanding of how OT works and its data flow; however, the Purdue model has some shortcomings:

1. The Purdue model is modeled after manufacturing data (IT) flow and hence cannot capture the physical implication of the electric power flow monitored in the OT network. From the cybersecurity perspective, the Purdue model can only identify cyberattacks that manipulate IT data. Such attacks can introduce measurement/data discrepancies between the compromised data source and its neighbor. Sophisticated cyberattacks that capture the physical laws of the power system cannot be detected by the Purdue model and can potentially lead to severe physical consequences for power systems. For example, a dissatisfied modeling engineer could maliciously adjust the network topology, state variables, or reasonability check based on his/her system knowledge.
2. The Purdue model that works well in a hierarchical organizational structure will not be applicable to power grids with new grid participants—such as DER aggregators, electric vehicle owners, and behind-the-meter consumers—that are beyond the control of the utility company. With the new participants, the grid data flow will no longer follow a hierarchical structure.
3. Given the Federal Energy Regulatory Commission Order 2222 [3], new grid participants, including customers and aggregators, will increase. The Purdue model cannot ensure the trustworthiness of the data collected from the new participants because these data do not belong to the utility. The remote monitoring and control of the devices of the new grid participants will introduce more vulnerabilities to the existing utility cybersecurity systems [4]. From the control center perspective, it is difficult to identify whether any data/measurement discrepancies are from the uncertainty of the DER aggregators or from FDIA. From the DER aggregator perspective, this increases the ambiguity on the credibility of the generation dispatch set points from the control center, therefore, it is critical to develop a new cybersecurity mechanism to replace the Purdue model and to assist the OT network in mitigating these new challenges in power grids.

Early identification and mitigation of cyberattacks are necessary to avoid catastrophic failure across the power grid. The current power system anomalous data can be categorized into anomalies in the OT and the IT networks. OT network anomalies are associated with the physics of the power flow, such as faults, switching of loads, capacitor bank switching, and other local events. IT network anomalies do not follow the physics of the power flow; they include intentional data manipulation, sensor malfunctioning, and communication failures, such as bad data, missing data, and cyberattacks related to FDIA. Currently, the EMS state estimation (SE) module is used to detect OT anomalous data by following North American Electric Reliability Corporation (NERC) standards IRO-018 and TOP-010 [5]; however, current SE and bad data

detectors perform better in filtering bad measurements to ensure normal system operation than detecting anomalous data and data manipulation attacks.

In this report, we propose a novel ensemble-based state prediction model, the Ensemble Cybersecurity Model (ECM). The ECM goes beyond the traditional SE-based residual-type bad data detector. By following the NERC reliability standards, the ECM will detect both system and anomalous data. Besides the bad data detector in the SE, the ECM will take advantage of the system prediction data and set additional boundaries to quantify the allowable operational range in real time. Any compromised measurements that can bypass the SE bad data detector but result in dramatic deviations from the predicted operational boundary can be identified and sent to the control center by the ECM. Moreover, ECM anomaly detection can also identify the network topology and command mismatches and share the measurements, topology, and command data with substations; therefore, the ECM can address insider threats or disgruntled employees who might launch high-impact cyberattacks.

2 Overview of SCADA-Assisted OT Networks

The utility transmission control center consists of people, procedures, and technology tools to monitor reliable electric grid operation 24/7 throughout the year. The control center technologies have evolved over the past decades from measuring sensors to decision support.

Figure 1 illustrates the modules of the OT system in a control center. The key modules include the front-end processor, which collects point-to-point remote terminal unit (RTU) measurements from the field; the SCADA and/or phasor measurement unit (PMU), which monitors the electric system in real time and operates the breakers and switches remotely to change the system configuration; the EMS, which runs the network application, such as the SE and contingency analysis; the engineering workstations; and other supporting tools.

