CYBER ATTACK RESILIENT MONITORING AND CONTROL FRAMEWORK FROM TRANSMISSION TO ACTIVE DISTRIBUTION POWER NETWORKS

A Thesis Submitted

in Partial Fulfilment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Sourav De

(2018EEZ0003)

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROPAR

RUPNAGAR-140001, INDIA

DECEMBER, 2024

Sourav De: *Cyber Attack Resilient Monitoring and Control Framework From Transmission To Active Distribution Power Networks*

Dedicated to My Beloved Family

- Whose unwavering love, encouragement and support have been the cornerstone of my journey, guiding me through every challenge and triumph.

Declaration of Originality

I hereby declare that the work which is being presented in the thesis entitled **CYBER ATTACK RESILIENT MONITORING AND CONTROL FRAMEWORK FROM TRANSMISSION TO ACTIVE DISTRIBUTION POWER NETWORKS** has been solely authored by me. It presents the result of my own independent investigation/research conducted during the time period from JULY, 2018 to MAY, 2024 under the supervision of Dr. Ranjana Sodhi, Associate Professor, Department of Electrical Engineering. To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted or accepted elsewhere, in part or in full, for the award of any degree, diploma, fellowship, associateship, or similar title of any university or institution. Further, due credit has been attributed to the relevant state-of-the-art and collaborations (if any) with appropriate citations and acknowledgments, in line with established ethical norms and practices. I also declare that any idea/data/fact/source stated in my thesis has not been fabricated/ falsified/ misrepresented. All the principles of academic honesty and integrity have been followed. I fully understand that if the thesis is found to be unoriginal, fabricated, or plagiarized, the Institute reserves the right to withdraw the thesis from its archive and revoke the associated Degree conferred. Additionally, the Institute also reserves the right to appraise all concerned sections of society of the matter for their information and necessary action (if any). If accepted, I hereby consent for my thesis to be available online in the Institute's Open Access repository, inter-library loan, and the title & abstract to be made available to outside organizations.

Sourav De

Signature

Name: Sourav De
Entry Number: 2018EEZ0003
Program: Doctor of Philosophy (Ph.D.)
Department: Electrical Engineering
Indian Institute of Technology Ropar
Rupnagar, Punjab 140001

Date: 17 December 2024

Acknowledgement

At the end of my Ph.D. research work, I would like to thank each and every person who has directly or indirectly helped me during the course of this work and contributed to the fulfillment of my thesis.

First and foremost, I would like to express my deepest sense of gratitude and profound feeling of reverence to my thesis supervisor **Dr. Ranjana Sodhi** who has introduced me to the world of research and provide me the opportunity to work in this interesting and challenging research area of recent times. It has been a great honor and privilege to work under her mentorship and be a part of her working research group. I profoundly acknowledge her constant guidance, wisdom, kindness, ideas and constructive suggestions that have been the cornerstone upon which my thesis stands.

Secondly, I am indebted to the establishment of Indian Institute of Technology Ropar, for providing me this precious opportunity, fellowship and academic resource to carry out my research work. I also extend my sincere thanks to the members of my Doctoral Committee (DC), **Prof. C.C. Reddy** (Chairperson), **Prof. J.S. Sahambi**, **Dr. J. Kalaiselvi** and **Dr. Sujata Pal** for their insightful comments, invaluable suggestions and thoughtful feedbacks during my Ph.D. examinations.

Furthermore, I am sincerely thankful to all my seniors, **Dr. M.V. Reddy**, NREL, USA, **Dr. Kapil Chauhan**, MNNIT Allahabad, India, **Dr. Yashasvi Bansal**, IIT Delhi, India, and **Dr. Shitikantha Dash**, NUS, Singapore for providing great research atmosphere, technical assistance and most importantly constant support and encouragement during my hard times. I also wholeheartedly thank to all my cooperating colleagues and juniors, **Mr. Balakrushna Sahu**, **Mr. Digamber Kumar**, **Ms. Milanpreet Kaur**, **Mr. Subal Beura**, **Miss Swati Agarwal**, **Mr. Manish Pandit**, **Mr. Nitish Kumar**, **Mr. Kondety Tony** for sharing their learnings, engaging in many fruitful discussions and also being a constant source of motivation and encouragement which make my Ph.D. journey and stay in IIT Ropar memorable and enjoyable.

Last but not the least, on a personal note, I am eternally grateful to my mother, **Mrs. Lovely De**, my father, **Mr. Pradip Kumar De**, my brother, **Mr. Subhayu De** and my sister-in-law, **Mrs. Rituparna De** for being the greatest strength of mine and for always believing in me and encouraging me through the highs and lows of this journey. I can never thank you enough for your incredible support, cooperation, inspiration, understanding and prayers. It would be impossible for me to complete this thesis without their unconditional love, unwavering trust and blessings. Above all, my utmost gratitude and offerings to the almighty God for being always with me and showering love, grace and strength.

Sourav De
209, SYnchrophasor Measurement And Research (SYMAR) Lab
J. C. Bose Block, Indian Institute of Technology Ropar
17 December, 2024

विद्या ददाति विनयं, विनयाद्याति पात्रताम्।
पात्रत्वाद्धनमाप्नोति, धनाद्धर्मं ततः सुखम्॥

Vidya Dadati Vinayam, Vinayad
Yati Patratam.
Patratvad Dhanamapnoti, Dhanad
Dharmam Tatah Sukham.

Certificate

This is to certify that the thesis entitled **CYBER ATTACK RESILIENT MONITORING AND CONTROL FRAMEWORK FROM TRANSMISSION TO ACTIVE DISTRIBUTION POWER NETWORKS**, submitted by **Sourav De (2018EEZ0003)** for the award of the degree of **Doctor of Philosophy** of Indian Institute of Technology Ropar, is a record of bonafide research work carried out under my guidance and supervision. To the best of my knowledge and belief, the work presented in this thesis is original and has not been submitted, either in part or full, for the award of any other degree, diploma, fellowship, associateship or similar title of any university or institution.

In my opinion, the thesis has reached the standard fulfilling the requirements of the regulations relating to the Degree.

Ranjana

Signature of the Supervisor(s)

Dr. Ranjana Sodhi
Associate Professor
Department of Electrical Engineering
Indian Institute of Technology Ropar
Rupnagar, Punjab 140001
Date: 17 December 2024

Abstract

In today's world, our power grids are becoming smarter, due to the rapid and wide spread integration of digital sensors, computers, Internet and Communication Technologies (ICTs) etc. The remarkable advancements of such modernized technologies and far-reaching use of various sophisticated remotely control devices have transformed our age old traditional energy sector from purely physical system to somewhat complex Cyber Physical Systems (CPS). While this CPS infrastructure has effectively prevented various disastrous scenarios like blackouts, uncontrolled shutdowns, unwanted frequency and voltage fluctuations, power loss and grid instability, nevertheless the close integration of power system's physical operation with that of unsecured cyber networks brings a new risk of cyber threats via unauthorized control access to the communication channels, exploitation of networking protocols, forced equipment outage and damage, manipulation of sensing and control signals and any kind of other sabotaging activities that jeopardize the normal monitoring and control functionality of power grid ranging from power transmission to power distribution. Thus this thesis aims to provide an overall comprehensive security solution towards developing a Cyber Attack Resilient Monitoring and Control (CARMC) framework by unveiling vulnerabilities across transmission **(T-System)** to distribution **(D-System)** power networks.

The research begins by identifying and addressing the key vulnerabilities introduced in the **T-System** networks. Following this, a comprehensive attack resilient framework is developed based on strategic placement of Phasor Measurement Units (PMUs) at such optimal locations that safeguard a minimal sets of measurements in order to make the system resilient against any kind of False Data Injection Attacks (FDIAs) on those selected vulnerable lines. After securing a set of critical meters by developing a secured metering infrastructure, the next research study of **T-System** focused on the detection and control technique of another simple but impactful attack named, Replay Attack (RA) that targets one of the core power system monitoring application of energy management system i.e Power System State Estimation (PSSE). In order to safeguard the PSSE against RA, the proposed technique leverages the secured phasor measurements obtained from the optimal PMU locations through a hybrid state estimator (HYB-SE) to correct the manipulated conventional meter readings.

The later half of the thesis focuses on detecting and mitigating vulnerabilities associated to the **D-System** networks, specifically Microgrids (MGs), where the Distributed Energy Resource (DER) controller's and its communication links are being targeted by the attacker to cause voltage and frequency instability to the grid. To this end, for the detection, classification and localization of cyber attacks, a statistical two-sample hypothesis test, called as Maximum Mean Discrepancy (MMD) index and a rule based algorithm coupled with XGBoost classifier is utilized respectively. After the attack being detected and classified successfully, the next study aims to develop a cyber-attack resilient control framework for the MG system based on designing Unknown Input Observer's (UIO) states and Back-stepping Integrated Sliding Mode Controller (BSMC) to mitigate

the overall effect of injected attack into the DER's secondary controller. Finally, with the aim of having a secured monitoring infrastructure in **D-System**, the last research study addressed the problem of accurate detection of islanding event in the presence of cyber-attacks.
The effectiveness of the proposed CARMC framework is validated through extensive offline simulation performed in MATLAB, PSCAD, RSCAD software and real-time testing incorporating various hardware platforms such as Real-Time Digital Simulator (RTDS) and dSPACE 1104 Research & Development controller board. The results demonstrate the ability of the CARMC framework to bolster the resilience of transmission and active distribution networks against diverse cyber threats.

Keywords: Topological Vulnerability; Phasor Measurements Units; Power System State Estimation; False Data Injection Attack; Replay Attack; AC Microgrid; Distributed Secondary Control; Maximum Mean Discrepancy; Unknown Input Observer; Back-stepping Integrated Sliding Mode Controller; Secured Passive Islanding Detection.

Contents

List of Figures

List of Tables

Glossary

Frequently used abbreviations, mathematical variables and symbols are listed below; other used terms are abbreviated in their respective chapters.

Key Abbreviations

ADN	Active Distribution Network
BR-PTDF	Branch - Power Transfer Distribution Factor
BSMC	Backstepping based Cliding Mode Controller
CAD	Cyber Attack Detector
CM	Confusion Matrix
CNT	Complex Network Theory
CPS	Cyber-Physical Systems
CSC	Cosine Similarity Coefficient
DDCD	Decaying DC Detector
DER	Distributed Energy Resource
DG	Distributed Generation
DMO	Distribution Management Operator
DMS	Distribution Management Systems
DSFC	Distributed Secondary Frequency Control
DSVC	Distributed Secondary Voltage Control
DoS	Denial of Service
FDI	False Data Injections
FDIA	False Data Injections Attacks
FRTL	Final Relay Trip Logic
HBC	Hybrid Betweenness Centrality
HIL	Hardware-in-Loop
ICT	Information and Communication Technologies
KF	Kalman Filter
KLD	Kullback-Leibler Divergence
MDDA	Multiple Data Dropping Attack
MG	Microgrid
MID	Mean-based coarse Islanding Detection
MMD	Maximum Mean Discrepancy
N-PTDF	Node - Power Transfer Distribution Factor
NDZ	Non-Detection Zone
OPP	Optimal PMU Placement
PCC	Point of Common Coupling
PIDM	Proposed Islanding Detection Method

PMU	Phasor Measurement Unit
PS	Power System
PSSE	Power System State Estimation
PTDF	Power Transfer Distribution Factor
RA	Replay Attacks
RDCA	Repetitive Data Cloning Attack
RKHS	Reproducing Kernel Hilbert Space
ROC	Receiver Operating Characteristics
ROCOV	Rate of Change of Voltage
RTDS	Real Time Digital Simulator
RTU	Remote Terminal Unit
SE	State Estimation
SG	Smart Grid
SRDL	Statistical Relay Digital Logic
UIO	Unknown Input Observer
WLS	Weighted Least Square
XGBoost	EXtreme Gradient Boosting

Sets, Vector, Matrices, and Indices

$\mathbb{V}$	A finite set of vertices
$\mathcal{E}$	A finite set of edges
$\mathcal{W}$	A finite set of Weights of $\mathcal{E}$
$\mathbb{G}$	Notion of undirected graph
$\mathcal{G}$	Notion of Directed Graph
$\mathbf{A}_{adj}$	Undirected adjacency matrix
$\mathscr{A}$	Adjacency matrix of directed graph
$\mathscr{D}$	Diagonal In-degree matrix
$\mathscr{L}$	Graph Laplacian matrix
$\bar{N}_i$	Set of Immediate Neighbors
$\boldsymbol{\mathfrak{D}}$	Notion of graph distance matrix
U	Set of non-negative eigenvectors
$C_E(v_i)$	Eigenvector centrality metric
$\mathbf{Q}$	Conductance matrix
$\vec{V}$	Voltage vector
$\mathbf{B}_b$	Bus incidence matrix
∇	Matrix with conductance at its diagonal element
$\mathbf{I}^{\mathbf{st}}$	Current flow matrix all possible source-target (st) pairs
F_E^i	Flow-energy of dominant node-i
$HBC(l_{i-j})$	Hybrid Betweenness Centrality Index of line l connecting node-i and node-j

$\bar{X}$	Vector of optimal PMU placement location
$\bar{\mathbf{A}}$	Binary connectivity matrix
$\mathfrak{L}$	Line serviceability indicator
$\mathfrak{B}$	Branch-to-node incident matrix
G_H^l	Current giant component of the initial graph after removing of line-l
RI	Resiliency Index
S	Giant component size
$\mathbf{H}$	Jacobian matrix
Z_{RTU}	Vector set of available sensors/RTU measurements
Z_a	Vector set of compromised RTU measurements
$\overset{*}{\mathbf{X}}$	Bus reactance matrix
$\mathbf{B_x}$	Branch network susceptance matrix
$\overset{*}{\mathbf{A}}$	Line incidence matrix
$\mathcal{Y}_i^r(t)$	Set of the capture data by i^{th} sensor in recording phase of RAs
$\mathcal{Y}_i^a(t)$	Set of malicious i^{th} sensor readings that fraudulently transmitted to CC
X_{RTU} and X_{HYB}	Vector of System States Based on Conventional and Hybrid SE respectively.
ε_{RTU}	Vector of RTU measurement Gaussian noise or errors
ε_{PMU}	Vector of PMU measurement Gaussian noise or errors
$\mathbf{R_{RTU}}$	RTU measurements error covariance matrix
$\mathbf{H_{RTU}}$	RTU measurement Jacobian matrix
(Z_{PMU})	Set of measurement vector incorporating PMU measurements only
(Z_{HYB})	Hybrid measurement vector incorporating RTU and PMU measurements
$\mathbf{R_{HYB}}$	Covariance matrix of hybrid estimator
$\mathbf{W_{HYB}}$	Weight of Hybrid measurement sets
$(\hat{Z}_{RTU}^{HYB})$	Reconstructed measurements based on the output states of hybrid estimator
$\tilde{Z}_{RTU}$	RTU sensor noisy reading which may or may not be compromised
Z_v	Sets of identified vulnerable RTU measurements
E_{en}	Entropy value
E_{se}	Shannon Energy value
$\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$ and $\mathbf{E}$	Known system matrices for MG network
$\mathbf{F}$,$\mathbf{J}$,$\mathbf{L}$ and $\mathbf{K}$	Designed matrices for UIO

$\hat{f}_\omega$	Vector of rough estimation of exogenous frequency attack input to DERs by UIO
B_p, B_u, B_f	Designed matrices for BSMC
X_ω	Vector containing frequency states of DERs
U_P	Injected active power input vector for DERs
$U_{s\omega}$	Desired control law output vector from BSMC to make the MG attack resilient
$\boldsymbol{\mathfrak{F}}$	State transition matrix for KF
$\bar{\mathbf{P}}$ and $\bar{\mathbf{Q}}$	Process covariance and model error covariance matrix
$\mathbf{R}$	Sensor measurement noise covariance matrix
$\bar{\mathbf{K}}$	Kalman Gain Matrix
Functions and Operators	
$\sum$	Sums of all elements of a vector/matrix
$(\cdot)^T$	Transpose Operator
$\mathbf{1}_{[\Xi]}$	An indicant function
$h_{RTU}(\cdot)$	Non-linear mapping function that relates measurements with states
$\mathfrak{J}$	Objective Function to for WLS State Estimator
$\\|*\\|_{\mathcal{H}}^2$	2-norm operation in RKHS
$\varphi(\cdot)$	Mapping function consist of a kernal matrix $\mathcal{K}(X, .)$ to map some feature sets to RKHS
$MMD[\mathscr{F}, \cdot, \cdot]$	Computing maximum mean discrepancy between two class of function in their feature space, $\mathscr{F}$
$\mathbb{E}[\cdot]$	Expectation Operator
$\sup_{\acute{f} \in \mathscr{F}}(\cdot)$	Finding supremum of a certain set of function $\acute{f}$, belongs to the feature space, $\mathscr{F}$
$\alpha.$	Attack modelling parameter
$\hat{F}$	Domain of classification and regression trees
$(\mathcal{L})$	Approximating loss functions
$\mathbb{L}$	Differential convex loss function
ϑ	Regularization Function
rank$(\cdot)$	Calculating matrix rank
V	Lyapunov Function
Constant, Variables and Symbols	
N	Cardinality of node set $\mathbb{V}$
M	Cardinality of edge set $\mathbb{E}$
v_i	i^{th} Node of a graph
$e_{i,j}$	Edge between Node-i and Node-j
$a_{i,j}$	Elements of adjacency matrix $\mathbf{A}_{adj}$
$w_{i,j}$	Weight of edge $e_{i,j}$

d_{ij}	Shortest path between two node pairs i.e v_i and v_j
λ	Eigenvalues of $\mathbf{A}_{adj}$
$\vec{\Omega}$	External injected/extracted current vector
I_s	Injected current at source node s
$\acute{P}$	Total no of source target pairs
$\sigma_i^{\acute{P}}$	Aggregate sum of injection currents for all $\acute{P}$
c_j	Cost of PMU placement at bus-j
ξ_j	Merit of far-ness
Ψ	Attack magnitude
$\hat{x}_{bad}$	Estimates of system states under attack condition
$\hat{x}$	Estimates of system states
V_i, θ_i	Voltage and Angle state variables of Bus-i respectively
t_{rec}	Recording window phase duration for replay attack
t_{rep}	Replaying window duration for replay attack
$\mathcal{T} \in [\mathcal{T}_s, \mathcal{T}_e]$	Recording interval of replay attack between starting instant $\mathcal{T}_s$ and stopping instant $\mathcal{T}_e$.
$\digamma$	A binary variable for triggering replay attack
$\acute{l}$	Window size of the attacker's recording phase
$\acute{n}$	Total number of replay attack sequence
$\tilde{\sigma}_{RTU}$	Standard Deviations of Z_{RTU}
$+\delta^{UB}$ and $-\delta^{LB}$	Upper and Lower bounds of the residue
Λ	Measurement residue
g_i	Pinning gain associated with DER-i
ω_i	Operating frequency of DER-i
$\mathscr{V}_{odi}$	Output Voltage of DER-i
ω_{ref}	Reference frequency set-point in p.u.
ω_j^i	Frequency communicated to DER-i from DER-j
P_i and Q_i	Filtered active and reactive power of DER-i
P_j^i and Q_j^i	Active and Reactive power communicated to DER-i from DER-j
m_{P_i} and n_{Q_i}	Frequency and voltage droop coefficient of DER-i respectively
δ_{ω_i}	Local neighbourhood frequency synchronization error
$\delta_{\mathscr{V}_i}$	Local neighbourhood voltage synchronization error
c_ω and $c_{\mathscr{V}}$	DER frequency and voltage control gain respectively
$\mathcal{U}_{\omega_i}$	Auxiliary frequency control input variables

ω_i^a	Corrupted frequency of DER-i secondary controller
$(\omega_j^i)^a$	Corrupted communicated frequency from DER-j to DER-i
Δ_i	Exogenous input injected to frequency synchronization error of DER-i due to attack on its controller
Δ_i^j	Exogenous input injected to the frequency synchronization error due to attack on incoming communication link between DER-j and DER-i
$\delta_{\omega_i}^a$	Corrupted local neighbourhood frequency synchronization error due to attack on DER_i controller
$\delta_{\omega_j^i}^a$	Corrupted local neighbourhood frequency synchronization error due to attack on incoming communication link to DER-i from DER-j
$\mathcal{U}_{\omega_i}^a$	Corrupted auxiliary frequency control input of DER-i due to attack on its controller
$\mathcal{U}_{\omega_j^i}^a$	Corrupted auxiliary frequency control input pertaining to incoming communication link to DER-i from DER-j
ζ_i^{attack}	Compromised frequency information recorded by the sensors after FDIA
f_i^a	Injected attack input by the attacker
Υ_i, η_i	Binary parameter to control launching of attack on DSFC and communication sensors respectively
γ	Threshold settings for MMD
$\mathcal{H}$	Hilbert space of real valued functions
μ_{P}	Embedded mean for the samples drawn from P
ν	Kernal Width
T_K	Total number of leaves in the tree
$\tilde{w}^*$	Optimal leaf Wwights
ϱ	Auto-pruning hyper-parameter
$beta$	Learning rate
η	Shrinkage parameter
k	k-fold cross validation
S_{ω_i}	Sliding surface for DER-i
k_n	Control gains for BSMC
V^{PCC}	Instantaneous 1-phase PCC voltage signal
V_{mean}^{PCC}	Mean of one cycle PCC voltage signal
τ	Decay time for decaying DC quantity

E_τ	Entropy of decaying DC quantity
$\breve{\nu}$	Process Noise
ℓ	Measurement Noise
r_s	Spearman Rank's Correlation Coefficient

Chapter 1

Introduction

1.1 General: Cyber Physical Integration of Smart Grid

In recent decades, the integration of advanced sensing, computing, Internet of Things (IoTs), Information and Communication Technologies (ICT) within the power sector has undoubtedly revolutionized the flexibility, reliability, efficiency and management of electrical grids to a considerable extent. This revolutionary and evolutionary changes have made a profound impact on national critical infrastructures, such as Power Systems (PSs); transforming them into the present-day complex Cyber-Physical Systems (CPS), known as Smart Grids (SGs) with offering numerous benefits such as, control of the two-way flow of electricity and information, efficient monitoring and control of real-time electricity generation and consumption, optimized resource utilization, reduced operational costs, increased renewable energy, empowering consumers through real-time visibility and control over their energy usage etc.

