Report submitted to the U.S. Department of Energy (DOE)
Office of Electricity Delivery and Energy Reliability
National Energy Technology Laboratory

Award # DE-OE0000830

Project Title: *Autonomous Tools for Attack Surface Reduction*

Final Project Report

Submitted by Iowa State University
March 2021

Key Project Personnel:

Manimaran Govindarasu, Iowa State University (ISU), Principal Investigator
Adam Hahn, Washington State University
Philip Hart, GE Global Research (GE-GR)
Aditya Ashok, Pacific Northwest National Laboratory (PNNL)
Hyekyung (Clarisse) Kim, Argonne National Laboratory (ANL)
Gelli Ravikumar, ISU
Vivek Kumar Singh, ISU
Pengyuan Wang, GE-GR
Bev Johnson, PNNL
Paul Krukow, Cedar Falls Utilities (CFU)
Jianzhe Liu (ANL)
Adam Peterson, CFU

Table of Contents:

1. Project Rationale

The electric power grid is a complex critical infrastructure that forms the lifeline of modern society, and its secure and reliable operation is of paramount importance to national security and economic well-being. However, recent findings documented in authoritative sources indicate the threat of cyber-based attacks growing in numbers and sophistication. However, securing the grid against stealthy cyber-attacks is a challenging task due to legacy nature of the infrastructure coupled with dynamic nature of threat landscape and ever growing sophistication of the adversaries. Additionally, the grid's attack surface continues to grow with the increased dependence on digital communications and control that now extends to each consumer through smart meters and distributed energy resources. Unfortunately, this expansive surface increases the grid's vulnerability and further exposes critical control systems in both substations and control centers.

To respond to this emerging need, we had successfully assembled an interdisciplinary team with academic- industry partnership to successfully conduct research, development, evaluation, demonstration, and commercialization of attack surface reduction tools, whose goal is to significantly reduce the cyber attack surface in the North American power grid. Our proposed project was a synergistic collaborative effort leveraging the synergistic expertise of the team members across power systems, cyber security and CPS security, testbeds, field deployments and demonstration, and successful commercialization. The team consisted of leading experts from two major universities – Iowa State University, Washington State University – complemented by reputed researchers from two DOE national laboratories – Pacific Northwest National Lab, and Argonne National Lab, one major utility vendor GE Global Research, and one utility partner – Cedar Falls Utilities (CFU). The team members have proven track record of successful academic-industry collaboration in interdisciplinary R&D projects, and bring onboard some of the best state-of-the-art testbed resources, industry-grade SCADA/EMS/DMS environment for experimentation and field demonstration.

2. Technical Approach

The following are the specific tasks have been successfully completed two phases (2016-2020).

Phase I:

Task 1: Developed and implemented a robust Project Management and Data Management Plan, coupled with a well thought out Risk Mitigation Plan.

Task 2.1: Developed *a comprehensive framework* that continually assesses and autonomously reduces the attack surface for the power grid control environment spanning across substations, control center and the SCADA network to significantly reduce the risks of cyber attacks.

Task 2.2: Developed *attack surface analysis techniques, metrics, and tools* that assess the attack surface at multiple levels including the control center, substations, and the SCADA network.

Task 2.3: Developed *attack surface reduction techniques and tools* that dynamically reduce attack surface and hence increase attacker's cost without interfering in the critical functions of the system.

Task 2.4: Prototyped, implemented, and quantitatively evaluated/validated the *techniques and tools* on a *realistic industrial CPS security testbed environment* by leveraging the unique resources of the team.

Task 3: *Developed Commercialization plan to transition* the developed tools into power system industry stakeholders for a broader adoption by leveraging the expertise of our industrial members.

Phase II:

Task 4: *Completed Field demonstration, verification, and evaluation* of the effectiveness of the attack surface analysis and reduction techniques on a realistic utility testbed environment. This also involved the development of realistic scenarios, sound metrics, data sets, evaluation criteria, and documentation.

3. Relevance to real-world scernarios

The recent cyber attack that targeted the Ukrainian power grid demonstrates the need for techniques to reduce system attack surfaces. This incident demonstrated pervasive vulnerabilities found within modern power systems, where critical control centers present a major attack target and whose compromise can directly cause a major blackout. The interconnectivity of these critical systems with corporate networks, neighboring utilities/ISO, remote substations, and increasingly deployment of smart meters and non- utility owned DERs will present an increased level of risk from future attacks. The proposed tasks will make modern power systems more resilient to sophisticated attacks, such as those exhibited in Ukraine. The following table identifies how the proposed tasks are aligned to prevent against pragmatic real-world threats such as those observed from the Ukraine incident 2015.

Ukraine Attack Element	Proposed Research Tasks
1) Compromise Corporate Network	Attack Surface Analysis Metrics and Tools **[*Task 2.2*]**
2) Accessed SCADA Network	Network-based Moving Target Defense (MTD) in Substation Networks **[*Task 2.3.2*]**
3) Infect Control Center	Anomaly Detection for EMS/DMS Network Application through Machine-learning @ Control Center **[*Task 2.3.3*]**
4) Inject Malicious Breaker Commands	CPS-based MTD @ Substations **[*Task 2.3.1*]**

4. Metrics and Performance Evaluation

The project included evaluation plans and associated metrics to provide evidence that the proposed techniques achieve their intended objectives and will transition to field demonstration and industry adoption. The evaluation plan includes benchmarks for the security, performance, fidelity, and reliability of the proposed technologies. Evaluations wwas be performed within cyber-physical security testbeds at ISU, WSU, GE-GR which provided realistic environments with Hardware-In-the-Loop power system simulations and SCADA/EMS/DMS software platforms. Furthermore, the second phase of the project showcased field demonstration and testing within Cedar Falls Utilities (CFU) power distribution grid environment and also integrating some of the attack reduction tools/analysis within GE Energy Management System (EMS) platform, and also OSI-Soft adopting attack surface analysis tool (Attack Host Analyzer) within their software development environment.

5. Outcomes and Impacts

The project had significantly advanced the state-of-the-art research and practice in improving the cybersecurity of our nation's power grid infrastructure against cyber threats. In particular, the proposed, designed, and deployed attack surface analysis and reduction algorithms and tools have contributed to significantly reducing the exposure and risk of the devices, substations, and the integrated SCADA/EMS/ DMS grid environment to cyber threat. Strong demonstration and evaluation techniques have verified the feasibility of the developed techniques on realistic cyber-physical testbeds and utility's real grid environment (CFU) and collaborative research and evaluation of attack surface reduction techniques (for wide-are monitoring and control) within a vendor (GE) EMS platform.

5.1 Distributed IDS Design, Implementation and Field Demonstration

A distributed, network-based IDS architecture (shown in Fig 1) was developed and deployed in the CFU network. The IDS Master (at Control Center) that collects all the alerts from the IDS Sensor nodes deployed at various substations, and visualizes those alerts on the *Sguil* dashboard, was deployed within the control center. There were 5 sensor nodes that were deployed at various locations where substations are present to detect alerts on the substation network traffic. Of the 5 sensor nodes, one was dedicated for testing purposes on a lab network that had a DNP3 device installed to simulate the actual network traffic. All the sensors trigger on the rule sets that are configured within them and send the alerts to the IDS Master. The alerts are propagated to the Master from the sensors using SSH. Having the sensor VMs communicate to the Master through SSH is a prerequisite for this IDS architecture to work.

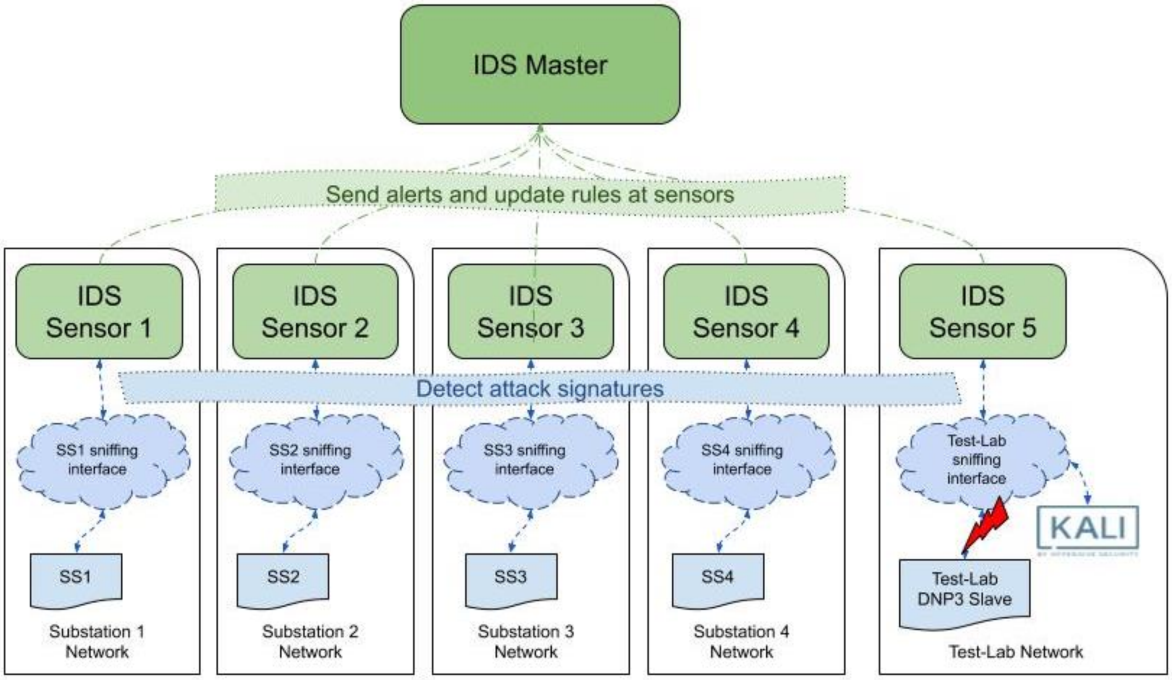


Figure 1: Architecture of Distributed IDS Deployment in CFU Distribution Grid Network

The tech transfer activity had its own challenges, a few of which are mentioned below: (1) Automating deployment of sensor-specific rules at individual sensors; (2) Dynamic behavior of analog measurement response packets as a result of change events reported; (3) Presence of CRC checksum for every 16 bytes; (4) Testing validity of analog measurement rules in the production network without injecting penetration testing attack scripts. This report details the tech transfer activity on IDS for DNP3 communication protocol between Iowa State University and Cedar Falls Utility carried out during the fall 2020. This field deployment showcased the successful deployment IDS architecture, the different IDS rule categories (network rules, status measurement rules, and analog measurement rules) and the various test phases that were carried out and displayed the triggered intrusions or anomalies at the IDS Master running at the Control Center.

5.2 Summary of Key Research Publications resulted from this Project

[1] P. Wang, M. Govindarasu, "Multi-Agent based Attack-Resilient System Integrity Protection for Smart Grid," IEEE Trans. on Smart Grid, vol. 11, no. 4, pp. 3447-3456, July 2020.

Abstract: Most System Integrity Protection (SIP) schemes deployed in smart gird today are centralized functions relying on wide-area communication. The highly centralized implementation makes SIP susceptible to the single point of failure induced by cyber attacks. In this paper, we present a novel multi-agent based design to enhance the cyber resilience of SIP while focusing on augmenting its situational awareness and self-adaptiveness. Specifically, we have investigated data-driven anomaly detection and adaptive load rejection within the decentralized SIP set-up. After attaining a comprehensive taxonomy of operation states of a power grid as a cyber-physical system, we are able to convert the anomaly detection to a multi-class classification problem. A supervised learning algorithm, named as Support Vector Machine embedded Layered Decision Tree (SVMLDT), is proposed as a possible solution. Anomaly detection is carried out by every agent separately, but the final decision depends on the consensus among all interconnected agents. Besides, we propose an adaptive load rejection strategy to mitigate the Denial of Service (DoS) attacks targeting the load shedding scheme. A real load rejection SIP scheme adopted by Salt River Project is modified to fit in the IEEE 39-bus model as a study case. Experiment results show that the proposed SIP can detect anomalous grid operation states and then adjust its remedial actions accordingly to adapt to the under-attack situations.

[2] G. Ravikumar, M. Govindarasu, "Anomaly Detection and Mitigation for Wide-Area Damping Control using Machine Learning," IEEE Trans. on Smart Grid, Early Access, 2020.

Abstract: In an interconnected multi-area power system, wide area measurement based damping controllers are used to damp out inter-area oscillations, which jeopardize grid stability and constrain the power flows below to their transmission capacity. The effect of wide-area damping control (WADC) significantly depends on both power and cyber systems. At the cyber system layer, an adversary can inflict the WADC process by compromising either measurement signals, control signals or both. Stealthy and coordinated cyber-attacks may bypass the conventional cybersecurity measures to disrupt the seamless operation of WADC. This paper proposes an anomaly detection (AD) algorithm using supervised Machine Learning and a model based logic for mitigation. The proposed AD algorithm considers measurement signals (input of WADC) and control signals (output of WADC) as input to evaluate the type of activity such as normal, perturbation (small or large signal faults), attack and perturbation-and-attack. Upon anomaly detection, the mitigation module tunes the WADC signal and sets the control status mode as either wide-area mode or local mode. The proposed anomaly detection and mitigation (ADM) module works in-line with the WADC at the control center for attack detection on both measurement and control signals and eliminates the need for ADMs at the geographically distributed actuators. We consider coordinated and primitive data-integrity attack vectors such as pulse, ramp, relay-trip and replay attacks. The performance of the proposed ADM algorithms was evaluated under these attack vector scenarios on a testbed environment for 2-area 4-machine power system. The ADM module shows effective performance with 96.5% accuracy to detect anomalies.

[3] V. K. Singh and M. Govindarasu, "A Cyber-Physical Anomaly Detection for Wide-Area Protection using Machine Learning," IEEE Trans. on Smart Grid, Early Access, 2021.

Abstract: Wide-area protection scheme (WAPS) provides system-wide protection by detecting and mitigating small and large-scale disturbances that are difficult to resolve using local protection schemes. As this protection scheme is evolving from a substation-based distributed remedial action scheme (DRAS) to the control center-based centralized RAS (CRAS), it presents severe challenges to their cybersecurity because of its heavy reliance on an insecure grid communication, and its compromise would lead to system failure. This paper presents an architecture and methodology for developing a cyber-physical anomaly detection system (CPADS) that utilizes synchrophasor measurements and properties of network packets to detect data integrity and communication failure attacks on measurement and control signals in CRAS. The proposed machine leaning-based methodology applies a rules-based approach to select relevant input features, utilizes variational mode decomposition (VMD) and decision tree (DT) algorithms to develop multiple classification models, and performs final event identification using a rules-based decision logic. We have evaluated the proposed methodology of CPADS using the IEEE 39 bus system for several performance measures (accuracy, recall, precision, and F-measure) in a cyber-physical testbed environment. Our experimental results reveal that the proposed algorithm (VMD-DT) of CPADS outperforms the existing machine learning classifiers during noisy and noise-free measurements while incurring an acceptable processing overhead.

[4] G. Ravikumar, B. Hyder, J. Rajan Babu, K. Khanna, M. Govindarasu, and M. Parashar, "CPS Testbed Architectures for WAMPAC using Industrial Substation and Control Center Platforms and Attack-Defense Evaluation," 5 pages, IEEE PES General Meeting, 2021.

Abstract: Advanced persistent threats and cyberattacks can impact wide-area monitoring, protection, and control (WAMPAC) system operation. Many cyber-physical system (CPS) testbeds have been developed for attack-defense experimentation and attack-resiliency tools evaluation for WAMPAC, but they are limited to a simulation-and-emulation based environment. This paper presents a quasi-realistic CPS attack-defense testbed based framework for WAMPAC applications using the industrial substation and control center platforms such as eTerra integrated with the hardware-in-the-loop CPS smart grid testbed available at Iowa State University. The proposed framework includes various combinations of industry-grade substation and control center platforms, communication topologies, real-time digital simulators, and a novel cyber-physical distributed intrusion-and-anomaly detection system (D-IADS) for WAMPAC applications. The D IADS includes a master at the control center and geographically distributed sensor devices at each substation. Each D-IADS sensor deployed at a substation or control center network monitors ingress and egress traffic, detect intrusions, and dispatch alerts to the D-IADS master. The D-IADS master centrally monitors and analyze the alerts and controls D-IADS sensors. We considered an EMP60 synthetic CPS grid as a case study to demonstrate the framework and proposed D-IADS for WAMPAC applications against cyberattack vectors such as Man-in-the-Middle DNP3 attack, denial-of-service, and data-integrity attacks.

[5] V. K. Singh and M. Govindarasu, "A Novel Architecture for Attack-Resilient Wide-Area Protection and Control System in Smart Grid," Resilience Week Symposium (RWS), 5 pages, 2020.

Abstract: Wide-area protection and control (WAPAC) systems are widely applied in the energy management system (EMS) that rely on a wide-area communication network to maintain system stability, security, and reliability. As technology and grid infrastructure evolve to develop more advanced WAPAC applications, however, so do the attack surfaces in the grid infrastructure. This paper presents an attack-resilient system (ARS) for the WAPAC cybersecurity by seamlessly integrating the network intrusion detection system (NIDS) with intrusion mitigation and prevention system (IMPS). In particular, the proposed NIDS utilizes signature and behavior-based rules to detect attack reconnaissance, communication failure, and data integrity attacks. Further, the proposed IMPS applies state transition-based mitigation and prevention strategies to quickly restore the normal grid operation after cyberattacks. As a proof of concept, we validate the proposed generic architecture of ARS by performing experimental case study for wide-area protection scheme (WAPS), one of the critical WAPAC applications, and evaluate the proposed NIDS and IMPS components of ARS in a cyber-physical testbed environment. Our experimental results reveal a promising performance in detecting and mitigating different classes of cyberattacks while supporting an alert visualization dashboard to provide an accurate situational awareness in real-time.

[6] S. Mohan, G. Ravikumar, and M. Govindarasu, "Distributed Intrusion Detection System using Semantic-based Rules for SCADA in Smart Grid," IEEE Transmission and Distribution (T&D) Conference, 5 pages, 2020.

Abstract: Cyber-physical system (CPS) security for the smart grid enables secure communication for the SCADA and wide area measurement system data. Power utilities world-wide use various SCADA protocols, namely DNP3, Modbus, and IEC 61850, for the data exchanges across substation field devices, remote terminal units (RTUs), and control center applications. Adversaries may exploit compromised SCADA protocols for the reconnaissance, data exfiltration, vulnerability assessment, and injection of stealthy cyberattacks to affect power system operation. In this paper, we propose an efficient algorithm to generate robust rule sets. We integrate the rule sets into an intrusion detection system (IDS), which continuously monitors the DNP3 data traffic at a substation network and detects intrusions and anomalies in real-time. To enable CPS-aware wide-area situational awareness, we integrated the methodology into an open-source distributed-IDS (D-IDS) framework. The D-IDS facilitates central monitoring of the detected anomalies from the geographically distributed substations and to the control center. The proposed algorithm provides an optimal solution to detect network intrusions and abnormal behavior. Different types of IDS rules based on packet payload, packet flow, and time threshold are generated. Further, IDS testing and evaluation is performed with a set of rules in different sequences. The detection time is measured for different IDS rules, and the results are plotted. All the experiments are conducted at Power Cyber Lab, Iowa State University, for multiple power grid models. After successful testing and evaluation, knowledge and implementation are transferred to field deployment.

[7] J. Ulrich, J. Drahos, and M. Govindarasu, "A symmetric address translation approach for a network layer moving target defense to secure power grid networks," Resilience Week (RWS), pp. 163-169, 2017.

Abstract: This paper will suggest a robust method for a network layer Moving Target Defense (MTD) using symmetric packet scheduling rules. The MTD is implemented and tested on a Supervisory Control and Data Acquisition (SCADA) network testbed. This method is shown to be efficient while providing security benefits to the issues faced by the static nature of SCADA networks. The proposed method is an automated tool that may provide defense in depth when be used in conjunction with other MTDs and traditional security devices.

[8] V. Kumar Singh, S. P. Callupe, and M. Govindarasu, "Testbed-Based Evaluation of SIEM Tool for Cyber Kill Chain Model in Power Grid SCADA System," North American Power Symposium (NAPS), 6 pages, 2019.

Abstract: Development of a smarter electric grid necessitates addressing the associated cyber security challenges. Since the interdependence between the legacy grid infrastructure and advanced information technology is growing rapidly, there are numerous ways advanced, motivated, and persistent attackers can affect the SCADA based critical infrastructure. Hence, developing a security information and event management (SIEM) is crucial for securing the SCADA power system. This paper presents the application of Security Onion (SecOn) to develop the network security monitoring (NSM) and intrusion detection system (IDS) in the context of SCADA cyber physical security. Initially, we have applied a cyber kill-chain model to demonstrate the different stages of attacks and associated mechanisms. Later, the rule based IDS (RIDS) is developed using Snort IDS, and tested in the cyber-physical SCADA environment. Furthermore, we have evaluated its performance in terms of accuracy and detection latency. Our experimental results reveal that the SecOn tool is efficient in monitoring and detecting attacks within an acceptable time frame with a high accuracy rate.

[9] V. Kumar Singh, H. Ebrahem, and M. Govindarasu, "Decision Tree Based Anomaly Detection for Remedial Action Scheme in Smart Grid using PMU Data," 5 pages, IEEE PES General Meeting, August 2018.

Abstract: The advanced and persistent cyber threats facing the critical infrastructure such as the smart grid are exponentially rising which require sophisticated defense strategy. Remedial Action Scheme (RAS), also known as Special Protection Scheme (SPS), relies on the interconnected cyber physical system for automated protection which is exposed to the multitude of vulnerabilities. In this paper, we propose an innovative approach to develop an Intelligent Remedial Action Scheme (IRAS) which can detect and distinguish cyber attacks from the physical disturbances in smart grid and later take smart corrective actions as required to minimize the impact on system reliability and economy. Specifically, we have proposed the decision tree based anomaly detection methodology which can distinguish between the normal tripping during power line faults and malicious tripping attack on the physical relays in the context of RAS. The classification model is developed using differential features of voltage and current phasors. Next, as a proof of concept, we have implemented and validated the proposed methodology in cyber physical environment at Iowa State's PowerCyber testbed. Finally, the proposed methodology is tested on modified IEEE 39 bus system in offline and real-time mode. Our experimental results show that the proposed method is efficient in detecting attacks and performing corrective actions within an acceptable time frame.

[10] P. Wang and M. Govindarasu, "Anomaly Detection for Power System Generation Control based on Hierarchical DBSCAN," North American Power Symposium (NAPS), 6 pages, 2018.

Abstract: The generation level of a generator in power grid is under control of multiple control loops such as governor action, Automatic Generation Control (AGC), and manual control. These control loops are vulnerable to cyber attacks and could become potential targets for the malicious adversary. In this paper, we investigated the detection of abnormal generation controls induced by cyber attacks. One data-driven anomaly detection methodology is presented based on the behavior conformity of generation units located in the same balancing authority. A semi supervised clustering algorithm with Hierarchical Density based Spatial Clustering of Application with Noise (HDBSCAN) is proposed for the detection model training. In our previous work, we obtained a synthetic dataset that covers scenarios including normal operation, generation control under ramp attack, switching attack, AGC integrity attacks, etc. and it is used to evaluate the proposed clustering algorithm. Experimental results show that the proposed algorithm provides better detection accuracy than K-means clustering and can distinguish not only between normal and abnormal generation controls but also among various anomaly scenarios.