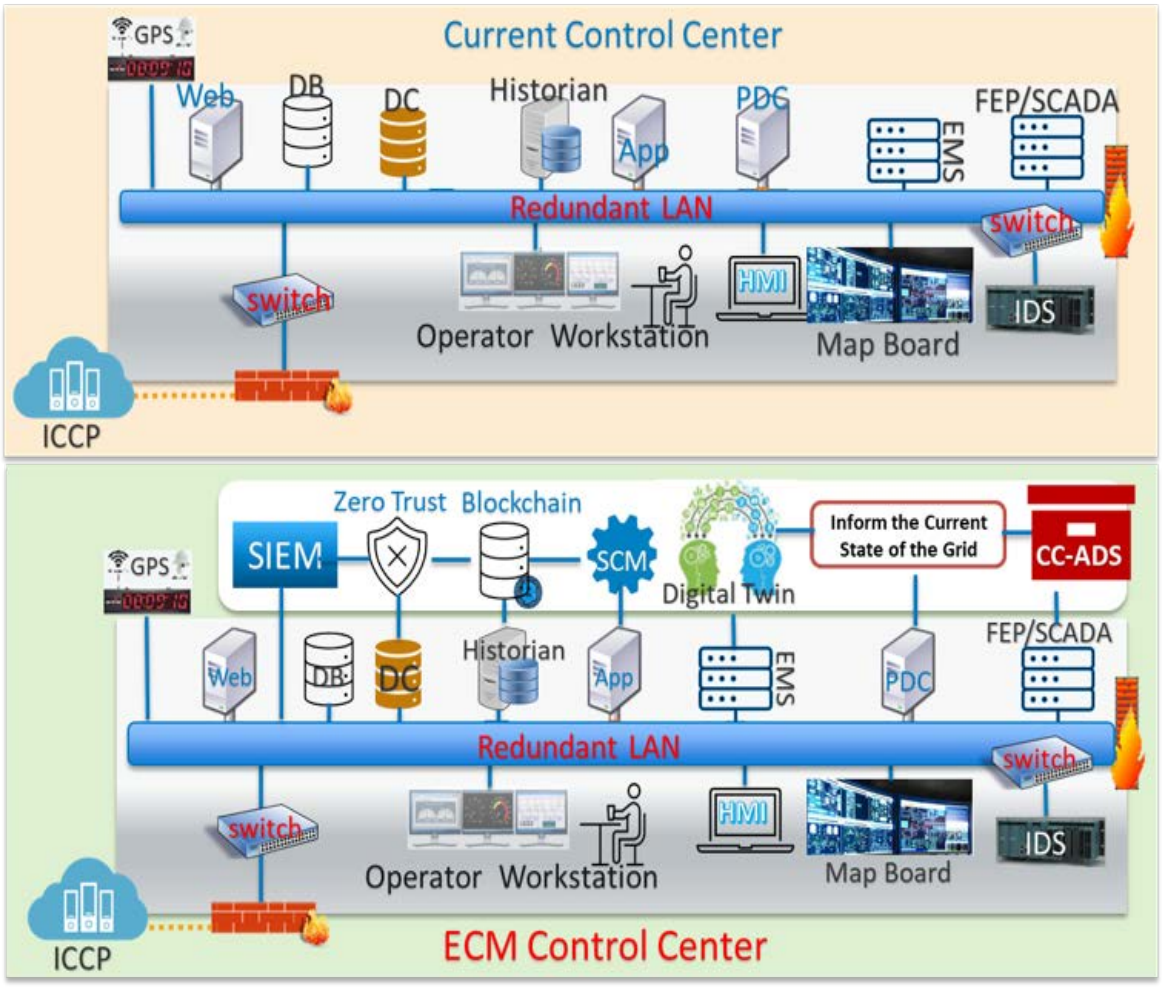

Figure 1. Comparison of the current control center and the ECM control center

*App: application server; PDC: phasor data concentrator; DB: database; SCM: system configuration management; DC: domain controller; SIEM: security information and event management; ICCP: Inter-Control Center Communication Protocol; CC-ADS: control center anomaly detection system; FEP: front-end processor.