While such involvements of CPS in monitoring and control operations of PS have effectively prevented various disastrous scenarios like blackouts, uncontrolled shutdowns, unwanted frequency and voltage fluctuations, power losses, and grid instability, they have also introduced new challenges to the PS operators in terms of device-level and network-level security [1, 2, 3]. Many remote devices located in the physical layer of SGs like digital sensors, actuators, smart meters, digital relays, Remote Terminal Units (RTUs), Intelligent Electronic Devices (IEDs), etc, and servers of communication and ICT interface layer like SCADA Server, Communication Server, Human Machine Interface Server (HMI) and, Database Server are connected to the open network via some corporate networks for being more flexible in management process which acts as a back-door access for the cyber attackers to get into operator's supervisory network control layer to disrupt various managerial decisions as shown in Fig. 1.1. Moreover, a large number of systems have been using third-party web-based applications for the monitoring of physical process and this direct connection to the internet could be an another possible path for the cyber attacker to penetrate into the enterprise network. Thus, the increasing dependence of the SG on the critical cyber networks and extensive interlaced with data communication layers at its various levels has exposed the power grid to potential vulnerabilities and persistent cyber threats such as forced equipment outage, manipulation of sensing and actuation signals by malicious actors, theft of intellectual property, exploiting financial arbitrage, and other kind of sabotage which ultimately hampers the normal grid functioning [4].

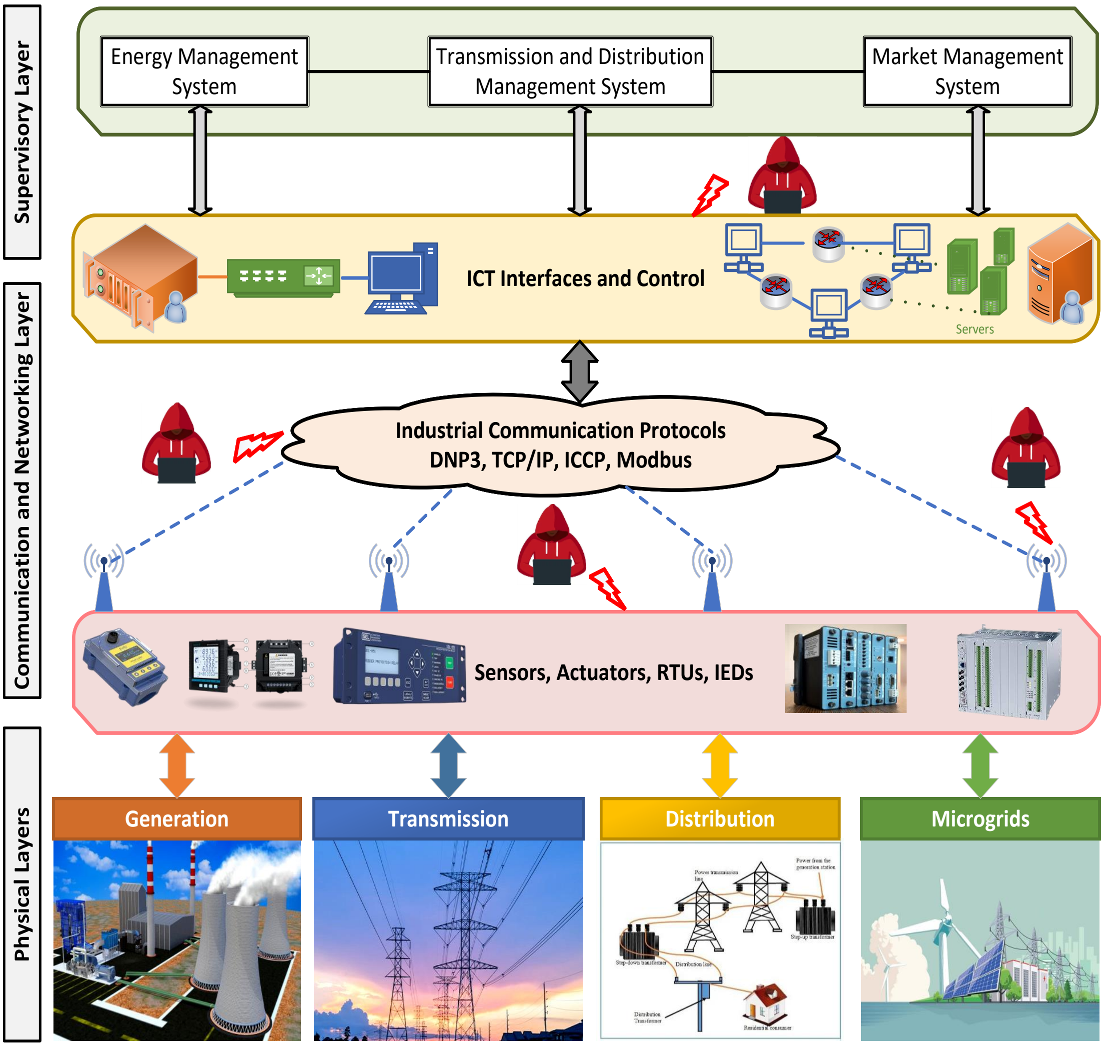


Figure 1.1: Generic illustration of a cyber-physical Smart Grid architecture and its vulnerable surface

1.1.1 Historical Events of Some Major Cyber Attacks in Smart Grid

As per the standards defined by the industry, a *Cyber-Attack* is an intrusion that can jeopardize the availability, confidentiality, or integrity of an information system or the data it processes, stores, or transmits. Additionally, it is noteworthy to mention that out of all network infrastructures, the energy sector has faced the greatest number of cyber intrusions and thus it is placed at significantly higher security risks in terms of attack severity and impact, according to vulnerability reports from the US ICS-CERT [5] and Kaspersky ICS-CERT [6]. One of its main reason is that the current communication networks for PSs and most SCADA network protocols e.g., Modbus, DNP3, TCP/IP, IEC 61850, etc, are not designed to be adequately protected from potential cyber threats. Hence, recent years have witnessed that power plants and electrical grids are becoming increasingly an attractive target for hackers due to the large number of individuals who could be impacted and the extent of damage that could be inflicted nationwide [7]. Notably, various reported historical incidents worldwide as listed in Table 1.1 have highlighted significant challenges stemming from such cyber-attacks in the cyber-layer of energy-sector infrastructures.

Table 1.1: Some reported major cyber attacks events targeted to energy sector

Year	**Attack Location**	**Attack Type/Method**	**Attack Target**	**Attack Impacts**
Oct, 2010	Natanz Nuclear Power Plant, Iran	Malware STUXNET	PLCs, ICS Network, and HMI	Manipulate the speed and rotation cycles of PLC controlled centrifuges
Dec, 2014	Nuclear and Hydroelectric Plant, South Korea	Malware, Social Engineering	ICT Interface	Steal sensitive information such as design documents, operation manuals, and employee data.
Dec, 2015 and 2016	Power Grid, Ukraine	FDIA, MitM	ICS, Relay and Circuit Breaker	Wide spread blackout took place in 3 major power distribution companies
Aug, 2017	Petro-chemical Plant, Saudi Aramco	Malware BLACKENERGY, Social Engineering	ICT Interface	Disrupt operations, compromise sensitive data, and inflict financial loss
Feb, 2019	Power Grid, Russia	Unconfirmed	ICS, Substations Equipment	Gained illegal access to technological control systems to affect dozens of settlements
Feb, 2020	Water treatment plant in Florida, US	Malware, Eavesdropping	Remote Access to HMI Software	Manipulates the chemical levels in the water supply.
May, 2021	Colonial Oil Pipeline, US	Ransomware (Via –Spear Phishing Email)	Fuel Storage Units, IT Systems	Led to fuel shortages, urgent shut down operations and economic loss
Feb, 2022	Multiple Oil terminals across Belgium and Germany, Northern Europe	Ransomware, DoS	Oil Refining Ports and Storage Facilities	Disable computers of Energy Department
Mar, 2023	Power Grid, South Africa	Eavesdropping attack	SCADA Servers, EMS, Substations	Compromised systems and establish remote connections on the electric utility to change the payloads

According to the statistics reviewed, China, Singapore, Russia, and the countries of the Commonwealth of Independent States (CIS) collectively account for the majority of cyber attacks. In July 2018, the U.S. Department of Homeland Security (DHS) and Industrial Control Systems Cyber Emergency Team (ICS-CERT) has issued the warning alerts against the international threat actors, who have constantly targeted the energy sector in the past. India, as a nation, is also undergoing rapid digitization across its various sectors, and is also not immune to the increasing number and severity of cyber threats. Figure 1.2 depicts the most affected countries which is very frequently been targeted by the cyber attacks where India continues to be one of the top-three most attacked countries by nation-state actors in the Asia-Pacific (APAC) region, accounting for 13% of cyber attacks. According to information reviewed by Mint [8], several high profile cyber attacks

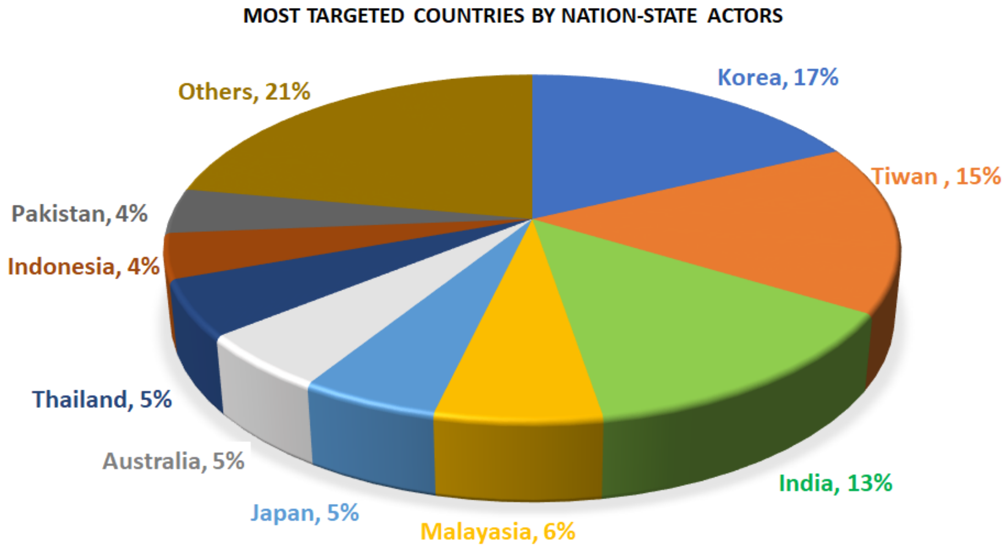


Figure 1.2: Global landscape of cyber attacks

incidents that Indian power sector has been encountered so far includes,

- The March 2018 attack on Haryana distribution company (DISCOM), which involved hacking into the commercial billing software of the highest-paying industrial customers.
- The November 2017 malware attack on THDC Ltd's Tehri dam in Uttarakhand, India, which targeted the critical infrastructure to steal sensitive operational data.
- The May 2017 ransomware attack on West Bengal State Electricity Distribution Co. Ltd (WBSEDCL), which affected the operations of the utility company and caused a prolong blackout.
- The February 2018 attack on a Rajasthan discom website, which disrupted online services and potentially compromised sensitive information.

Figure. 1.3 shows the statistical record of total number of cyber intrusions happened in India as reported in the annual report of CERT-IN. The data reveals an exponential surge in reported cyber incidents throughout last 6 years span upto year 2022 where year 2021 and 2022 has been found as the worst year so far for India when it comes to cyber attacks [9]. It has also been noticed that Indian power sector are facing such cyber threats with at least 30 events reported daily. All these aforementioned incidents highlight the growing threat of cyber attacks on power sector and the need for enhanced and sustainable cyber security measures to protect it against such threats. Therefore, to cope up with those cyber physical challenges, countries like the U.S. Government, Department of Energy (DoE) and NATO nations with the Cyber Defense Center of Excellence (CDCOE), are actively investing in research, development, and guidelines to enhance the cyber security of their power infrastructure.

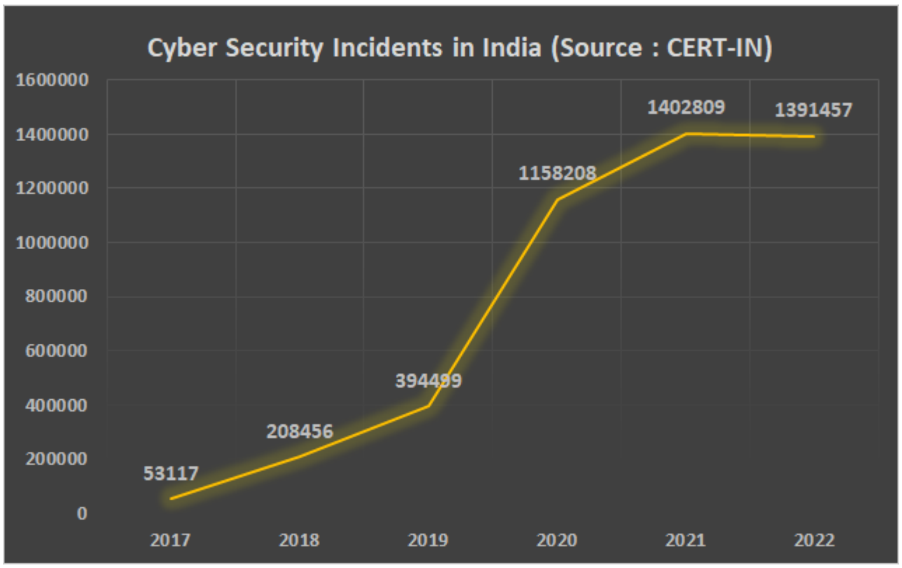


Figure 1.3: Statistics of reported cyber attack incidents in India since 2017-22

1.1.2 Types of Cyber Attacks in Power Grid

There are numerous types of attacks that could be carried out in SG. The most commonly occurring attacks in SG are constructed with each aimed at compromising one or more of the three security objectives: Confidentiality, Integrity, and Availability as shown in Fig. 1.4.

Integrity attacks [10, 11, 12, 13] target the legitimacy and consistency of information within the system. Unauthorized individuals may gain access to the operator network to modify or destroy legitimate data, compromising its accuracy and trustworthiness. Integrity attacks seek to illicitly delay and alter the original data's content, including customer account and billing information, voltage and sensor readings, control commands and device status to obscure the limited visibility of PS. In the context of power substation networks, integrity attacks may involve broadcasting fake Address Resolution Protocol (ARP) network packets to induce malfunction or disconnect RTUs and IEDs from the substation gateway. Attackers typically penetrate the system's network security through methods such as password cracking, wiretapping, or exploiting software vulnerabilities to gain authentication and access control. Once inside, they may inject false data, manipulate code, obscure identity of legitimate devices or replay malicious data packets to mislead system operators into making incorrect decisions regarding system operations.

On the other hand, *Confidentiality and Availability attacks* [14, 15, 16] aim to disclose or steal intellectual property, personal privacy, and proprietary information, while also restricting timely and reliable access to relevant information. These attacks can result in loss of network availability, leading to adverse consequences such as the loss of real-time monitoring of the power grid and major power blackouts. Attackers may conduct port

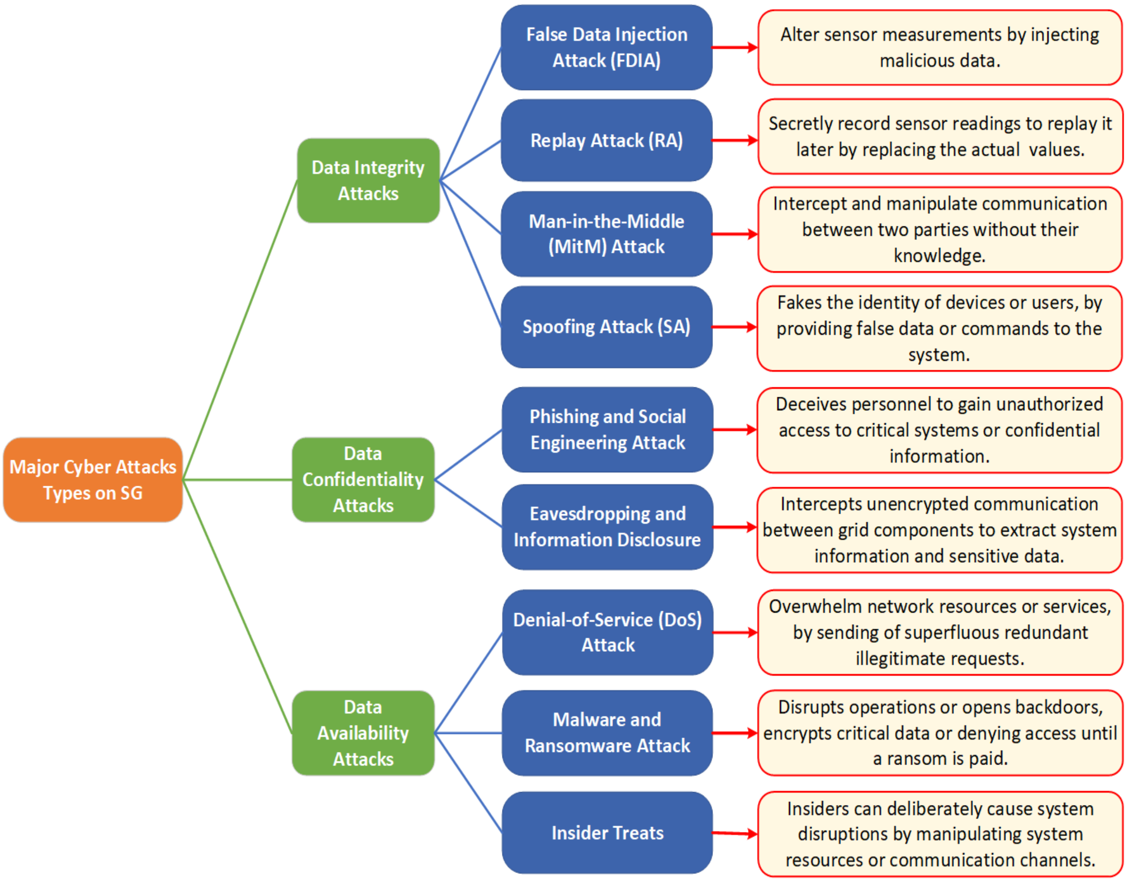


Figure 1.4: Most commonly occurring attacks in SG

scanning and packet sniffing to capture network tariff information and execute Denial of Service (DoS) attacks by flooding the network with illegitimate requests or jamming communication channels, rendering them inaccessible for intended use. IP-based protocols such as TCP/IP and IEC 61850 are mainly vulnerable to such availability attacks. Additionally, confidentiality attacks like phishing, social engineering involve hackers impersonating authorized users and surreptitiously enters a communication channel between two parties. After then, attacker has gained the ability to eavesdrop on the conversation, potentially steal data, or even spoof the messages between source and destination.