[11] V. Kumar Singh, H. Ebrahem, and M. Govindarasu "Security Evaluation of Two Intrusion Detection Systems in Smart Grid SCADA Environment," North American Power Symposium (NAPS), 6 pages, 2018.

Abstract: The increased complexity and interconnectivity of SCADA infrastructure in the power system have exposed it to the multitude of vulnerabilities. There is a growing emphasis towards developing an efficient intrusion detection system (IDS) to strengthen the security of the SCADA control system. This is

a research-in-progress paper which presents the application of two anomaly-based intrusion detection systems (AbIDS) in detecting the stealthy cyber-attack on the SCADA control system. We have applied the IDS tools Snort and Bro, in designing the IDS and later, compared their performances in terms of detection rate and latency in the alert packets with a motive of selecting better IDS for the SCADA security. Specifically, the timing-based rule is applied to identify the malicious packets based on the high temporal frequency in the network traffic. For the case study, we have implemented the SCADA based protection scheme which performs an autonomous protection to mitigate the system disturbances. We first implemented the stealthy cyber-attack which compromised the SCADA controller followed by data integrity attack on the system generator. Next, we perform the impact analysis during the attack followed by performance evaluation of IDS tools. Our experimental results show that the IDS tools are efficient in detecting cyber-attacks within an acceptable time frame for different sizes of network packets.

[12] V. Kumar Singh, M. Govindarasu, "Evaluation of Anomaly Detection for Wide-Area Protection Using Cyber Federation Testbed," 5 pages, IEEE PES General Meeting, August 2019.

Abstract: Cyber physical security research for smart grid is currently one of the nation's top R&D priorities. The existing vulnerabilities in the legacy grid infrastructure make it particularly susceptible to countless cyber-attacks. There is a growing emphasis towards building interconnected, sophisticated federated testbeds to perform realistic experiments by allowing the integration of geographically-dispersed resources in the dynamic cyber-physical environment. In this paper, we present a cyber (network) based federation testbed to validate the performance of an anomaly detector in context of a Wide Area Protection (WAP) security. Specifically, we have utilized the resources available at the Iowa State University Power Cyber (ISU PCL) Laboratory to emulate the substation and local center networks; and the US Army Research Laboratory (ARL); to emulate the regional control center network. Initially, we describe a hardware-in-the loop based experimental setup for implementing data integrity attacks on an IEEE 39 bus system. We then perform network packet analysis focusing on latency and bandwidth as well as evaluate the performance of a decision tree based anomaly detector in measuring its ability to identify different attacks. Our experimental results reveal the computed wide area network latency; bandwidth requirement for minimum packet loss; and successful performance of the anomaly detector. Our studies also highlight the conceptual architecture necessary for developing the federated testbed, inspired by the NASPI network.

[13] G. Ravikumar, A. Nicklaus, and M. Govindarasu, "Cyber-Physical Smart Light Control System Integration with Smart Grid using Zigbee," IEEE Innovative Smart Grid Technology (ISGT) Symposium, 5 pages, 2020.

Abstract: This paper presents a hardware-in-the-loop cyber-physical system architecture design to monitor and control smart lights connected to the active distribution grid. The architecture uses Zigbee-based (IEEE 802.15.4) wireless sensor networks and publish-subscribe architecture to exchange monitoring and control signals between smart-light actuators (SLAs) and a smartlight central controller (SLCC). Each SLA integrated into a smart light consists of a Zigbee-based endpoint module to send and receive signals to and from the SLCC. The SLCC consists of a Zigbee-based coordinator module, which further exchanges the monitoring and control signals with the active distribution management system over the TCP/IP communication network. The monitoring signals from the SLAs include light status, brightness level, voltage, current, and power data, whereas, the control signals to the SLAs include light intensity, turn ON, turn OFF, standby, and default settings. We have used our existing hardware-in-the-loop (HIL) cyber-physical system (CPS) security SCADA testbed to process signals received from the SLCC and respond suitable control signals based on the smart light schedule requirements, system operation, and active distribution grid dynamic characteristics. We have integrated the proposed cyber-physical smart light control system (CPSLCS) testbed to our existing HIL CPS SCADA testbed. We use the integrated testbed to demonstrate the efficacy of the proposed algorithm by real-time performance and

latency between the SLCC and SLAs. The experiments demonstrated significant results by 100% real-time performance and and low latency while exchanging data between the SLCC and SLAs.

[14] V. Kumar Singh, A. Ozen, and M. Govindarasu, "A hierarchical multiagent-based anomaly detection for wide-area protection in smart grid," 6 pages, Resilience Week (RWS), 2018.

Abstract: Future smart grid capabilities provide assurance to expand the advanced information and communication technologies to evolve into densely interconnected cyber physical system. Remedial Action Scheme (RAS), widely used for wide-area protection, relies on the interconnected networks and data sharing devices, which are exposed to the multitude of vulnerabilities. This paper presents our proposed approach to developing multi-agent based RAS scheme against the system aware stealthy cyber-attacks. Specifically, we propose the two-level hierarchical architecture which consists of distributed local RAS controllers (RAScs) as local agents, operating at different zones/ areas, which are constantly monitored by an overseer, the central agent. The local controllers receive local and randomly changing outside zonal measurements and cyclically forwards to the overseer. The overseer identifies the corrupted controller using the anomaly detection algorithm which processes the measurements coming from the local controllers, compute measurement errors using local and outside zonal measurements, perform validation checks, and finally detect anomalies based on the two-step verification. Next, as a proof of concept, we have implemented and validated the proposed methodology in cyber physical environment at Iowa State's PowerCyber testbed. We have also implemented the coordinated attack vectors which involve corrupting the local controller and later performing stealthy attacks on the system's generator. We have evaluated its performance during the online testing in terms of detection rate and latency. The experimental results show that it is efficient in detecting different classes of attacks, including ramp and pulse attacks.

[15] G. Ravikumar, B. Hyder, and M. Govindarasu, "Efficient Modeling of HIL Multi-Grid System for Scalability & Concurrency in CPS Security Testbed," North American Power Symposium (NAPS), 6 pages, 2019.

Abstract: Cyber-event-triggered power grid blackout compels utility operators to intensfy cyber-aware and physics-constrained recovery and restoration process. Recently, coordinated cyber attacks on the Ukrainian grid witnessed such a cyber-event triggered power system blackout. Various cyber-physical system (CPS) testbeds have attempted with multitude designs to analyze such interdependent events and evaluate remedy measures. However, resource constraints and modular integration designs have been significant barriers while modeling large-scale grid models (scalability) and multi-grid isolated models (concurrency) under a single real-time execution environment for the hardware-in-the-loop (HIL) CPS security testbeds. This paper proposes a meticulous design and effective modeling for simulating large-scale grid models and multi-grid isolated models in a HIL real-time digital simulator environment integrated with industry-grade hardware and software systems. We have used our existing HIL CPS security testbed to demonstrate scalability by the real time performance of a Texas-2000 bus US synthetic grid model and concurrency by the real-time performance of simultaneous ten IEEE-39 bus grid models and an IEEE-118 bus grid model. The experiments demonstrated significant results by 100% real time performance with zero overruns, low latency while receiving and executing control signals from SEL Relays via IEC-61850 protocol and low latency while computing and transmitting grid data streams including stability measures via IEEE C37.118 synchrophasor data protocol to SEL Phasor Data Concentrators.

[16] Philip Hart, Sowmya Acharya, Honggang Wang, "Coherency-Based Detection Algorithm for Synchrophasor Cyberattacks, North American Power Symposium (NAPS), 6 pages, 2019.

Abstract: The wide area monitoring system (WAMS) is critical for power system situational awareness, but represents a growing cybersecurity vulnerability. Malicious adversaries may seek to compromise one or more PMUs in order to effect control decisions that unnecessarily disrupt typical grid operations. One example of a particularly pernicious attack vector is the spoofing or replaying of a fault event using one or more compromised PMUs. This work documents the development and validation of a coherency-based cyberattack detection algorithm that integrates a sliding-window singular value decomposition (SVD) with physics-based partitioning analysis to achieve accurate classification of events. Special consideration is given to discerning a sophisticated fault-replay or fault spoofing attack from actual faults. A software-based cybersecurity testbed has been developed for rigorous testing of the algorithm. The algorithm is further validated using simulated synchrophasor datasets obtained from a MinniWECC 63-bus test system. Results show that the algorithm can successfully detect fault replay attacks even when over half of the PMUs are compromised.

[17] Pengyuan Wang, Honggang Wang, Philip Hart, Xian Guo, and Kaveri Mahapatra, "Application of Chebyshev's Inequality in Online Anomaly Detection Driven by Streaming PMU Data,'' 5 pages, IEEE PES General Meeting, August 2020.

Abstract: The day-to-day operation of modern power systems is highly reliant on prompt and adequate situational-awareness. This can be achieved via various system monitoring functions such as anomaly detection, in which static thresholds are commonly utilized to distinguish the normal and the abnormal system states. However, a predetermined static threshold usually lacks the flexibility to adapt to unobserved scenarios. In this paper, we propose two self-adaptive synchrophasor data driven anomaly detection approaches based on Chebyshev's Inequality. The proposed approaches have been evaluated with Kundur's 2- area system and Mini-WECC system. Experimental results verify that the proposed approaches can dynamically adapt to unprecedented scenarios, and detect anomalous events with lower false alarm rate compared to static threshold based detection.

5.3 Technical Report from GE Global Research

The rest of this document contains GE-GR's final technical report for this project.

Autonomous Tools for Attack Surface Reduction

WAMS Cyberattack Detection Algorithm: Final Report

Contract:
DE-OE0000830

Prepared for:
Dr. Manimaran Govindarasu
Iowa State University
2520 Osborn Dr.
Ames, IA 50011-1046
gmani@iastate.edu

Philip Hart[1]
Pengyuan Wang[1]
Honggang Wang[1]
Xian Guo[1]
Sowmya Acharya[1]
Kaveri Mahapatra[1]

[1]GE Research

Version 1
June 15, 2020

Acknowledgement This material is based upon work supported by the Department of Energy National Nuclear Energy Technology Laboratory under Award Number(s) DE-OE0000830.

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Contents

List of Figures

List of Tables

Nomenclature

CEDS	Cybersecurity for Energy Delivery Systems
DOE	Department of Energy
EMS	Energy Management System
GER	General Electric Research
ISU	Iowa State University
PNNL	Pacific Northwest National Laboratory
PMU	Phasor Measurement Unit
SVD	Singular Value Decomposition
WAMS	Wide Area Monitoring System

1 Phase I: WAMS Algorithm Development and Evaluation

1.1 Introduction and Motivation

Widespread deployment of phasor measurement units (PMUs) in the transmission system has been instrumental in achieving unprecedented levels of situational awareness. Synchrophasor data from PMUs in the Wide Area Monitoring System (WAMS) facilitates the operation of a growing number of control room Energy Management System (EMS) tools, including, but not limited to, rapid and robust linear state estimation, online monitoring of voltage stability, and analysis of poorly-damped oscillations. However, since important control decisions are executed based on input from this monitoring system, synchrophasor measurements represent a large and growing cybersecurity vulnerability. It is conceivable that an adversary with malicious intent will seek to compromise one or more PMUs in order to effect control decisions that unnecessarily disrupt typical grid operations. For a system targeted by a malicious entity, the systems attack surface is comprised of (i) methods, (ii) channels, and (iii) data items [1]. Historically, attack surface metrics have been proposed that exhaustively quantify these elements and weight them according to potential impact on system operation. In context of the power grid, the control center and its EMS applications and control infrastructure comprise the system to be defended by the proposed algorithm. WAMS synchrophasor measurements, standardized according to IEEE C37.118.1, provide a key channel for illicit entry of the malicious adversary into the control center.

1.2 Objective

The objective of this work is to successfully develop and validate an online cyberattack detection algorithm for synchrophasor data that reduces the attack surface of the control center and its WAMS-dependent software applications. This work focuses, in particular, on developing an online tool that detects cyberattacks on the physical layer of the WAMS network, i.e. the content within the synchrophasor measurements. It is assumed that any information technology (IT) and operational technology (OT) cyber defense mechanisms have been compromised by the adversary, and that the content of the synchrophasor measurements, possibly including voltage & current phasors and/or frequency from one or more PMUs, has been altered by a malicious adversary in order to negatively impact power systems operation. A strong emphasis has been placed on the capability of the algorithm to discern actual physical fault events (e.g., short circuits, line faults, etc.) from spoofed synchrophasor that mimics fault-like behavior, including replay attacks or spoofed fault signals.

Figure 1 shows a high-level overview of the developed algorithm. Whether implemented as a component of the phasor data concentrator (PDC) data quality reporting service or as a preconditioning step within WAMS applications at the control center, the algorithm is intended to operate nearly autonomously from the control center and has minimal interaction with the EMS. The algorithm reduces the attack surface of the control center by cyber-hardening a key channel into the EMS that would otherwise be available to the malicious adversary. In this sense, the proposed algorithm differs from many other synchrophasor cyberattack detection techniques in the literature which necessitate close interaction with the EMS tool, such as state-estimation-based anomaly detection methods. It is assumed that the algorithm will be applied to aggregated synchrophasor dataset from two or more PMUs. It is also assumed that at least one–but not all–PMUs may have been compromised by the adversary.

In contrast to many other state-of-the-art, power grid cyberattack detection algorithms, the focus here is on the measurement and control system associated with an entire power grid, rather than on the measurement and control system associated with a single component (e.g., a power plant) within the power grid. Additionally, since this attack surface reduction tool can operate autonomously on PMU channels outside

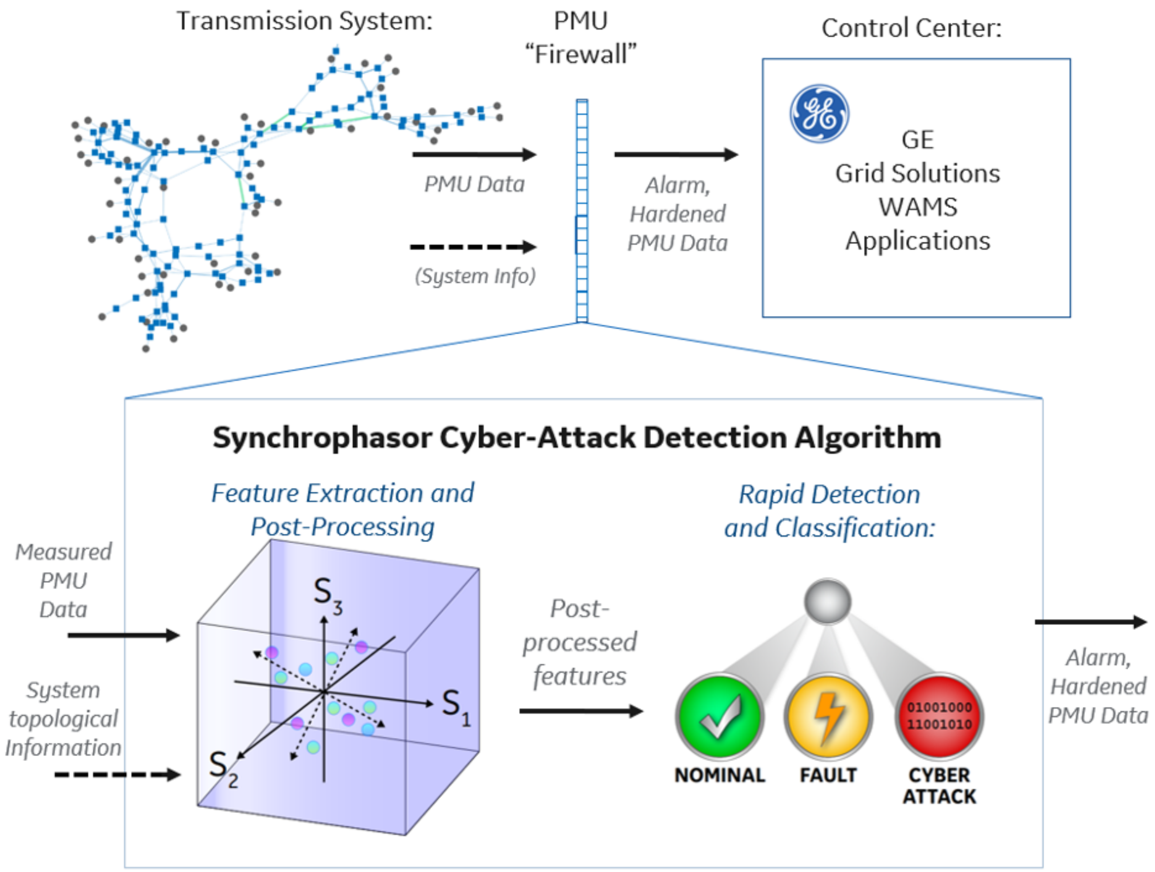


Figure 1.1: Developed algorithm reduces the attack surface of WAMS-dependent EMS applications.

of or at the extreme periphery of the EMS, communication of system information between the algorithm and the EMS is inherently limited. Details of the power system commonly available to control room operators, such as the estimated network topology and location of sensors, generators, and loads, may not be made continuously available to the algorithm for purposes of feature engineering and training. Therefore, feature engineering will instead make extensive use of the known physical properties of power networks, depending upon spatiotemporal correlations and physics-based principles that broadly apply to generic 3-phase, sparsely-connected, reactive power grids of arbitrary size and interconnection, under various states of time-varying load and generation.

Considering the above discussion, it is desired that the algorithm have the following characteristics:

1. High accuracy: High true positive and true negative rates for cyberattack and fault classification.

2. Autonomy: it is desired to minimize the requirement for continuous interaction between the proposed algorithm and the EMS, and minimize the effort needed for training of the algorithm, allowing the algorithm to conveniently reside at the PDC or at the extreme periphery of EMS applications.

3. Ease of Implementation: Due to the rarity of severe physical fault events within the power grid that can be used for online training, complete reliance on purely data-driven methods for online training may limit the utility of the algorithm. Due to Characteristic #2, model-based training of the algorithm using control center data should be limited and infrequent.

1.3 Review of PMU Cyberattack Detection Techniques

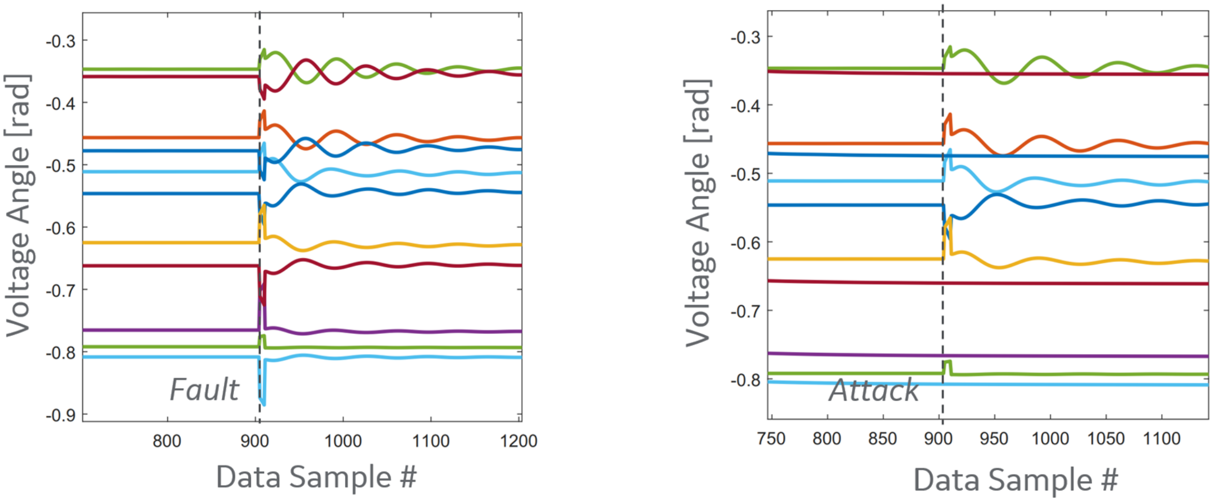
The detection and elimination of bad data from synchrophasor measurements in the context of power system state estimation has received considerable attention in literature [2]. The detection of cyberattacks warrants dedicated approaches, as they are hard to detect using the conventional methods for detecting bad data [3]. The survey in [4] lists different types of cyberattacks and defenses against them in the context of power systems.

Intrusion detection techniques can be broadly classified into three categories: (i) signature-based (or misuse-based), (ii) anomaly-based, and (iii) specification-based techniques [5]. Whereas signature-based techniques compare the current activity to signatures of known attacks, anomaly detection focuses on characterizing normal behavior, and seeks to identify any departures from normality. Specification-based techniques compare ongoing events to patterns of events that have been specified, through manual effort, to be normal. Alternatively, from a different standpoint, detection techniques can also be classified using a dichotomy of model-based versus data-driven. Model-based approaches exploit system-specific information like network parameters and topology information in the detection algorithm, whereas the data-driven models only rely on the characteristics of measurement data. A state-estimation based algorithm is presented in [6] for identifying angle biases and current scaling errors in the phasor data using the augmented state vector approach. These errors can arise from issues with the global positioning system (GPS), timing circuits, instrument channels, and/or data channel scaling. Pre-processing of PMU data before state-estimation to eliminate data quality issues is achieved with the help of Kalman filtering in [7]. The algorithm is able to identify instances of bad data like data dropouts, repeated values and replace them with values generated using the filter. A mathematical model of cyber-physical attacks is developed in [8] by modelling the power system as a linear time-invariant control system. Centralized and distributed attack detection and identification monitors are designed with the help of the proposed model. The algorithms in [6], [7], and [8] are model-based and depend on availability of accurate models for the system under consideration. The authors of [9] recognize the low-rank property of PMU data matrices and formulate a convex optimization problem for the detection of data substitution attacks. The matrix factorizations and optimization render this data-driven algorithm computationally expensive, therefore not amenable for online attack detection. The spatio-temporal correlations inherent in phasor measurements are exploited in [10] for online detection of bad data including cyberattacks. The bad data points are formulated as spatio-temporal outliers as compared to other measurements, and a density-based outlier detection technique is used to detect them. A neural network-based method is proposed in [11] which associates 108 features to each transmission line where PMUs are installed. Deep auto-encoders tuned with normal data present higher reconstruction errors in the presence of manipulated data, and a threshold test is used to detect the anomalies. For successful identification, multiple such encoders monitoring small subsets of PMUs should be installed across the network.