Identifying cyberattacks at an early stage is usually difficult for control centers until after the utility networks are compromised and hence affect millions of customers and worsen upstream

connections. To a certain extent, private vendors—such as SIGA, Dragos, Darktrace, Honeywell, SCADAfence, and Omicron—have developed multiple software and hardware platforms to address cybersecurity challenges [6]. For operational support, various EMS tools are available in the market, such as General Electric's E-Terra, Siemens's Spectrum Power, Emerson's OSI Monarch, and ABB's Ability Network Manager. Although consistent efforts from researchers and vendors have been put into developing various approaches in the automatic detection of cyberattacks, system vulnerability assessments, asset inventories, and optimal corrective actions, there are still some challenges that need to be addressed.

2.1 IT-OT Convergence

Traditional power system control centers request that the OT networks be physically isolated from unsecured IT networks to ensure the system's cybersecurity. With the rapid power grid transformation and communication technology evolution, IT and OT networks are gradually converging. This trend, however, introduces new cybersecurity risks to the power grid, such as unauthorized access to OT networks via IT. Existing cyber defense mechanisms, including firewalls and intrusion detection system [7], have been developed to protect the IT and OT networks against such cybersecurity risks; however, the 2010 Stuxnet and 2015 Ukraine [8] cyberattacks indicate that the existing defense schemes are insufficient for the increasing cybersecurity threats.

2.2 Insider Threat

Existing cyberattack defense mechanisms in IT and OT networks can hardly protect control centers from insider attack threats. Insider attackers can be adverse system operators or external hackers that compromise access to the control center via social engineering attacks. Such attackers usually have more attack resources and relatively complete knowledge of the IT defense schemes as well as the OT operational modules; therefore, these attacks can result in more severe operational issues and attack consequences.

2.3 System Vulnerability Resulting from DERs

The increasing share of DERs introduces new vulnerabilities to control centers. Compared with the substation and traditional RTUs, the security protection level of the DER aggregator—which now has bidirectional communication with the control center—is much lower.

3 Proposed Ensemble Cybersecurity Model

This section describes the proposed ECM model that can address the shortages of the Purdue model, described in Section 1, and the new cybersecurity challenges, stated in Section 2. The model structure is illustrated in Figure 1 (ECM control center). Figure 2 demonstrates the key functions in the ECM, including the IT cyber defense, the OT operational out-of-band cyber defense, and the resilience support for autonomous recovery for the EMS. The following subsections describe the key functions in detail.

3.1 IT Cyber Defense

Compared with the Purdue model, a major innovation of the ECM control center is to ensure zero trust on each cyber layer. Such a scheme will make the new model "segmented." Conventionally, in a Purdue model-based SCADA network, if trusted, the network accepts at the lower layer (Layer 3–Layer 0). The hierarchical nature of the Purdue model makes the lower layer vulnerable because the trusted agent (or user) at Layer 3 could be a hacker with the compromised agent's credentials. The National Institute of Standards and Technology explained that "perimeter-based network security has also been shown to be insufficient since once attackers breach the perimeter, further lateral movement is unhindered." This conventional trust, i.e., no more verification once trusted, becomes more problematic because the pandemic has increased remote access. Zero trust architectures implement "don't trust, always verify" at each layer. This means that each layer should be segmented; however, zero trust architectures lack the NERC Critical Infrastructure Protection (CIP) system configuration change management requirement, which specifies who can make changes to the system and what changes are approved.

Figure 2. ECM key functions

3.2 OT Operational Out-of-Band Cyber Defense

For the OT network, the ECM employs an operational out-of-band cyber defense strategy based on a control center anomaly detection system (CC-ADS). Figure 3 illustrates the SE prediction model in CC-ADS and how CC-ADS can restrict the out-of-bound operation.