The aforementioned attacks underscore the significance of implementing strong cyber security methods to safeguard SG systems from malevolent acts. From the network layer perspective, the risks presented by integrity, confidentiality, and availability threats in the SG domain can be reduced by putting measures like encryption, access control, intrusion detection, and security awareness training into practice. However, to keep SG infrastructure secure and resilient in the face of challenging cyber threats from PS applications or physical layer perspective, proactive attack detection, correction and resilient monitoring and control framework is of urgent need for the overall defense to secure our power grids.

1.1.3 Taxonomy of False Data Injection Attacks in Smart Grid Based on its Attack Model, Target and Impact

False data injection (FDI) attack, has emerged as a most sophisticated form of cyber attacks in recent times and seems to be a topic of great interest for both, the SG industry and research. Therefore, this subsection presents a comprehensive taxonomic overview of FDI attacks particularly focusing on SG based on several key dimensions such as attack construction model, end target applications, and its impact as shown in Fig. 1.5.

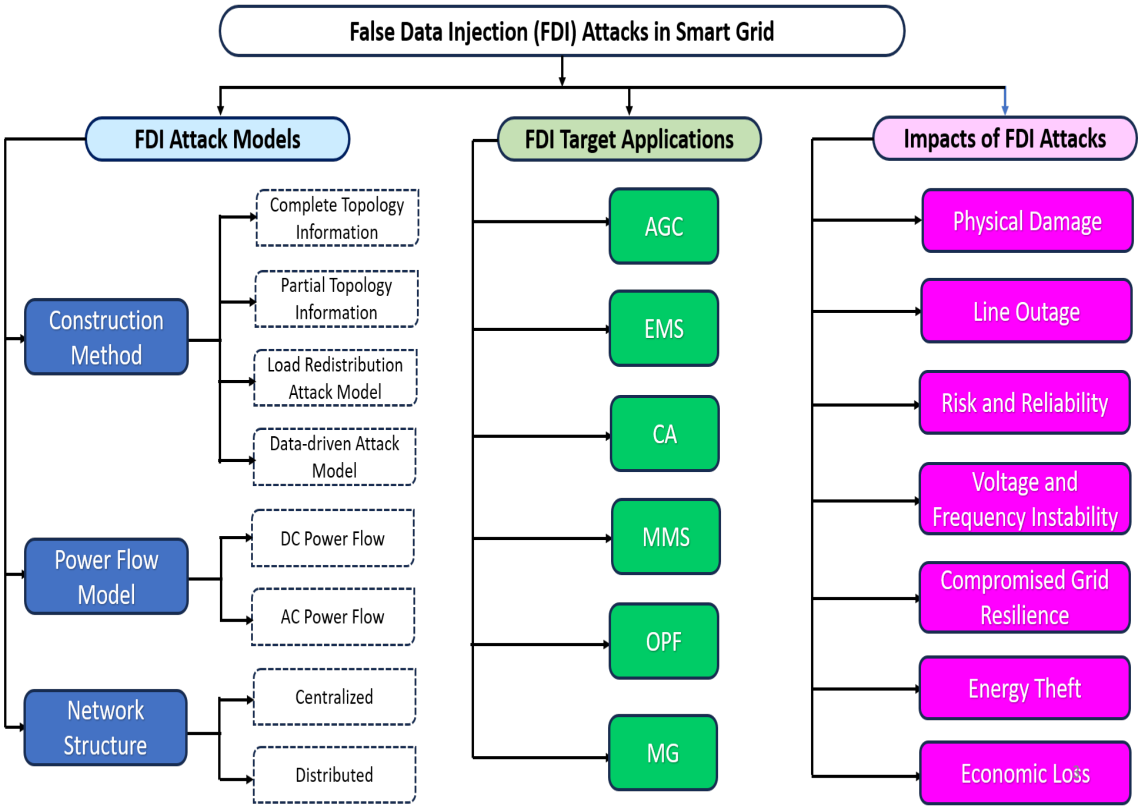


Figure 1.5: Taxonomy of FDI attacks in Smart Grid

FDI Attack Models: Various threat models exist for FDI attacks in Smart Grids, aiming to disrupt grid operations. As per the existing literature, earlier the main assumptions for the method of attack vector construction was attackers possessed complete knowledge of system topology to remain undetected [17, 18]. However, this assumption is not always realistic, as attackers often lack resources for extracting topological information or face restricted physical access to sensitive meters. Later, it was well established, that based on obeying some observability criteria between the attacking and non-attacking region of a larger PS, a stealthy attack can also be constructed by exploiting the partial topological information [19, 20]. Besides, recently alternative approaches of employing data-driven model-free techniques and load redistribution adversarial models are also gaining significant popularity. However, the stealthiness of an attack construction model mainly depends on the nature of power flow (DC or AC) method considered in the construction process [21]. This usually may differ in terms of attacker's key motives such as minimum allowable threshold selection criteria for successful attack launch, minimum investment of attack resources and degree of complexity and attack efforts involving in formulation of stealth attack vector. Although DC-based attacks are less complex and stealthier, they're less practical due to the majority of PS applications being nonlinear with AC dynamics [22, 23]. Now, in the distribution side specially multiagent microgrid (MG) system, another aspect is added with FDI attack model formation which defines the centralized or distributed structure of MG network. Both of the structures associated with low bandwidth communications channels for exchange of information among Distributed Energy Resources (DER) units, equally vulnerable to FDI attacks.

FDI Target Applications: Cyber attackers typically target specific SG applications, components, or functional entities across power generation, transmission, distribution, and active distribution networks. As far as the generation and transmission sides are concerned, Automatic Generation Control (AGC) is critical for maintaining grid frequency by regulating load generation balancing. Attackers may target AGC sensor measurements or control commands to cause frequency violations, triggering remedial actions leading to system-wide blackouts [24]. Similar to this, State Estimation (SE) [10], another vulnerable application in the Energy Management System (EMS), is also a prime target for the cyber adversaries because its malfunction or large error in the estimated output could seriously impair all other crucial ancillary services and applications, like economic dispatch, Optimal Power Flow, and Contingency Analysis (CA). Attackers may deceive the CA process by deliberately adding a transmission line contingency to the standard contingency list by introducing fictitious data into the SE process which ultimately causes line overloading and cascaded blackouts [25]. Aftermath of such event could also impact the Market Management System (MMS) which aims at facilitating the electricity market operation by setting location marginal price along with managing transaction between electricity service providers and utility consumers [26]. Last but not the least, because DERs are connected with insecure and unencrypted communication protocols on Active Distribution Network (ADN) side, they are also most likely to be targeted by the PS hackers [27].

Impacts of FDI Attacks: The threat of cyber-attacks primarily impacts the stability, reliability, dependability, economy, consumer privacy, and social welfare of Smart Grids. First of all, FDI attacks can manipulate sensor data or control signals to cause voltage and frequency instabilities in the grid which thereby initiates wrong control commands that potentially leads to equipment malfunction, overload and physical damage. Injecting false data disrupt normal power flow calculations, which by mistake, triggers unnecessary outages on transmission lines [28, 29]. Reliability is also compromised by erroneous data in decision-making processes, increasing the risk of equipment failures and service interruptions. Successful FDI attacks weaken the grid's ability to withstand disturbances, making it more susceptible to blackouts. Attackers accessing MMS can misuse SCADA advanced metering infrastructure to manipulate meter readings or falsify customer billing information, leading to electricity theft [30]. Moreover, by manipulating the real-time and day ahead pricing signal of electricity market operation through biased transmission line congestion, attacker can exploit huge financial profit in virtual bidding process. Finally, the cumulative impact of FDI attacks, including equipment damage, service interruptions, compromised grid resilience and energy theft, can result in significant economic losses for utilities, businesses, and consumers, affecting productivity, revenue, and overall economic stability of SG [2].

1.2 Literature Review

The aim of the present thesis work is to develop a holistic attack-resilient monitoring and control framework. To this end, both, the transmission as well as the active distribution network is considered in the study. A detailed literature survey is carried out in this section, exploring the existing research, focusing on the cyber attack vulnerabilities, current monitoring and control solutions, identifying the research gaps and emerging challenges in the current cyber security frameworks.

1.2.1 Vulnerability Assessment and Its Resiliency Analysis

One of the most important tasks when it comes to power grid cyber security is knowing how to identify and evaluate the system's weaknesses as well as how resilient it is to external events. The foundation of any cyber security strategy lies in vulnerability assessment, which offers a methodical way to find flaws, openings, and other points of entry that could be used by a perpetrator. As was previously mentioned, the modern installations of a variety of equipment, including distributed sources, digital relays, phasor measurement units (PMUs), RTUs, and IEDs, have become essential parts of this vital infrastructure, spanning from the generation of electricity to its transmission and distribution which makes the grid large, sophisticated, interconnected, and complex. This interconnected nature results in a scenario where a single failure can have severe consequences, ranging from medium-scale to large-scale blackouts and the destruction of major power equipment such as transmission lines, transformers, and generators [31]. These failures can result from two main categories of extreme events.

- Natural, having medium to high impact, high frequency events e.g., violent weather condition, floods, earthquakes, etc.[32, 33];
- Synthetic, having high impact, low frequency events like cyber and physical attacks, blended attacks, and human made accidents, etc.[34, 35].

Due to the wide-spread deployment of distributed sensors based technologies and highly integrated nature of cyber-physical control systems, various critical infrastructures are now a days becoming targets of various synthetic attack i.e man made attack. These attacks pose a significant threat to the resilience and security of these infrastructures as it is witnessed by 2015 Ukrainian power grid attack [36, 37]. A review of current trends suggests that such attacks are expected to increase in the near future, and thus it is imperative to focus on prevention, mitigation, and detailed vulnerability assessment in a holistic manner. By taking proactive measures to identify and address vulnerabilities within the power grid, it is possible to enhance their resilience against potential synthetic attacks. Therefore, the next two subsequent subsection will broadly discuss about the literature pertaining to existing vulnerability assessment and the framework for resiliency improvement.

1.2.1.1 Vulnerability Assessment

Reference [38] investigated how coordinated cyber-physical attacks can exploit vulnerabilities in power systems that follow the N-1 security standard, potentially overloading transmission lines, resulting cascading failures. A tri-level model is proposed in to analyze these attacks by utilizing semi-definite programming relaxation and primal-dual formulation for optimization. Case studies show that in N-1 secure systems, the proposed attack can trigger the tripping of additional lines, or can creates N-1-1 contingency. However, the attack's impact is severely constrained by the load measurements' bound of change. A dynamic risk assessment model for CPS against cyber attacks is constructed in [39] considering both software vulnerabilities in cyber devices and physical consequences in power systems. It estimates the physical effects of minimum shedding loads in N-1 circumstances brought on by a maliciously controlled SCADA system in a substation. Another multi contingency vulnerability algorithm is proposed in [40, 41], using graph theory and DC power flow based linear sensitivity factors. As an aspect of a novel vulnerability assessment model, [42] utilized a stochastic counterfactual risk analysis method to get around data limitations of topological information. The research conducted in [43] evaluates power system cyber vulnerabilities incorporating both physical failures and cyber security risks by developing a statistical framework. This framework was built upon human dynamic theory where attacker versus defender interactions are modeled via static and Markov decision model. Reference [44] addressed the vulnerability of CPS considering the impact of cyber layer failures on cascading failures. Vulnerability indices are established based on network structure and power flow properties under different interface and attack strategies, which helps to analyze the CPS performance before and after cascading failures. The findings demonstrate that malicious attacks and critical cyber nodes significantly increase vulnerability. Similarly in [45], a clustering-based vulnerability evaluation framework is proposed adopting a mixed-integer linear programming (MILP) approach for searching minimum combination of the most vulnerable communication channels under certain extreme operational constraints.

Aforementioned literature survey reveals that majority of the reported cyber attack strategies, and hence, their defense frameworks heavily rely on an aprior detailed system studies. However, such a detailed analysis of the system along with so much of real-time data might not be accessible to the attacker for devising an attack. Thus, it will be beneficial if a vulnerability evaluation approach can be developed which exploits the topological structure of the system, and thereafter, develops a cyber attack resilient framework against such power grid structural vulnerabilities. Recently, the concept of

Complex Network Theory (CNT), has gained considerable attention from the research community for evaluating the structural aspects of the network's system vulnerabilities because of its prominent features and simpler approach of solving various large scale problems in the domain of different complex networks i.e social networks, biological networks, citation networks, brain network etc. [46, 47]. Centrality measures are the essential tools of CNT, which estimate the significance of certain features of complex networks according to the structural properties of nodes, edges, and their level of connectedness. There are essentially two categories of centrality metrics exits that are frequently used to assess the effectiveness of any real world network dynamics and analyze its influential nodes or edges depending upon the network structure: those that use local information and those that use global information [48]. Since it only looks for local information, the local metric has the advantage of a somewhat faster computing speed, whereas the global metric has a moderate to high computational complexity but measure network's overall performance from a wider angle.

Numerous interdisciplinary studies have been conducted by modeling conventional power grid within the framework of CNT, which used various fundamental traditional centrality indices, such as degree centrality, betweenness centrality, closeness centrality etc., to assess the structural vulnerabilities of power network by quantifying the structural importance of any nodes or edges [49, 50, 51, 52, 53]. The results have shown that electric power networks not only have the characteristics of small-world networks [52], but also have the crucial characteristics of scale-free networks [51, 53], which make it vulnerable to deliberate attack and sturdy to random attack or accidental failure of transmission lines. In [54], a two-step screening-and-ranking approach is proposed to assess the vulnerability of transmission grids under extreme contingencies i.e natural and synthetic attack event. At its first step, vulnerable transmission lines are selected based on critical eigenvalue sensitivities and topology analysis that searches for the cutsets in the system leading to islanding. In the next ranking step, time domain simulation are performed to rank those screened out transmission lines according to their actual dynamic impacts. Since most cyber criminals will only possess a limited amount of system information, a standard power grid N–1 security analysis cannot be expanded to fully evaluate the risk. Therefore, authors of [55] make use of graph theory based closeness and edge betweenness centrality metric to investigate cyber physical vulnerabilities for N-X contingencies with limited resources. The simulation results of the method shows that pertaining to the loss of bus or node injection, closeness centrality seems to be a superior vulnerability assessment tool for identifying high impact event than the edge betweenness centrality which aim to assess the loss of multiple line outages. However this centrality based methods have a limitation on selecting maximum number of top contingencies upto three. In reference [56], an extended betweenness centrality metric is used by incorporating some electrical parameters in the formulation of traditional centrality to identify vulnerable components of the network. But in that ranking scheme topological attributes are completely ignored and dynamic ranking are also not incorporated for power system vulnerability analysis. Therefore, it will not be as good as contingency ranking metric to identify critical components. In [57], an improved betweenness is proposed over adjacent graph based on mapping of topological parameters to electrical network to assess the vulnerable features of the transmission networks. But this proposed approach, includes the effect of overload mechanism only due to spontaneous fault but not the impact of attack. The authors of [58] developed various node-attacking strategies and conducted an empirical analysis of their effects on the structural perspective and operational performance of the power grid using a number of conventional centrality metrics. But the only attacks shown are node attacks, which are less probable than line attacks.

1.2.1.2 PMU Equipped Secured Metering Framework

One way of utilising the vulnerability assessment results is to exploit the vulnerable points in designing of a secured metering groundwork to prevent or mitigate the effect of the cyber-physical attack on the grid infrastructure. Identification and then protection of basic measurements set with advanced information technology (IT) security measure or safeguarding certain measurements with the deployment of PMU are the key ideas to improve the resiliency in attack detection frameworks [59, 60]. A greedy approach was presented in [17, 61] to choose a small subset of measurements that must be protected against data integrity attacks by strategically placing secured PMUs so that the attacker's attack resources rise several times higher than they would if there was no protection. However, the greedy approach based methods may stuck in local optima and thus not able to assure the best optimal choices of PMU location all the times. Finding the bare minimum number of connected lines in a system architecture to maintain observability requirement is also crucial for executing vulnerability evaluations and enhancing online security monitoring, in addition to identifying critical PMU deployment locations. An observability recovery problem is thus formulated as a MILP problem in [62] to find the locations for sequential restoration of PMUs after a massive cyber attack on the grid that affects the situational awareness and cyber physical resilience. However, a priori information of optimal sets of PMU based on greedy and random strategies are required in advance to accelerate the recovery process. Reference [63] have effectively addressed this issue through the development of a bi-level optimization problem where the lower level problem carries out the job of finding traditional Optimal PMU Placement (OPP) for some selected combination of secured lines and the upper level problem is used to determine the optimal allocation of specific lines, allowing utilities to make better decisions regarding power system monitoring based on available information of critical buses and transmission lines. The major difficulties that arise out of the optimization problem are lack of transparency between two levels which obscures interpretability of accurate line selections, convergence issues over a large scale system and computationally expensive and complex solutions. Another methodology for optimizing the placement of PMUs considering both system and topology aspects of disturbances are presented in [64]. The proposed approach intends to improve the accuracy of pre- and post-disturbance monitoring, especially for single transmission line outages, by articulating the PMU placement problem to assure full observability of power systems and integrating post-disturbance variations. The authors of [65] demonstrate an ingenious pre-deployment PMU technique that works in harmony with the current PMU deployment approach. This is capable of deterring an attacker from successfully launching a linear FDI attacks. In order to combat different injection in wide-area monitoring and control systems, a multi-sensor temporal prediction based wide-area control method is developed in [66]. This method collects real-time measurement data from available PMUs, modifies those estimates using a temporal prediction filter to find any discrepancies. However, the effectiveness of the temporal prediction filter heavily relies on its ability to identify patterns specific to malicious injections thus it might lead to false positives, where discrepancies are identified even when no malicious injection is present. This can trigger unnecessary control actions and disrupt normal grid operations. Modified Teaching-Learning Based Optimization has been used to solve this multi-objective PMU placement problem. An end to end PMU-based attack resilient cyber physical framework in Wide Area Monitoring and Control System has also been portrayed in [67]. In this work, an in-depth resilient architecture, various attack resilient algorithms are developed which effectively denominate the aspects of cyber-security risk assessment, attack detection, prevention and mitigation approaches. The above mentioned literature survey reveals that majority of the existing approaches are computationally expensive, complex, requiring apriori system information, struggling to find global optimal PMU locations, often settling for suboptimal solutions and mainly focusing on observability without considering the order in which they installed in face of multi-layer of line outages which mainly caused due to cyber attacks. Thus, a simple yet effective smart metering placement framework is still of interest which can secure an optimal set of measurements using minimum investment.