A cyberattack on the physical layer of the WAMS network can have a number of manifestations. More than one PMU may be compromised by the attacker. The synthetic synchrophasor data injected into a compromised PMU channel may appear as zero values, or as dropped packets, or sudden changes in magnitude and/or phase angle. Depending upon the level of sophistication of the attack, the signals may or may not respect the physics of the power system (e.g., maintain consistency with Kirchhoffs voltage and current laws). It is acknowledged that in the case that all PMUs are compromised by the attacker, it can be become nearly impossible to detect the attack. Nonetheless, detection algorithms should be developed to detect all but the most sophisticated and comprehensive attacks, particularly in the case that the injected signal is highly likely to impact control decision making. From this standpoint, a particularly pernicious attack vector is one that spoofs or replays the transient that occurred during a fault event, from one or more compromised PMUs. Once a PMU, or a communication channel from a PMU, has been compromised, a hacker may compel this unit or channel to transmit to the phasor data concentrator (PDC) pre-recorded phasor data resembling that which would have been observed during a real fault event. This particular attack vector could easily trick a human operator, or even an automated controller, to implement protective actions, potentially resulting in lost load. In the case of a fault replay attack on a large number of PMU channels, the injected signals may even display a superficial consistency with network physics, making it difficult for an experienced human operator to detect–at little additional cost to the attacker. Fig. 1.2(a), below, shows the transient that ensues at as a result of a fault event, at 11 buses at which PMUs are deployed. A replay attack on just over half of the PMUs is shown in Fig. 1.2 (b). If given only a brief time period following the event in Fig. 1.2(b) to make a control decision, an operator may easily be convinced that an actual fault event has occurred. Automated decision-making tools may be also be susceptible to this attack.

The focus of this work is on replay attacks where the attacker modifies the synchrophasor data by replacing measurements with data which correspond to past system state(s). The authors of [12] propose a detection algorithm which periodically injects a randomized signal to the measurements and use a linear time-invariant model of the system to detect replay attacks on smart-meter data. The data-driven approach presented in [13] introduces a metric called self-correlation coefficient for detection of replay attacks in power systems. The approach leverages the fact that the replay attack portions of the measurements show more periodicity than normal measurements. Reference [14] uses data mining techniques to detect cyberattacks including replay attack. The key concept used is called common path mining, which works by checking if a certain sequence of events has occurred in the prescribed manner. The detection system is first given a representative set of paths which serve as signatures for normal operating scenarios. The actual measurements are then compared to these signatures to flag anomalies, if any. Reference [15] proposes a method which monitors the intra-PMU and inter-PMU correlations between different measurement quantities. Injection of spoofed data alters the normal values of these correlations and can be detected. The true-positive rate detected by the authors ranges between 77% and 86% for different scenarios.

Since PMUs can transmit new samples 60 times a second, this data can quickly accumulate and grow to unmanageable proportions. The sheer volume of sensor data from PMUs, SCADA sensors, and smart



(a) Physical fault

(b) Fault replay cyberattack

Figure 1.2: Simulated PMU node angles in the aftermath of a fault transient (a) and after a fault replay cyberattack (b) for 11 buses within a power system (angles are unwrapped and detrended).

meters can be a hindrance to grid operators who require timely and actionable information. Over 2000 PMUs have been installed in the United States [16], and MISO has access to 3500 synchrophasor measurements from 489 PMUs [17]. Using conventional sampling rates, the MISO PMU network alone might generate over 100 megabytes of synchrophasor data every minute, and multiple terabytes over the course of a month. Improved techniques are still needed to distill the sensor data and reduce it to timely, actionable information, in the context of real-time cybersecurity monitoring. Regardless of the type of intrusion detection algorithm used, features of the PMU data must be down-selected by the algorithm for it to successfully classify an event.

A powerful method for reducing a large dataset into a small number of important features is principle component analysis (PCA). Authors in [18] and [19] use principle component analysis of PMU data for event detection. A primary goal of their work is to reduce the dimensionality of PMU data in order to more conveniently and efficiently identify changes in the operating condition of the system, such as changes in the system topology or changes in generator inputs. It was observed that a low-dimensionality system underlies a PMU data matrix containing just voltage magnitude and phase angle information. The algorithm for event detection requires training using historical PMU data to determine the PMUs that have the most impact on the singular values of the system: these pilot PMUs subsequently play a predominate role in event detection, in the online application of the algorithm.

One important conclusion from this review is that there is still a significant research gap associated with WAMS cyberattack detection algorithms that can operate autonomously from the control center, that can leverage both data-driven and model-based training, and that can accurately detect fault replay attacks or spoofed fault signals.

1.4 Algorithm Development

To achieve high classification accuracy, autonomy, and ease of implementation, the algorithm developed by GE Global Research (GEGR) integrates data-driven and model-based approaches for rapid and convenient training. The algorithm is capable of being trained online and can function entirely in a purely data-driven manner, if necessary. Alternatively, online training can be supplemented or replaced entirely by model-based training, using an infrequent, expedient physics-based analysis of the power system, if topological data is made available to the algorithm by the EMS. The following sections provide background regarding two techniques that are instrumental in the operation of the cyberattack detection algorithm, including (i) the singular value decomposition (SVD); and (ii) an example model-based power system partitioning analysis tool.

1.4.1 Background on PCA and SVD

Let $\mathbf{X}$ represent an $n \times m$ data matrix, where m represents the number of variables and n represent the number of observations. It is assumed that $\mathbf{X}$ is a ‘centered’ data matrix; i.e., to obtain $\mathbf{X}$, subsequent to the initial gathering of the data, the mean of each column of values has been subtracted from each element in that column.

The covariance matrix is denoted by $\mathbf{C_x}$:

$$\mathbf{C_x} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X} \tag{1}$$

In $\mathbf{C_x}$, the diagonal element (k, k) represents the variance of the samples of variable k; the off-diagonal element i, j represents the co-variance between variables i and j.

In the case that off-diagonal element (i, j) of $\mathbf{C_x}$ is large, this indicates that an obvious linear relationship has been observed between the two variables i and j. In a sense, it also indicates a redundancy in the measurements: with high confidence, variable i can be easily calculated from variable j, and it isn’t necessary to measure both variable i and variable j (or vice-versa). One use of PCA is dimensional reduction: it can help eliminate this redundancy and distill the number of variables into those that are most critical.

To eliminate this redundancy, and eliminate the off-diagonal terms, the covariance matrix $\mathbf{C_x}$ can be diagonalized. Since $\mathbf{C_x}$ is symmetric by construction, we have:

$$\mathbf{C_x} = \mathbf{V}\mathbf{D}\mathbf{V}^T \tag{2}$$

where the columns of $\mathbf{V}$ contain the eigenvectors of $\mathbf{C_x}$, also known as principal axes or principle directions. The covariance matrix eigenvalues found along the diagonal of $\mathbf{D}$ are ordered by magnitude from largest to smallest.

The principle components are the new variables that result after a linear transformation on the original centered data matrix $\mathbf{X}$. This linear transformation uses the eigenvector matrix $\mathbf{V}$ obtained in the diagonalization of the covariance matrix:

$$\mathbf{F} = \mathbf{X}\mathbf{V} \tag{3}$$

where the values of the principle components, also known as factor scores or PC scores, are denoted by $\mathbf{F}$. These values can also be interpreted as the projections of the observations on the principle components. The eigenvector matrix $\mathbf{V}$ is also sometimes referred to as a projection matrix, as it allows for the projection of the observations onto the principle axes.

The singular value decomposition (SVD) is a factorization of the centered data matrix $\mathbf{X}$. The SVD factorization is given by:

$$\mathbf{X} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^T \tag{4}$$

where $\mathbf{\Delta}$ is a diagonal matrix containing singular values (positive numbers, ordered from largest to smallest) and $\mathbf{U}$ is an $n \times n$ matrix whose columns are the eigenvectors of $\frac{1}{n-1}\mathbf{X}\mathbf{X}^T$. The matrix $\mathbf{V}$ contains the right singular vectors and is equivalent to the matrix of right eigenvectors of the covariance matrix, $\mathbf{C_x}$, defined previously.

By construction of $\mathbf{U}$ and $\mathbf{V}$, it can be shown that $\mathbf{X}$ multiplied by a column vector $\mathbf{v_i}$ of $\mathbf{V}$ results in a output vector u_i multiplied by singular vector σ_i.

The row space of $\mathbf{X}$ is given by the columns of $\mathbf{V}$. The column space of $\mathbf{X}$ is given by the columns of $\mathbf{U}$. However, if $\mathbf{X}$ is rank r, then $m-r$ additional vectors are required in $\mathbf{V}$ and $n-r$ vectors are required in $\mathbf{U}$ in order to make these square matrices; these extra vectors form the nullspace $N(\mathbf{X})$ and left nullspace $N(\mathbf{X}^T)$ and can be chosen orthonormal.

Since $\mathbf{U}$ and $\mathbf{V}$ are symmetric by construction, they each have sets of orthogonal eigenvectors. Due to this orthogonality, it can be shown that $\mathbf{U}^T\mathbf{U} = 1$ and $\mathbf{V}^T\mathbf{V} = 1$: in each case, the matrix is unitary, i.e. the inverse is equivalent to the transpose.

Geometrically, the SVD represents a rotation by $\mathbf{V}^T$, a subsequent scaling by $\mathbf{\Delta}$, and, finally, a rotation by $\mathbf{U}$.

1.5 Link between PCA and SVD

It can be easily shown that $\mathbf{XV} = \mathbf{U\Delta}$, i.e. that the the principle components obtained in PCA can be determined using the SVD decomposition. More specifically, the principle components can be calculated using the product of the left singular vector matrix and the singular values:

$$\mathbf{XV} = \mathbf{U\Delta V}^T\mathbf{V} = \mathbf{U\Delta} \tag{5}$$

Additionally, it can be shown that the singular values obtained from SVD are closely related to the eigenvalues of the covariance matrix obtained from PCA:

$$\begin{aligned}\mathbf{C_X} &= \frac{1}{n-1}(\mathbf{U\Delta V}^T)^T\mathbf{U\Delta V}^T && (6)\\ &= \frac{1}{n-1}\mathbf{V\Delta U}^T\mathbf{U\Delta V}^T && (7)\\ &= \frac{1}{n-1}\mathbf{V\Delta}^2\mathbf{V}^T && (8)\end{aligned}$$

So that:

$$\mathbf{D} = \frac{1}{n-1}\mathbf{\Delta}^2 \tag{9}$$

Figure 3 provides an annotated overview of the SVD transformation using notation and conventions (e.g., the orientation of the data matrix) that are adopted by the algorithm. The singular values and singular vectors serve as critical inputs to a post-processing stage of the anomaly detection and event classification subroutines.

1.6 Network Partitioning and the SVD

In the electromechanical timescale, the transient behavior of voltages and currents at each node of power system is contingent upon a set of well-known nonlinear differential equations and algebraic constraints (1). Over the decades, the network structure that underpins this differential-algebraic equation set has been closely studied, and tools have been developed that leverage this underlying structure to partition the network into clusters of nodes [20],[21],[22]. Such partitioning strategies have historically been applied in the

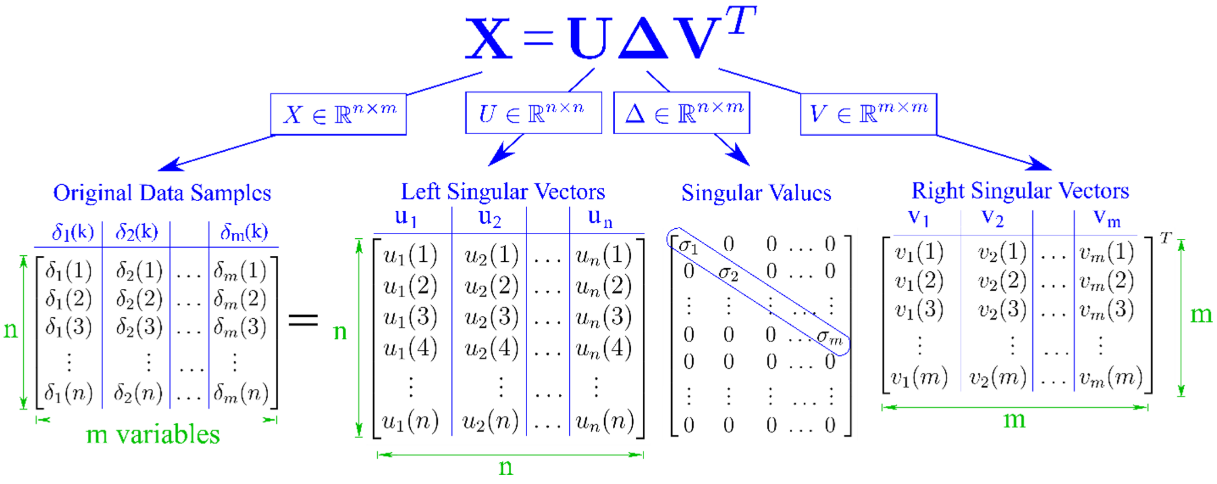

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^T$
$X \in \mathbb{R}^{n \times m}$
$U \in \mathbb{R}^{n \times n}$
$\Delta \in \mathbb{R}^{n \times m}$
$V \in \mathbb{R}^{m \times m}$
Original Data Samples
Left Singular Vectors
Singular Values
Right Singular Vectors
$\delta_1(k)$ $\delta_2(k)$ $\delta_m(k)$
u_1 u_2 u_n
v_1 v_2 v_m
σ_1 σ_2 σ_m
n
m variables
n
n
m
m

Figure 1.3: Annotated overview of the singular value decomposition.

development of reduced-order models that represent clusters of nodes as a single aggregate node, drastically reducing the number of differential and/or algebraic equations and thereby improving computational efficiency of the model.

For a simple three-node power system, Fig. 1.4 demonstrates the impact of electrical distance between nodes on the behavior of voltage angles at those nodes, in the aftermath of a significant event. In the aftermath of any significant event, unique signatures of power system behavior can usually be observed that, in general, are more consistently associated with the network structure than they are with the location or severity of a particular fault. The SVD can be applied to a time window of synchrophasor data, and the output from the SVD can potentially provide insight into such signatures. In [23], the intent is to use the SVD of measured sensor data as a data-driven means to inform an aggregation strategy, ultimately resulting in the optimized construction of a (reduced-order) physics-based model–similar to the objective in [20],[21], and [22]. The cybersecurity algorithm described here is not concerned with reduced-order modeling, but rather intends to leverage (i) the SVD, (ii) partitioning analysis, or (iii) a combination of both the SVD and partitioning analysis, for purposes of feature extraction, algorithm training, and signature identification. The extracted and post-processed signatures ultimately drive the algorithms main event detection and classification subroutines. While a variety of different partitioning analysis techniques could be employed within the developed cyberattack detection algorithm, the Generalized Eigenvalue Perturbation method [20] has several desirable qualities, including expediency and scalability.

The GEP algorithm [20] is now briefly reviewed, borrowing heavily from the notation presented in [20]. For a power system of arbitrary size and interconnection, vector-valued equation set (1) represents the nonlinear set of differential equations and algebraic constraints that dictate the dynamic behavior of a power system in the electromechanical timescale. This simplified model is most appropriate, and most commonly used, for transient stability analysis. Prime mover dynamics, exciter dynamics, and power system stabilizers are not represented in the model. Loads are assumed to be constant power, the network is assumed to be symmetric and balanced, and transfer conductances of the network are neglected.

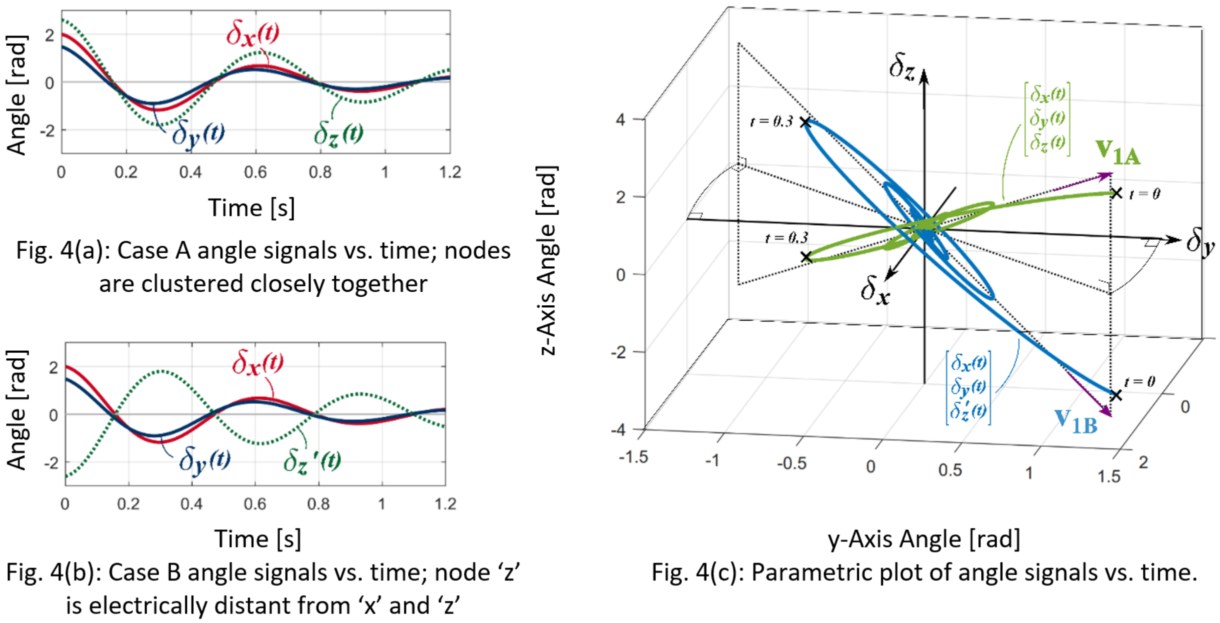


Fig. 4(a): Case A angle signals vs. time; nodes are clustered closely together

Fig. 4(b): Case B angle signals vs. time; node 'z' is electrically distant from 'x' and 'z'

Fig. 4(c): Parametric plot of angle signals vs. time.

Figure 1.4: For a three-node system, two different cases of time-domain PMU angle signals (left) and associated 3-dimensional parametric curve.

$$\begin{aligned}
\dot{\vec{\omega}} &= \mathbf{M}^{-1}(\mathbf{N_1}(\vec{P}_I - \vec{P}_N(\vec{\delta},\vec{V})))\\
\mathbf{N_1}\dot{\vec{\delta}} &= \vec{\omega}\\
0 &= \mathbf{N_2}(\vec{P}_I - \vec{P}_N(\vec{\delta},\vec{V}))\\
0 &= \vec{Q}_L(\vec{V}) - \vec{Q}_{NL}(\vec{\delta},\vec{V})
\end{aligned}$$

Where:

$$\vec{P}_N(\in\mathbb{R}^n)\ := \mathbf{A}\ \text{diag}(\overrightarrow{bl})\ \text{diag}(\exp(|\mathbf{A}|^\intercal ln(\vec{V})))\ \sin(\mathbf{A}^\intercal\vec{\delta}) \tag{10}$$

$$\vec{Q}_{N,L}\ (\in\mathbb{R}^n\ ) := [\text{diag}(\overrightarrow{V_L})]^{-1}\mathbf{L_2}|\mathbf{A}|\text{diag}(\overrightarrow{bl})\text{diag}(exp(|\mathbf{A}|^\intercal ln(\vec{V})))\cos(\mathbf{A}^\intercal\vec{\delta}) \tag{11}$$

And where:

$\overrightarrow{bl}\ \in\mathbb{R}^l$:= vector of network branch admittances, including the transient reactance of synchronous machines.

$\mathbf{A}\in\mathbb{R}^{nxl}$:= the full bus incidence matrix for the graph associated with the transmission network.

$\vec{P}_L\ \in\mathbb{R}^{n-m}$:= vector of constant active power loads, at load buses. A positive value denotes that active power is being injected into the network.

$\vec{Q}_L\ \in\mathbb{R}^{n-m}$:= vector of constant reactive power loads, at load buses. A positive value denotes that reactive power is being injected into the network.

$\vec{V}\in\mathbb{R}^n$:= vector of all bus voltages, including synchronous machine buses and load buses.

$\vec{\delta}\ \in\mathbb{R}^n$:= vector of all bus voltage angles, including synchronous machine buses and load buses, relative to the synchronous reference frame.

$\vec{P}_f \in \mathbb{R}^m :=$ vector of synchronous machine active power output signals measured by each synchronous machine and filtered using a single-pole low-pass filter.

$\vec{Q}_f \in \mathbb{R}^m :=$ vector of synchronous machine reactive power output signals measured by each synchronous machine and filtered using a single-pole low-pass filter.

$\vec{P}_0 \in \mathbb{R}^m :=$ vector of power setpoints of the prime mover

$\vec{V}_0 \in \mathbb{R}^m :=$ vector of generator internal voltages.

$\mathbf{N_1} \in \mathbb{R}^{mxn} :=$ rows 1 through m of an nxn identity matrix.

$\mathbf{N_2} \in \mathbb{R}^{(n-m)xn} :=$ rows m through n of an nxn identity matrix.

diag(y) gives a diagonal matrix with the components of the vector y forming its diagonal entries.

$\overrightarrow{V_L} \in \mathbb{R}^{n-m} :=$ a vector of algebraic states representing the voltages at the load buses: $\overrightarrow{V_L}(i)$ represents the voltage at node$m + i$.

$\vec{P}_I \in \mathbb{R}^n := [\vec{P}_0^{\mathsf{T}}, \vec{P}_L^{\mathsf{T}}]^{\mathsf{T}}$

$\vec{Q}_{N,L} : \mathbb{R}^{n-m} \to \mathbb{R}^{n-m}, \vec{Q}_{N,L} :=$ a vector-valued function giving reactive power absorbed by load buses, normalized by voltage magnitude:

$\mathbf{M} \in \mathbb{R}^{mxm} :=$ diagonal matrix with generator inertias along the entries, ordered by bus number

When linearized at an equilibrium operating point, (1) is proven in [20] to be closely associated with the symmetric differential algebraic equation (2):

$$\mathbf{E}\dot{\vec{x}} = \mathbf{R}\vec{x} \tag{12}$$

where:

$$\mathbf{E} = \begin{bmatrix} \mathbf{I}_{mxm} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{2(n-m)x2(n-m)} \end{bmatrix} \tag{13}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{M}^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{2(n-m)x2(n-m)} \end{bmatrix} \tag{14}$$

$$\mathbf{R} = \mathbf{U}^{\mathsf{T}}\mathbf{J}\mathbf{U} \tag{15}$$

And $\mathbf{J}$ represents the system Jacobian:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \vec{P}_N}{\partial \vec{\delta}} & \frac{\partial \vec{P}_N}{\partial \overrightarrow{V_L}} \\ \frac{\partial \vec{Q}_{N,L}}{\partial \vec{\delta}} & \frac{\partial (\vec{Q}_{N,L} - \vec{Q}_L)}{\partial \overrightarrow{V_L}} \end{bmatrix} \tag{16}$$

It can be shown that the smallest non-zero generalized eigenvalue of (2) is closely associated with the dominant inter-area dynamics between two clusters of nodes. Due to the symmetry of the matrices within (2) that results from its admittance-matrix-like structure, an iterative procedure identifying the clusters was shown in [20] to always converge. In the GEP approach, the sensitivity of this smallest non-zero generalized eigenvalue of $(\mathbf{E}$,$\mathbf{R}_{\mathsf{S}})$ to perturbations in the branch susceptances guides the selection of a cut-set of branches. This branch cutset forms the optimal boundary between two clusters.

The GEP partitioning algorithm [12] is reviewed below and linked to the SVD, with the ultimate goal of training a simple cyberattack detection algorithm described later. Notation is borrowed from [12]. The GEP approach can be implemented by following the step-by-step instructions described in Section IV of [12]. These instructions are summarized in steps 1 through 4, below. Step 5 connects the partitioning result to the detection algorithm.