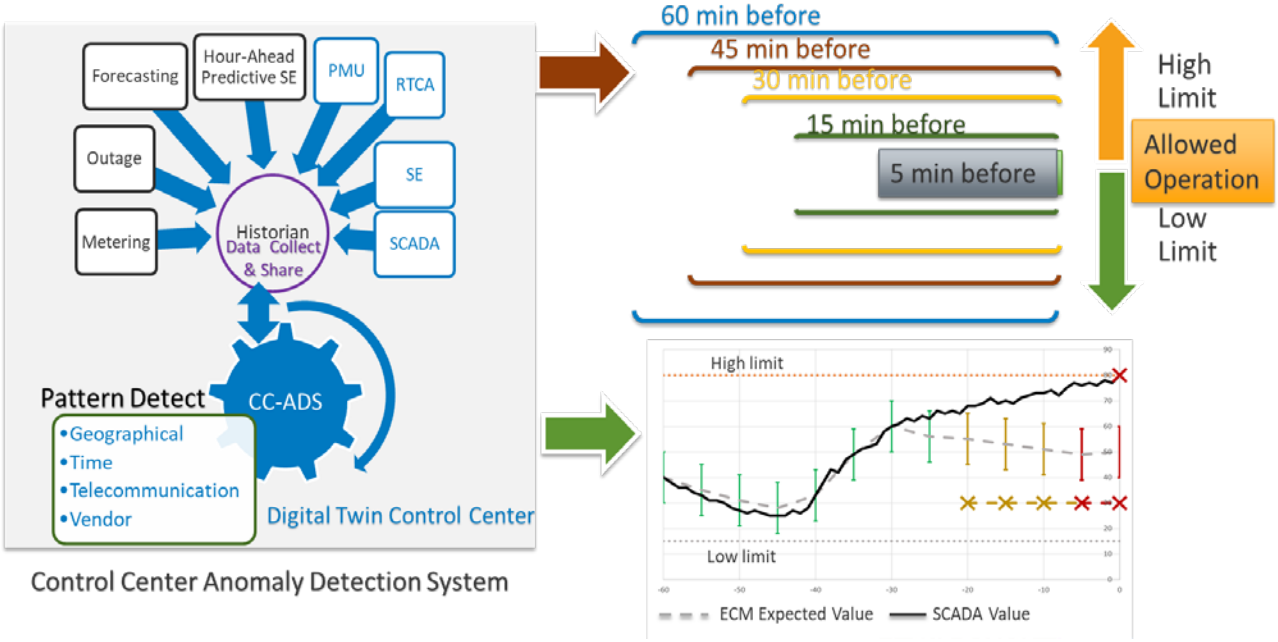

Figure 3. Proposed ensemble-based state prediction model of the operational feasibility of the SCADA measurement range

CC-ADS aims to identify the anomalies of the system configuration, measurement, and operational command in the raw data before the alarm system is triggered. By doing so, the advanced functionalities in the EMS framework, such as the SE and the real-time contingency analysis, will be shielded from the attacks. To achieve this goal, CC-ADS proposes a state prediction method by comparing the current SCADA value to the expected operational measurement range (EOMR), which includes five steps:

- **Step 1:** Synchronize the system data, including data from the SCADA, PMU, and SE estimated states.
- **Step 2:** Achieve EOMR by (1) forecast-based SE—such as future hour-ahead unit commitment, load forecast, scheduled interchange, and scheduled outages—given the current network topology and (2) the artificial intelligence/machine learning prediction from the historical data.
- **Step 3:** Check the consistency of the data. The consistency among the measurements will be checked and compared with the EOMR. If a SCADA value is outside the EOMR, then the ECM will transfer the data to the digital twin to evaluate three mismatches: (1) network topology mismatches at the utility area from the expected to the current, (2) measurement mismatches at each substation from the expected to the current, and (3) command mismatches at each substation comparing a number of commands issued from the control center to a number of commands received at the substation.

- **Step 4:** Perform blockchain immutability checks to ensure that the system configuration and operation command are valid and not manipulated by man-in-the-middle-attacks or insider attackers. Hyperledger Fabric blockchain, a network of the database, is used to host the system configuration change management records. With blockchain, the approval record for any system change must be verified before it takes place. Such approval records include "who is allowed," "what changes can be made," "on which device," and "what time frame should be executed before implementing the change." For the command mismatches, the ECM will provide a blockchain-based ledger that records the operator's study of why the command is issued, when to start, when to end, its impact, and at what substation. Because the operator must follow "select-and-operate" procedures to issue commands, substations will verify the commands via the blockchain-based ledger.
- **Step 5:** Restrict out-of-bound operational commands. Once the anomalies are detected, the control center will drop the suspicious data and operate the system within the upper and lower limits identified by the EOMR. Suspicious operational commands will be suspended and reassessed.