1.2.2 Replay Attack Resilient State Estimation Framework at Transmission-Level

Once a secure PMU infrastructure is established throughout the system, the next subsequent critical task is leveraging these secure measurements in EMS applications, notably Power System State Estimation (PSSE). However, the PSSE is highly susceptible to various data integrity attacks, specially Replay Attack (RA), emphasizing the need for an attack-resilient framework to safeguard PSSE, upon which many other EMS applications like optimal power flows, economic dispatch, contingency analysis rely. Therefore, the following subsection will delve into existing literature on available security measures for attack-resilient PSSE against RAs. It has been acknowledged by various research studies that sophisticated data integrity attack such as FDI attacks, RAs etc are designed with the intention to fool the traditional state estimator and remains undetected from the defender's surveillance which ultimately arises the concerns of detection and control of attack extremely challenging. There mainly three broad categories of cyber attack that very frequent to been seen in CPS are: *denial of service attacks*, where superfluous illegitimate requests are being sent to the host machine to temporarily or indefinitely disrupts it services; *data injection attack*, where a carefully designed synthetic unknown value is injected with the original measurement set to falsify it without being noticed by defender's security mechanism; and replay attack, where a valid transmission of data is first fraudulently recorded and then maliciously repeated or delayed with the use of a record-with-replay attack script features as shown in Fig. 1.6. The first two categories of attacks can be somewhat minimized or prevented to some extent by using anomaly identification and multi factor based authentication tools, implementing strong firewall, intrusion detection and data loss prevention mechanism, advanced statistical and signal processing based attack detection methods [68]. However, as compared to above two types, on one hand replay attacks are in general very easy to be executed in real practice and on the other hand a bit difficult to be spotted due to maintaining statistical similarities of the replayed signal with the original observations and thereby having capability of passing examination of cryptographic keys, resulting interrupting the power delivery and degrade system performances. Thus, this subsection explores those literature that intend to replay attack (RA) detection and its secured isolation from the CPS.

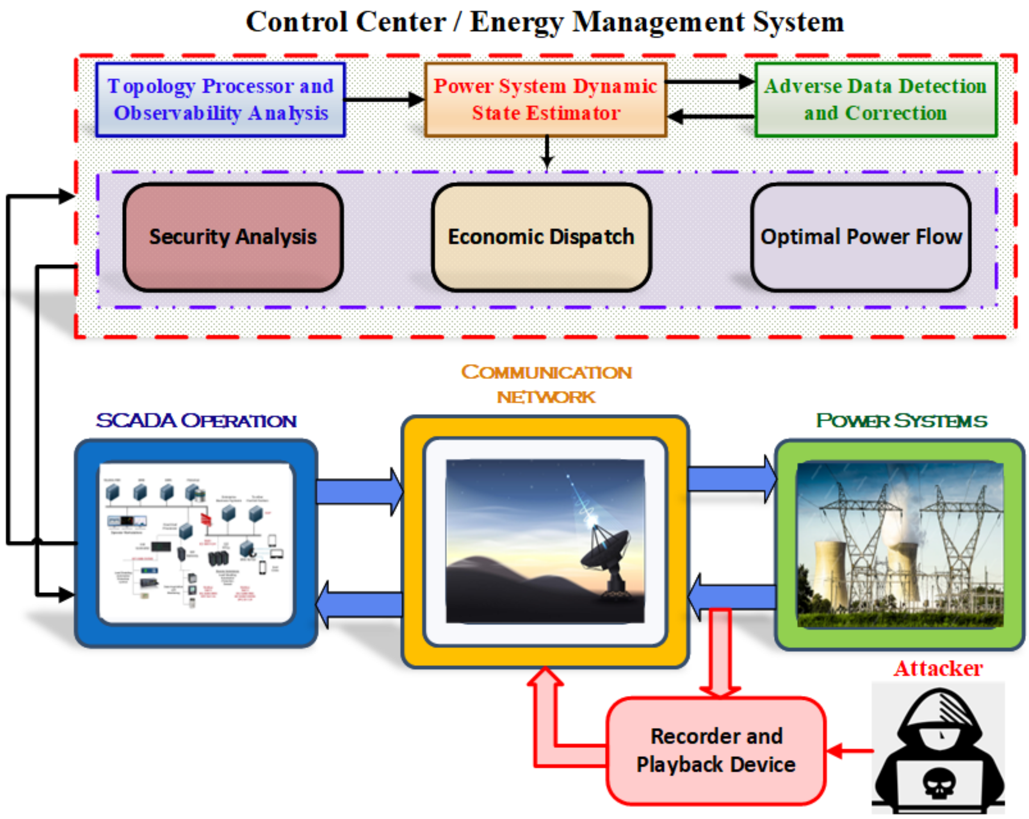


Figure 1.6: Execution of Replay Attack and its impact on Power System State Estimation

To ensure the security and reliability of CPS in the face of RAs, various defense mechanisms have been proposed out of which attack detection is considered as a fundamental element of security measures. The existing literature classifies these defense approaches into two main categories: *Passive Approaches* and *Active Approaches*.

1.2.2.1 Passive Approaches:

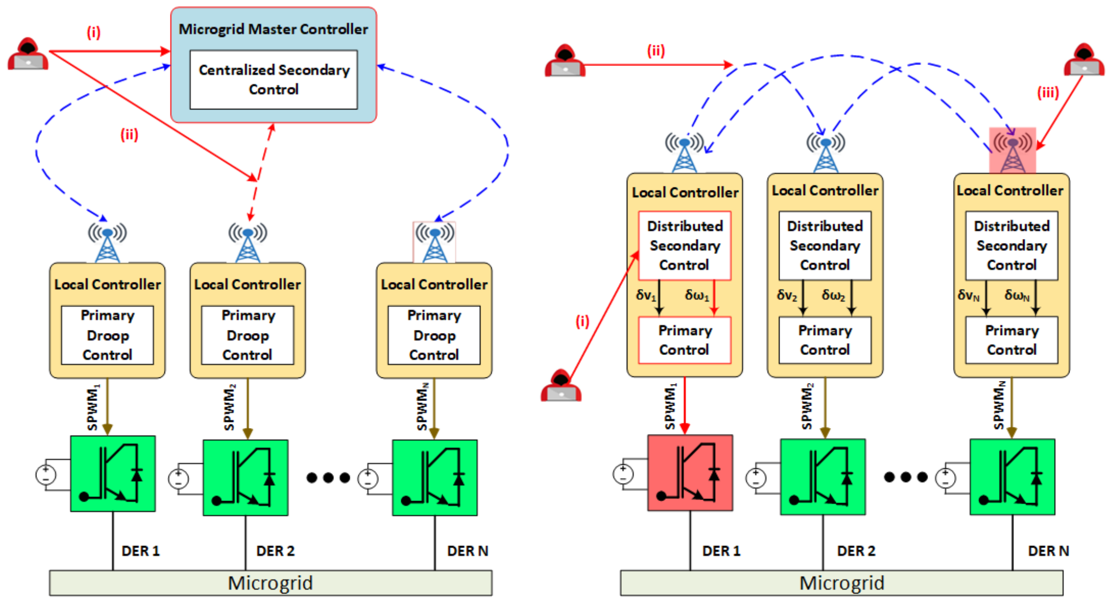
The Passive Approaches are the ones which utlise some observatory parameters such as model residuals, measurement estimates, computational and sensory resources etc. References [69, 70] investigate the issue of RAs in securing distributed state estimation using a multisensor approach. At first, the compromised sensors are identified and then with the help of designed distributed observers the adverse effects of the attacks can be eliminated. Another approach for secure state estimation in CPS is presented in [71], emphasizing multisensors information fusion. However, the above-mentioned studies needs for further development in improving the detection rates of RAs. Additionally, this multi-sensor based RA detection method assumes proactive protection of some sensors by operators, making them immune to manipulation but involving computationally expensive operations, limiting real-time implementations. Kalman filter integrated with linear quadratic Gaussian regulator is used in [72] to develop a data driven methodology to detect RAs in SCADA sensor or meters. At first, the state space modelling of plant and measurement function is derived and then statistical measure over the model residuals are utilized to detect RA. However the proposed approach is limited to liner-time invariant system only. Reference [73] introduces a modified receding-horizon control approach for discrete-time linear time invariant system to address RAs and assesses its impact on the system performance. In sensor networks, the issue of distributed $H\infty$ filtering in discrete-time nonlinear systems susceptible to RAs is examined in [74]. It establishes a pattern to explain the temporal behavior of RAs and adds an indicator variable to detect them. A real-time, PMU-based data-driven cyber-attack detection mechanism is proposed in [75] to detect continuous RAs in wide-area monitoring, protection, and control system employing Autoregressive Integrated Moving Average modeling and Kullback Leibler divergence analysis. However, a major disadvantage of this time-series model based replay attack detection method is the substantial amount of data preprocessing and tuning work required. It is highly dependent on precise topology knowledge and system parameters, which may vary according on the load conditions, as well as a precise understanding of the probability distribution of time-series data. In order to successfully identify RA, the authors in [76] use the relative changes in eigen value information derived from singular value decomposition and Pearson correlation of PMU measurements. However, there are growing concerns over the method's effectiveness and a significant computing overhead, especially when dealing with RAs that exhibit delays or long playback duration. A new control strategy is proposed in [77] utilizing standard Model Predictive Control (MPC) scheme to detect replay attacks and take corrective actions. It leverages the receding horizon nature of MPC and the concept of controllable sets to identify inconsistencies caused by replayed data. An idea of using blockchain-based decentralized framework to addresses the challenge of detecting cyber attacks in large-scale power systems with real-time sensor data is first conveyed in [78]. This approach focuses on detecting coordinated RAs based on locally reported alarms and associated statistics, while preserving data privacy. In general, the major limitations of the above-mentioned passive approaches lies in the fact that, firstly some of the methods are not suitable for bulk power delivery systems as in such case it is very difficult to get suitable linearized model without much approximations. The accuracy of some of the model-based techniques heavily depends on the topology information and knowledge of system parameters which may change in loading condition.

1.2.2.2 Active Approaches:

Active approaches to RA detection involve intentionally injecting an external signal, such as noise or a specific watermarking signal, into the system to monitor its response. These methods rely on analyzing the system's output in relation to the injected input to detect RAs effectively. The receiver compares the system's response to the injected challenge with a pre-defined expected response based on the challenge design. Unlike passive approaches that rely solely on analyzing existing data patterns, active methods create a dynamic environment where deviations from expected behavior can reveal anomalous sensor readings caused by replay attempts of the adversary. A popular strategy for RA detection in active techniques is the addition of watermarking signals [79, 80, 81, 82, 83], i.e., encrypt the measurement signals and control inputs with embedding watermark. Subsequently, a range of statistical tests are conducted in an attempt to potentially detect RA signals and suggest countermeasures for attacks. This watermark can be a random sequence, a cryptographic hash, or any signal designed to be easily detectable by the receiver. For example, an intriguing approach for periodical injections of independently Gaussian noise or any harmonic oscillation to the control signal which is only known to defender was carried out in [79, 81]. This carefully chosen "noise" introduced into the system is commonly known as watermarking, which deliberately creates discrepancies between the genuine system states and the compromised ones in the event of RA. Unlike the previous method, authors in [82] proposed a frequency-based detector for RA detection in CPS where a sinusoidal watermarking signal with a time varying frequency as an authentication signal was injected in the closed-loop systems and then checks if the frequency components in the output signal match the time profile of authentication signal or not. However, these strategies are effective in achieving high accuracy in detection, but it necessitates meticulous design and safeguarding of the authentication signal with compromises in controller performance. In reference [83], the authors explored a dynamic variation of watermarking which involves embedding unalterable patterns into a medium, capable of detecting any manipulation of sensor measurements. Apart from these, there are other variants of watermarking methods can also be found out in literature such as additive [84], multiplicative [85], time-varying [12] and optimal watermarking [86]. Additional complexity with increased cost in the design and implementation of those watermarking signal generator is one of its major hindrance with this class. Also the assumption of taking full access of all the available sensors in the CPSs followed by some of the papers seems not to be so realistic in practical sense. Nevertheless, this approach has also certain weakness such as: Need of improved detection rate, RAs identification may gets failed if signal be encrypted by the attacker before coded and limited tolerance for extended-duration RAs. The preceding literature surveys have effectively underscored the critical need for secure metering and resilient infrastructure at the transmission (T-system) level, along with existing approaches to address these challenges. However, as the focus of this thesis extends to provide an end-to-end cyber attack resilient monitoring and control framework, it is crucial that power system researchers should also consider the distribution (D-system) side's cyber attack vulnerabilities as well. This is particularly relevant in the context of microgrid systems, which are currently been experiencing rapid growth in distributed renewable energy penetration but often lack robust cyber security measures. Therefore, the subsequent literature survey sections will now delve into the detection and security measures of D-systems, necessary to maintain reliable and resilient distributed energy infrastructure. Thus, to start with, the next literature review section will now explore methods for accurately detecting, classifying, and localizing cyber attacks within MG, aiming to enhance the overall cyber security posture of critical energy systems.

1.2.3 Cyber Attack Detection and Classification Techniques in Islanded Microgrid at Distribution-Level

The concept of microgrid brings up a new dimension in monitoring, protection and control of information and power flow to the existing distribution management system (DMS) which fulfills the gap between the reliability and sustainability requirements and manages diverse power demand issues economically with the advantages of significant reduction of pollution margin, higher energy utilization rate, lower power transmission loss, etc with the effective and coordinated integration of (DERs). All of these multifaceted benefits makes MG to be a promising solution for future self-reliant autonomous power delivery networks which can operate in either grid-connected or islanded (autonomous) mode via three-level hierarchical control architecture: primary, secondary, and tertiary [87, 88]. Among these hierarchies, secondary control is the main critical component which guarantee the reliable operation of MG by compensating any deviation in the voltage and frequency parameter with the help of exchanging global information among the neighboring DERs. In that perspective, MG is also be a part of complex CPS and familiar to be known as an networked control multiagent system where each DER is treated as an agent and they have the privilege to communicate among themselves to serve certain system-level objectives and thus reach to a global consensus agreement [89]. Figure 1.7(a) and 1.7(b) shows two commonly used communication architecture of DMS i.e (a) Centralized and (b) Distributed, where the co-operative secondary control mechanism for both of the above control structures are equally vulnerable to high risk of data manipulation attacks due to widespread use of remote sensing, transmitting and computing devices such as sensors, actuators, controllers, and vulnerable communication links. Alongside the high volume deployment of power electronic converters with their cloud-based software-intensive controllers integrated with unencrypted susceptible communication protocols such as Modbus, DNP3, IEC 61850, TCP/IP, make the sensors to be easily compromised, alter the data of communication ports via packet sniffing and thus creates a treacherous channel for the intruders that allows them to take full access of DMS for hijacking the DER's controllers and thereby perform numerous malevolent activity in the grid [90, 91]. The injection of corrupted data either in the secondary controller of DERs or in its neighboring communication links forcefully introduces significant errors in the voltage and frequency distributed secondary consensus law that drives the whole system towards instability. Under these circumstances, timely detection of malicious attacks with correct identification of either the misbehaving agents or the corrupted incoming communication link is very important in the aspect of cooperative network-based MG system. The existing literature



(a) Centralized (b) Distributed

Figure 1.7: Microgrid control structures, its communication network and its arena that's prone to be attacked. (i) Controller hijacking attacks, (ii) False data injection to communication channels, and (iii) Sensor compromised.

on cyber security issues in islanded MGs can be broadly categorized into three main groups. The first group focuses solely on cyber attack detection, classification, and localization, techniques aiming to quickly identify and isolate the affected DERs to restore normal MG operation. However, this approach may result in compromised DER utilization and impede maximum power serving capabilities of the MG system. The next two research groups focus on developing resilient frameworks to mitigate the effects of injected attacks, ensuring continued reliability and functionality despite persistent potential security threats. Thus, this current subsection highlights the works belonging to the first group only, while the next subsection will delve into the details of the next two groups.

In that reference, to address the issues related to cyber attack detection, the authors in [92] exploit the difference in output of secondary voltage sublayers of DERs to propose a cooperative vulnerability factor index that can differentiate the attacks on the voltage sensors of the corrupted agents. At the onset of attack this factor will converge to some non-zero steady state value, indicating the possibilities of cyber manipulation. A unified anomaly detection approach is presented in [93] to calculate a data integrity index using the Laplacian disagreement function which is nothing but the consensus error of the communication graph topology following which agents are communicating among each other within a MG cluster. The False Data Injection Attacks (FDIA) detection problem in [94] was formalized as monitoring a change in the set of some MG candidate invariant parameters and identified DERs are then removed from the system in order to regain MG stability. A discordant-based detection method is proposed in [95] using the local and neighboring measurements to distinguish the attack nodes under deception and destabilization attacks on the current sensors. In reference [96], parametric time–frequency logic was employed to classify cyber attacks and fault-based anomalies in MG. The proposed detector utilizes time–frequency information extracted from training datasets, which include anomalous data. During the testing phase, this information is applied to identify abnormal elements among the normal inputs. Authors in [97], present an idea to detect FDIA and DoS attacks in the sensors or communication networks based on signal temporal logic which basically works by evaluating output voltage and currents of a MG against desired specifications and comparing it with some operational bounds. A robust cyber-attack detection algorithm is proposed in [98, 99] based on the safe estimation of states using the Kalman Filtration techniques while modeling the whole MG in state space. Apart from the above-mentioned model-based methods, a few more existing works had used various data driven neural network and machine learning-related model-free methods, such as Artificial Neural Networks [100, 101], Recurrent Neural Networks [102], Reinforcement Learning [103, 104, 105] and Deep Learning [106, 107] to detect, classify and localized the attacked node i.e., DERs in MG. However, the issues with the above-mentioned methods seem to be quite complex and need accurate model verification tools. Alongside these model-free approaches also increase the computational burden and system complications due to requirement of voluminous labeled data to train the detectors.

It has been observed from the above literature that most of the existing works are implemented keeping voltage regulation and current sharing of DC MGs as a target application and only few works among [108, 109, 110, 111, 112, 113, 114, 115, 116] are primarily focused on detection of cyber attacks over consensus-based distributed secondary control of AC MGs. In [108], a time-dependent cryptographic technique, named as Link Error Counter is used to detect any data manipulation on the communication links.

A chi-square test enabled, residual-based cyber attack detection method is proposed in [109] which utilizes a cyber physical real-time reference model built upon a digital real-time simulator. Authors in [110], used subspace identification method of control theory as an integrated approach for the detection of attacks in AC MG's sensors, controller's input and local loads via modelling the MG as a state space representation of CPS. The overall detection framework however is based on centralized secondary control infrastructure which is computationally expensive and prone to be affected by single point failure. Additionally, although the above proposed detectors able to quickly spotted the multiple attacks with different level of accuracy, but such chi-square test based attack detection method failed to work faithfully for stealthy attack in the MG [117]. A fully distributed synchronous detection method is proposed in [111] to detect the attack DERs by relying only local neighboring communication information. Reference [112] introduced a novel noninvasive anomaly diagnosis mechanism for inverter based resources, aiming to ensure reliable and secure operations by classifying grid faults and cyber-attacks. After collecting locally measured voltage and frequency inputs and mapping them in a XY-plane, the proposed approach can able to draw the characterization process by utilizing physics-informed empirical laws and sample-based trajectory analysis. It still needs qualitative data for training related to various fault scenarios even though it does not require complicated mathematical models. Designing mechanisms for such high-accuracy anomaly diagnosis is further restricted by the availability of such qualitative data. Authors in [113] proposed a framework for developing cyber-security test bed that counteracts with emerging vulnerable cyber intrusions in power distribution systems which can jeopardize the safe and reliable operation of DERs. Various types of non-parametric and parametric attacks, their effects and the proposed countermeasure are demonstrated via HIL simulation. An attack detection and identification method based on dynamic state estimation using an unknown input observer (UIO) can be found in [114, 115]. This method estimates MG states and generates a residual function to detect FDI attacks, triggering a detection alarm for attack isolation and mitigation. In order to achieve the objectives of correct identification of attacked agent in the distributive communication network, a relative entropy based attack detection technique utilizing kullback-liebler divergence (KLD) criterion is presented in [116]. As per this method, by exploiting the statistical properties inherited by KLD of auxiliary secondary frequency control input of a DERs under normal and compromised situation, attack can be detected. However, the technique described imposes a considerable communication overhead and computational burden on distributed secondary control. Recent research, such as [118], has also highlighted concerns regarding the complete detectability of the KLD method. It has been demonstrated that attackers can still achieve stealthiness by exploiting mathematical expectations, covariance, and probability distribution knowledge on normal innovation and designed upper bound parameters of KLD statistics.