1. $\mathbf{E}$ and $\mathbf{R}_\mathrm{S}$, matrices defined in [12], are closely related to a linearized version of (1). As a first step, these two matrices are computed, assuming a particular equilibrium operating point of (1). For the differential-algebraic system comprised by $(\mathbf{E}, \mathbf{R}_\mathrm{S})$, the smallest generalized eigenvalue and its associated eigenvector are identified, denoted by λ and v, respectively.

2. The gradient of the generalized eigenvalue λ with respect to the vector of network susceptances, $\overrightarrow{bl}$, is then computed. Under simplifying assumptions, index i of gradient $\vec{g}$ is given by a readily-identifiable analytical expression: $\vec{g}_i = \partial\lambda / \partial \overrightarrow{bl}_i = \frac{v^\intercal \frac{\partial \mathbf{R_S}}{\partial \overrightarrow{bl}_i} v}{v^\intercal \mathbf{E} v}$

3. Gradient $\vec{g}_i$ is used to identify the element of the network chosen for deletion, by finding the smallest positive and real constant γ such that for some index r: $[\overrightarrow{bl} - \gamma\vec{g}]_r = 0$. Delete the entry at index r of the susceptance vector $\overrightarrow{bl}$.

4. Repeat Steps 1 - 3 for the new system with the missing branch, until a sufficient number of branches have been eliminated such that the nodes of the original network are fully partitioned into two disjoint subsets. A single iteration of the GEP algorithm has now been completed.

5. Once the network has been partitioned into two disjoint subsets of nodes (Subset 1 and Subset 2) using the GEP algorithm, a synthetic right singular vector (RSV) is created. This synthetic RSV constitutes a prediction of a vector within the $\mathbf{V}$ matrix obtained from the SVD. This synthetic RSV is first instantiated as a zero vector with length equal to the number of PMUs associated with the synchrophasor angle data matrix, with the first element of the synthetic RSV corresponding to the first column of data (i.e., the first PMU), the second element of the synthetic RSV corresponding to the second PMU, and so on. The elements of this synthetic RSV are populated with a 1 or -1 depending on whether the node that hosts each PMU is located in Subset 1 or Subset 2, respectively. If PMUs are not included within the boundaries of the portion of the network subject to the above GEP partitioning analysis, the element associated with that PMU remains zero. Once all elements are appropriately populated, the synthesized RSV is then normalized to have a magnitude of 1.

6. To generate additional synthetic RSVs for the network that correspond to faster inter-area system modes across smaller areas of the network, steps 1-5 can be iteratively repeated for the sub-networks of the power system associated with Subset 1 and Subset 2 determined in Step 4.

1.7 High-Level Algorithm Description

Fig. 1.5 shows a block diagram of the proposed cyberattack detection algorithm, including event detection and classification steps. A key component of the algorithm includes the online application of the singular value decomposition (SVD) to pre-processed synchrophasor data. Among other benefits, the SVD provides dimension reduction. The features and signatures obtained from the application of the SVD upon synchrophasor data are post-processed and utilized within both the event detection and classification subroutines. The event detection runs continuously, examining the SVD features for evidence of an event. Upon detection of an event, the detection algorithm invokes the event classification subroutine.

The event classification subroutine uses the features and system signatures obtained from the SVD of the time window of data corresponding to the new event. To classify the event as a fault or a cyberattack,

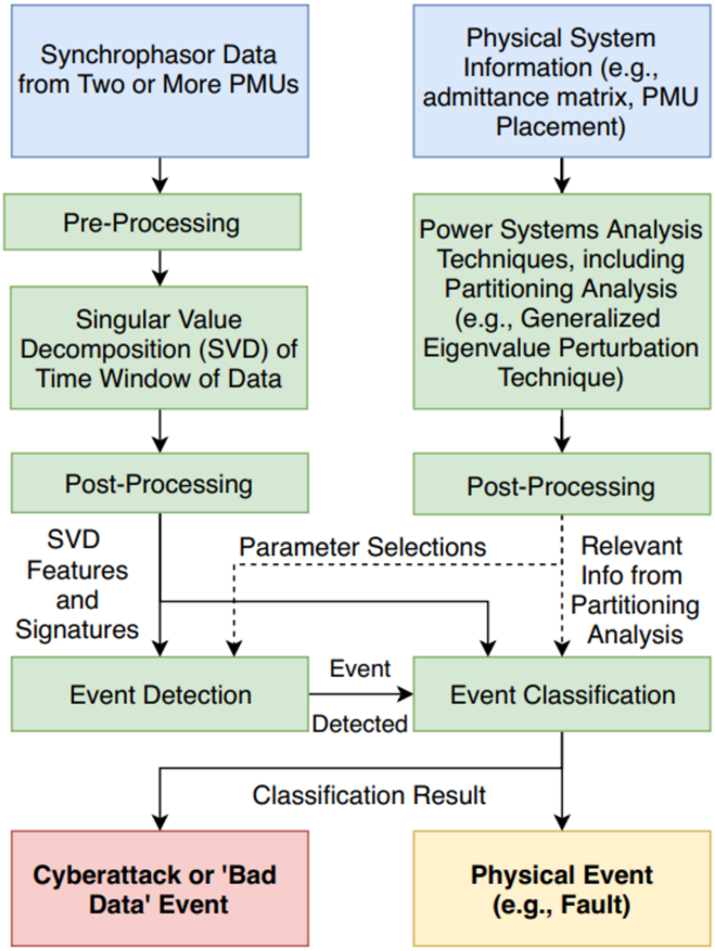

Synchrophasor Data from Two or More PMUs
Physical System Information (e.g., admittance matrix, PMU Placement)
Pre-Processing
Singular Value Decomposition (SVD) of Time Window of Data
Power Systems Analysis Techniques, including Partitioning Analysis (e.g., Generalized Eigenvalue Perturbation Technique)
Post-Processing
Post-Processing
SVD Features and Signatures
Parameter Selections
Relevant Info from Partitioning Analysis
Event Detection
Event Detected
Event Classification
Classification Result
Cyberattack or 'Bad Data' Event
Physical Event (e.g., Fault)

Figure 1.5: High-level block diagram of developed cybersecurity algorithm.

the classification subroutine also examines relevant information obtained from partitioning analysis and/or from historical events.

The dotted line within Fig. 1.5 indicates that the use of parameters and other relevant information from the partitioning analysis is optional, contingent upon availability of recent system topological information. Likewise, physics-based model analyses can optionally be used to inform the selection of parameters within the event detection algorithm. While test results indicate that parameter selection is relatively straightforward, and that a common set of parameters may be successfully used for a wide range of different power systems, model-based analyses can improve confidence in the selection of optimal algorithm parameters, such as certain decision thresholds and the size of the data window used for the SVD.

1.8 GEGR Software-Based Testbed

A flexible, software-based testbed has been realized in the MATLAB/Simulink platform to facilitate the development and rigorous evaluation of the candidate WAMS cyberattack detection algorithm. Within this testbed environment, a first-principles, physics-based model of a standard 4-machine, 2-area power system, (Fig. 1.6) described in [24], is used to generate realistic time-dependent trajectories of system states in response to dynamic events.

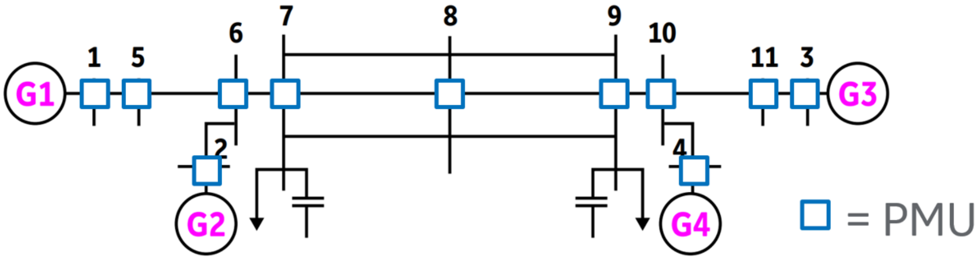


Figure 1.6: The Kundur two-area test system from [24] used for algorithm development and evaluation.

The test environment is capable of automatically simulating a large number of randomized transient events using the two-area test system, including actual fault events and spoofed data injections from compromised PMU(s) that resemble fault events. For the studied power system, the test environment collects PMU data throughout the duration of each event, resembling typical behavior of a Wide Area Monitoring System. The PMU data is comprised of time-stamped, synchronized samples of voltage and current phasor magnitude and phase angles, collected at a user-defined sampling rate, under realistic off-nominal frequency conditions. Automated MATLAB scripts provide the following functionality:

1. Specification of parameters relevant to the physical characteristics of the power network, including fault events within the network.
2. Specification of parameters relevant to the construction and testing of a cyberattack spoofing detection algorithm.
3. Construction of a nested cell variable of sliding-window PMU data matrices and cyberattack events using simulation results obtained from running the Simulink model or from import of existing simulation results.
4. Training of the cyberattack detection algorithm using a model-driven Generalized Eigenvalue Perturbation partitioning script.
5. Training of the cyberattack detection algorithm using SVD features and signatures calculated from previous events.
6. Application of the cyberattack spoofing detection algorithm to the nested cell variable containing PMU data matrices.
7. Three-dimensional visualization of the test matrix results.

Regarding specification of parameters relevant to the physical characteristics of the network, the simulation environment is highly customizable, allowing the user a high degree of control over a large number of relevant parameters, including:

1. The nature, placement and severity of physical fault events, including short circuit and line faults.
2. The number of PMUs within the system.
3. Placement of PMUs within the system.
4. Specification of which PMUs have been compromised.

5. Specification of the type of cyberattack on each PMU unit. Currently, the choices are limited to the following:
6. A compromised PMU replays historical fault data.
7. A compromised PMU erroneously reports a user-defined constant magnitude and phase angle.
8. A compromised PMU erroneously reports a magnitude and phase angle of zero.
9. Specification of line impedances and generator properties, including inertia and damping.
10. Specification of the different features of the simulation data and experiment results to be stored for future analysis.

Many of the of the above environmental parameters (e.g., 1-5) can be randomized, facilitating large scale studies across a range of potential environmental conditions. Additionally, parameters such as number of PMUs in the network, or number of compromised PMUs, can be swept across a range of values, allowing greater insight into the impact of the parameter on the performance of the algorithm.

The software-based testbed also allows for the specification of parameters relevant to the operation of the algorithm, including all relevant decision thresholds and size of the SVD data window.

The testbed environment permits most of the above algorithm parameters to be swept across a range of values in large-scale test matrices to empirically determine the trend of classification performance as a function of the value of the algorithm parameter.

Fig. 1.7 demonstrates example post-event synchrophasor waveform results derived from the software-based testbed for four different combinations of environmental parameters. Severity and location of a physical fault is varied between the four tests, and waveforms of post-processed synchrophasor voltage angle are plotted for three PMUs.

1.9 Algorithm Performance Evaluation Results

This section documents the performance of the algorithm following its application to windows of PMU data recorded using the 11-bus simulation testbed described earlier. Additionally, the algorithm is applied to more realistic simulation results obtained using PNNLs testbed.

Fig. 1.8 shows example windows of simulated voltage angle waveforms that were collected using the software-based testbed. The waveforms were collected shortly after various events, including a new fault event, a historical fault event, and a fault replay cyberattack. Fig. 1.8(a) and (c) both correspond to the same new fault event, but Fig. 1.8(c) shows the behavior of the voltage angle only for nodes 1,3,5,11. Likewise, Fig. 1.8(b) and (d) both correspond to the same historical fault event. The SVD was applied to time windows of data shown in (a), (b), (e), and (f), and the detection algorithm was invoked. The algorithm used the SVD features and signatures in an attempt to properly classify the new fault event and the cyberattack event. Synchrophasor information from nodes 1,3,5,11 were found to be particularly critical for algorithm performance. Using the post-processed feature and signature information obtained from new and historical events, the algorithm correctly concluded that the transient in Fig. 1.8(a) was a real fault, and that the transient in Fig. 1.8(e) did not result from a real fault event but was instead either bad data or a cyberattack.

Results to follow demonstrate a thorough investigation of the performance of the algorithm for a range of power system conditions, and across multiple test matrices of algorithm parameter combinations. For each tested condition and combination of parameters, the cyberattack detection algorithm is applied to 600 randomized fault event simulations and 600 randomized fault-replay cyberattack event simulations. Random short circuit faults and line faults are applied to the power system, and the transients that result are also

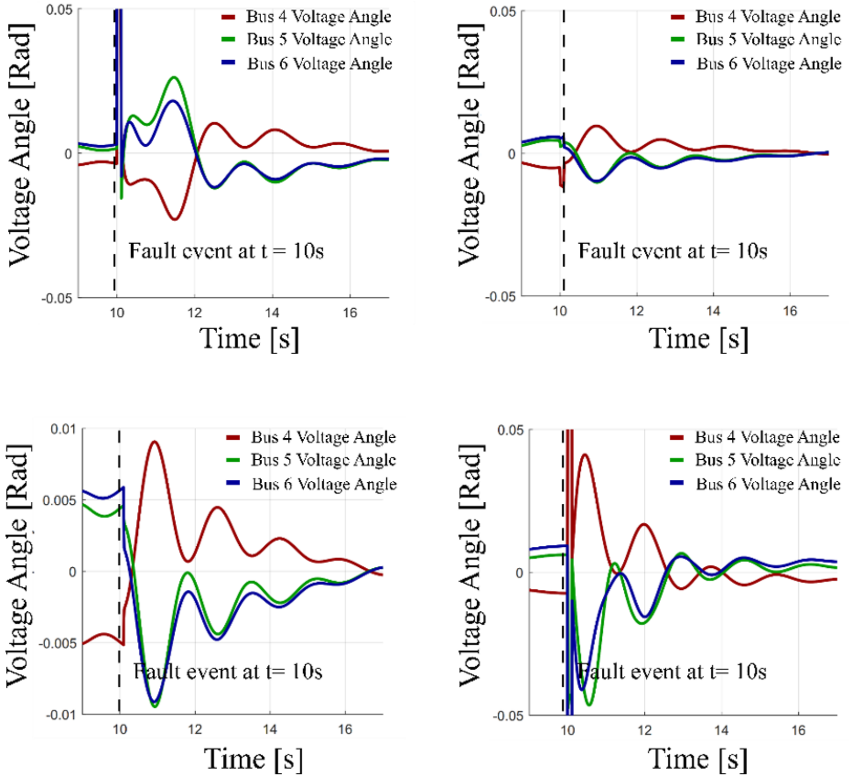
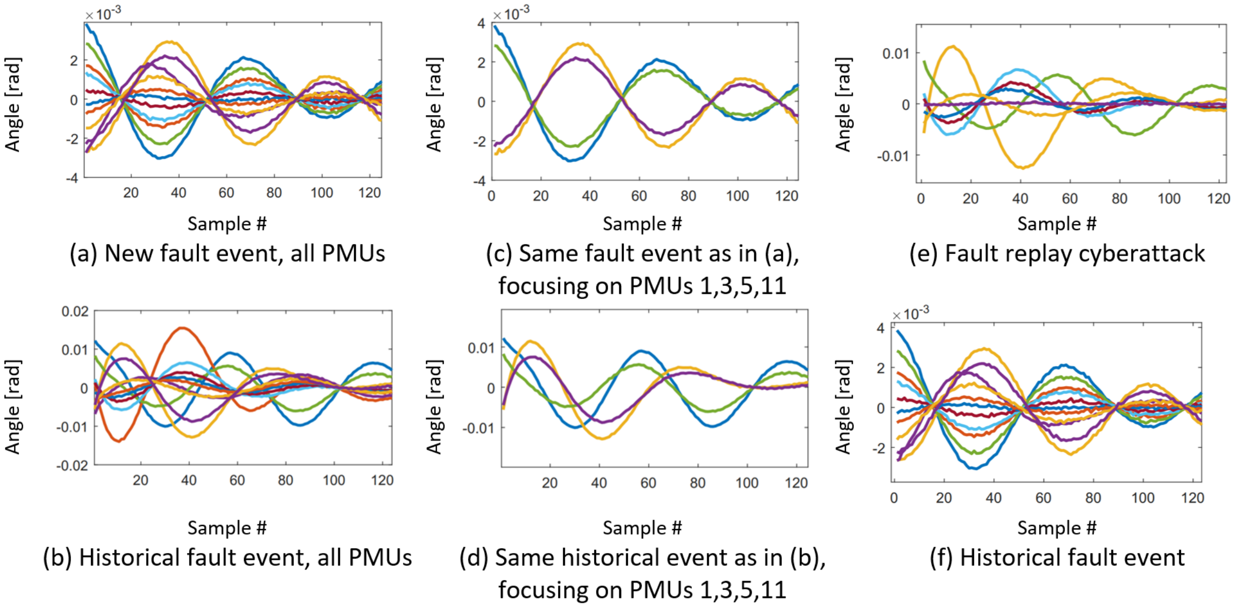


Figure 1.7: Waveforms of post-processed synchrophasor voltage angle for three PMUs, and four random fault events.



(a) New fault event, all PMUs

(c) Same fault event as in (a), focusing on PMUs 1,3,5,11

(e) Fault replay cyberattack

(b) Historical fault event, all PMUs

(d) Same historical event as in (b), focusing on PMUs 1,3,5,11

(f) Historical fault event

Figure 1.8: Simulated voltage angles for all nodes within a 4-machine, 11-bus network, capturing behavior subsequent to a new fault event and a cyberattack, and comparing both with a historical fault event.

recorded for use in a fault-replay cyberattack. For a given fault-replay cyberattack, the transients replayed at the compromised PMUs were all recorded from the same event, and the fault transient is replayed at the same PMU from which it was recorded. A matrix of different algorithm parameter selections, number of PMUs, and number of compromised PMUs is studied. In most of the investigations to follow, the data window length is swept from 10 to 200 samples, and a particularly important decision threshold (referred to here as rotation threshold angle) is swept from 0.5 degrees to 60 degrees.

Outputs of a classification subroutine can be: (i) true-positive, which is a correct classification of an abnormal event as abnormal; (ii) false-positive, which is an incorrect classification of a normal event as abnormal; (iii) true-negative, a correct classification of a normal event as normal; and (iv) false-negative, an incorrect classification of the abnormal event as normal. In this section, true positive rates are recorded as the algorithm is invoked to classify fault events and cyberattack events. In the case of fault event classification performance, the true positive rate will be calculated, for purposes of this report, by tallying the number of correct fault classifications and dividing by the total number of events that were classified as either a fault or a cyberattack. Similarly, in the case of cyberattack classification performance, the true positive rate is given by the number of correct cyberattack classifications divided by the total number of events that were classified as either a fault or a cyberattack. It should be noted that often, a randomized event–either a fault event or a cyberattack event–was not significant enough to surpass the event detection threshold and invoke the classification subroutine. For each analysis of a particular set of parameter selections, number of PMUs, and number of compromised PMUs, these minor events constituted approximately a quarter or less of the 1200 events studied.

For tests in which PMUs are not deployed at every bus, Table 1.1 shows which buses contain PMUs (marked with 1), and which buses omit PMUs (0). Bus numbers are labeled in the one-line diagram shown in Fig. 1.6. Locations for PMU deployment were arbitrarily selected.

Table 1.1: PMU locations used for tests, in the case that fewer than 11 PMUs were deployed in the system.

	Bus 1	Bus 2	Bus 3	Bus 4	Bus 5	Bus 6	Bus 7	Bus 8	Bus 9	Bus 10	Bus 11
9 PMUs	0	1	1	1	1	1	1	1	1	1	0
7 PMUs	0	1	0	1	1	1	1	1	0	1	0
5 PMUs	0	1	0	1	0	1	0	1	0	1	0
3 PMUs	0	0	0	1	0	1	0	1	0	0	0

1.9.1 Performance under Nominal Power System Parameters

In the case of the performance results in this section, the test system used to generate simulations is characterized by two distinct areas separated by a long transmission line corridor. Parameters for the test system can be found in [24]. This configuration is particularly conducive to inter-area oscillations between two clusters of nodes.

Fig. 1.9 demonstrates a collection of 3-D surface plots that show the true positive rate of the classification algorithm as a function of an important decision threshold parameter and length of the sliding window (the number of rows in the data matrix, X). For each element of the surface plot–i.e., each combination of the threshold parameter and data window size–the true positive rate shown is an average value calculated over 600 fault or cyberattack events. True positive rate surface plots are thus created for three different numbers of PMUs deployed in the network (11, 9, and 7).

For the cyberattack events used to generate the right column of true positive surface plots in Fig. 1.9, five of the deployed PMUs are compromised by the attacker. A purely data-driven approach to algorithm training is pursued (GEP partitioning is not used), and features & signatures from only one historical event are used within the classification subroutine.

Table 1.2 shows samples from each of the surface plots in Fig. 1.9, for a particularly promising selection of parameters: a rotation threshold of 18 degrees and a data window size of 150. In the case of this selection of parameters, and for many other neighboring elements of the test matrix, the algorithm has exceptional performance, demonstrating high true positive rates for both fault and cyberattack event classification for 11, 9 and 7 PMUs. Surprisingly, even when there are only 2 uncompromised PMUs left, Fig. 1.9(f) demonstrates that a high true positive rate can still be achieved for cyberattack events, provided that the threshold angle parameter is chosen to be small. Fortunately, results in Fig. 1.9(e) show that it is acceptable to maintain a very low threshold angle provided that the data window size is high enough.

Fig. 1.9(a),(c), and (e) show how a very small threshold angle value of less than 10 degrees results in poor true positive rates for fault event classification, especially for small window sizes.

There is a clear dependency of the fault true positive rate on the data window size: for high classification performance, it seems as if it would be best for the data window to be long enough to include at least 70 synchrophasor samples. As observed in Fig. 1.9(a),(c), and (e), data windows larger than approximately 100 samples provide diminishing returns. Notably, in the immediate aftermath of the fault, the spatiotemporal correlation of different synchrophasors is dominated more by electrical proximity to the fault than it is by the electromechanical interaction between generators.

For the PMUs that are compromised and subject to a sophisticated replay attack that uses a real historical fault signal, a small data window may reduce the likelihood that characteristic system signatures will be identified in a replay attack. Is therefore intuitive that cyberattack classification true positive rate is slightly bolstered by a small data window, as seen in Fig. 1.9(b),(d), and (f).

Table 1.2: True-Positive Selections for Angle Rotation Threshold of 18 Deg. and Window Size of 150 Samples, from Experiment Results in Fig. 1.9.