In addition to the EOMR introduced in Step 2, CC-ADS adds a new detection mechanism, the balancing authority area control error (ACE) limit (BAAL). BAAL is one of NERC's key reliability standards. For example, BAL-001 R2 states, "Each Balancing Authority shall operate such that its clock-minute average of Reporting ACE does not exceed its clock-minute Balancing Authority ACE Limit (BAAL) for more than 30 consecutive clock-minutes" [9]. CC-ADS compares the SCADA measurement against the BAAL limit to see whether the measurement is strengthening or harming the grid. Once the ACE is within the operational limits, by moving the interconnection frequency around 60 Hz [10], CC-ADS checks grid reliability by determining the balancing direction based on the generation and load direction.

3.3 Resilience Support: EMS Auto-Recovery

Another goal of the ECM is to quickly restore EMS operations from failure when a cyberattack disrupts both the primary and backup EMS. The autonomous recovery feature of an EMS is supported by digital-twin technology and predictive operational simulation on the cloud-based Dispatcher Training Simulator (DTS) under the operator's supervision and in compliance with the approved EMS recovery operating procedure, NERC CIP-009-3. A new control center EMS architecture is proposed to incorporate the digital-twin technology and to make the new model "collaborative." The architecture enables the key features of the EMS anomaly detection system and self-recoverability under fatal cyberattacks and/or massive Inter-Control Center Communications Protocol (ICCP) interruptions. Additionally, the ECM shares the operationally feasible measurement range with a substation where it compares the SCADA value with the control center's expected measurement.

3.4 Special DER Support Function

For DER support at the distribution level, the utility can make its private cloud available to each layer. Traditionally, in the Purdue model, the lower layer must pass data to the higher layer, e.g., from Layer 0 to Layer 1; however, the utility can operate a private cloud service that allows each layer direct access to a higher level, e.g., from Layer 0 to Layer 2, bypassing Layer 1. This will make the ECM nonhierarchical. Consumers, aggregators, and smart device vendors use their

public networks to share their data with an X.509 certificate. This public exposure attracts hacker activities. By providing the utility with a private cloud service, the utility can securely access nonprivate utility operational data by consumers, their aggregators, and smart device vendors and provide the data to an advanced distribution management system (ADMS).

4 Implementation, Benefit, and Impacts of ECM

4.1 Implementation

The demonstration project of the ECM will be implemented via NREL's Western Interconnection reliability coordinator control room simulator, Real-Time Analytics for Grids (RTAG) platform. RTAG was developed to evolve control room solutions for emerging challenges because of increased renewable generation penetrations and cybersecurity threats. It has adopted a production-grade West-Wide System Model (WSM) [11] and is integrated into the commercial DTS [12]. The proposed resilient EMS introduces a digital-twin technology [13], [14] approach for system integration between the EMS and RTAG to secure the EMS resilience and recovery capability from major cyberattacks. Artificial intelligence/machine learning will be applied to compare the EMS and RTAG DTS simulation results for the detection of control room operational anomalies.

4.2 Benefits and Impacts of Proposed Framework

The proposed ECM will greatly support system operators to obtain situational awareness of cybersecurity threats and increase trust in third-party measurement data using the control room operational knowledge base with historical EMS data, the digital-twin platform on a full Western Interconnection operation model, and a novel human-in-the-loop artificial intelligence technology integrated into a control room setting. The new EMS system with built-in CC-ADS will help U.S. utilities defend control rooms against various cyberattacks. Additionally, this will allow cybersecurity professionals and researchers to study the impacts of the ECM in EMS, ADMS, and DER management systems via simulation tools. Finally, the control room operational knowledge and the physics of electrical laws to detect anomalies pave the way for OT cybersecurity defense.