While the first group of methods is able to detect attack and is able to mitigate it by only restricting its spread through the complete isolation of infected DERs, they may suffer with following major drawbacks: (1) Decreasing system utilization efficiency, (2) Requiring upgradation of existing secondary control hardwares, and (3) Undermining consensus agreement in case of severe attacks. Moreover, majority of the above literature is based on two pessimistic assumptions [119]: (1) The information communicated from the leader DER is not attackable as the operator strongly safeguards such links through some expensive advanced security mechanisms, and 2) the resiliency of the network controlled multiagent system depends on the characteristic features of the network topology, i.e., the network should always posses graph connectivity of at least more than $2f$. That means, to prevent f malicious DERs from disrupting the operation of its neighbouring DERs in distributed consensus rule, the minimum connectivity requirement for detecting attacks and achieving resiliency with disseminating information reliably is to make at least $(2f+1)$ neighbours DERs to be intact. Both these assumptions are not always practically possible. Thus, an efficient cyber attack detection scheme is of interest which can work accurately even if these assumptions are relaxed.

This motivated to explore the following literature that solves the issues highlighted above, in the first group by designing a suitable framework where the resiliency against persistent attack penetration into the MG system can be achieved through successful mitigation of the attack effect.

1.2.4 Cyber Attack Resilient Control and Mitigation Techniques in Islanded Microgrid

The second group of research literature involves incorporating additional data signals alongside commonly transmitted signals in the secondary control layers. This necessitates the implementation of an extra communication layer commonly called as *Hidden Layers* to design a special control laws to mitigate the attack and thus achieve resilient operation. For instance, in [120] a robust hidden layer-based attack resilient distributed secondary frequency control (DSFC) scheme is proposed to eradicate attacker's effort. This scheme is based on designing a suitable hidden-layer of a virtual system placed on the top of communication layer and securely coupled with the entities of physical and communication nodes. However, this method has limitation in selecting some carefully designed virtual system parameters which if being accessible by the attacker can cause loss of synchronization among DERs. A similar kind of attack containment based control method is presented in [121] where, a virtual control layer is designed with hidden networks in a cooperative and adversarial multi-inverter MG network. The concept of proposing resilient cooperative distributed secondary control scheme with integration of original MG system to a virtual system via interconnection of some virtual communication layers, shielded from being subjected to cyber attack is also fostered in [122]. Reference [123] introduces another cross-layer resilient control strategy for an islanded AC MG against FDIA and DoS attack where the bottom layer comprises of physical inverters and load, middle layer contains communication network for relative information exchange among inverters and the top layer represents a masked virtual parallel control network that execute resilient control commands in face of cyber attacks. A few more resilient methods that contain an original system, a virtual network and its associated hidden layers can also be found in [124, 125]. This competitive design criteria of those methods are verified by the Lyapunov-based stability theorem to ascertain guaranteed consensus dynamics while the attackers either intercept the DERs communication networks and corrupt its local state feedback input. In [126], a novel intrusion mitigation approach is proposed based on a weighted mean subsequence reduced (WMSR) technique to control the information flow of corrupted DER via a virtual communication graph. However, the applicability WMSR technique is limited by the minimum algebraic connectivity requirement of the communication graph for providing consensus among the DERs. To enhance the maximum resilience of the AC MG in the scenarios where DSFC of all DERs are compromised, the authors in [127] proposed a resilient co-operative frequency control framework. This framework introduces a few auxiliary state and resiliency index variables into the conventional secondary control scheme which basically counterbalance the negative effects of the attack penetration by regulating auxiliary control inputs. However, aforementioned mitigation methods are effective in dealing FDIAs in DSFC controllers and communication links, but the additional cost and computational burden due to incorporation of additional hidden communication layer is a mater of concern for its practical implementation, specially for large systems.

The third group of work mainly focuses on using some observer-based finite time control scheme or computation of some off-set compensatory terms to control the impact of the attack on MG secondary control action. For instance, the detection and isolation problem of FDI attacks are effectively dealt in [128] by introducing an interval observer to estimate the interval state of the physical system accurately and then utilizes interval residuals as a detection threshold. Additionally, an attack signature logical judgment matrix-based isolation algorithm is also proposed to isolate sensors where FDI attacks may be injected. Another residual observer based attack detection and mitigation method is discussed in [129] which exhibits it effectiveness in dealing with intermittent integrity attack in MG while satisfying network and stability constraints. Using distributed observer, an improvised attack detection and compensation method based on confidence and trust factor with faster convergence speed are proposed in [130, 131]. In [132], conventional secondary control is replaced by a novel integrated distributed control for frequency and voltage regulation to make the controllers resilient to cyber-attacks. A distributed adaptive algorithm is proposed in [133, 125] by combining a distributed state observer and H_∞ controller to mitigate the deception attacks on the MG controllers and sensors. In [134, 135], the authors employed distributed sliding mode control to estimate false signals and calculate cyber-resilient offset compensation terms, aiming to effectively mitigate cyber attacks in AC MGs. However, such methods necessitate heavy tuning efforts to properly scale hyper-parameters of the models and hence dynamic load sharing performance and retaining system stability against persistent attack situation is questionable.

Table 1.2 presents a comparative state-of-the-art summary of the existing literature on cyber-attack detection and mitigation strategies, emphasizing its key differentiating factors such as type of system considered, requirement of additional resources, computational overhead, resiliency capacity, response against natural events, and detection and mitigation capability. Furthermore, it highlights the validation tools adopted by each approach, providing insights into their effectiveness and applicability in real-world scenarios.

Table 1.2: Comparative Performance of various Cyber Attack Detection and Mitigation Schemes for Islanded MGs

Ref	Grid Types	Processing Overhead	Resilience Capability	Additional Resources	Effectiveness During Fault and Load Change	Attack Types	Localization of Attacked DERs	Attack Detection and Isolation	Mitigation	Validation Tools
[136]	AC	High	N/2	VHCL	×	FDIA/ DoS	×	✓	×	Matlab
[137]	AC	High	N/2	×	Load	FDIA	×	✓	×	PSCAD/ EMTDC
[138, 98]	DC	High	NS	LO+UIO	×	FDIA	✓	✓	×	Matlab
[131, 130] [139]	AC	High	N/2	DO	Load	FDIA	×	✓	Case Dependent[1]	Matlab
[116, 108]	AC	High	N/2	×	×	FDIA/ DoS	✓	✓	Case Dependent[1]	Opal-RT HIL+ Raspberry Pi
[126, 120]	AC	Medium	N	VHCL	×	FDIA	×	×	Case Dependent[1]	Matlab
[121]	AC	High	N/2	VHCL	×	FDIA	×	×	✓	Typhoon HIL+ dSPACE 1202
[123]	AC	High	NS	VHCL	Load	FDIA/ DoS	×	×	✓	PSCAD/ EMTDC
[132]	AC	Medium	N	DO+AC	Both	FDIA	✓	×	✓	PSCAD/ EMTDC
[133, 140]	AC	High	N	DO+AC	Load	FDIA	×	×	✓	Matlab
[141]	AC	Low	N-1	×	Both	FDIA	✓	Detection	✓	Matlab
[142]	AC-DC	Medium	N/2	DO	×	FDIA	×	×	Case Dependent[1]	Opal-RT+ B&R PLC
[143, 144] [110]	DC	Medium	NS	×	×	FDIA/ DoS	✓	✓	×	Matlab
[145]	DC	Medium	N/2	DO	×	FDIA	✓	✓	Case Dependent[1]	Matlab
[101]	DC	High	NS	Training Data	Load	FDIA	×	×	✓	Matlab
[146]	AC	High	NS	Training Data	Load	FDIA	×	✓	×	Matlab
[147]	DC	High	NS	Training Data	×	FDIA	×	✓	×	RTDS
[111]	AC	Low	NS	×	Load	FDIA	✓	✓	×	Matlab
[148]	AC	Medium	N/2	×	Load	FDIA	✓	Detection	✓	Matlab

[1]Depends on the nature of algebraic graph connectivity, **N:** Number DERs unit in MG, **NS:** Not Specified, **VHCL:** Virtual hidden Control Layer, **DO:** Distributed Observer, **LO:** Luenberger Observer, **UIO:** Unknown Input Observer, **BSMC:** Backstepping Sliding Mode Control, **AC:** Adaptive Controller/Compensator, **HIL:** Hardware-in-Loop

1.2.5 Cyber-Secured Islanding Detection

The rise of DG systems has brought about the critical issue of islanding, wherein a section of the distribution network becomes electrically isolated from the main power grid but remains powered by DGs. It is imperative for DG systems to possess robust islanding detection capabilities to mitigate the risks associated with islanded operation. Failure to promptly disconnect islanded generators within 2 second of its formation can lead to various complications for both the generators and the connected loads, including safety hazards and equipment damage. This issue however is now becoming more complicated due to recent rapid digital innovations happening within the grid which enables physical and remote access to the sensor measurements and local controls to be easy and unsecured, makes widesperad use of communication servers and networks immensely vulnerable and heightened the infidelity about the quality and integrity of received islanding data. To this end, the first segment in this section explores various conventionally available methods for detecting islanding and non-islanding events, while the second segment delves into efforts to fortify islanding detection schemes against cyber manipulation or preemptively detect cyber interference prior to islanding decisions.

1.2.5.1 Conventional Islanding Detection Methods:

In the event of any abnormal condition or fault at utility side, the accurate identification of islanding is very crucial for changing the operating control modes of DGs to autonomously operate in islanded condition and ensure reliability of power supply. Thus, to provide guidelines and requirements for the grid interconnection reliability and performance requirements for DGs under unintentional islanding scenarios, various islanding detection standards are prepared as listed in Table. 1.3. However, as far as the detection is concerned, the existing literature of conventional islanding detection methods (IDMs) are primarily classified into three categories: (1) Remote or Communication-based Methods, (2) Active Methods and (3) Passive methods.

Table 1.3: Islanding detection standards

Parameters	**IEEE-1547**	**IEEE-929-2000**	**IEC-62116**
Quality Factor	1	2.5	1
Detection time (Sec)	t<2s	t<2s	t<2s
Frequency range (Hz)	$59.3 \leq f \leq 60.5$	$59.3 \leq f \leq 60.5$	$(58.5) \leq f \leq 61.5$)
Voltage range (p.u.)	$0.88 \leq V \leq 1.10$	$0.88 \leq V \leq 1.10$	$0.85 \leq V \leq 1.15$

Remote or Communication-based Methods: This method need a dedicated communication infrastructure between control unit of DGs and utility grid to determine the islanding state [149]. In [150], a power line carrier communication signal is broadcasted from the utility substation to a designated distribution feeders path. In case the signal is lost by the receiver of DGs, an islanding scenario is suspected. There is an idea of using transfer trip which majorly monitors the status of all the circuit breakers and reclosers of an area to take decision over islanding events. This scheme needs standard data communication for the practical implementation [151, 152] . There are some other types of remote-based method also exists which necessitates installation of some electronic component such as inductor or capacitors which are normally connected in open state. However as soon as islanding occurred, the circuitry of the component gets closed due to significant changes of impedances detected at Point of Common Couplings (PCC) side [153, 154]. Nevertheless the remote or communication-based IDMs are very fast with having almost zero Non Detection Zone (NDZ), but the major obstacles of the practical implementation of such methods is huge capital investment over specialized hardware and expensive communication infrastructure requirements which might gets failed in any adverse system condition.

Active Methods: This philosophy of this method is to inject some minor disturbances in to the DGs control unit to observe how that influence the power system parameters. This observances introduced a large impact against such a modest disturbance injected when the DGs are actually islanded, whereas the impact can be usually negligible in grid connected condition. In [155], a disturbance is generated by injecting a signal equivalent to 1% of the d-axis current reference at a frequency of 20 Hz. Additionally, in [156], a disturbance is introduced into the maximum power point tracking controller when the absolute deviation of the output voltage exceeds a predefined threshold. In addition, various other considerable contributions to the active islanding detection techniques are: harmonic current injection and harmonic distortion-based technique [157, 158], impedance based active frequency drift [159], sandia frequency shift [160], active slip frequency [155], phase angle shift [161, 162], Active correlation [163, 164], negative sequence current injection [165] etc. Although active methods have a zero Non Detection Zone (NDZ), injecting disturbances into the inverter control circuit can degrade power quality, cause unwanted transients, and result in reduced performance, particularly in scenarios with multiple DG systems.

Passive Methods: Continuous monitoring of power system quantities and electrical signals at the PCC and thereafter compare its natural variations with a predefined thresholds are the key working ideology of this method. The conventional passive islanding detection methods such as over/under voltage (OUV), and over/under frequency (OUF) are now modified or combined with some other signal-processing and machine learning based techniques as reported in various literature. For example, support vector machine [166, 167], principal component analysis [161], Decision Tree [168], Wavelet Transform [169] and other neural network based [170, 171] approaches are the example of few modernized ML techniques commonly familiar for islanding detection. Reactive power control based method [172] is another very popular passive anti-islanding detection technique which rely on monitoring reactive power variations within the power system to detect islanding events passively, without the need for actively injecting signals or modifying system parameters. Reference [173] utilize the rate of change of dynamic load behavior to devise an passive islanding scheme consisting of both synchronous and inverter based generation. Apart

from this, a few more passive islanding techniques such as Rate of Change of Frequency, Rate of Change of Voltage Phase angle, mathematical morphology, modal decomposition, transient event detection along with positive sequence superimposed current angle at PCC has been found in [174, 175, 176, 177, 178].
A detailed literature review of the most popular islanding detection scheme under all the available IDMs and their comparative performance on various parameters are summarized in Table 1.4.

1.2.5.2 Cyber Secured Islanding Detection Methods:

The islanding detection problem is highly susceptible to cyber attacks in following three ways [189].

1. It is feasible that cyber physical attacks of any type can target the grid, changing a crucial signal parameters (voltage, frequency, phase angle, current etc) before it reaches to an IDM and creates a false, inadvertent islanding scenario. This fake action not only compromise the reliability and resiliency of the microgrid but also disrupts the power supply to critical loads, posing potential safety hazards and economic losses.

2. On the contrary, if a genuine islanding scenario occurs, but the attacker manipulates the results to mask it, the microgrid may experience severe fluctuations in frequency and voltage. This can lead to a significant reduction in power quality, causing damage to electrical equipment, particularly sensitive loads.

3. Furthermore, some islanding detection techniques heavily rely on proliferation of communication channels and gateways to transmit measurement data to the control center for processing via software-defined algorithms. However, this reliance poses significant risks of unauthorized data manipulation, DoS attacks, or damage to the communication channel, potentially disrupting actual islanding detection algorithms.

Currently, there is a scarcity of literature available addressing this crucial issue of cyber-secured islanding, which are now being discussed below.
Literature suggests that the first crucial step towards enhancing the D-System's resiliency is to integrate synchrophasor technologies. The advent of μPMUs has emerged as a valuable tool in the MG environment, offering high precision sampling rates for monitoring purposes which in turn increased visibility and situational awareness to DMS. Thus, owing to high resolution measurements obtained with high speed data acquisition, it is evident from the literature that there is an ongoing trends of making use of μPMUs in various pioneering work of unintentional islanding detection with the aim of achieving fastest response time, accurate identification of point of disconnection, lower cost, negligible NDZ and finally maintaining grid stability and prevent potential safety hazards [190, 181]. In terms of cyber security as well, the inherent intelligence of µPMUs can contribute to reducing the risk of cyber attacks aimed at manipulating islanding scenarios [191]. In

Table 1.4: Review of AID Methods

A-TS: Automatic Threshold Selection; **NDZ:** Non Detection Zone; **CS:** Capacitor Bank Switching; **NIL:** Non-linear Load; **PQ:** Power Quality; **QF:** Quality Factor; **FS:** Fault Studies; **DT:** Detection Time (ms) ; **AS:** Ancillary Services (Such as LVRT/FRT)

Review of Active Islanding Methods									
Methodology	**Parameters**	**A-TS**	**NDZ**	**CS**	**NIL**	**PQ/QF**	**FS**	**DT**	**AS**
Active Frequency Drift [159]	Injection of distorted current waveform	N	0	Y	N	High/ Y	Y	< 600	N
Sandia Frequency Shift [160]	Injection of distorted current waveform	N	0	Y	N	High/ Y	Y	< 600	N
Harmonic resonance [157, 158]	Injection of harmonic current disturbance	N	0	Y	N	High/ Y	Y	< 200	N
Active Slip Frequency [155]	Adaptive reactive power injection	N	0	Y	N	Low/ Y	Y	280	N
Active Cross-Correlation [163, 164]	Second-order harmonic current injection	N	0	Y	N	High/ Y	Y	< 400	N
Review of Passive Islanding Methods									
Methodology	**Parameters**	**A-TS**	**NDZ**	**CS**	**NIL**	**PQ/QF**	**FS**	**DT**	**AS**
ROCPAD Relay [179]	Phase angle between V & I	N	less	Y	N	N/N	Y	< 13	N
Rate of Change of Frequency Relay [174]	Receiver Operating Characteristics	N	less	Y	N	N/Y	Y	< 40	N
Mathematical Morphology [176]	3-Phase V & I	N	less	Y	N	N/ Y	Y	15	N
Transient Component [178]	Transient and positive sequence superimposed I angle	N	less	Y	N	N/ Y	Y	< 20	N
Modal Components [177]	Voltage phasor from MDFT based model transformation	N	less	Y	N	N/ Y	Y	< 130	N
Variational Mode Decomposition (VMD) [180]	Singular Value Decomposition based mode singular entropy	N	less	Y	Y	N/ Y	N	< 10	N
Pearson's Correlation Coefficient [181]	correlation between F & phase angle	N	less	Y	N	N/ Y	Y	< 250	N
Review of Hybrid Islanding Methods									
Methodology	**Parameters**	**A-TS**	**NDZ**	**CS**	**NIL**	**PQ/QF**	**FS**	**DT**	**AS**
Harmonic Profile Injection (HPI)[182]	$d(HPI)/dt$ as active & passive binary tree (BT) to detect	Y	0	Y	N	high/ Y	Y	200	N
Optimized Sandia frequency shift & dF/dt [183]	dF/dt to suspect & SFS as active method to conform	N	less	Y	N	low/ Y	Y	< 500	N
Gibbs Phenomenon based hybrid method [184]	$d(THD)/dt$ & Frequency shift based active method	N	less	Y	N	Low/ Y	Y	< 400	N
Q-f droop and VV, UV, dF/dt [185]	Q-f disturbance, under V, Unbalance V, dF/dt	N	less	Y	N	low/Y	Y	< 200	N
dV/dt & the dP/dt[186]	dV/dt as passive to suspect & dP/dt as active AID	N	Small	N	N	Low/ Y	Y	$20 - 30$	N
Review of Signal Processing and Machine Learning based Passive Islanding Methods									
Methodology	**Parameters**	**A-TS**	**NDZ**	**CS**	**NIL**	**PQ/QF**	**FS**	**DT**	**AS**
Auto-regressive and Support Vector Machine [167]	AR coefficients of voltage and current	Y	0	Y	N	N/ Y	Y	50	N
S-Transform and Support Vector Machine [187]	energy and standard deviation from ST matrix	Y	0	Y	Y	N/ Y	Y	25	N
Decision Tree based Intelligent Relay [168]	NDZ-based training/testing strategy using V& I	Y	0	Y	N	N/ Y	Y	100	FRT
Wavelet Packet Transform & BPNN [169]	Energy and entropy of V_{PCC} at different levels of WPT	Y	0	Y	N	N/Y	Y	< 40	N
Phase space and Adaptive ensemble classifier [188]	Normalized Euclidean norm based feature vector	Y	0	Y	N	low/ Y	Y	< 100	N
Adaptive Neuro-Fuzzy Interface System [171]	V_{RMS}, I_{RMS}, Total Harmonic Distortion, F, P, Q	Y	0	Y	N	N/Y	Y	< 40	N

extreme conditions when the communication channels between the control center and the μPMU may be compromised, μPMUs equipped with such intelligent autonomous upgraded module will be of great assistance to reduce the cyber security risk and bolster the resilience of islanding detection in MGs. This similar concept is also utilized in [192, 193] for reducing the risk of cyber attack, where μPMUs are first deployed in each busses of MG to obtain the voltage information. Then an intelligent separate subroutine is implanted within μPMU architecture that exploits the angle difference between positive and negative sequence component for generating decision over an islanding event. The data transmission channel employed in conventional μPMU applications for island detection is not used in the suggested method which reduced the risk of cyber attacks. However, cost of placing μPMUs at every bus will be a major hindrance of its practical implementation. The difference in positive sequence superimposed impedance angle between PCC and DG ends are used in [194] for devising a secured islanding detection method. The distinction between non-islanding and islanding event is confirmed based on observing the changes of the index value in positive and negative direction respectively. However, the proposed method need the PCC signals to be transmitted in each cycle to the DG end which is designed to be secured by integrating advanced security measures and encryption protocols.