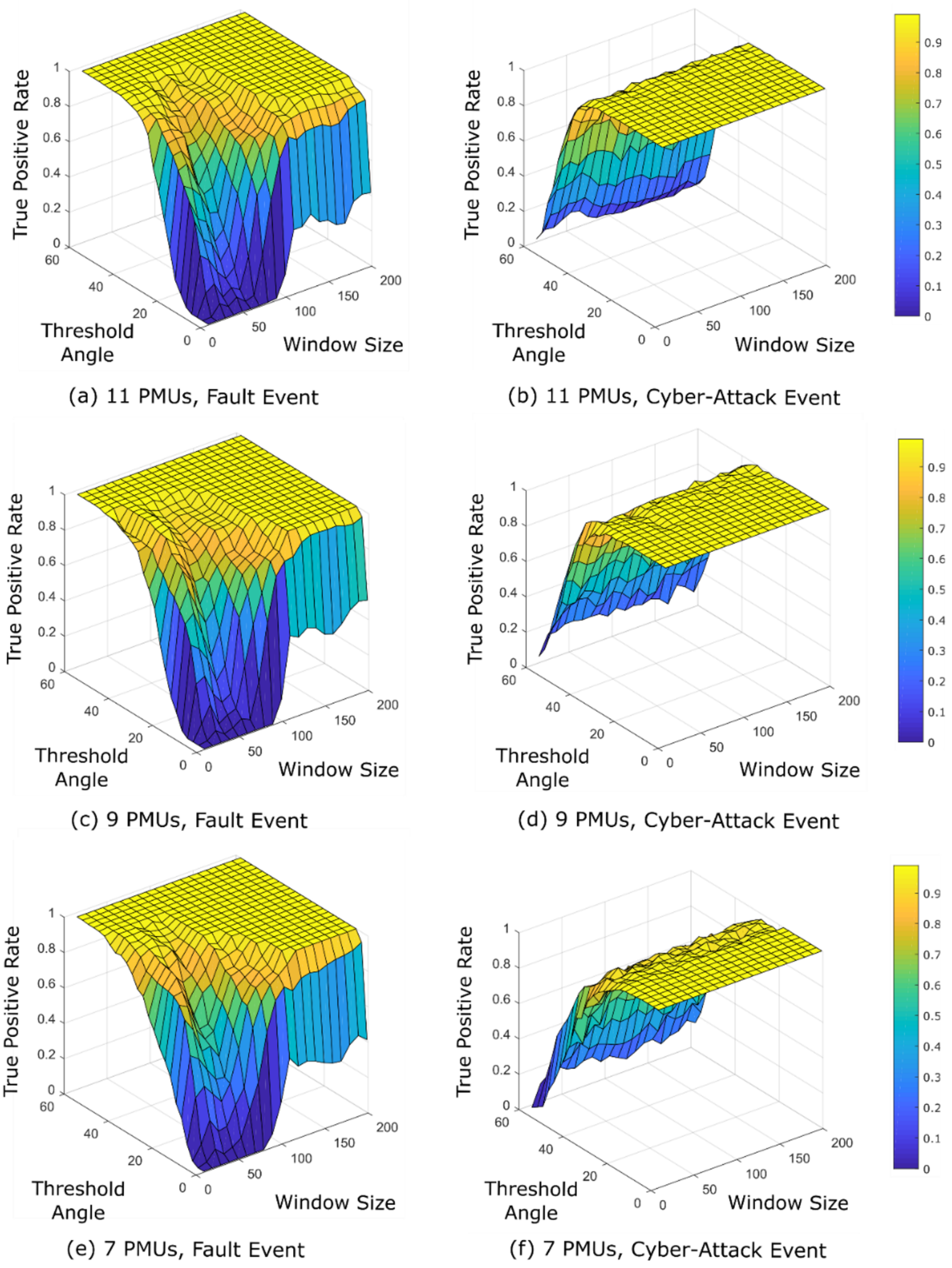
	Fault True Positive Rate (TPR)	**Cyberattack True Positive Rate (TPR)**
11 PMUs	0.83	0.33
9 PMUs	0.78	0.29

1.9.2 Performance under Augmented Power System Conditions

It should be acknowledged that the system measured by the PMUs will not always contain two distinct areas separated by a long transmission line. In the results shown in this section, in the interest of thoroughly testing the algorithm under a wider range of challenging circumstances, the length of the long transmission line corridor in the testbeds Simulink model is divided by 10. Subsequently, 1,200 randomized events are newly simulated.

As in the case of Fig. 1.9, Figs. 10-15 show collections of 3-D surface plots that demonstrate true positive rate of the classification algorithm as a function of the main decision threshold parameter and length of the sliding window (the number of rows in the data matrix, X). Different numbers of PMUs are compromised in each of the Figs. 10-15. For each element of the surface plot–i.e., each combination of the threshold parameter and data window size–the true positive rate shown is an average value calculated over 600 fault or cyberattack events.

One PMU Compromised, 1 Historical Event (Fig. 1.10): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.10, only one of the deployed PMUs is compromised by the attacker. A purely data-driven approach to algorithm training is pursued (GEP partitioning is not used), and features & signatures from only one historical event are used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

(c) 9 PMUs, Fault Event

(d) 9 PMUs, Cyber-Attack Event

(e) 7 PMUs, Fault Event

(f) 7 PMUs, Cyber-Attack Event

Figure 1.9: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network (11 to 7). Nominal 2-area system parameters. For the cyberattacks, five random PMUs are compromised. One historical event is used within the classification subroutine.

True positive rate surface plots are created for three different numbers of PMUs deployed in the network (11, 7, and 3). Under the augmented network conditions, Fig. 1.10 and Table 1.3 show that the performance of the classification is moderately acceptable when 11 PMUs are deployed, and that performance suffers as fewer PMUs are deployed. A direct comparison to the results of Fig. 1.9 is not possible, but it can likely be surmised that the new system conditions result in the deterioration the performance of the algorithm.

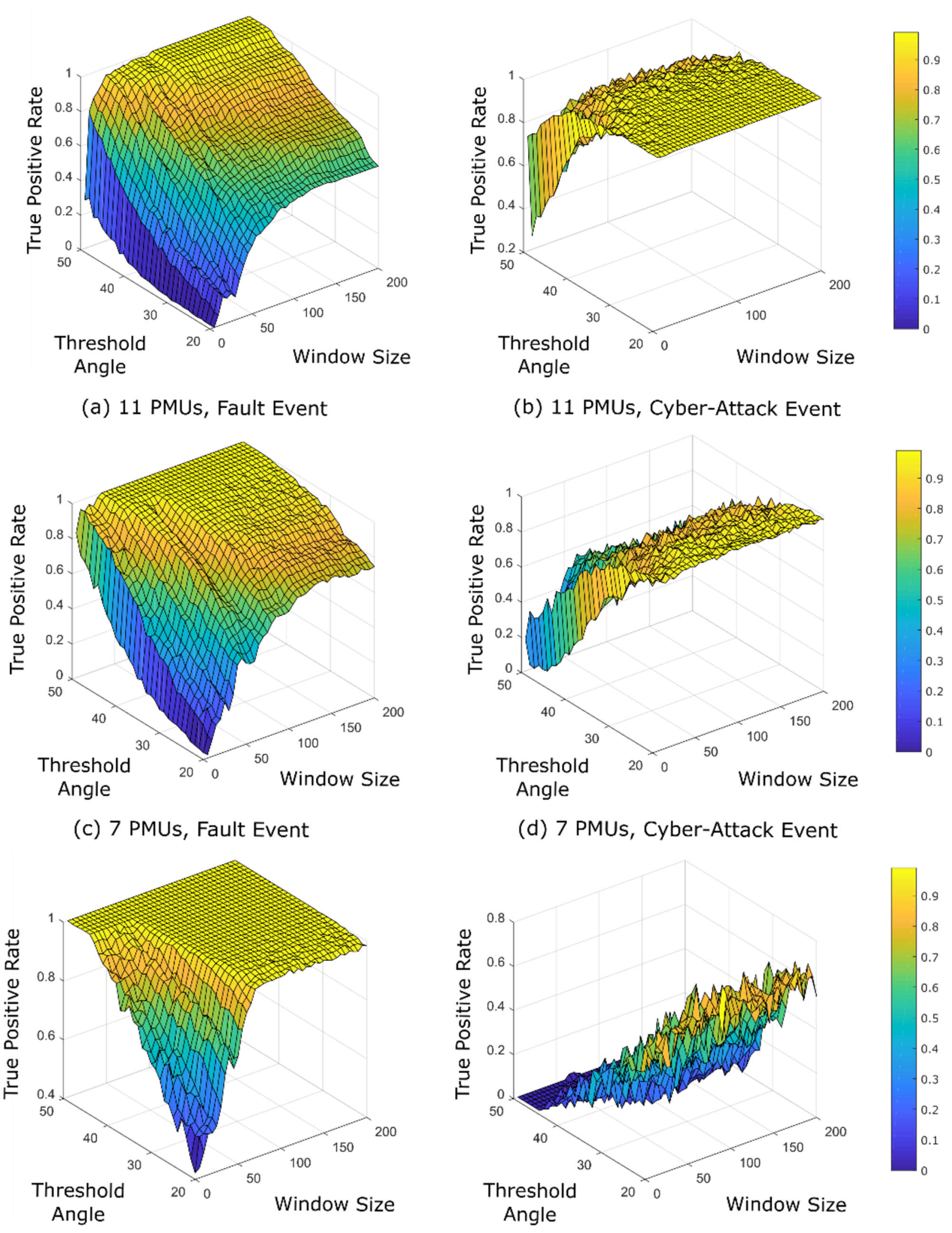
Two PMUs Compromised, 1 Historical Events (Fig. 1.11): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.11, two of the deployed PMUs are compromised by the attacker. A purely data-driven approach to algorithm training is pursued (GEP partitioning is not used), and features & signatures from only one historical event are used within the classification subroutine. True positive rate surface plots are created for three different numbers of PMUs deployed in the network (11, 7, and 3). Under the conditions of the new system, Fig. 1.11 and Table 1.4 show that the performance of the classification has deteriorated with respect to Fig. 1.10, in the case that 11 PMUs are deployed.

Two PMUs Compromised, 3 Historical Events (Fig. 1.12): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.12, two of the deployed PMUs are again compromised by the malicious adversary. The same conditions were used to produce Fig. 1.12 as those that were used to produce Fig. 1.11, with the exception that features & signatures from three historical events are used within the classification subroutine, instead of from just one historical event. Under these conditions, Fig. 1.12 and Table 1.5 show that the performance of the classification has significantly improved with respect to Fig. 1.11. The true positive rates for fault and cyberattack events is excellent when 11 PMUs are deployed, but still deteriorates substantially when 7 PMUs are deployed.

Two PMUs Compromised, 5 Historical Events (Fig. 1.13): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.13, two of the deployed PMUs are again compromised by the malicious adversary. The same conditions were used to produce Fig. 1.13 as those that were used to produce Fig. 1.12, with the exception that features & signatures from five historical events are used within the classification subroutine, instead of from just one historical event. Under these conditions, the results in Fig. 1.13 and Table 1.6 show that the classification performance does not show any improvement over the results in Fig. 1.12.

Five PMUs Compromised, 1 Historical Event (Fig. 1.14): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.14, five of the deployed PMUs are now compromised by the malicious adversary. The same conditions were used to produce the plots in Fig. 1.14 as those that were used to produce the plots in Fig. 1.9, with the exception that the system used for generating the events in Fig. 1.14 lacks long inter-area transmission lines. This set of surface plots therefore provides an opportunity to directly assess the impact of the length of the inter-area transmission lines on algorithm performance. In the case of the augmented system, Fig. 1.14 and Table 1.7 show that the performance of the algorithm still approaches a satisfactory level when 11 PMUs are deployed, and the performance is only marginally worse than in the case of Fig. 1.11. As before, the performance deteriorates substantially with a decrease in PMU deployment. Comparing Figs. 9 and 14, it is easy to conclude that the algorithm enjoys a substantial advantage in performance under the original system conditions.

Nine PMUs Compromised, 1 Historical Event (Fig. 1.15): For the cyberattack events used to generate the right column of true positive rate surface plots in Fig. 1.15, nine of the deployed PMUs are now compromised by the malicious adversary. Otherwise, the same conditions were used to produce Fig. 1.15 as those that were used to produce Fig. 1.11. Under these conditions, the results in Fig. 1.15(a) and (b), and in Table 1.8, show that the classification performance has deteriorated significantly. With only two uncompromised PMUs and one historical event in the repository, under the conditions of the augmented system, the algorithm is not capable of discerning fault from cyberattack.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

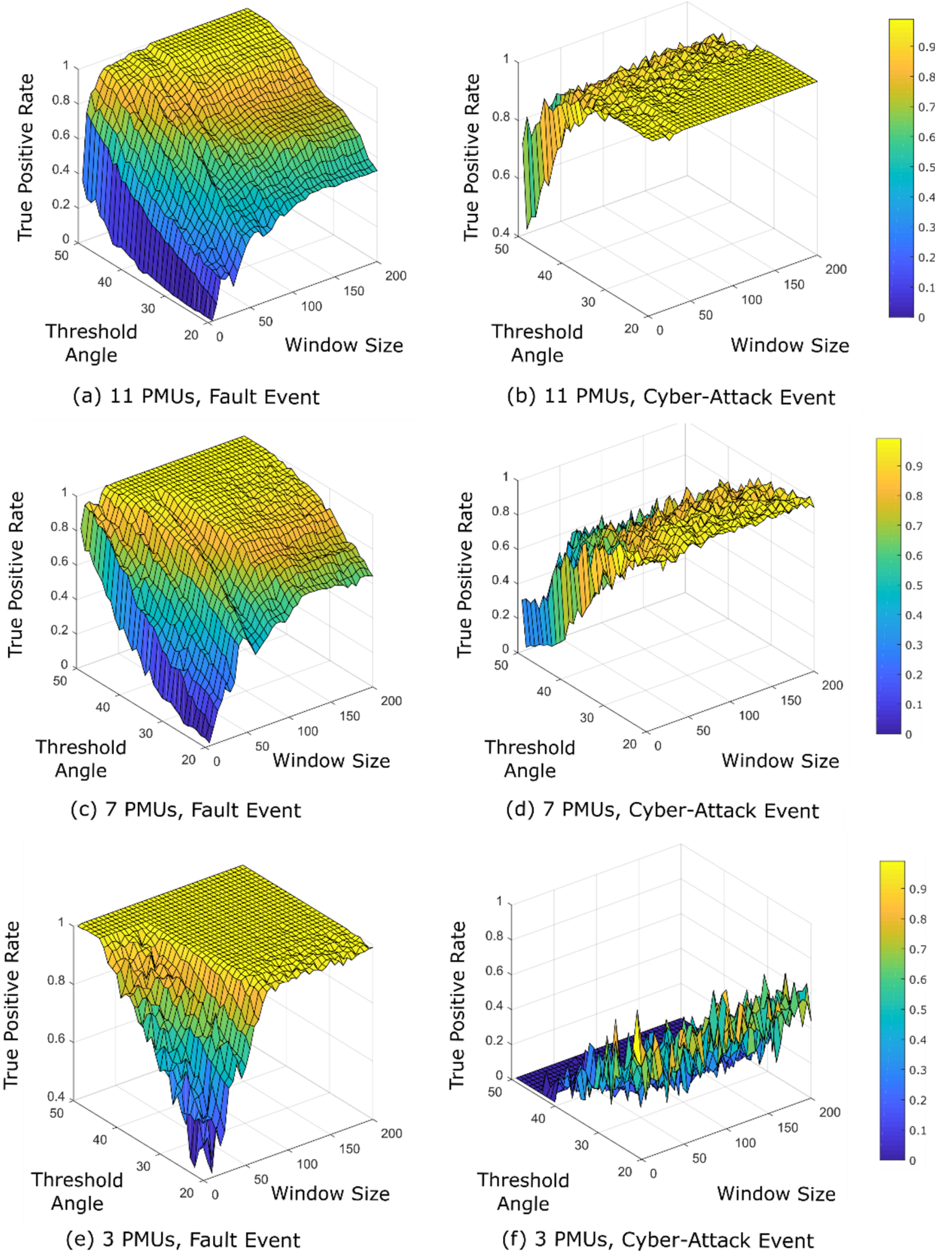
(c) 7 PMUs, Fault Event

(d) 7 PMUs, Cyber-Attack Event

(e) 3 PMUs, Fault Event

(f) 3 PMUs, Cyber-Attack Event

Figure 1.10: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, only one PMU is compromised. One historical event is used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

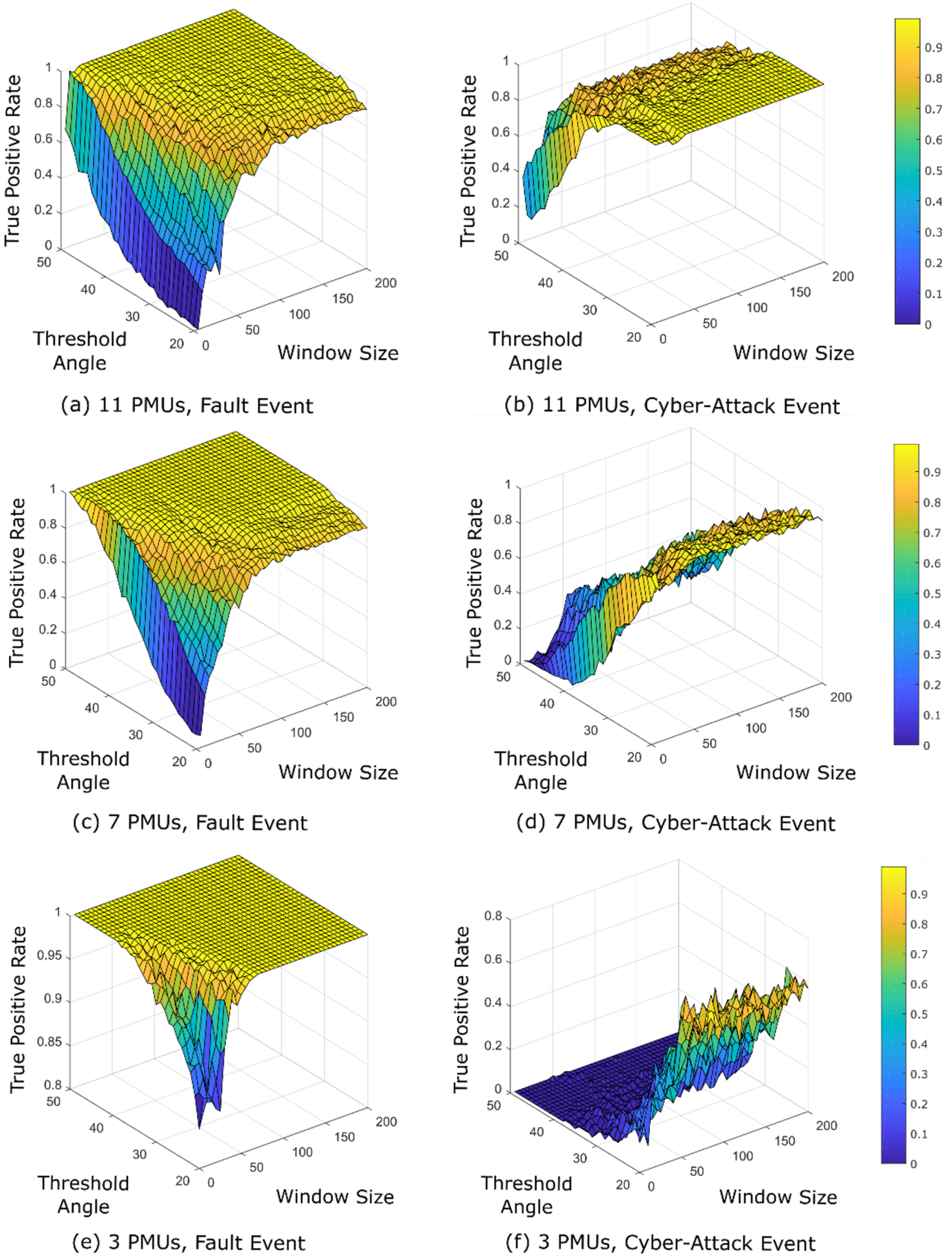
(c) 7 PMUs, Fault Event

(d) 7 PMUs, Cyber-Attack Event

(e) 3 PMUs, Fault Event

(f) 3 PMUs, Cyber-Attack Event

Figure 1.11: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, two PMUs are compromised. One historical event is used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

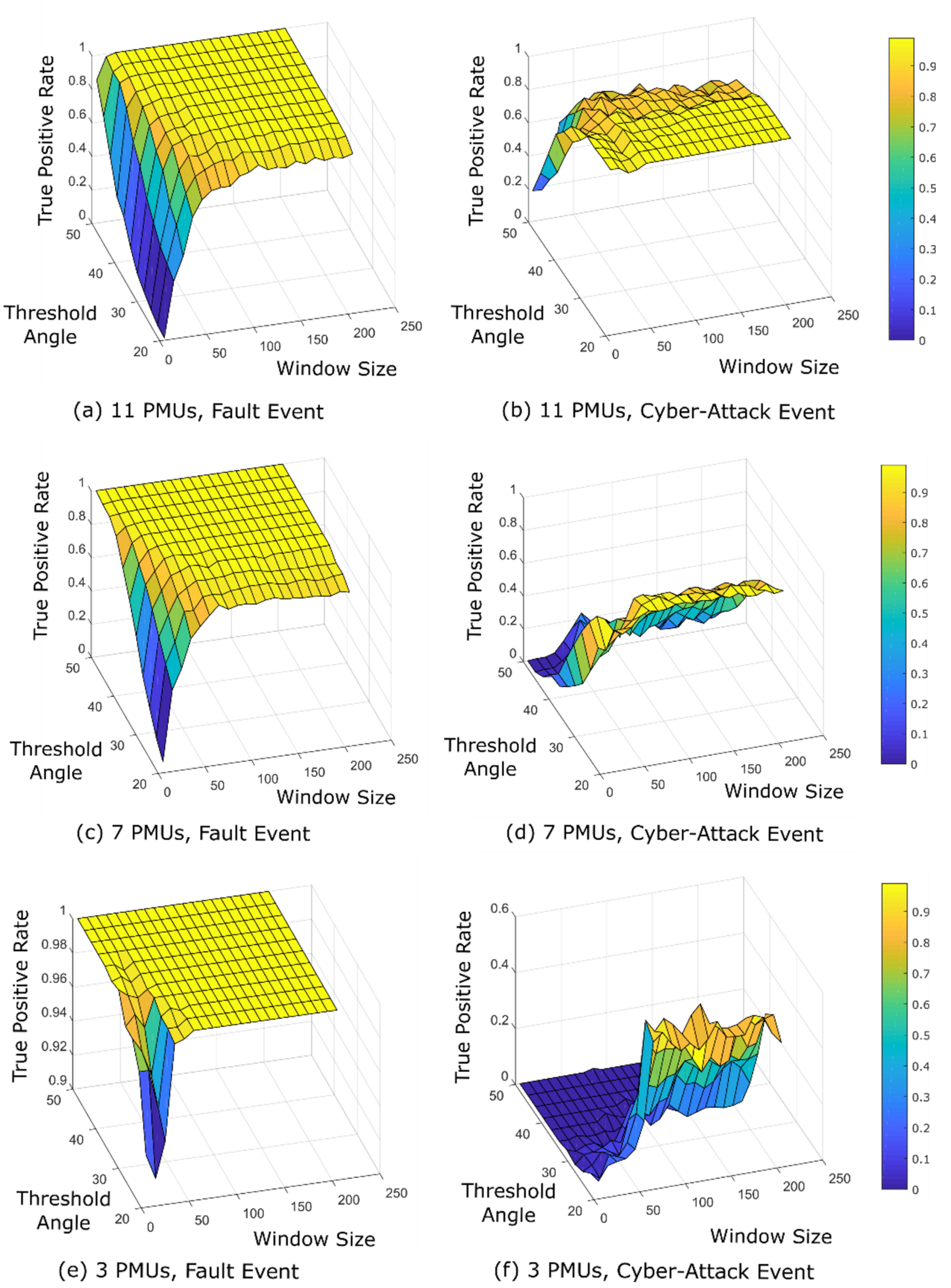
(c) 7 PMUs, Fault Event

(d) 7 PMUs, Cyber-Attack Event

(e) 3 PMUs, Fault Event

(f) 3 PMUs, Cyber-Attack Event

Figure 1.12: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, two PMUs are compromised. Three historical events are used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