5 Concluding Remarks

This paper introduces an ensemble-based state prediction model to support power system OT networks for anomaly detection, cyberattack mitigation, and OT service autonomous recovery. The proposed ECM employs an IT and OT cybersecurity defense strategy embedded in a strong resilience support service to address the shortages of the Purdue model and hence protect power system OT networks against various cybersecurity threats. In addition to real-time measurements, other system data—including forecast-based SE, load forecast, scheduled interchange, scheduled outages, and artificial intelligence/machine learning prediction from historical data—are exhaustively used to generate EOMR to detect anomalies in measurements, system configurations, and operational commands. Advanced technologies—including zero trust architecture, digital twins, blockchains, and private clouds—have been used in the ECM to ensure comprehensive support for the OT network.

The implementation of the ECM via NREL's RTAG platform is in progress. With support from the ECM, power system control centers can better manage the deployment of massive amounts of renewable generation and protect power systems against the increased threat of cyberattacks.

References

1. "IEEE Standard Cybersecurity Requirements for Substation Automation, Protection, and Control Systems," in *Proc. 2015 IEEE Std C37.240-2014*, pp.1-38, Jan. 30, 2015.

2. G. Rathwell, "PERA master planning guide," Enterprise Consultants, Ltd., http://www.pera.net/Pera/PERA_Guide/PERA_Guide.pdf.

3. E. Zhou, D. Hurlbut, and K. Xu, "A primer on FERC Order No. 2222: Insights for international power systems," National Renewable Energy Laboratory, Golden, CO, Tech. Rep., NREL/TP-5C00-80166, 2021, https://www.nrel.gov/docs/fy21osti/80166.pdf.

4. U.S. Department of Energy, "Cybersecurity considerations for distributed energy resources on the U.S. electric grid," Office of Cybersecurity, Energy Security, and Emergency Response and Office of Energy Efficiency and Renewable Energy, Washington, D.C., https://www.energy.gov/sites/default/files/2022-10/Cybersecurity%20Considerations%20for%20Distributed%20Energy%20Resources%20on%20the%20U.S.%20Electric%20Grid.pdf.

5. D. Dolezilek and L. Hussey, "Requirements or recommendations? Sorting out NERC CIP, NIST, and DOE cybersecurity," in *Proc. 2011 IEEE 64th Annual Conference for Protective Relay Engineers*, pp. 328-333, doi: 10.1109/CPRE.2011.6035634.

6. J. Cirelly, "7 best OT security vendors," *Comparitech*, Aug. 11, 2022, https://www.comparitech.com/net-admin/ot-security.

7. G. M. Makrakis, C. Kolias, G. Kambourakis, C. Rieger, and J. Benjamin, "Industrial and critical infrastructure security: Technical analysis of real-life security incidents," *IEEE Access*, vol. 9, pp. 165295-165325, 2021.

8. "Lessons learned from a forensic analysis of the Ukrainian power grid cyberattack," Interchange, https://blog.isa.org/lessons-learned-forensic-analysis-ukrainian-power-grid-cyberattack-malware.

9. NERC, "Real power balancing control performance," April 16, 2015.

10. NERC, "BAL-001-2 – Real power balancing control performance standard background document," Feb. 2013.

11. H. Zhang et al, "Extending MOD-033 methodology to create chronological WECC basecase for NTRR analysis," presented at the WECC Model Validation Subcommittee (MVS) meeting, April 2021. https://www.wecc.org/_layouts/15/WopiFrame.aspx?sourcedoc=/Administrative/Zhang%20-%20MOD-033%20Methodology%20to%20Create%20Chronological%20WECC%20Basecases.pdf&action=default&DefaultItemOpen=1.

12. S. V. Vadari, M. J. Montstream, and H. B. Ross Jr., "An online dispatcher training simulator function for real-time analysis and training," *IEEE Trans. Power Syst.*, vol. 10, no. 4, pp. 1798-1804, Nov. 1995.

13. A. M. Madni, C. C. Madni, and S. D. Lucero, "Leveraging digital twin technology in model-based systems engineering," *Systems*, vol. 7, no. 1, pp. 7, 2019, https://www.mdpi.com/2079-8954/7/1/7.

14. K. M. Alam and A. El Saddik, "C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems," *IEEE Access* 5, pp. 2050–2062, 2017.