With the research gaps identified through the comprehensive literature survey done in previous section, several potential areas for further investigation are revealed which sets up the motivation for the undertaken research. Thus, the next section is dedicated for delineating the motivation to elucidate how these identified gaps serve as the driving force for the research endeavors and highlight the significance of addressing them in advancing the field.

1.3 Motivation

Traditional power systems operation primarily focused on creating strategies to manage physical faults or disturbances in the system, such as outages, deviations from normal frequencies, and voltage imbalances. But present-day smart grids are getting equipped with more sensing, communication, and distributed control techniques to accommodate renewable generations, electric vehicle loads, storage, demand response, and other emerging technologies. This substantially increases the data transfers at both the transmission **(T-system)** as well as the distribution-level **(D-system)** grids, and makes the grid more vulnerable to cyber attacks. This thesis has, therefore, undertaken the task of developing an end-to-end cyber-attack resilient monitoring and control framework, considering both the T and D systems.

The transmitted data when reaches at the CC, is first used to carry out the observability analysis of the PS. In case, the received set of measurements are sufficient enough to estimate all the system voltages, the real-time measurements are fed to the very crucial application , i.e., the Power System State Estimation (PSSE). The output of the PSSE thereafter serves to many critical applications' decisions such as optimal power flow, economic dispatch, contingency analysis, etc. Clearly, the very first step towards building an attack-resilient framework demands the system operators to perform *vulnerability analysis* in order to pinpoint the weak points in their system that need to be protected from the possible data breaches, or have a fallback in order to make it more resilient to external threats. Vulnerability analysis is associated with the physical behaviour of PS which mainly has two aspects: (i) topological structure and (ii) operational states. Thus, there are two types of vulnerability analysis in power systems: structural vulnerability analysis and conventional vulnerability analysis. Large-scale power systems face challenges for conventional vulnerability analysis, which is based on complete operational data, topological information, and standard engineering models. Conversely, a power system's physical behaviour and its topological structure are closely related because a structural alteration may affect a power system's operating conditions, which, in turn, may affect the system's physical behaviour. Unfortunately, *current research often overlooks this structural vulnerability perspective, leaving a critical gap in understanding the full range of threats and unobserveability issues posed by such attacks.* Subsequently, this gap also leads to a potential solution which can make the grid resilient from cyber threats. Making some of the meters immune to attack so that the observeability of the system is maintained even in attacked scenario can be one of the effective approach towards devising potential remedies.

After safeguarding a set of critical meters in the system, the next vulnerable point to be strengthened in an attack-resilient framework naturally becomes the heart of the EMS, i.e., Power System State Estimation. The manipulation of operational states of the system via injection of false data into the unprotected sensor measurements or replaying them with previously recorded data after some alteration can greatly impact the outcome of the PSSE, and thereby, of the subsequent critical decisions. Commonly observed attacks in the PSSE are data injection attacks, where a synthetically designed value is injected to falsify measurements without detection; and replay attacks, where valid data transmission is fraudulently recorded and maliciously repeated or delayed with the use of a record-with-replay attack script features. The first category of attacks has been well researched in the literature, and can be minimized or prevented by using anomaly identification and multi factor based authentication tools, implementing strong firewall, intrusion detection and data loss prevention mechanism and advanced statistical and signal processing based attack detection methods. *The impact of RAs, however, has not been rigorously analysed on the PSSE.* The execution of RAs is very simple and straightforward, but it is difficult to be spotted due to maintaining statistical similarities of the replayed signal with the original observations and thereby having capability of passing data intrusion detection test. So, in order to safeguard the PSSE against RAs, first, (*from an attacker's perspective*) the modelling and injection of the RAs on a limited (but impactful) number of PS sensors needs to be carried out. Next, (*from defender's* *perspective*) a defensive correction approach needs to be developed for the PSSE to identify and resist different kinds of RAs by utilizing the secured sensor measurements in the estimation process.

Conventionally, cyber attacks are perceived as primary threats to transmission systems (**T-system**) owing to huge power flows associated with the transmission network, and thereby, the amount of impact these attacks could create in **T-systems**. It is, however, crucial to acknowledge vulnerabilities in **D-systems** as well, particularly in the light of heavy integration of the DERs at D-level, in order to develop end-to-end defense strategies. In **D-systems**, MG involves the communication and networking architecture for the efficient monitoring and control operation of DER units. To maintain the voltage and frequency stability throughout the grid, DERs need to communicate their information to its neighbouring DERs, or master controller unit through a prescribed communication architecture i.e., either centralized or distributed, both of which are equally vulnerable to attacks. Cyber criminals focus their attacks on components of the power system that heavily rely on information technology. These include the controller network, responsible for executing DER's distributed secondary control algorithms and programmed logics, the sensors network, which comprises software-based RTUs and IEDs, and the communication network, utilizing cables and diverse protocols for efficient data transmission. Consequently, cyber attacks can manipulate data transmitted within the smart distribution grid architecture, affecting parameters such as power injection, voltage measurement, line flow, and the operational state of relays, breakers, and switches. This motivates to establish a coherent attack resilient unified framework for stable functioning of MG. As per the literature review, *the existing knowledge gap fuels two compelling research pursuits: (1) precise attack detection, localization and classifications to identify the vulnerabilities exploited, and (2) the design of robust mitigation measures to counter the attacks on DERs controllers.* As far as the attack detection is concerned, most of the study focused in contributing to identify any observed abnormal event is FDIA or not. A critical gap still remains i.e, to dissect and pinpoint these attacks because absence of this information hinders the effective mitigation strategies. So, in **D-system** domain, it is envisaged to first accurately detect an attack, followed by the precise attack classification and localization. Finally, this information is exploited to develop a novel attack mitigation scheme.

Another pressing challenge within MG networks is the islanding operation of DERs, wherein these resources persist in operation even after disconnecting from the power grids. This isolation can have disastrous consequences in terms of power system stability and quality, as DERs may struggle to maintain stable voltage and frequency within safe operating ranges. Therefore, detecting islanding conditions becomes imperative to prevent catastrophic damage to sensitive loads in MG. However, the smart-active distribution grid, constituting a cyber-physical system with various components such as Renewable Energy Resources, ICTs, IoTs, and IEDs, faces coordination and security challenges. These issues make islanding detection methods significantly challenging in the presence of emerging cyber threats. Unauthorized access to islanding detection data, encompassing internal voltage, phase angles, and power output of inverters, can disrupt the actual islanding detection algorithm, leading to adverse consequences for all grid components. However, *it has been perceived from the literature that current islanding detection methods often fall short in the face of these sophisticated cyber threats.* This observation propels the motivation to conduct research and devise a novel islanding detection scheme integrated with a cyber attack detection method. This integrated approach aims to synergistically address the issues of unauthorized access and data manipulation, empowering system operators with accurate information and preventing false decisions regarding suspected islanding events caused by cyber attacks.

Hence, the present-day smart grids requires an end-to-end cyber attack resilient monitoring and control framework considering the challenges at both, the transmission-level (**T-system**) as well as distribution-level (**D-system**). The framework must encompass the essential stages of developing a cyber attack resilient strategies, viz., vulnerability assessment, attack detection, localization and mitigation, extending seamlessly from the transmission to the microgrid level.

1.4 Aim and Objectives of Thesis

The primary aim and main research objectives for this thesis are meticulously designed in light of the above-identified research gaps and motivational background.
The overall workflow and the holistic view of thesis organization is depicted in the Fig. 1.8. This thesis comprehensively investigates cyber security vulnerabilities for both Transmission Systems (**T-Systems**) and Distribution Systems (**D-Systems**) within a unified framework. For **T-Systems**, the focus of research contribution is on securing meter monitoring infrastructure and developing reliable attack detection and control mechanisms. For **D-Systems**, the work addresses challenges through resilient control and mitigation schemes, along with secured islanding detection monitoring. Each aspect of cyber security challenges is meticulously examined from both the attacker's and defender's perspectives, ensuring a comprehensive understanding of the threats and the corresponding protective measures.
Aligned with the comprehensive workflow described above, the thesis is designed to meet the following objectives:

Objectives:

1. To carry out a topological vulnerability assessment, and to develop a cyber-attack resilient secured metering infrastructure for the T-system.
2. To devise a novel framework for safeguarding the power system state estimation from replay attacks by exploiting the limited secured measurements obtained from objective-1.

3. To accurately detect, classify and localise cyber attacks in an islanded AC microgrid system.

4. Following the detection and attack localization information from objective-3, to devise a robust attack resilient control framework to compensate the effect of attack over DER's secondary control and retain MG's stability.

5. To develop a cyber attack-immune islanding detection Schemes (IDS) in MGs.

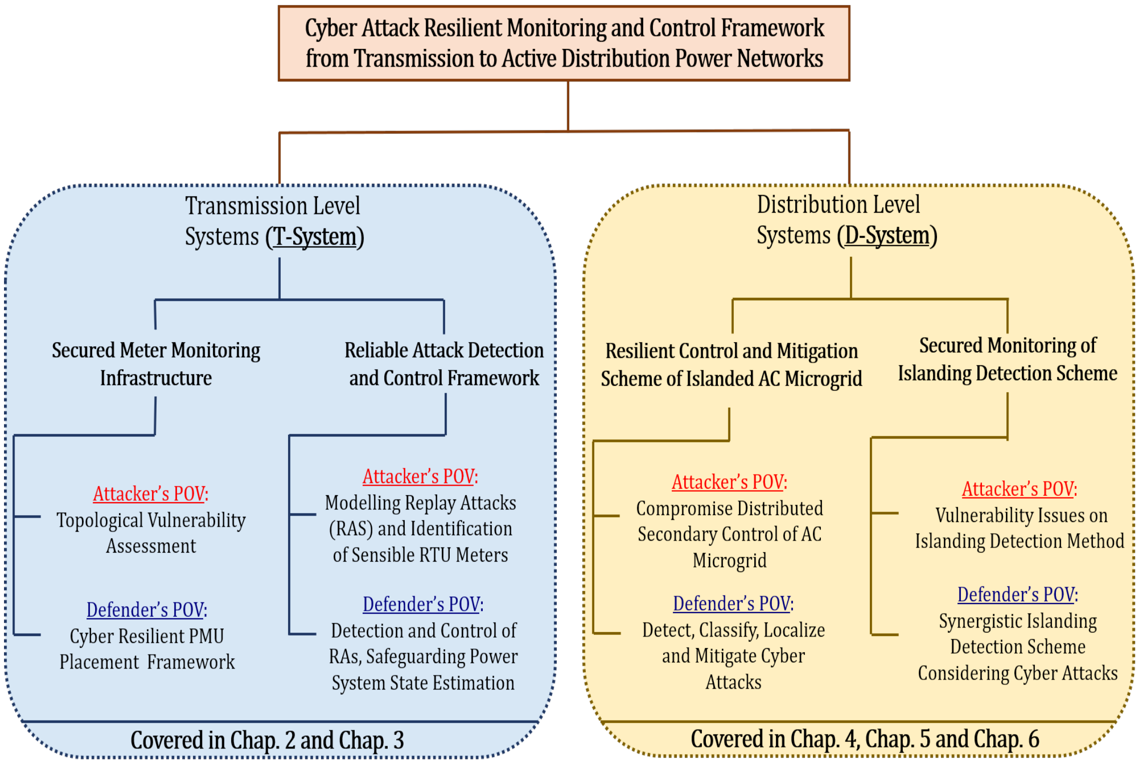


Figure 1.8: Flowchart of overall research work reported in thesis

1.5 Assumptions Considered in the Thesis

1. PMU is considered to be calibrated, error free, resilient and cyber secured device due to its advanced cyber security features. Information flow through the PMU channels are safeguarded with proper authentication and key encryption mechanisms.

2. Meter measurements that are kept under the strict surveillance of PMU's observability coverage range are assumed to be faithful and tamper-proof.

3. Attackers are assumed to have knowledge and access to critical components such as substation, human machine interface, communication port and network switches, enabling them to execute cyber physical attacks on the grid. However it is realistic to assume that they are always bounded by the limited number of sensor's manipulation due to their finite amount of attack budget or resources.

4. In the context of attack localization and mitigation in D-Systems, only one DER is assumed to be targeted at a time by the adversary.

1.6 Thesis Organization

Finally, the thesis work has been organized in the following seven chapters. The brief summary of each chapter is given as follows,

Chapter 1: Introduction

This chapter starts with the basic overview of cyber-physical critical infrastructure of the smart grid along with the brief explanation on most commonly seen cyber attack and its security challenges arising due to modern digitization of the grid. After presenting the literature survey on the proposed research work, it highlights the key research gaps in the literature. Then it sets the motivation behind carrying out the vulnerability analysis and attack detection and mitigation within contemporary power systems, particularly in the face of intricate cyber-attacks both at transmission (**T-system**) level as well as distribution level (**D-system**). Finally, it outline the thesis's proposed objectives and presents the preface of whole thesis work.

Chapter 2: Cyber Attack Immune Metering Framework

This chapter proposes a robust cyber-attack resilient framework, designed to address structural vulnerabilities in SGs. The proposal is mainly divided into two parts. In the first part, a novel and effective attack strategy known as Hybrid Between-ness Centrality (HBC) is proposed from the attacker's point of view to rank the most vulnerable transmission lines whose malicious tripping causes severe structural damage to the power system, losing system's observability and situational awareness. With the line ranking information obtained from the HBC, in the next part, a unique objective function is developed for the strategic placement of PMUs, aiming to safeguard those resulted vulnerable lines against FDIAs. The outcomes of this strategic PMU placement yields minimal sets of secure measurements that need strong protection to guarantee state variable integrity. This helps in enhancing system's resiliency index and to remain observable even in presence of a data integrity attack i.e., FDIA in some top vulnerable lines. The effectiveness of this framework is demonstrated through case studies on the IEEE 14-bus and New England (NE) 39-bus test systems.

Chapter 3: Novel Replay Attack Detection and Mitigation Framework for Power System State Estimation

This chapter deals with the problem of detection and correction of a very stealthy cyber threat i.e., Replay Attacks which can pose significant risks in various monitoring and control applications of **T-systems**. The work proposed in this chapter first explores the exploitation of topographical information and Power Transfer Distribution Factor to identify the most vulnerable SCADA meters whose compromised state could jeopardize the system stability. These vulnerable meters are sniffed and compromised for launching the replay attacks. Next, based on the secured phasor measurements from optimally placed PMU locations (as obtained from Chapter 2), a Hybrid state estimation algorithm is proposed which successfully detects and mitigates any replay attacks, if launched, from the PSSE measurement set. The effectiveness of the proposed scheme is demonstrated through simulating the model using the Real Time Digital Simulator (RTDS) on the IEEE 14-bus and NE 39-bus test systems.

Chapter 4: Detection, Classification and Localization of Cyber Attacks in Islanded AC Microgrid

This chapter focuses on the cyber attacks in **D-systems**, more specifically in AC islanded MG systems. Unlike the existing methods, the precise *localization* of the source of the attack is also explored as a key strategy in this chapter, thereby, allowing for the isolation of affected DERs from the system and minimizing the overall impact. In order to achieve this goal, two comprehensive detection approaches are presented. The first method relies on a two-sample distance-based probabilistic measure called the maximum mean discrepancy in a distributed cooperative secondary control of islanded MG for the timely detection of any malicious attack with correct identification of the misbehaving DERs. In addition to this statistical measures based detection, this chapter also introduces a second approach utilizing a machine learning based classifier, specifically the XGBoost algorithm to detect, classify and locate attacks. Once the attack is detected, two statistical inconsistency measure i.e. shannon energy and entropy are calculated and is utilized to introduce a novel rule-based attack classification approach integrated with the same XGBoost classifier to classify various types of injection attacks in the DER's controllers. Having detected the type of cyber-attacks, lastly a multi-class attack localization schemes after exploiting a few more statistical features to be incorporated in the XGBoost classifier, which aids in pinpointing the specific attacked DERs, streamlining the process of isolating compromised components from the system in worst-case scenarios. The proposed scheme is validated on a modified IEEE 13-bus islanded AC MGs systems modelled in the Real Time Digital Simulation environment.

Chapter 5: Unknown Input Observer and Back-stepping Integrated Sliding Mode Control based Cyber Attack Mitigation Framework

This chapter focuses on developing an attack resilient control framework that has the ability to mitigate the impact of the attack inflicted onto DER units. The framework is built upon the secondary control layer functionalities of the MG as it is most prone to be targeted by the attackers; aftermath of which leads to cascaded blackouts, endangering system stability. The proposed resilient controller first assesses the output of the attack detector obtained from Chapter 4 for the designing of an unknown input observer that can keep track the system states which, in turn, helps in computing the injected amount of attack bias by the perpetrator under compromised situation. Later, the coarse estimated bias obtained via previous step is further utilized in the backstepping-based sliding mode controller design approach to generate a suitable control law that enforces the injected attack to be compensated by finer adjustments of the compensation signal. Notably, this is achieved without necessitating any modifications to the existing hardware of the Distributed Secondary Frequency Controller or the addition of extra communication channels. To thoroughly validate the effectiveness of the proposed mitigation scheme, the modified IEEE-13 bus distribution test feeder operating in an islanded mode is modelled in detail with RSCAD software of RTDS. Furthermore, a Hardware-in-Loop (HIL) simulation control environment incorporating RTDS and dSPACE 1104 R&D controller is set up for the real time implementation of this proposed scheme to demonstrates its accelerated convergence speed and superior performance in handling unknown disturbances, uncertainties, and potential stealthy attacks.

Chapter 6: Synergistic Islanding and Cyber Attack Detection Scheme

Accurate and timely detection of an unintentional islanding in **D-systems** heavily depends upon the quality of data and its precise time of arrivals. FDIA, if launched on the key signal parameter before it being fed to any islanding detection algorithm may trigger a false unintentional islanding alarm. Cyber attacks can also delay the islanding decision, prolonging the exposure of vulnerable equipment and increasing the risk of catastrophic failures. This chapter, thus, proposes a cyber-attack resilient Islanding detection scheme. Initially, a Kalman filter based cyber attack detector (CAD) is utilized as a first layer of defense to check the integrity of the measurement of interest before being used as an input for the novel islanding detection scheme. The proposed detector, CAD is constructed based on the absolute difference between two very popular statistical correlation measure named Spearman's rank correlation and Cosine-Similarity respectively. As soon as the estimated data obtained from Kalman filtering are found to be contaminated due to any kind of attack, the proposed CAD quickly tries to trace the estimated change in the observed data pairs. Apparently when the error difference i.e., CAD approaches to zero, a flag is generated to identify the event as cyber attacks. Next, if the proposed CAD confirms absence of any cyber attack, a novel statistical property inherited passive islanding detection technique is activated to detect the unintentional islanding. The voltage mean value and the entropy information is exploited to develop a Mean based Islanding Detector (MID) along with an entropy-based Decaying DC Detector (DDCD). The MID and DDCD information is finally utilized to design a statistical relay digital logic (SRDL) that accurately distinguishes the islanding and non-islanding events. The proposed scheme is rigorously tested on a real life small scale industrial facility i.e., Banshee's industrial microgrid test system, modelled in the RTDS, on the basis of the IEEE-1547, UL 1741 standards.

Chapter 7: Conclusions

This chapter finally summarizes the key findings of the research work carried out in this thesis, along with a few areas for the future research.

Appendix A: Test System Data

Appendix A discusses the details of various test networks used in the thesis.

Chapter 2

Cyber Attack Immune Metering Framework

2.1 Introduction

As discussed in Chapter-1, since our future smart grids are going to be integrated with more monitoring, communications and distributed sensor-based technologies, they are becoming large, sophisticated, interconnected and complex Cyber Physical Systems (CPS) which is now being targeted by man-made or synthetic attacks. Therefore, assessment of vulnerability in the power grid network subject to extreme contingencies, cyber-physical attack or an unwanted natural disaster becomes a paramount importance to improve power grid's safety and resiliency against such unprecedented events. In order to assess the vulnerability of any network purely from a graph theoretical perspective, this chapter exploits some of the centrality metrics to investigate power system vulnerabilities from purely topological perspective by introducing a novel attacking mechanism to attack the most critical lines of transmission network which leads to major loss of system integrity, network structure and efficiency. The aim of the proposed work is to equip a given system with minimum number of Phasor Measurement Units (PMUs) such that the system remains observable even in presence of a data integrity attack i.e., False Data Injection Attack (FDIA) in some top vulnerable lines.