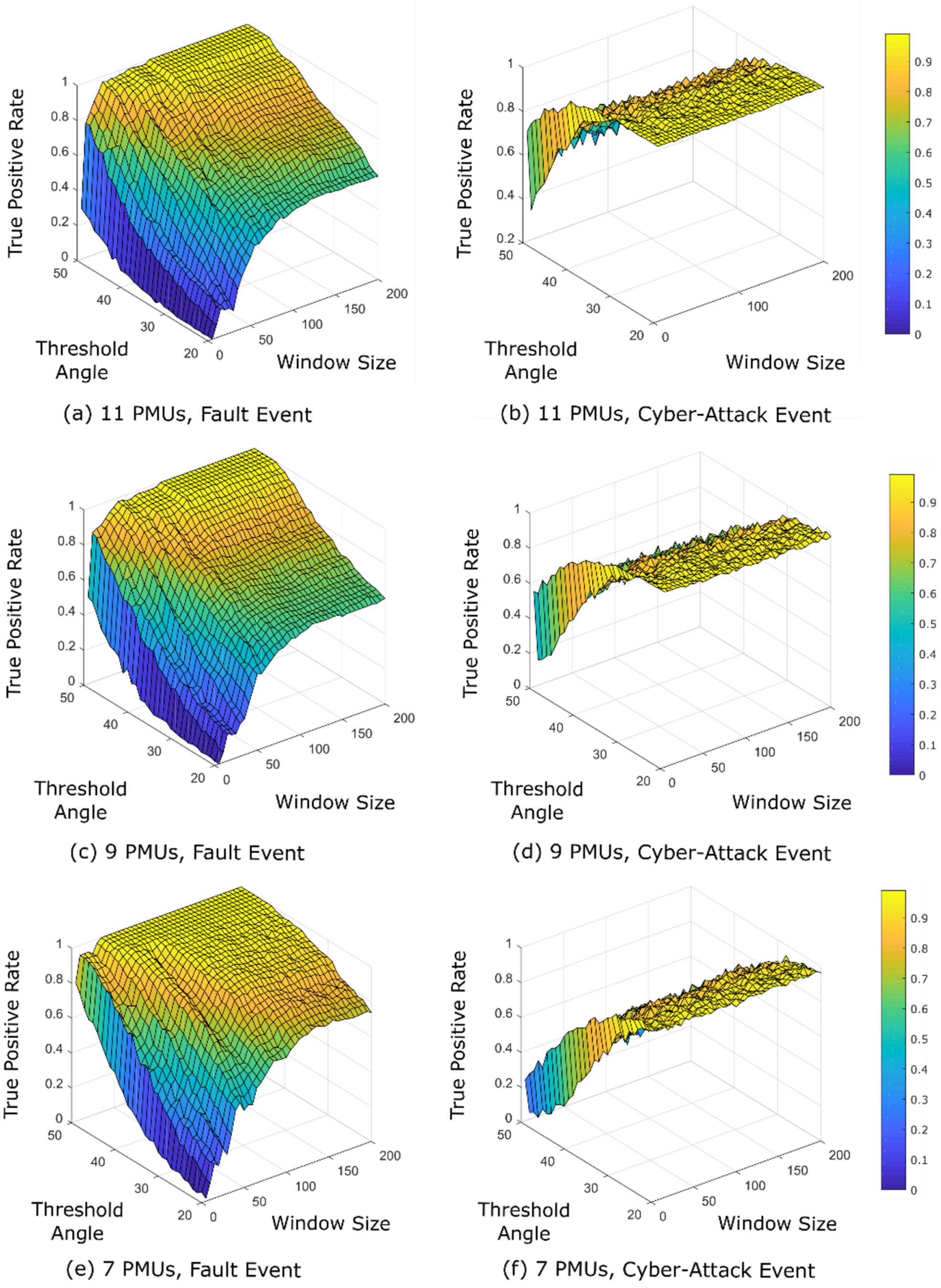
(c) 7 PMUs, Fault Event

(d) 7 PMUs, Cyber-Attack Event

(e) 3 PMUs, Fault Event

(f) 3 PMUs, Cyber-Attack Event

Figure 1.13: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, two PMUs are compromised. Five historical events are used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

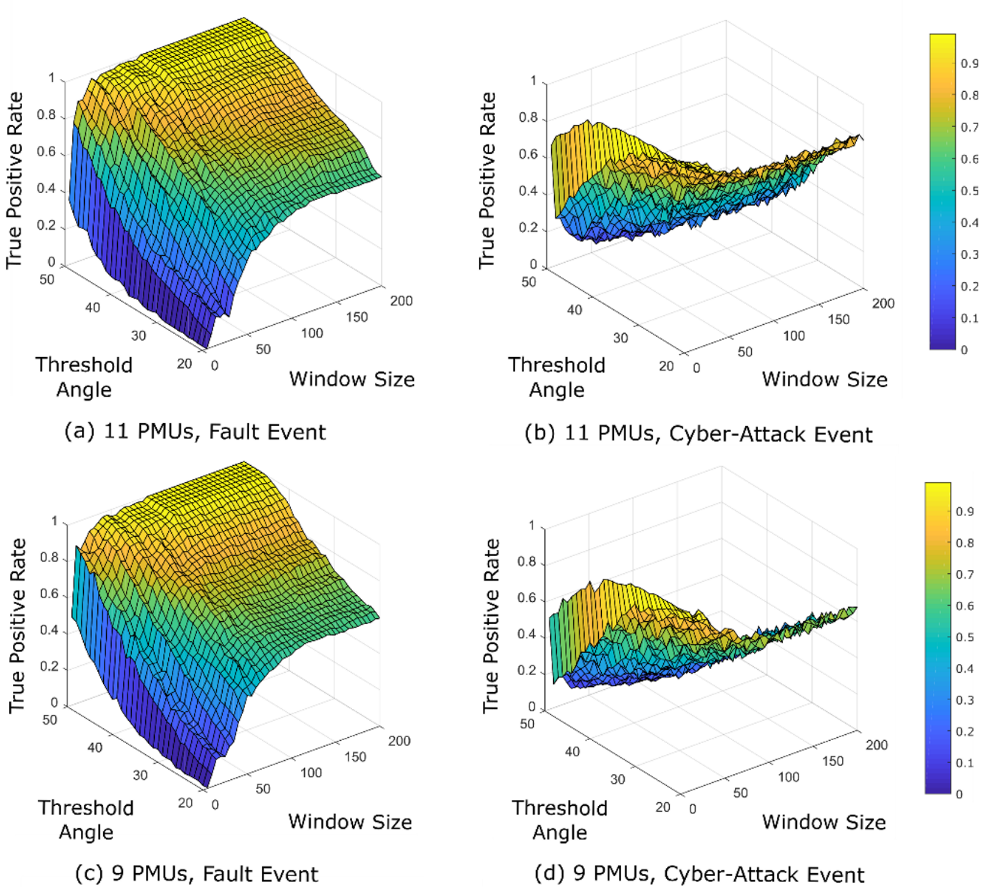
(c) 9 PMUs, Fault Event

(d) 9 PMUs, Cyber-Attack Event

(e) 7 PMUs, Fault Event

(f) 7 PMUs, Cyber-Attack Event

Figure 1.14: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, five PMUs are compromised. One historical event is used within the classification subroutine.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

(c) 9 PMUs, Fault Event

(d) 9 PMUs, Cyber-Attack Event

Figure 1.15: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network. Augmented network system parameters are used, and the two areas are 10x closer together. For the cyberattacks, nine PMUs are compromised. One historical event is used within the classification subroutine.

Table 1.3: True-Positive Selections for Angle Rotation Threshold of 39 Deg. and Window Size of 150 Samples, from Experiment Results in Fig. 1.10.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	0.93	0.98
7 PMUs	0.98	0.77
3 PMUs	1.0	0.05

Table 1.4: True-Positive Selections for Angle Rotation Threshold of 37 Deg. and Window Size of 105 Samples, from Experiment Results in Fig. 1.11.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	0.85	0.95
7 PMUs	0.97	0.83
3 PMUs	1.0	0

Table 1.5: True-Positive Selections for Angle Rotation Threshold of 31 Deg. and Window Size of 130 Samples, from Experiment Results in Fig. 1.12.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	0.99	0.98
7 PMUs	1.0	0.8
3 PMUs	1.0	0.28

Table 1.6: True-Positive Selections for Angle Rotation Threshold of 29 Deg. and Window Size of 105 Samples, from Experiment Results in Fig. 1.13.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	1.0	0.98
7 PMUs	1.0	0.77
3 PMUs	1.0	0.25

Table 1.7: True-Positive Selections for Angle Rotation Threshold of 38 Deg. and Window Size of 160 Samples, from Experiment Results in Fig. 1.14.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	0.91	0.87
7 PMUs	0.83	0.76

1.9.3 Performance with Partitioning Analysis used for Training

In this section, cyberattack detection algorithm performance is investigated in the case that the GEP partitioning tool from [20], described in an earlier section, is used to train the algorithm. In the case of the

Table 1.8: True-Positive Selections for Angle Rotation Threshold of 35 Deg. and Window Size of 160 Samples, from Experiment Results in Fig. 1.15.

	Fault True Positive Rate (TPR)	Cyberattack True Positive Rate (TPR)
11 PMUs	0.83	0.33
7 PMUs	0.78	0.29

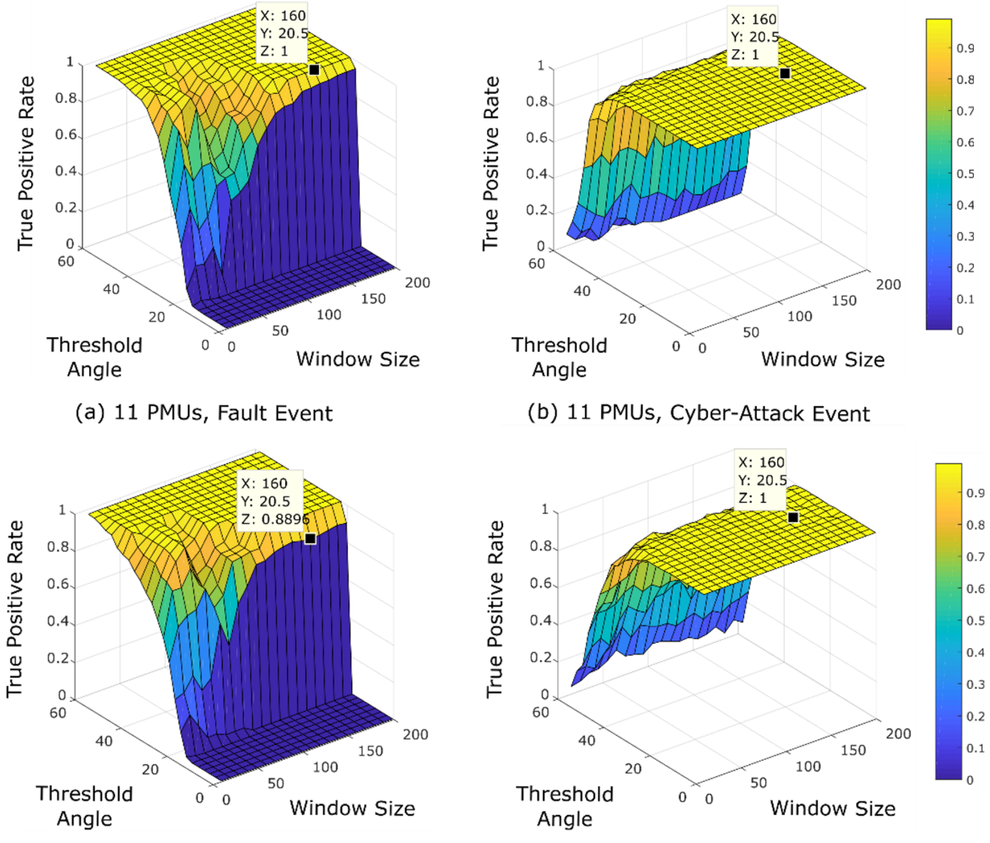
cyberattack events, five PMUs are compromised by the attacker. Knowledge of the physical characteristics of the 2-area, 11-bus system in Fig. 1.6, is assumed to be available to the cyberattack detection algorithm. This knowledge includes network susceptances and the network incidence matrix, machine location and inertias, and PMU location. The reduced, symmetric differential algebraic equation set (2) was determined, and $\mathbf{E}$ and $\mathbf{R}_\mathrm{S}$ calculated. To address the linearized equilibrium point implicit in the construction of $(\mathbf{E},\mathbf{R}_\mathrm{S})$, conditions of zero load, zero power injection, uniform voltage magnitudes and 0 phase angle for all buses were assumed. Steps 1-6 of the partitioning algorithm described earlier were applied. At the termination of the GEP algorithm, the parallel transmission lines defining the one half of the long transmission line corridor were identified as the optimal branch cutset that separates two clusters of nodes. It should be noted that the partitioning tool could have been applied iteratively, to further divide the clusters identified by the first iteration of the tool. The result of the GEP partitioning algorithm was post-processed, and the post-processed result was then applied to the classification subroutine. As shown in Fig. 1.16, a threshold angle of approximately 23-25 degrees would allow for perfect performance: true positive rates of 1.0 for both fault and cyberattack classification, if 11 or 7 PMUs are deployed.

1.9.4 Validation of Algorithm Performance using PNNL Testbed

GEGR collaborated with PNNL to test the developed cyberattack detection algorithm with more-realistic synchrophasor data using one of PNNLs software-based testbeds. PNNLs testbed simulated test system is comprised of a state-space representation of the MinniWECC model, linearized around an equilibrium operating point. A brief overview of the MinniWECC model and its implementation within PNNLs testbed is provided in [25]. The MinniWECC model includes 115 ac transmission lines, 34 generators, 19 load buses, and 2 dc transmission lines, and is aggregated in such a manner so as to preserve the two most significant inter-area modes of the true WECC system. Sixty-three buses are available for PMU measurement. Fig. 1.17 demonstrates voltage angle transients from 30 PMUs distributed throughout the MinniWECC system, after preconditioning.

To test GEGRs algorithm, 5 realistic physical events were simulated using PNNLs software testbed. The resulting 63 voltage phasor signals obtained during the simulation were down-selected to match the chosen number of PMUs assumed to be deployed in the system. To process the PNNL data, scripts were developed at GEGR that performed a randomized down-selection to the chosen number of PMUs, which effectively randomized the deployment of PMUs within the system.

For a given deployment configuration for the PMUs, a random selection of the 5 available events was used as the historical event for training purposes, and, in the case of fault events a second random selection of the 5 available events was used as the new event. In the case of the cyberattack event, normal, nonevent data is substituted for the new event. Five fault and cyberattack tests are performed, each with a new random selection of compromised PMUs. These 5 fault and cyberattack tests were repeated for 10 different PMU deployments configurations, for a total of 100 tests (50 fault and 50 cyberattack). Finally, the 100 tests were repeated for each element within a larger test matrix, which varied the number of PMUs deployed and the number of compromised PMUs.



(a) 11 PMUs, Fault Event

(b) 11 PMUs, Cyber-Attack Event

(c) 7 PMUs, Fault Event

(d) 7 PMUs, Cyber-Attack Event

Figure 1.16: True positive rates for fault event and cyberattack event classification, across a range of threshold angle parameters, window sizes, and number of PMUs in the network (11 and 7). Nominal 2-area system parameters are used. The GEP partitioning tool is used to train the algorithm. For the cyberattacks, five random PMUs are compromised.

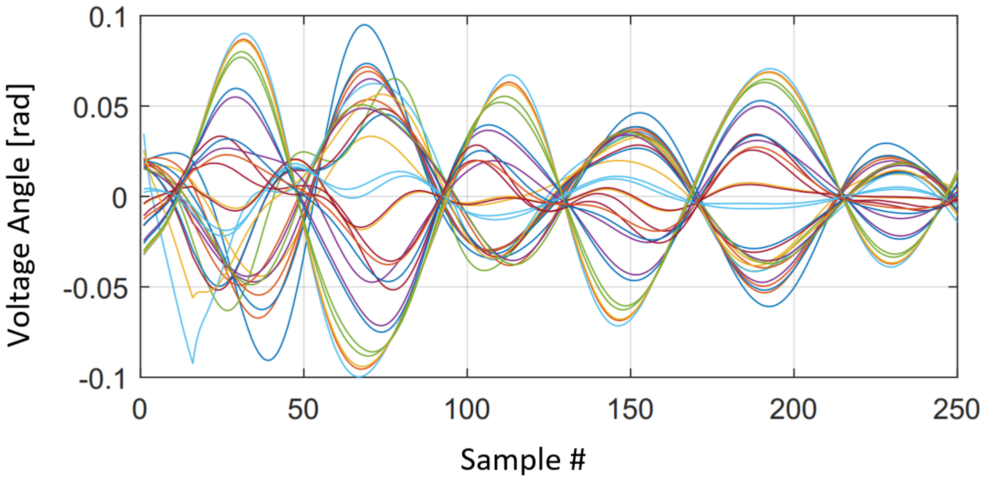


Figure 1.17: Example voltage angle transient from post-processed, simulated synchrophasor data, obtained at 30 nodes within the MinniWECC system (raw phasor data generated at PNNL).

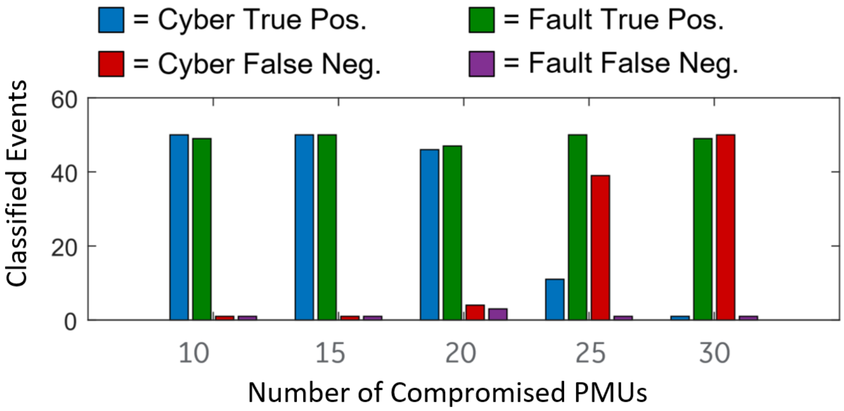


Figure 1.18: Performance of GEGRs algorithm when applied to GEGRs testbed data. 30 PMUs are deployed.

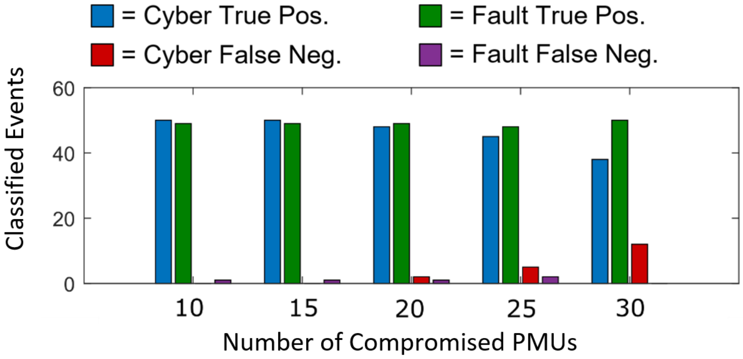


Figure 1.19: Performance of GEGRs algorithm when applied to GEGRs testbed data. 40 PMUs are deployed.

Algorithm parameter selections were informed by the results from rigorous experiments with the GEGR software-based testbed, described earlier. The rotation angle threshold was chosen to be 30 degrees, and the window size chosen to be 100 samples. Manual training using the GEP partitioning algorithm was not undertaken; historical events are used to train the algorithm, and only one historical event was used in the event repository.

If PMUs are deployed at 30 of the buses in the system, Fig. 1.18 shows the relative amounts of correctly- and incorrectly-classified fault and cyberattack events for different quantities of compromised PMUs. The algorithm was found to perform satisfactorily, as close to 50 of the fault events and close to 50 of the cyberattack events were classified properly if 10 or 15 PMUs are compromised in a system with 30 PMUs. The algorithm still performs in a relatively robust manner even if 20 PMUs are compromised by the malicious entity. However, there is a sharp cut-off phenomenon: performance degrades significantly if only 5 or less PMUs are left uncompromised. If 40 PMUs are deployed, Fig. 1.19 shows how performance of the algorithm improves, for the same numbers of compromised PMUs. If 50 PMUs are deployed, and 30 are compromised in a cyberattack, Fig. 1.20 shows perfect classification for both faults and cyberattacks.

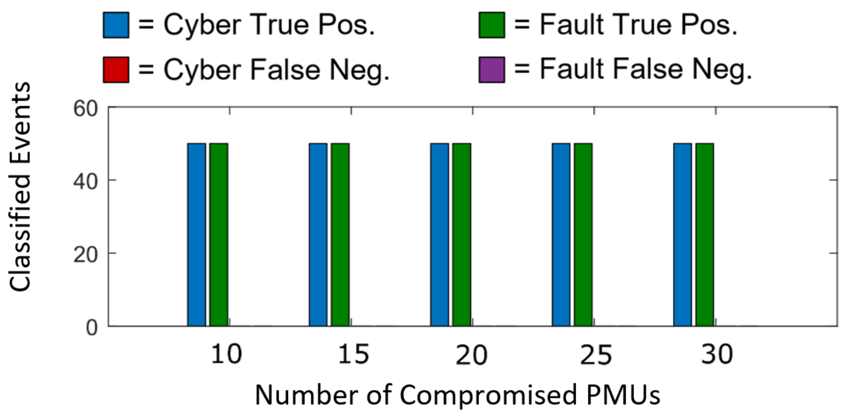


Figure 1.20: Performance of GEGRs algorithm when applied to GEGRs testbed data. 50 PMUs are deployed.

1.10 Conclusions: Algorithm Development

Rigorous testing using GRCs software-based testbed showed that the WAMS cyberattack detection algorithm demonstrates either satisfactory or exceptional performance under a broad range of conditions. However, results did show that the algorithm may have a weakness when applied to power systems that do not contain long inter-area transmission lines. Therefore, it can be surmised that performance of the algorithm is dependent upon the physical characteristics of the system included within the boundaries of the WAMS network, including the existence (or lack there-of) of dominant, inter-area swing modes within the WAMS network footprint.

A thorough investigation was made of the sub-optimal condition in which the inter-area transmission lines were significantly reduced in length. In this sub-optimal condition, the algorithm appears to require a larger number of PMUs and knowledge of more historical events to achieve satisfactory performance. Refinements will be made to the algorithm to address the classification performance under these conditions.

In the case of both the original, unaltered Kundur two-area test system from [24] and in the realistic MinniWECC model–the algorithm demonstrates exceptional ability to discern between fault and replay attacks. Excellent performance was observed even under the circumstances in which features & signatures from only one historical event are available to the classification subroutine and a majority of the PMUs have been compromised in the cyberattack. Results also showed that even a single iteration of the GEP partitioning tool is sufficient to fully train the algorithm such that it achieves high classification performance. Finally, the collaboration with PNNL presented the opportunity to successfully test the algorithm in a larger test system, with a larger quantity of data, and more-realistic system behavior.