The rest of this chapter is organized as follows. Section 2.2 describes the preliminary concept of graph theory pertaining to the application of Complex Network Theory (CNT) on a power grid network. The proposed hybrid betweenness centrality attack strategy is discussed in detail in Section 2.3. Taking the vulnerable lines locations into account, Section 2.4 proposes a new optimization formulation to obtain optimum PMU locations with improved redundancy and complete topological observability under attacking condition. Secured measurements resulting from Optimal PMU Placements (OPP) are used for resiliency assessment of the network, as discussed in Section 2.5. Simulation results of the proposed method for vulnerability assessment and attack-resilient PMU placement are illustrated on the IEEE 14-bus and New England 39-bus systems in Section 2.6. Finally, the chapter is ended with concluding remarks in Section 2.7.

2.2 Preliminaries to Graph Representation of Power Grid

To access the power system vulnerabilities, the power network is seen as a complex, scale-free network modelled using graph theoretic approach. This graphical approach is useful to obtain the information of the nature of interconnection and the topological behaviour of the network. The abstract graph is formally defined as $\mathbb{G} = (\mathbb{V}, \mathcal{E}, \mathcal{W})$, where $\mathbb{V}$ is a finite set of vertices, $\mathcal{E}$ is a finite set of edges which is analogous to bus node and transmission lines of a power network respectively and $\mathcal{W}$ represents the set of weights of edge set $\mathcal{E}$.

1. The edges in the set $\mathcal{E}$ can be expressed as $\mathcal{E} = \{e_{i,j}|v_i, v_j \in \mathbb{V}\}$, where v_i and v_j are the extreme nodes of $e_{i,j}$. $e_{i,j} = \pm 1$, if node i and j are interconnected, otherwise $e_{i,j} = 0$.

2. The weights of the edge set $\mathcal{E}$ of a weighted graph $\mathbb{G}$, can be symbolized as $\mathcal{W} = \{w_{i,j}|v_i, v_j \in \mathcal{E}\}$, where $w_{i,j} \in \mathbb{N}$ and $w_{i,j} \subset \mathcal{E}$. If $w_{i,j} = 1$, the graph is called to be unweighted.

3. The cardinality of node set $\mathbb{V}$ and edge set $\mathcal{E}$ is N and M, respectively which constitutes an undirected adjacency matrix $\mathbf{A}_{adj}$, of size $N \times N$ of a graph network $\mathbb{G}$, where the elements of the matrix $\mathbf{A}_{adj}$, $a_{i,j} = w_{i,j}.e_{i,j}$, if node i and j are interconnected, otherwise $a_{i,j} = 0$.

4. Let d_{ij} be the shortest path between two pairs of node i.e v_i and v_j and the square matrix $\mathfrak{D}$ is termed as graph distance matrix that takes all-pairs of shortest path of node set $\mathbb{V}$ into the account.

5. $\mathbb{G}$ is said to be undirected if $e_{i,j} \in \mathcal{E}$ is same as that $e_{j,i} \in \mathcal{E}$ and therefore $a_{i,j} = w_{i,j}.e_{i,j} : a_{j,i} = w_{j,i}.e_{j,i}$, where $e_{i,j} = 1$.

In this chapter, for the analysis of vulnerability in power network, the graph representation is considered to be connected, undirected, unweighted with no loops and parallel transmission lines is associated.

2.3 Hybrid Betweenness Centrality: A Novel Vulnerable Link Identification Metric

The vulnerability of a transmission line towards cyber-attack is assessed based on the proposed Hybrid Betweenness Centrality (HBC) index. The proposed HBC index considers the combined effect of two centrality indices, (a) Eigenvector Centrality and (b) Current Flow-based Centrality. The first centrality index provides more robust comprehensive information about most influential set of nodes in a network that gives a better insight into the dynamical view of the network. Whereas the second centrality index gives an intermediate measures of global and local characterization of nodes. This doesn't follow the exact shortest path philosophy; therefore, it bears more valuable information spread from one node to another. That make it a suitable measure of application in a power network or any other network where information flows in a random direction based on network parameters properties.

2.3.1 Eigenvector Centrality Metric

Computing eigenvector centrality is a popular tool for measuring the influence of a node based on the importance of its neighbours, as well as its 2-hop and 3-hop neighbouring nodes. It works with the philosophy of assigning relatively higher importance to all the high-index nodes that contribute more to the score than lower-indexed nodes of the network. Therefore, the global eigenvector centrality of a node is the summation of centralities of its adjacent nodes. For a graph $\mathbb{G}$, the eigenvector centrality is computed as mentioned below:
Let say $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_N$, are the eigenvalues of $\mathbf{A}_{adj}$, mentioned in Section 2.2, of the network $\mathbb{G}$ such that for each $\{\lambda_i, i \in \mathbb{N}\}$, there exists a non-negative eigenvector U which satisfy the relation $\mathbf{A_{adj}}U = \lambda U$. Now, the non-zero solution of the equation $(\mathbf{A_{adj}} - \lambda\mathbf{I})U = 0$ gives a principal eigenvector ($U = [u_1, u_2, u_3, ..., u_N]^T$) i.e eigenvector corresponding to the largest eigenvalues (λ_{max}) are considered here to measure the centrality denoted as $C_E(v_i)$ of a node v_i as shown below:

$$C_E(v_i) = \frac{1}{\lambda_{max}} \sum_{k=1}^{N} a_{ik} u_k \tag{2.1}$$

To find the absolute relative score of a node, the results need to be normalized such that sum over all the nodes is equal to 1.

2.3.2 Current Flow-based Centrality Metric

There are some real-world scale-free networks (e.g Power Networks, Water-flow Networks, Mechanical or Thermal Networks etc.), where conventional shortest (geodesic) path betweenness concept fails to extract the accurate dynamics and actual behaviour of the system as information can flow in such network in any direction efficiently through the various available paths. Due to this drawback, conventional centrality indices based on the pure topological concepts will not work well in power network as it disregards the real physical properties and the operative constraints of power grids. This is primarily because the flow of electricity in power network can pass through various available path governing the physical law of KCL and KVL, unlike the other commonly seen networks. The second differentiating factor between power networks and other commonly seen networks (such as, Biological networks, Transport road networks, Communication networks etc.) is that, in other network each vertex is used to either function as a source node or sink node though which some physical quantities are transmitted, but in power network we already designated nodes based on their unique functionality such as some nodes are called as generation nodes, and some are load nodes. And power flow can only take place from generation nodes to load nodes. Thus, influence of high flow degree centrality node is found to be most destructive to the network. These issues are now taken into the consideration where a flow-based centrality as a measure of betweenness is now introduced in to the picture with the two following steps.

2.3.2.1 Computing potential (v_i) of any node-i under a unit current injection at source node-s and unit current extraction at target node-t:

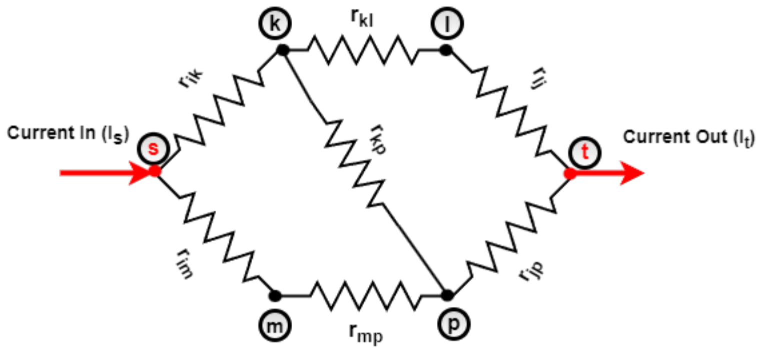


Figure 2.1: An example of electrical circuit

Let I_s and I_t is the external current injected at node s and extracted at node t as shown in Fig. 2.1. Let v_i and v_j is the potential of two random nodes i and j. Applying Kirchhoff's current law i.e summation of all incoming and outgoing currents at any particular node is equal to zero, which implies the following equation need to be satisfied.

$$\sum_{j=1}^{d_i} I_{ij} = \sum_{j=1}^{d_i} \frac{v_i - v_j}{r_{ij}} \tag{2.2}$$

where d_i is the degree of the node-i and $v_i - v_j$ is the potential difference between i and j. Applying Eq. (2.2) to every node of G yields

$$\mathbf{Q}\vec{V} = \vec{\Omega} \tag{2.3}$$

where $\mathbf{Q} \in \mathbb{R}^{N\times N}$ is the conductance matrix with elements $q_{ij} = 1/r_{ij}$. $\vec{V} \in \mathbb{R}^{N\times 1}$ is the voltage vector and $\vec{\Omega} \in \mathbb{R}^{N\times 1}$ is external injected/extracted current vector and it can be defined as follows.

$$\vec{\Omega} = \begin{cases} +I_s & \text{for } i = s, \\ -I_t & \text{for } i = t, \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

In Eq. (2.3), $\mathbf{Q}$ is termed as laplacian matrix and using Eq. (2.4), Eq. (2.3) can be rewritten as following matrix form:

$$\mathbf{Q} = \mathbf{B}_b \nabla \mathbf{B}_b^{\mathbf{T}} \tag{2.5}$$

where, $\mathbf{B}_b \in \mathbb{R}^{N\times M}$ is the Bus incidence matrix and $\nabla = (\phi_{ij}) \in \mathbb{R}^{M\times M}$ is a diagonal matrix with conductance specified to its graph edges.

$$\phi_{ij} = \begin{cases} q_{ij} & \text{for } i = j, \\ 0 & \text{for } i \neq j \end{cases} \tag{2.6}$$

Now, it is to be noted that we can't directly take the inverse of $\mathbf{Q}$, from Eq. (2.3), as this graph Laplacian matrix $\mathbf{Q}$ is singular. It basically indicates one of the equation among N^{th} number of algebraic equations is redundant. So, Kirchhoff's current conversion law is violated. To get rid of this problem, consider any one node of the test electrical circuit say, r to be as reference node and therefore, potential of this node is zero i.e $v_r = 0$. This assumption allows us to remove the r^{th} row and r^{th} column from the Laplacian matrix $\mathbf{Q}$ and thereby reduce the dimension of $\mathbf{Q}$ by one i.e $(N-1)\times(N-1)$. The newly obtained reduced matrix denoted as $\tilde{\mathbf{Q}} \in \mathbb{R}^{(N-1)\times(N-1)}$ is now a non-singular, invertible matrix. Later, to maintain the cardinality of set $\mathbb{V}$, the r^{th} row and r^{th} column of the matrix $\mathbf{Q}$ is reintroduced in matrix $\tilde{\mathbf{Q}}$ after inversion process is completed. The resultant matrix is denoted by $\breve{\mathbf{H}} = \breve{h}_{ij} \in \mathbb{R}^{N\times N}$.

$$\breve{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{Q}}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \tag{2.7}$$

This $\breve{\mathbf{H}}$ matrix of Eq. (2.7) is now going to be used for computation of voltage vector $\vec{V}$ using the relation below:

$$\vec{V} = \breve{\mathbf{H}}\vec{\Omega} \tag{2.8}$$

Now put the value of external current vector Ω into Eq. (2.8) yields

$$\vec{V} = I_s\vec{H}_s - I_t\vec{H}_t \tag{2.9}$$

where $\vec{H}_s$ and $\vec{H}_t$ are the s^{th} and t^{th} column vector of matrix $\breve{\mathbf{H}}$ respectively. As $I_s = I_t$, therefore Eq. (2.9) can rewritten as follows:

$$\vec{V} = I_s(\vec{H}_s - \vec{H}_t) \tag{2.10}$$

Thus, the voltage (v_i) of a node i of the graph $\mathbb{G}$ under an amount of current injection at source node-s and current extraction at target node-t can now be expressed as:

$$v_i^{st} = I_s(\breve{h}_{is} - \breve{h}_{it}) \tag{2.11}$$

2.3.2.2 Computation of current flow matrix ($\mathbf{I^{st}}$) over all possible source-target (st) pairs:

Using Eq. (2.12), one can compute potential drop across any node for all st pairs, i.e $\{s, t \in V\}$. Thus current-flow I_i^{st} over all st pairs can also be calculated as follows:

$$I_i^{st} = \frac{1}{2} \sum_{j=1}^{d_i} \frac{|v_i - v_j|}{r_{ij}} \tag{2.12}$$

Now replacing Eq. (2.11) in Eq. (2.12) and put $1/r_{ij} = q_{ij}$, then

$$I_i^{st} = \frac{I_s}{2} \sum_{j}^{N} q_{ij} \; |\breve{h}_{is} - \breve{h}_{it} - \breve{h}_{js} + \breve{h}_{jt}|, \quad \text{for } i \neq s, t \tag{2.13}$$

It is also to be noted that as the currents are specified in source node-s and target node-t, therefore, one can simply write

$$I_s^{st} = I_t^{st} = \begin{cases} I_s & \text{if } s, t \in |f^i_{s \to t}(j)|, \\ 0 & \text{if } s, t \notin |f^i_{s \to t}(j)| \end{cases} \tag{2.14}$$

where, $f^i_{s \to t}(j)$ denotes the current flow routes between node i and node j.

2.3.3 Proposed Hybrid Betweenness Centrality Metric

So, it is realized from the above two centrality indices that while eigenvector provides information about the presence of the most influential nodes from a structural perspective, current flow-based centrality, fundamentally based on the idea of inclusion of non-geodesic paths, aids in identifying the presence of such potential super-spreader nodes that involves those current flow paths and also the quantity of information that passes through those specific nodes. Thus, to identify those critical paths of the network through which optimal information flows while considering the weight of those super influential nodes, a novel Hybrid Betweenness Centrality is proposed in this work as a measure of centrality. Therefore, in simple words the HBC index will screen out the most vulnerable lines that the attacker might target to impose maximum structural damage to the system.

While defining HBC, it is assumed that the power can flow from generating nodes to load nodes through all the probable paths, and all the generating nodes (g) belong to the source set (s) and all the load nodes (d) belong to the target node set (t). Next, if $\acute{P}$ defines the total number of combinations of source-target (st) pairs, the first step of HBC formulation is to create the external current injected matrix ($\mathbf{I^{st}} \in \mathbb{R}^{N \times \acute{P}}$) over all ($st$) pairs using Eq. (2.13). Thereafter, the flow-energy of dominant node-i, (F_E^i) is calculated as,

$$F_E^i = C_E(v_i) \times \sigma_i^{\acute{P}} = C_E(v_i) \times \sum_{s \neq t \in \mathbb{V}} I_i^{st} \tag{2.15}$$

$$F_E^j = C_E(v_j) \times \sigma_j^{\acute{P}} = C_E(v_j) \times \sum_{s \neq t \in \mathbb{V}} I_j^{st} \tag{2.16}$$

where, $\sigma_i^{\acute{P}}$ denotes the aggregate sum of injection currents at node-i over all st pairs. Finally, the proposed HBC index is defined for each transmission line (l), associated with two end nodes, i.e, "From Node" (i) and "To Node" (j), as,

$$HBC(l_{i-j}) = \frac{F_E^i - F_E^j}{\sum_{i=1}^{N} \mathbf{1}_{[i \in g]} \times \sum_{j=1}^{N} \mathbf{1}_{[j \in d]}} \tag{2.17}$$

where, $\mathbf{1}_{[\Xi]} = 1$, if the condition $[\Xi]$ is true, otherwise $1_{[\Xi]} = 0$. Denominator of Eq. (2.17) is used for normalization process which represents the number of (st)-pairs considered for the calculation. The HBC index represents the relative drop of the flow-energy between two extreme nodes of each lines, and acts as a vulnerability indicator for a line. Once all the lines are ranked as per their HBC values, clearly, the top-ranked lines associated with maximum drop of flow energy are most likely to be chosen first by the attacker to launch an attack as it indicates the most influential path of information flow in a power system graph. The rest of the lines will thereafter be selected based on their vulnerability index value. The overall formulation steps of HBC in graph $\mathbb{G}$ using Eq. (2.17) for a unit current injection (i.e.,$I_s = 1$) is outlined in Algorithm 1.

2.4 Development of PMU Assisted Cyber-attack Resilient Framework

A literature survey reveals that PMUs are one of the best candidate devices to detect many unobservable attacks [59],[195]. Therefore, the developing of various placement algorithms for identifying strategic locations of PMUs to defend any kind of data integrity attack has now become a growing interest to the researcher community due to its higher synchronization rate, advanced security measure and impervious communication [196] and networking architecture [197]. All the measurements with greater accuracy collected by PMUs from different locations of a geographically dispersed area are correctly labelled with real-time stamps, which increases inherent robustness against attacks. Therefore, if the System Operator (SO) can get the complete system observability through phasor measurements alone, they can make an informed decision even in the event that any of the data on the vulnerable lines is compromised. Therefore, it becomes a natural step forward to include the information of vulnerable links while equipping a power system network with PMUs to provide it a maximum protective covering against cyber-physical attack. To achieve this goal, in this chapter, a novel objective function is formulated that incorporates the line serviceability index when the system is under attack, which is then utilized in a Mixed-integer Linear Programming model to guarantee complete topological observability.

Algorithm 1: Hybrid Betweenness Centrality

Input: A is the Adjacency Matrix of an Unweighted and Undirected Graph $\mathbb{G}(\mathbb{V}, \mathcal{E}, \mathcal{W})$.
$\mathcal{E} = \{e_{i,j} | v_i, v_j \in \mathbb{V}\}$, the Edge-list.
$\mathbf{Q} \in \mathbb{R}^{N\times N}$, be the graph laplacian matrix of G.

Output: Hybrid Betweenness Centrality, $HBC(l_{i-j})$

begin
- $H\tilde{B}C(l_{i-j}) \in \mathbb{R}^m \longleftarrow \mathbf{0}$;
- $C_E(v_i) \longleftarrow \emptyset$;
- $value \longleftarrow 0$;
- Solve for eigenvector $U \in \mathbb{R}^N$ corresponding to λ_{max}, by $|\mathbf{A} - \lambda\mathbf{I}| = 0$;
- **for** $i \leftarrow 1$ *to* m **do**
  - **for** $j \leftarrow 1$ *to* m **do**
    - **if** $a_{i,j} \neq 0$ **then**
      - value = value + $a_{i,j} * u(j)$
    - **end**
  - **end**
  - $C_E(v_i) = \frac{1}{\lambda_{max}}$*value;
  - value = 0
- **end**
- $\breve{\mathbf{H}} \in \mathbb{R}^{N\times N} \longleftrightarrow \tilde{\mathbf{Q}} \in \mathbb{R}^{(N-1)\times(N-1)}$;
- **for** $t \leftarrow d$ *to* $t \geq 1$ **do**
  - **for** $s \leftarrow g$ *to* $s \leq t$ **do**
    - **for** *each* $e_{ij} \in E$ **do**
      - **if** $i \neq s, t$ **then**
        - $I_i^{st} \leftarrow I_i^{st} + (1/2)|\breve{h}_{is} - \breve{h}_{it} - \breve{h}_{js} + \breve{h}_{jt}|$
      - **end**
    - **end**
  - **end**
- **end**
- **for** $i \leftarrow 1$ *to* m **do**
  - $\Delta F_E^l = (C_E(v_i) \times \sum_{s\neq t\in V} I_i^{st} - C_E(v_i) \times \sum_{s\neq t\in V} I_i^{st})$
  - $HBC(l_{i-j}) = \frac{\Delta F_E^l}{|s|\times|t|}$
- **end**

end

2.4.1 Conventional Optimal PMU Placement

In graph theoretic approach, a power system network equipped with PMUs are said to be topologically observable if all the nodes of the graph are directly or indirectly traversed by optimally placed PMUs. Such problem can be solved by integer Linear Programming (ILP) [198] by minimizing the conventional optimization formulation from cost minimization perspective as shown below:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) = \sum_{j=1}^{N} c_j x_j, \qquad\qquad x_j \in \mathbb{R}, \mathbb{Z} \\ \text{subject to} \quad & \bar{\mathbf{A}}\bar{X} \geq 1 \equiv (\mathbf{A_{adj}} + \mathbf{I})\bar{X} \geq 1 \\ & x_j = (0/1), \quad \textit{a binary variable} \end{aligned} \tag{2.18}$$

where,

- c_j is the cost of PMU placement at bus-j. In this study, PMU installation cost at all buses are assumed to be unity. It is also assumed that PMU has equipped with sufficient channels to measure voltage and current phasor at its installation bus and all its incidents lines respectively.
- x_j represents an element of the output vector $\bar{X}$, where a value of 1 indicates the presence of a PMU at the j^{th} bus location.
- $\mathbf{I}$ is the identity matrix of size N $\times$ N.
- $\bar{\mathbf{A}}$ is the binary connectivity matrix of size N $\times$ N with entries $\breve{a}_{ij}$ as:

$$\breve{a}_{ij} = \begin{cases} 1 & \text{if } i = j \text{ or, if i and j are connected,} \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

The above objective function is used to solved using conventional ILP for the values of x_j.