Additional refinements in Phase II and included in the subsequent sections include augmentation of the initial event detection subroutine of the algorithm. Additional topics of investigation may include, but are not limited to: (1) the exploration of how power system stabilizers and other yet-unmodeled components impact the algorithm performance; and (2) the continued validation of the algorithm using even larger test systems, potentially including 1000s of buses.

GE Digital runs EMS software for 10 of the 14 largest power grid operators, 170 systems in USA, 70% of Middle East utilities, and 70% of Africa, and is a leading WAMS software provider. In Section 4 of this report, the prototype algorithm will be demonstrated in a more realistic environment by integrating it with the GE Digital PhasorAnalytics software platform used for synchrophasor visualization.

1.11 Bibliography: Algorithm Development

[1] Pratyusa K. Manadhata, An Attack Surface Metric, Ph.D. dissertation, Carnegie Mellon University, 2008.

[2] R. Deng, G. Xiao, R. Lu, H. Liang and A. V. Vasilakos, ”False Data Injection on State Estimation in Power Systems–Attacks, Impacts, and Defense: A Survey,” IEEE Transactions on Industrial Informatics, vol. 13, no. 2, pp. 411-423, 2017.

[3] Y. Liu, P. Ning, Reiter and M. K., ”False Data Injection Attacks Against State Estimation in Electric Power Grids,” Proceedings of the 16th ACM Conference on Computer and Communications Security, pp. 21-32, 2009.

[4] H. He and J. Yan, ”Cyber-physical attacks and defences in the smart grid: a survey,” IET Cyber-Physical Systems: Theory & Applications, vol. 1, no. 1, pp. 13-27, 2016.

[5] R. Mitchell and I. R. Chen, Behavior-Rule Based Intrusion Detection Systems for Safety Critical Smart Grid Applications, IEEE Transactions on Smart Grid, vol. 4, no. 3, pp. 12541263, Sep. 2013.

[6] S. G. Ghiocel, ”Phasor-measurement-based state estimation for synchrophasor data quality,” IEEE Transactions on Power Systems, 2014.

[7] K. D. Jones, A. Pal and J. S. Thorp, ”Methodology for Performing Synchrophasor Data Conditioning and Validation,” IEEE Transactions on Power Systems, 2015.

[8] F. Pasqualetti, F. Drfler and F. Bullo, ”Attack Detection and Identification in Cyber-Physical Systems,” IEEE Transactions on Automatic Control, 2013.

[9] M. Wang, ”A Low-Rank Matrix Approach for the Analysis of Large Amounts of Power System Synchrophasor Data,” in 2015 48th Hawaii International Conference on System Sciences, 2015.

[10] M. Wu and L. Xie, ”Online Detection of Low-Quality Synchrophasor Measurements: A Data-Driven Approach,” IEEE Transactions on Power Systems, 2017.

[11] J. Wang, D. Shi, Y. Li, J. Chen, H. Ding and X. Duan, ”Distributed Framework for Detecting PMU Data Manipulation Attacks With Deep Autoencoders,” IEEE Transactions on Smart Grid, 2018.

[12] T. Tran, O. Shin and J. Lee, ”Detection of replay attacks in smart grid systems,” in 2013 International Conference on Computing, Management and Telecommunications (ComManTel), 2013.

[13] M. M., Z. P., D. D., P. C., F. M. and A. H.M., ”Detecting Replay Attacks in Power Systems: A Data-Driven Approach,” in Advanced Computational Methods in Energy, Power, Electric Vehicles, and Their Integration, Springer, 2017.

[14] S. Pan, T. Morris and U. Adhikari, ”Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems,” IEEE Transactions on Smart Grid, 2015.

[15] J. Landford, ”Fast sequence component analysis for attack detection in smart grid,” in 2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS), 2016.

[16] P. Overholt, D. Ortiz, and A. Silverstein, Synchrophasor Technology and the DOE: Exciting Opportunities Lie Ahead in Development and Deployment, IEEE Power and Energy Magazine, vol. 13, no. 5, pp. 14-17, Sept. 2015. See: https://www.midwestreliability.org/MRODocuments/MISO%20PMU%20Presentation%20to%2

[17] Y. Chen, L. Xie, and P. R. Kumar, Dimensionality reduction and early event detection using online synchrophasor data, in 2013 IEEE Power Energy Society General Meeting, 2013.

[18] L. Xie, Y. Chen, and P. R. Kumar, Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis, IEEE Transactions on Power Systems, vol. 29, no. 6, pp. 27842794, Nov. 2014.

[19] C. DeMarco and J. Wassner, A generalized eigenvalue perturbation approach to coherency, in Proceedings of the 4th IEEE Conference on Control Applications, Sep. 1995, pp. 611617.

[20] J. Chow, Time-Scale Modeling of Dynamic Networks with Applications to Power Systems. SpringerVerlag Berlin and Heidelberg GmbH, 1982

[21] R. Podmore, Identification of Coherent Generators for Dynamic Equivalents, IEEE Transactions on Power Apparatus and Systems, vol. PAS-97, no. 4, pp. 13441354, Jul. 1978.

[22] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill and B. C. Pal, ”Coherency identification in power systems through principal component analysis,” in IEEE Transactions on Power Systems, vol. 20, no. 3, pp. 1658-1660, Aug. 2005.

[23] P. Kundur, N. J. Balu, and M. G. Lauby, Power system stability and control. New York: McGraw-Hill, 1994, pp. 813-816.

[24] James D. Follum, Electromechanical Mode Estimation in the Presence of Forced Oscillations, Ph.D. dissertation, University of Wyoming, 2014.

[25] K. Mahapatra, N. R. Chaudhuri, R. Kavasseri, and S. Brahma, Online Analytical Characterization of Outliers in Synchrophasor Measurements: A Singular Value Perturbation Viewpoint, IEEE Transactions on Power Systems, vol. PP, no. 99, 2017.

2 Phase IIA: Event Detection Subroutine Refinement

2.1 Introduction

Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior. These nonconforming patterns are often referred to as anomalies, discordant observations or peculiarities in different application domains [1]. For a modern power grid, an anomaly could manifest as physical disturbances including transmission line faults, fault induced protective relay action, a sudden loss of generation or load, oscillation, etc. or malicious cyber activities such as false data injection, spoofing attack, or critical information availability attack. Prompt detection of these two types of anomalies in power grid is vital in order to prevent economic losses and protect utility assets and even human lives. The availability of real-time, spatially-distributed system operation data of high accuracy is indispensable for system-level anomaly detection. Synchrophasor data collected by Phasor Measurement Units (PMU) have brought unprecedented observability into the operation of power grids. The integration of streaming data from a group of synchronized PMUs provides real-time ”snapshots” of the grid under monitoring with high fidelity, and the critical information that those data carry should be fully leveraged in anomaly detection. Meanwhile, the vast amount of PMU data renders the manual analyses time consuming, and the conventional rule-based detection schemes ineffective, making it almost impossible to capture unobserved anomalies with sufficient expediency. Therefore, researchers have been attempting to develop automated offline and online anomaly detection methods to cope with streaming PMU signals.

In the literature, there are many examples of synchrophasor anomaly detection algorithms that use static thresholds determined in an offline manner [2]-[6]. In [2], a density-based local outlier factor approach is proposed to detect false data injection attacks, and the decision threshold is determined using offline training and a historical dataset. In both [3]and [4], the static decision threshold is determined in an offline, brute force manner, by repeatedly applying the proposed algorithm to a sample PMU dataset while sweeping the threshold across a wide range. A user guide composed by NREL [5] introduces a few event detection methods that can be used to analyze streaming PMU data. 3-σ criteria is utilized to distinguish the sliding windows that contain major events from those which don't. These methods can be used to extract features from synchrophasor sliding windows, but it is worth noting that the 3-σ confidence interval deployed in these methods is static and determined by the complete offline observation set. Similarly, Aditya et al. utilizes a static confidence interval in online anomaly detection, which is derived offline based on historical observances of the metrics under consideration [6]. Most of the aforementioned anomaly detection methods leverage either static threshold or static confidence interval based criteria to determine if a specific system state is abnormal or not, and thus lack adaptiveness to unobserved scenarios.

In this section, we propose a PMU data driven online anomaly detection solution for power grids, which adopts Chebyshev's Inequality using two different approaches: (i) accumulative feature sample set and (ii) sliding feature sample set. Both approaches replace the static threshold or confidence interval with an adaptive confidence interval in the anomaly detection. Use of Chebyshev's Inequality in anomaly detection reduces the massive efforts required to determine a static detection criterion. The proposed solution is capable of automatically updating the knowledge base of the detection engine such that it guarantees the detection decision is always made according to an up-to-date confidence interval. In the following subsections, a feature engineering method entitled Minimum Volume Enclosing Ellipsoid (MVEE) is reviewed, Chebyshev's Inequality and the two approaches for its application in online anomaly detection are introduced, and experimental evaluation and result analysis are elaborated upon.

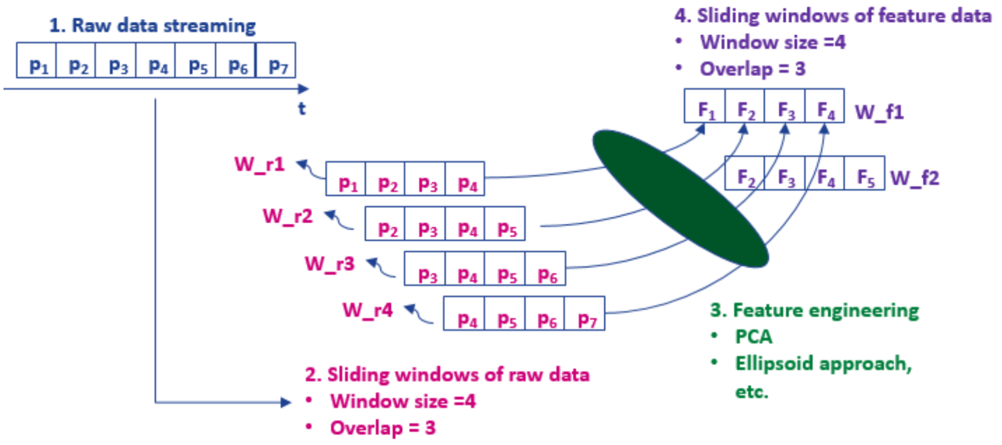


Figure 2.1: Sliding window schema.

2.2 Streaming PMU Data Feature Extraction

Prior to application of the anomaly detection algorithm, it is often required to carry out feature engineering in order to extract the most critical information and/or reduce the dimension of the raw data. In this section, we apply multidimensional Minimum Volume Enclosing Ellipsoid (MVEE) on the PMU data scoped by every sliding window to attain the features needed for anomaly detection.

2.2.1 Application of Sliding Window

Critical power system event information is embedded in raw PMU data sets. For streamed data such as synchrophasor measurements, industry practitioners often utilize a sliding window for data analysis [8] such that the feature extraction and detection can be performed in the context of the most recent raw data segment. Various features can be derived to capture the correlation among multiple PMU channels. Such features can serve as holistic system state indicators. It should be noted that the sliding window concept can be applied not only to raw data, but also to feature values repeatedly derived from this data, since the feature values also become a conventional time-series dataset. Fig. 2.1 depicts the main approach to applying sliding windows used in this section. The label ”pi” in the figure denotes a sample of the multidimensional PMU data at time instance ’*i*’, and ”F*i*” represents the features value attained from a raw data sliding window ”W_r*i*”. ”W_f*i*” is a second level sliding window, being applied in the feature space. The length of the sliding window and the overlap between two adjacent windows usually is determined according to the time scale of the specific targeted events empirically or as a tuning parameter by conducting repetitive offline trainings. For online application, the window length dictates the applicable feature extraction techniques, due to computational complexity concerns.

2.2.2 MVEE based Feature Engineering

In reference [7], the authors propose a feature extraction strategy entitled Minimum Volume Enclosing Ellipsoid (MVEE). It defines a problem of searching for a minimum volume ellipsoid enclosing a set of PMU data points within a sliding window. The data within a sliding window can be denoted by a matrix $X_{N\times M}$, i.e., this data set contains M samples from N PMU measurement channels. The sliding window matrix X can also be denoted as $x_1, x_2, x_3, ..., x_M, x \in R^N$. An enclosing ellipsoid is defined as in (17).

$$E(A, c) = \{x \in R^N | (x - c)^T A(x - c \leq 1\} \tag{17}$$

In (17), vector $c \in R^N$ represents the center of the ellipsoid, and the positive definite matrix A contains information regarding the shape and orientation information of the ellipsoid. Both A and c can be determined by solving the optimization problem (18), and A and c together determines the MVEE enclosing all the data points within the sliding window X. The objective function of (18) reflects the volume of the ellipsoid.

$$\min_{A,c} \frac{1}{\sqrt{\det A}} s.t.(x_i - c)^T A(x_i - c \leq 1), i = 1, ..., M, A \prec 0 \tag{18}$$

After the MVEE is obtained by solving (18), different features associated with the ellipsoid can be extracted and utilized for anomaly detection. For instance, the semi-axis lengths of the ellipsoid corresponding to sliding window j can be calculated after applying singular value decomposition on matrix A^j. In this calculation, $u_i, v_i \in R^n$ are the left and right singular vectors respectively, and λ_i are the singular values. The semi-axes lengths r_i of the MVEE can be obtained as

$$r_i^j = \frac{1}{sqrt\lambda_i} \tag{19}$$

In this section, we propose to use the log sum of λ_i^j as defined in (20) as the anomaly detection feature. Please refer to [7] for more features that can be derived from MVEE.

$$feat^j = log \sum_{i=1}^{N} \lambda_i^j \tag{20}$$

2.3 Application of Chebyshev's Inequality

Critical information embedded in the streaming PMU data can be obtained after feature extraction. But in order to detect anomalous operating states of a power grid, a decision has to be made given a specific feature value. Typically, when the observed feature values deviate from the ”normal pattern”, one may claim the occurrence of anomalies in the system.

2.3.1 Static Threshold vs. Chebyshev's Inequality

In order to tell if the feature value deviates from the normal value, we can compare the real time feature values with a static threshold. Fig. 2.2 provides an example. The bottom plot depicts the voltage magnitude measurements from 11 PMUs, and the top one shows the feature values corresponding to sliding windows. We can qualitatively observe that the absolute feature value increases when the events occur. While a static threshold of '5' would provide satisfactory detection performance for the events shown here, a clear deficiency is that the static threshold lacks self-adaptiveness, and it will require manual adjustment when facing other events to maintain the desired performance in a dynamic system. The traditional static threshold often remains unchanged for too long, resulting in more frequent false positives and false negatives as the system state evolves.

Instead of determining a static threshold through laborious offline analyses or qualitative assessments, we can reformulate this problem as seeking for the confidence interval for the feature value with a predetermined confidence level by utilizing Chebyshev's Inequality.

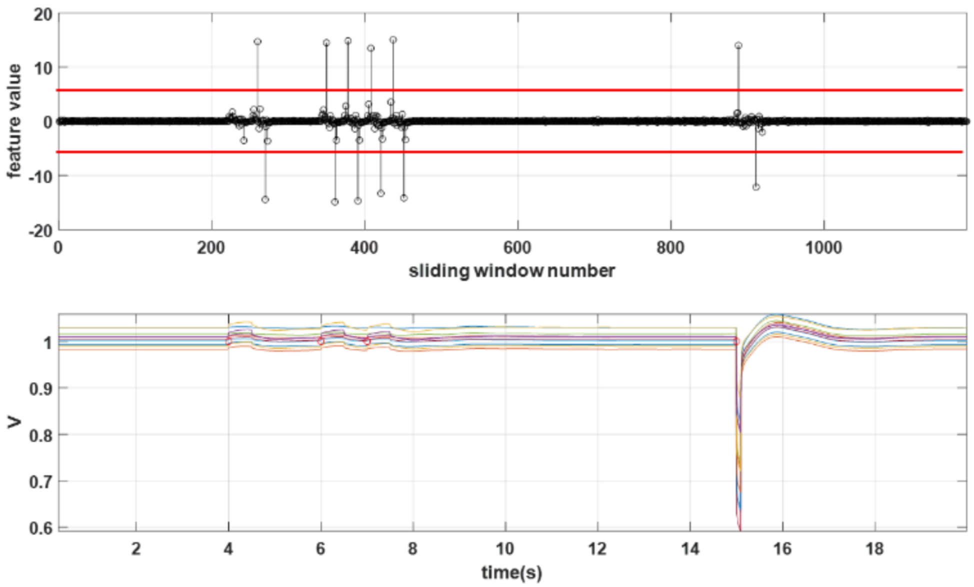


Figure 2.2: Anomaly detection based on static threshold.

For a random variable x with population mean μ and standard deviation σ, and any $k > 0$, Chebyshev's Inequality states that

$$P(|x - \mu| \geq k\sigma \leq \frac{1}{k^2}) \tag{21}$$

Given a confidence level $\frac{1}{k^2}$, the confidence interval in anomaly detection can be denoted as [μkσ,μ+kσ]. The expression indicates that the probability of x$\notin$ [μkσ,μ+kσ] is less than $\frac{1}{k^2}$. It's noteworthy that, to apply Chebyshev's Inequality, we need either the knowledge of the population mean and variance or a very large number of samples so that the sample mean and variance can be also utilized. However, in practice, operators can only have access to limited number of sliding windows.

To cope with this disadvantage of Chebyshev's Inequality, statisticians have developed an extended version of (21) [9]-[10]. With a given sample set of limited size, the extended Chebyshev's Inequality defines the confidence interval with sample mean m, sample standard deviation s, and the sample size N as in (22), where $Q^2 = \frac{N+1}{N}s^2$. The constant k can be set as 3 to 5 to achieve a satisfactory confidence level. The fact that there is very low probability that the new feature value will fall outside of the confidence interval, in a statistical sense, can be used to determine if an anomaly has occurred.

$$P(|x - \mu| \geq kQ) \leq \frac{1}{N+1}[\frac{N+1}{N}(\frac{N-1}{k^2} + 1)]) \tag{22}$$

In the remainder of this section, considering that the application of the extended Chebyshev's Inequality relies on a feature value sample set, we will propose two different approaches by which to scope the feature value sample set.

2.3.2 Approach 1: Accumulative Feature Sample Set (CI 1)

The first proposed approach by which to implement Chebyshev's Inequality is to use an accumulative feature sample set to determine the confidence interval. As shown in Fig. 2.3, the continuous online detection

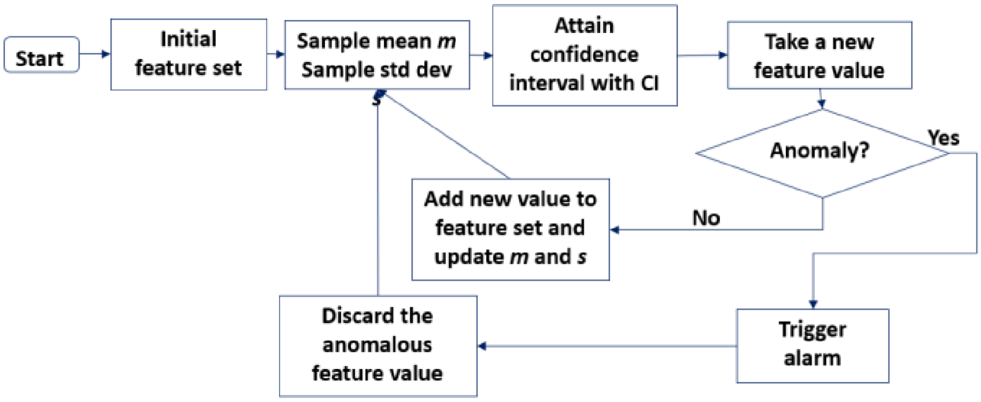

Start
Initial feature set
Sample mean *m*
Sample std dev *s*
Attain confidence interval with CI
Take a new feature value
Anomaly?
Yes
No
Add new value to feature set and update *m* and *s*
Discard the anomalous feature value
Trigger alarm

Figure 2.3: Flowchart of CI 1.

starts with an initial feature sample set. The sample mean and standard deviation can be calculated and then further utilized to determine the confidence interval under a certain confidence level. When a new feature value is obtained based on the latest sliding window, it is compared against the confidence interval. If the new feature value falls out of the interval, it is treated as an anomaly. Otherwise, the system state is determined as normal. Only if the new feature value is normal will it be added into the feature sample set, and a re-computation of the sample mean and standard deviation is needed. In this method, the feature sample set continues growing. This method can determine the duration of events by continuously triggering alarms during the entirety of the event since it does not involve any anomalous feature values in the sample set.

$$S_N^2 = \frac{1}{N-1}\sum_{j=1}^{N}(f_j - \bar{f}_N)^2 \tag{23}$$

$$S_N^2 = \frac{N-1}{N-1}S_{N-1}^2 + \frac{1}{N}(f_N - \bar{f}_{N-1}) \tag{24}$$

When applying Chebyshev's Inequality with an accumulative feature sample set, the confidence interval needs to be dynamically updated after a new sample is observed. To improve the computation efficiency, equation (24) is utilized instead of (23) in the sample variance update. In both equations, S_N^2 and $\bar{f}_N$ is the sample variance and sample mean of the augmented sample set respectively, which contains N feature values $f_j, j = 1, ..., N$. $\bar{f}_{N-1}$ is the sample mean before the new feature value is observed. With (24), the sample variance computation complexity decreases from O(N) to O(17).

2.3.3 Approach 2: second-level sliding window (CI 2)

Another method by which to prepare the feature sample set is to keep updating the sample set along with the observance of new sliding windows. As shown in Fig. 2.4, whether the feature value derived from the latest sliding window is normal or not, it will be added to the feature sample set. In the meantime, the oldest feature value will be discarded from the feature sample set. In such a way, the feature sample set becomes the second level sliding window, as shown in Fig. 2.1. This detection method updates its sample set whenever a new feature value arrives,and thus the feature sample set remains the same length. This approach can reduce the alarms since the feature sample set will adapt to transient processes.

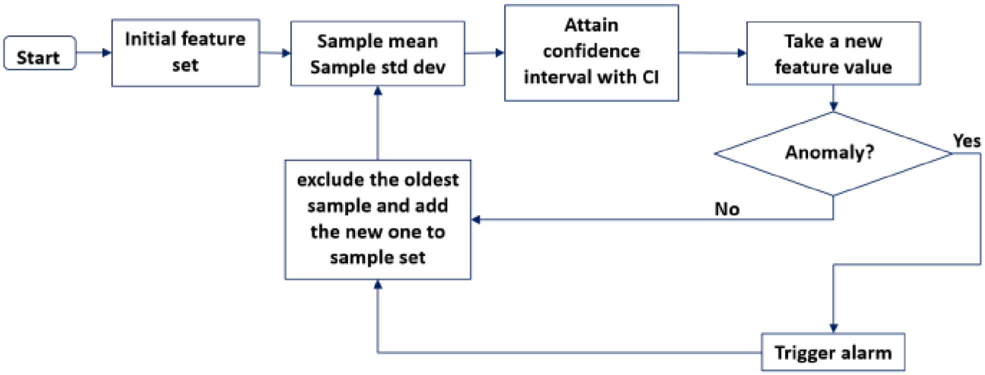


Figure 2.4: Flowchart of CI 2.

2.4 Experimental Evaluation

2.4.1 Experimental Setup

In this section, we evaluate the proposed solution with two power system models. The first is the Kundur's 2-area 4- machine system introduced in [11], as depicted in Fig. 1.6. This system is simulated using Matlab/Simulink, and we assume that all the 11 buses have PMUs installed. The second model is the Mini-WECC system [12], which is a reduced-order dynamic model of the WECC system with 120 buses, 172 lines/transformers and 34 generators. Mini-WECC system is modeled and simulated with the MATLAB toolbox PST (Power System Toolbox).

Both physical disturbances and data integrity attacks are simulated for the 2-area system. In total, 200 physical event simulations are conducted, each lasting for 20 seconds. Each simulation involves a load step change event at t=4s and a short circuit fault at t=15s. 200 more simulations are conducted when the system is under data integrity attacks. We assume the adversary has recorded historical PMU measurements, and he will replay the recorded data on a randomly selected PMU channel in real-time. In the simulation, the cyber attack is equivalent to replacing the data from one PMU channel corresponding to the event A with the data from the same channel recorded for a historical event B. Three-phase to ground faults of a transmission line and load step changes are simulated with Mini-WECC model. For each type of event, twelve 20s-long simulations are conducted. The three-phase fault occurs on a pre-determined critical line at 15s, and the fault clears at 15.1s. For the simulation of load step change, all loads in the system change at 1.4s simultaneously. The PMU sampling rate for all the simulations are configured to be 60 samples per second, and the voltage magnitude and angle are recorded for each bus. In the application of the MVEE feature extraction, the sliding window size is selected as 20 samples, and the overlap between two adjacent windows covers 19 samples. In this section, we assume that the PMU data are ”cleaned” by data preprocessing such as data filtering, and therefore, the interference influences due to data noises, harmonics, etc. have not been incorporated in the synthetic data. We will investigate such conditions in our future work.

2.4.2 2-Area system case study

Plots in Fig. 2.5 provides some insights into the performance of the static threshold method, and the CI 1 and CI 2 approaches. Fig. 2.5 (a) and (b) represent a physical event simulation and a cyber event simulation respectively. Subplots in each from bottom to top are the positive sequence voltage magnitudes of 11 PMUs, feature values, detection results of static threshold and that of CI 1. The detection outcome for each sliding window is either 1 or 0. A pulse taking value 1 represents an anomalous system state, while 0 means the system condition is normal. From (a) and (b), it can be seen that the CI 1 is able to detect all the anomalies in time (note the 4s and 15s time points). The static threshold based detection fails to detect the load step change event in (a) at 4s, and for it to work for cyber attack scenarios in (b), the static threshold needs to be changed to 0.5. Otherwise, nothing will be detected for (b). Realistically, one cannot change the static threshold for every specific event. For the same two events, Fig. 2.5 (c) and (d) depicts the detection results with CI 2. If one compares the top subplot in (a) with (c), it is obvious that the CI 2 causes less alarms than CI 1. The reason is that CI 2 updates its base feature sample set by incorporating every feature value regardless if it is normal or abnormal. That is, CI 2 makes the detection based on the information from the ”localized” time interval with predefined length while decision making in CI 1 depends on all the information in the past. If the feature sample set size is too small, such as 50 feature values are utilized in (d), CI 2 can frequently induce false positives. Statistically, Table 1 lists the average detection time for the three different anomaly detection approaches. If a true event doesn’t get detected, the corresponding detection time is noted as 20s, which is the end time of the simulation. The performance of CI based approaches is better than the static threshold in terms of detection promptness. Table 2 provides the recall values (recall = $\frac{TruePositive}{AllPositive}$) for the three approaches. Recall is adopted here since a false negative is much more undesirable compared to a false positive in power system. The detection recall for both CI based methods are satisfactory in physical events detection. However, for the stealthy replay attack that we simulate, the recall is not as good with the single MVEE feature selected. Other features that are more sensitive to the cyber attacks should be leveraged in the detection. But with the given set-up, we can see that CI 2 is more sensitive than CI 1 and delivers better recall. Fig. 2.6 shows the detection results when multiple anomaly events take place in quick temporal succession. Three load changes occur at 4s, 6s, and 7s respectively. CI 1 (50 initial feature values) roughly scopes the anomaly as two time intervals, while CI 2 (80 feature values) doesn’t trigger any alarms during the third event. This is because after the secondload change event, the sample set is full of ”abnormal” samples and therefore, feature values during the third load change are labeled as normal. This can be regarded as a disadvantage intuitively, but in fact it might become an advantage when we want to reduce the alarm numbers, and there is no need to distinguish among frequent events.

Table 2.1: Average detection time (s).

Event Type	**CI 1**	**CI 2**	**Threshold**
Load Change	4.0634	4.0333	12.8
Fault	15.1976	15.2044	16.3
Cyberattack	17.4160	17.3443	20.0

Table 2.2: Average detection recall (%).

Event Type	**CI 1**	**CI 2**	**Threshold**
Load Change	100	100	46.5
Fault	100	100	78.0
Cyberattack	54.0	67.5	0

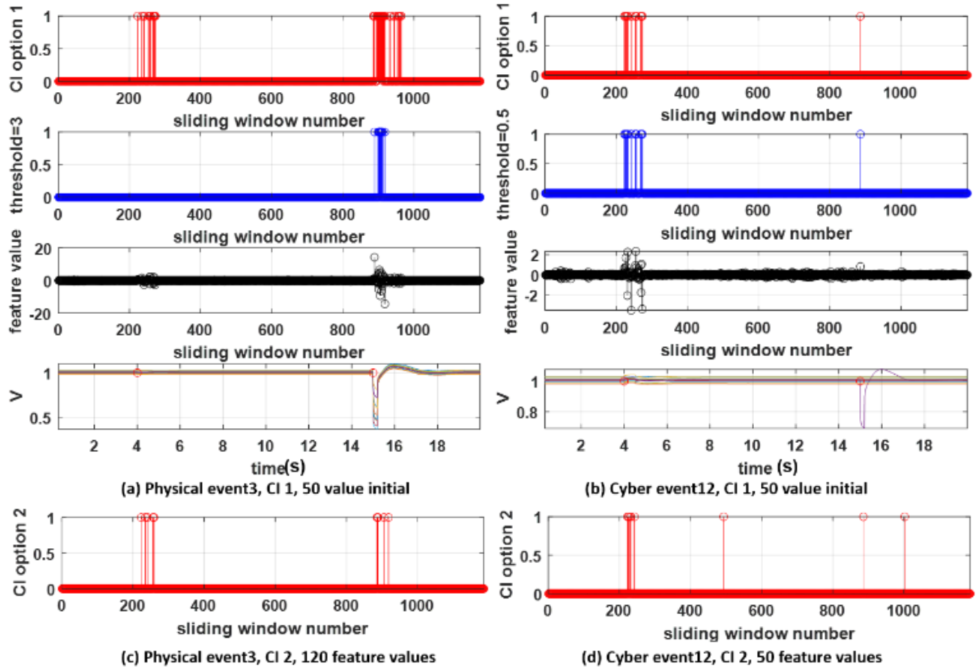


Figure 2.5: Holistic comparison of CI 1, CI2 and static threshold.

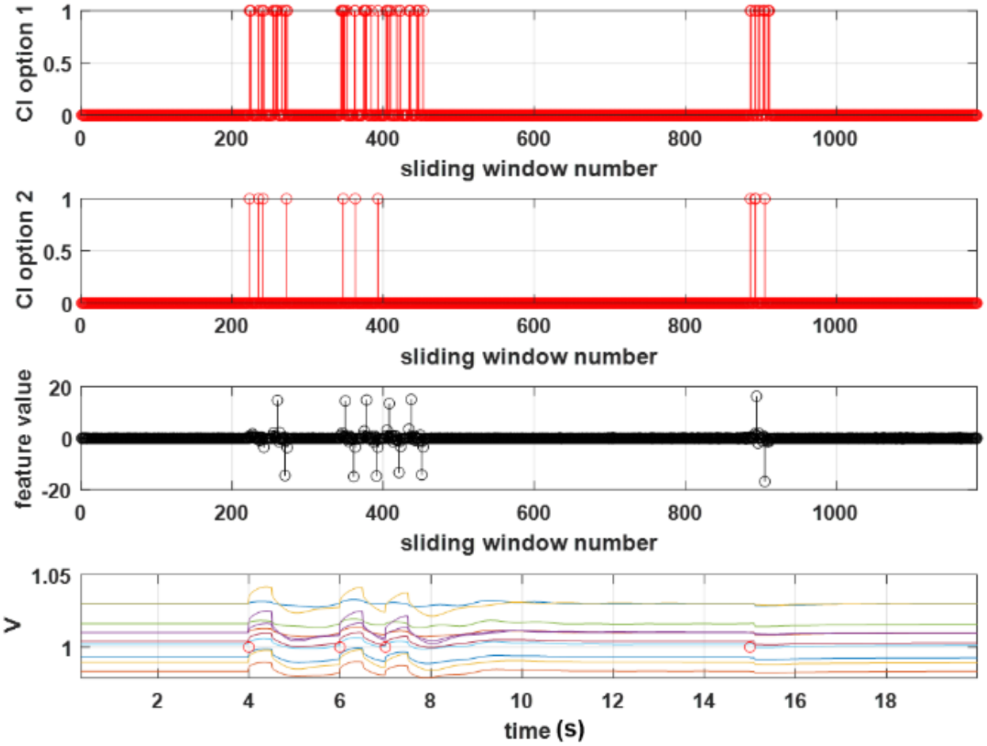
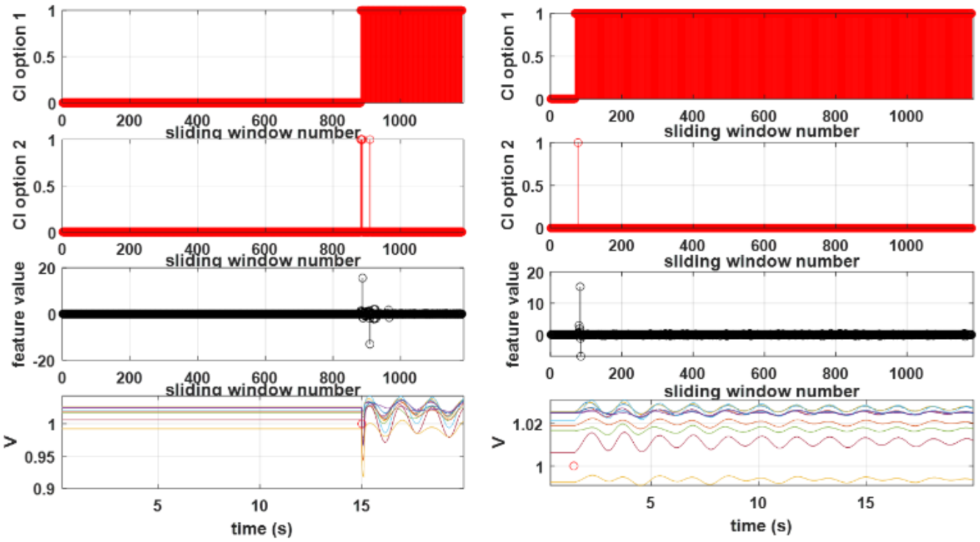


Figure 2.6: Frequent events detection.



(a) Short circuit at 15s

(b) Load change at 1.4s

Figure 2.7: CI based methods applied on Mini-WECC data.

2.4.3 Mini-WECC case study

In Fig. 2.7, (a) represents a short circuit event occurs in MiniWECC system and (b) a load step change. CI 1 keeps sending alarm from the time when the event starts for both cases, while the CI 2 only triggers a few alarms in the beginning. This further demonstrates the difference between CI 1 and CI 2. Fig. 2.7 CI based methods applied on Mini-WECC data

2.5 Conclusions: Anomaly Detection Subroutine

This section investigates the application of Chebyshev's Inequality in power system anomaly detection. Two approaches are proposed: CI 1 utilizes accumulated feature samples to determine the dynamic confidence interval; while CI 2 relies on a fixed-size feature sample set. Compared to CI1, CI 2 is more sensitive to anomalous events and it can also reduce the number of alarms. Meanwhile, CI 2 is more likely to commit false positive errors, and its performance highly depends on the selected feature sample size. In practice, multiple features should be utilized in anomaly detection, and CI 1 and 2 can be combined to achieve better accuracy and adaptiveness.

2.6 Bibliography: Anomaly Detection Subroutine

[1] V. Chandola, A. Banerjee, and V. Kumar, ”Anomaly detection: A survey,” ACM Comput. Surv., vol. 41, no. 3, pp. 158, 2009.

[2] M. Wu and L. Xie, ”Online Detection of False Data Injection Attacks to Synchrophasor Measurements: A Data-Driven Approach,” in Hawaii International Conference on System Sciences, Jan. 2017.

[3] J. M. Lim and C. L. DeMarco, ”Bad data detection and estimation in high dimensional measurement data,” in 2017 IEEE Power Energy Society General Meeting, 2017, pp. 15.

[4] P. J. Hart, S. Acharya, H. Wang, ”Coherency-Based Detection Algorithm for Synchrophasor Cyberattacks,” in North American Power Symposium, Oct. 2019.

[5] A. Allen, M. Singh, E. Muljadi, and S. Santoso, ”PMU Data Event Detection: A User Guide for Power Engineers,” October, 2014.

[6] A. Ashok, M. Govindarasu, and V. Ajjarapu, ”Online Detection of Stealthy False Data Injection Attacks in Power System State Estimation,” IEEE Trans. Smart Grid, vol. 9, no. 3, pp. 16361646.

[7] J. Ma, Y. V. Makarov, C. H. Miller, and T. B. Nguyen, ”Use multidimensional ellipsoid to monitor dynamic behavior of power systems based on PMU measurement,” IEEE Power Energy Soc. 2008 Gen. Meet. Convers. Deliv. Electr. Energy 21st Century, PES, pp. 18, 2008.

[8] P. Wang, M. Govindarasu, A. Ashok, S. Sridhar, and D. McKinnon, ”Data-driven anomaly detection for power system generation control,” IEEE Int. Conf. Data Min. Work. ICDMW, vol. 2017-Novem, pp. 1082 1089, 2017.

[9] J. G. Saw, M. C. K. Yang, and T. S. E. C. Mo, ”Chebyshev Inequality with Estimated Mean and Variance,” The Amerian Statistician, vol. 38, no. 2, pp. 130132, 1984.

[10] A. Kabn, ”Non-parametric detection of meaningless distances in high dimensional data,” Stat. Comput., vol. 22, no. 2, pp. 375385, 2012.

[11] P. Kundur, N. J. Balu, and M. G. Lauby, ”Power system stability and control,” New York: McGraw-Hill, 1994, pp. 813-816.

[12] Kosterev, Dmitry N., Carson W. Taylor, and William A. Mittelstadt. ”Model validation for the August 10, 1996 WSCC system outage.” IEEE transactions on power systems 14, no. 3 (1999): 967-979.

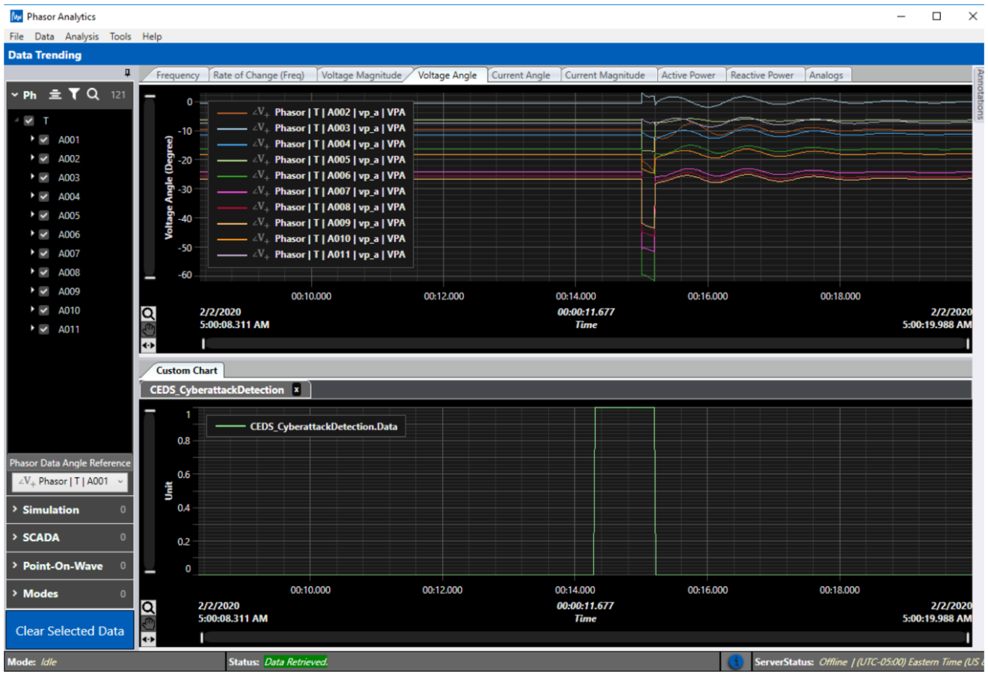


Figure 3.1: Fault Event T3.

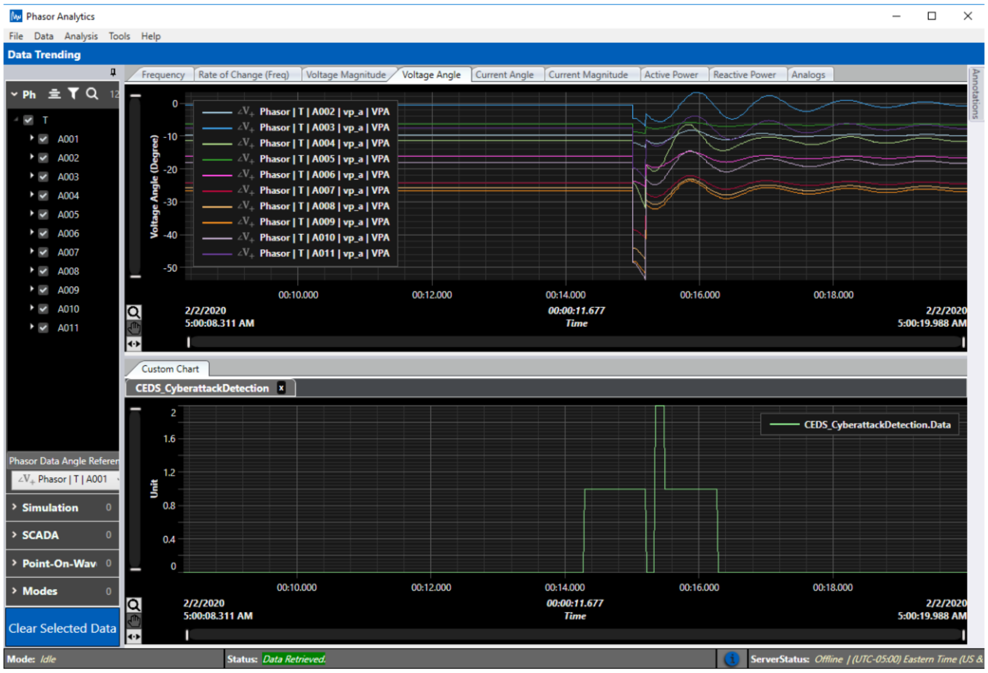
3 Phase IIB: Integration of Prototype Algorithm into GE PhasorAnalytics Synchrophasor Visualization Platform

A finalized version of the synchrophasor cyberattack detection algorithm has been ported from its MATLAB-based development / validation environment into a set of fully-functioning Python modules, which enables integration of the algorithm into a commercial synchrophasor visualization platform–PhasorAnalytics, from GE Digital–using an embedded Python interface. It was verified that the standalone Python version performs identically to MATLAB version. The Python interface within PhasorAnalytics allows for rapid prototyping of new software features, such as the WAMS cyberattack detection algorithm, within the software framework.

Synthetic synchrophasor event datasets were pre-assembled by GEGR. These simulated test and training datasets were converted to the JSIS format and ingested into PhasorAnalytics. Using these synthetic datasets, the performance of the cyberattack detection algorithm was demonstrated in PhasorAnalytics for both replay cyberattacks and physical faults.

The WAMS cyberattack detection algorithm was trained on random two fault events from a set of 20 fault events. New fault events ‘T3’ and T6 were simulated, and their datasets were converted to JSIS csv format and loaded into PhasorAnalytics. In Fig. 3.1 and 3.2 show the resulting visualization within PhasorAnalytics, respectively (synchrophasor voltage angles are relative to bus 1). It can be seen that the algorithm successfully detects fault events T3 and T6 and classifies them correctly.

Addtionally, cyber replay attack event T1 was simulated, it’s dataset converted to JSIS csv format and loaded into PA. Fig. 3.3 shows how the algorithm successfully detects and classifies the event as a cyber-attack.

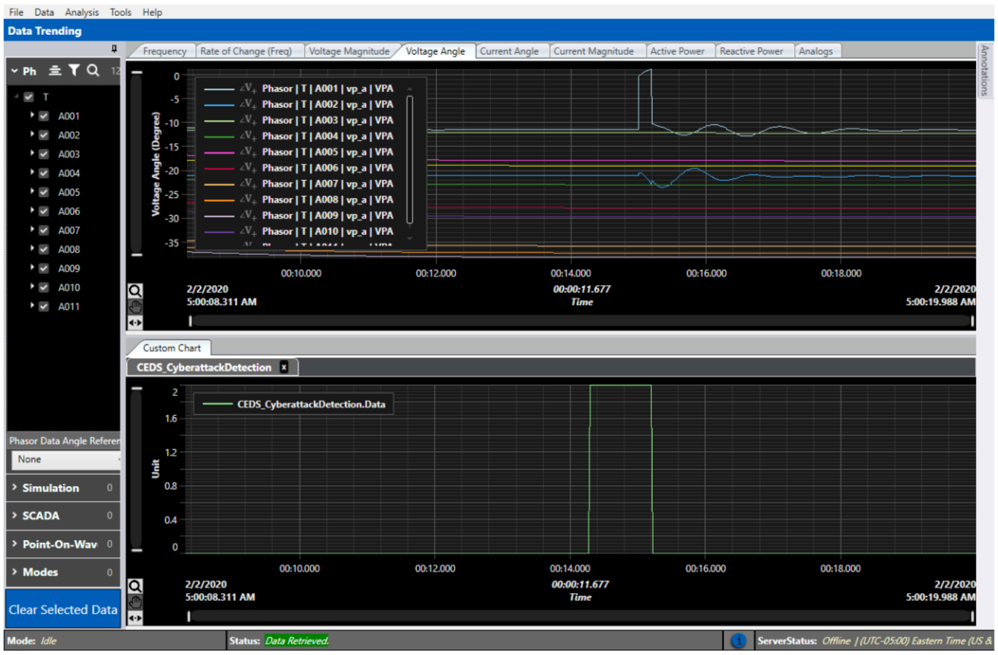


Mode: *Idle*

Status: *Data Retrieved.*

ServerStatus: *Offline | (UTC-05:00) Eastern Time (US &*

Figure 3.2: Fault Event T6.



Mode: *Idle*

Status: *Data Retrieved.*

ServerStatus: *Offline | (UTC-05:00) Eastern Time (US &*

Figure 3.3: Cyber Replay Event T1.