2.4.2 Proposed Modified Objective Function for ILP-based Attack Resilient PMU Placement

Minimization of the objective function in Eq. (2.18), results in multiple PMU locations with same cost and only valid when the system is in intact condition i.e free from any attacks. When the system is subjected to any attacks the lines are going to be removed as per the attacking strategy, then the system topology gets modified and therefore earlier ILP-based PMU placement technique will give incomplete observability of the systems. Moreover, as PMU deployment process is an offline procedure and is considered in power system planning stage only, therefore modification of constraints at each round of PMU placement is not feasible. Considering this problem as a motivation, in this chapter a novel objective function is formulated which will take care of this issues associated with HBC attacking strategy. As the attackers are always constrained with limited resources in terms of time and information to outline an attack, the proposed work assumes that it is sufficient to safeguard the system against FDI attack in top 20% of the most vulnerable lines in the system in order to enhance the system resilience. To this end, the strategic selection of vulnerable lines is carefully orchestrated by the HBC ranking based attack model in Section 2.2. Top 20% vulnerable lines are, thereafter, utilized in the proposed PMU placement problem as an input in line serviceability indicator ($\mathfrak{L}$) and branch-to-node incident matrix ($\boldsymbol{\mathfrak{B}}$), defined in Eq. (2.22) and Eq. (2.23) respectively. These '$\mathfrak{L}$' vector and '$\boldsymbol{\mathfrak{B}}$' matrix modifies as these critical lines are attacked, and go out of service. Thus, the resilient Optimal PMU Placement (OPP) problem is formulated as -

$$\begin{aligned} \underset{x}{\text{Minimize}} \quad & f(x) = \sum_{j=1}^{N} \xi_j \; x_j, \qquad\qquad x_j \in \mathbb{R}, \mathbb{Z} \\ \text{subject to} \quad & \bar{\mathbf{A}}\bar{X} \geq 1 \equiv (\mathbf{A_{adj}} + \mathbf{I})\bar{X} \geq 1 \\ & x_j = (0/1), \text{a binary variable} \end{aligned} \tag{2.20}$$

ξ_j in Eq. (2.20) is termed as the merit of far-ness which is analogous to cost function of PMU placements, and is defined as

$$\xi = \mathfrak{L}_{1\times M} \; \boldsymbol{\mathfrak{B}}_{M\times N} \; \boldsymbol{\mathfrak{D}}_{N\times N} \tag{2.21}$$

where,

- $\mathfrak{L}$ is the line serviceability indicator for N bus system as: $\mathfrak{L} = [a_1, a_2, a_3, \ldots, a_M]$, $M \in \mathbb{R}$. Where the entries $\mathfrak{L}$ as:

$$a_l = \begin{cases} 1 & \text{if line } l \text{ is in service,} \\ 0 & \text{if line } l \text{ is in out of service} \end{cases} \tag{2.22}$$

- $\boldsymbol{\mathfrak{B}}$ is the undirected branch-to-node incident matrix of size M $\times$ N with entries l_{i-j} as:

$$l_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are connected, } v_i, v_j \in V \text{ and } l_{i-j} \in E \\ 0 & \text{otherwise} \end{cases} \tag{2.23}$$

- The matrix $\boldsymbol{\mathfrak{D}}$ represents the graph distance, with dimensions N $\times$ N, containing the shortest path between all pairs of vertices in a network. In this context, considering the network as unweighted and undirected ensures that the distances are always positive. The values of the elements d_{ij} in matrix $\boldsymbol{\mathfrak{D}}$ are computed using established techniques such as the Floyd-Warshall algorithm. The distance metric examined in this study adheres to three fundamental axioms.
  - Main diagonal entries of $\boldsymbol{\mathfrak{D}}$ are set to 0 corresponding to v_i, i.e $d_{ii} = 0 \; \forall \; 1 \leq i \geq N$.
  - Off-diagonal entries of $\boldsymbol{\mathfrak{D}}$ are non-negative i.e $d_{ij} \geq 0$, if $v_i \neq v_j$. d_{ij} will be 0,

if there is no shortest route available between the node v_i and v_j.

- As the graph is undirected here, the matrix $\mathfrak{D}$ should be symmetric in nature i.e $d_{ij} = d_{ji}$ for $v_i, v_j \in V$.

The weight ξ_j assigns significance to each node in the network based on the average far-ness between all pairs of source and target nodes. The first term of the vector $\mathfrak{L}$ considers the influence of outage of lines according to their ranking in the HBC schemes. The matrix $\mathfrak{B}$ captures the effect of the network breaking into islands following an attack, indirectly reflecting the size of the system's giant component through branch and node connectivity. Furthermore, the evaluation of the distance matrix $\mathfrak{D}$ matrix involves searching for best optimal path by reducing the far-ness between pair of nodes when the system is subjected to an attack. At the same time, it also takes care of the total system topological observability in account with improved redundancy. Minimizing this objective function aims to maximize the system's full topological observability while identifying optimal locations for a minimal number of PMU placements in case of line outages due to attack.

As depicted in Fig. 2.2, the IEEE 5-bus system serves as an example. According to the proposed HBC ranking, Line 4-5 (L7) and Line 2-5 (L5) (highlighted in red) emerge as the top two critical lines. This vulnerable line information is then used in the proposed attack resilient OPP problem formulation as summarized in the Table 2.1. The results in Table 2.1 demonstrate that the proposed PMU placement approach effectively incorporates line vulnerability information from the HBC ranking into its objective function. It conducts a search for optimal PMU locations while not altering the existing system constraints. Clearly, if the bus-2 and bus-5 are equipped with PMUs, then even in the event of FDIA in the conventional measurements of these lines and lines being out, the system still remains observable with phasor measurements.

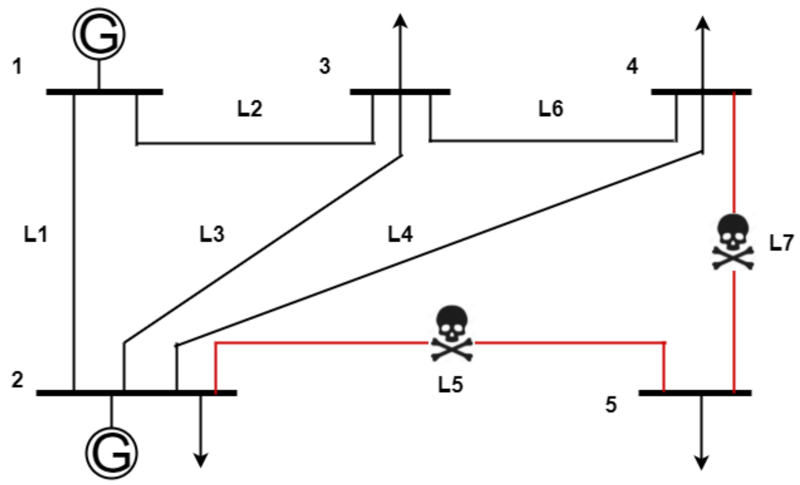


Figure 2.2: Single line diagram of IEEE 5-bus system identifying the critical lines

Table 2.1: Attack resilient OPP solution details in the IEEE 5-Bus network

Vulnerable Links	Line Serviceability Indicator ($\mathfrak{L}$)	Branch-to-Node Incident Matrix ($\mathfrak{B}$)	Distance Matrix ($\mathfrak{D}$)	Attack Resilient OPP Results
Line 4-5	$\mathfrak{L} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$	$\mathfrak{B} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathfrak{D} = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 0 & 2 \\ 2 & 1 & 2 & 2 & 0 \end{bmatrix}$	Bus-2
Line 4-5 and Line 2-5	$\mathfrak{L} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$	$\mathfrak{B} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathfrak{D} = \begin{bmatrix} 0 & 1 & 1 & 2 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	Bus-2 and Bus-5

2.5 Evaluating Cyber-Attack Resilience Using Secure PMU Measurements

After developing a PMU-equipped network in section 2.4, this section proposes a performance-based metric to assess the network's resilience against false data injection attack on AC state estimator [199]. It is presumed that (1) the resilience of a power system network to attacks is contingent upon several factors, including the number of PMUs, number of measurements, number of attacked buses, and attack magnitude (Ψ). (2) Because PMUs are extremely complex devices with cutting-edge security features, an attacker cannot access PMU measurements because of their encrypted, secure communication protocols [59].

In order to survive an attack, a system should have sufficient number of secured measurements so that following an attack on a vulnerable link, the attack is still detectable. The resiliency, in the proposed work, is quantified as the fraction of the secured phasor measurements obtained from the proposed PMU placement strategy in the total measurement vector of the state estimator, i.e.,

$$Critical\ Measurement\ Ratio\ Index = \frac{(Number\ of\ secured\ phasor\ measurements)}{(Total\ number\ of\ measurements)} \tag{2.24}$$

The higher the Critical Measurement Ratio Index (CMRI) is the more cyber-attack resilient system is.

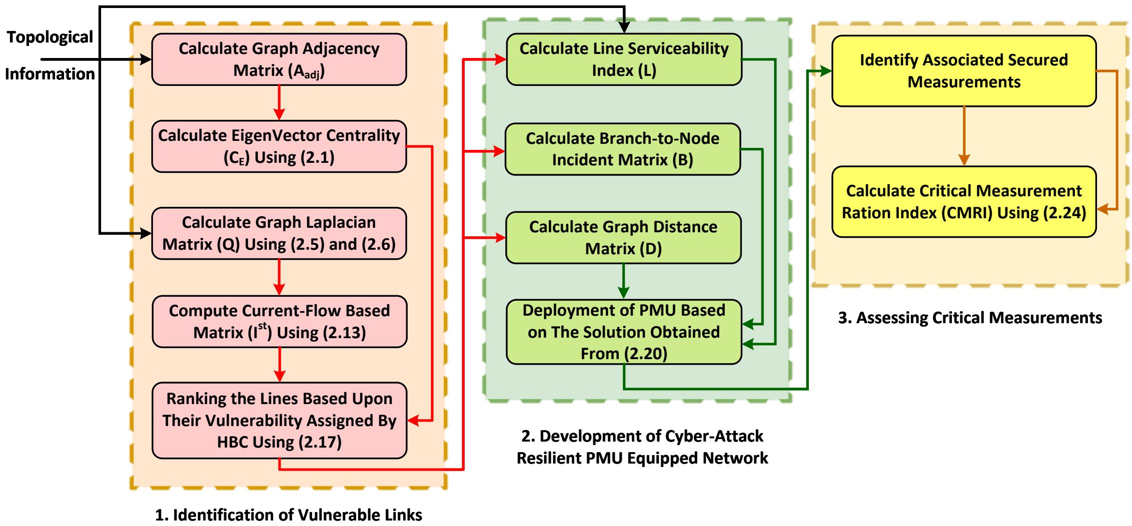


Figure 2.3: Flowchart of the proposed PMU-assisted cyber-attack resilient framework

The flowchart in Fig. 2.3 illustrates the entire proposed framework for enhancing cyberattack resilience using PMU measurements. This framework comprises three main stages viz., 1) Identification of vulnerable links in the network, 2) Developing a cyber attack-resilient network with PMU integration, and 3) Evaluating cyber-attack resiliency using the proposed CMRI. These stages are clearly delineated in the flowchart of Fig. 2.3. In *Stage 1*, the topological properties of the power system is analyzed to device a new attacking strategy and also estimate the impact of line outages due to the proposed attack over the topological breakdown of the studied power networks. Initially, the graph's connectivity is determined by computing the adjacency matrix using network topological parameters. Subsequently, a HBC attack strategy is formulated based on a combination of (i) eigenvector centrality and (ii) current flow-based matrix information. The former takes the contribution of influential nodes based on their adjacent nodes' importance, from purely topological perspective while the latter considers current flow information into account to bring-in the dynamic behaviour of the system too. Finally, all system lines are ranked according to their HBC values to assess their vulnerability to cyber-attacks.

Given the attackers' limited resources in terms of time and information, the proposed approach focuses on safeguarding the system against cyberattacks by targeting the top 20% of the most vulnerable lines. To this end, the information of the top 20% vulnerable lines, obtained from *Stage 1*, is utilized in the proposed PMU placement problem of Stage 2 as an input in line serviceability indicator ($\mathfrak{L}$) and branch-to-node incident matrix ($\mathfrak{B}$) of the Eq. (2.20) as detailed in Section-2.4. The '$\mathfrak{L}$' vector and '$\mathfrak{B}$' matrix modifies as these vulnerable lines are attacked, and goes out of service. The attack-resilient optimal PMU placement conducted in *Stage 2* ensures system observability even in the event of a data integrity attack, such as a FDIA, on these vulnerable lines. The outcome of this attack resilient OPP provides twofold benefits- i) the state estimation results become more accurate and reliable, and ii) due to the secure communication infrastructure of PMUs, a set of safe and secured measurements is obtained which is difficult to manipulate by the adversary. *Stage 3* validates these observations through the proposed Resiliency Index, which assesses the system's resilience from the perspective of obtained secured measurements among total number of available measurements to run reliable state estimation followed by the successful detection of data injection attacks based on largest normalized residual (LNR) testing.

2.6 Results and Discussion

The test results of proposed PMU-assisted cyber-attack resilient framework is presented on the standard IEEE 14-bus and the New England 39-bus test systems. The simulation work was performed in MATLAB R2018a, loaded on a laboratory PC (an HP ProDesk 600 G4 SFF) with specifications including a 64-bit Windows-10 operating system, an Intel(R) Core(TM) i7-8700 CPU running at 3.20 GHz, and 16 GB of RAM.

To evaluate the qualitative measures of the vulnerability assessment on the power grid subjected to attack, a well-known popular performance parameter named as "Giant-Component Size (S)", is used in the Chapter. This metric quantifies both the topological and operation characteristics of the power grid. In the case, removal of lines strategically causes the network to be partitioned into several components, which are basically the disjointed version of original graph. The graph statistics, giant component mechanism seeks for the largest connected components of the graph that contains maximal fraction of nodes of the parental graph's nodes. It is usually calculated by the ratio of current giant size of the network after attack to the initial network size N. Mathematically its is expressed as shown below:

$$S^l(\%) = \frac{\sum_{i=1}^{N} \mathbf{1}_{i \in G_H^l}}{N} \times 100 \tag{2.25}$$

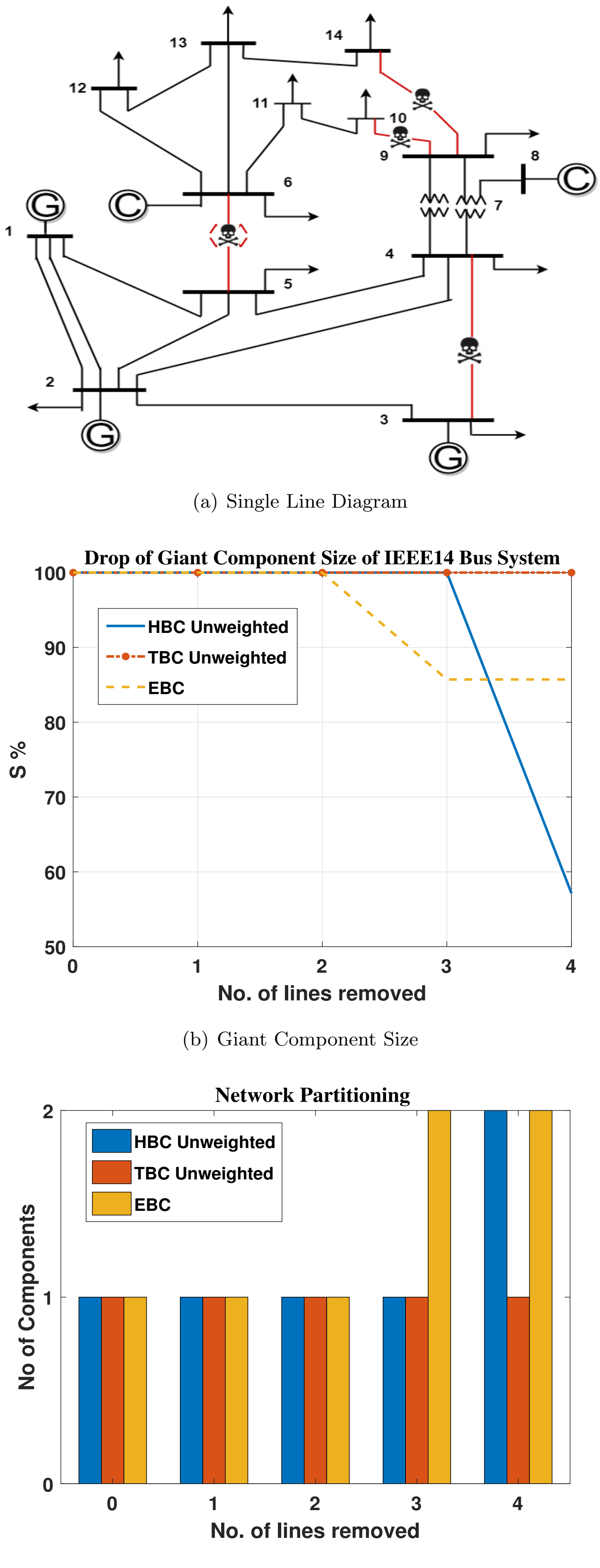
where,$\mathbf{1}_{[\Xi]}$ is an indicant function as explained in Section 2.3. G_H^l is the current giant component of the initial graph after removing of line-l. After calculating the proposed centrality value, giant component is evaluated after every attack to show how it cause structural damage to the network. A steeper fall of giant-component metric, signifies substantial damage of the grid with higher degree.

2.6.1 IEEE 14-bus Test Systems

2.6.1.1 Vulnerable links identification

The IEEE 14-bus system consists of 5 generators and 20 transmission lines. The top-4 vulnerable lines identified as per the proposed HBC based method are tabulated in Table 2.2 and also being shown in Fig. 2.4(a). Figure 2.4(b) depicts the drop of giant component size of the system when the attacks lines are physically out from the system in the subsequent order of the attack strategy.

The test result is also compared with two other conventional attack approaches i.e



(a) Single Line Diagram

(b) Giant Component Size

(c) No. of Components

Figure 2.4: Effectiveness of HBC on IEEE 14-bus test system

Table 2.2: Ranking of Lines of IEEE 14-Bus System

Hybrid Betweenness Centrality		
Line No	Line Info	Value
16	Line 9-10	1.0
6	Line 4-3	0.9304
17	Line 9-14	0.7809
10	Line 5-6	0.4346

Topological Betweenness Centrality (TBC) [200] and Electrical Betweenness Centrality (EBC) [56]. Attacks on vulnerable links based on TBC ranking do not result in the creation of islands in the system, as shown in Fig. 2.4(b) where the Giant Component Size stays at 100%. In contrast to EBC, which was 87%, the proposed HBC-based attacks, however, caused the largest reduction in the system's Giant component size, reaching 57% following the simultaneous attack on all four of the most vulnerable links. This demonstrates the supremacy of the HBC-based attack technique above all others pertaining to severe structural fragmentation and substantial damage of the systems. Therefore, HBC-based attack has more potential to create cascades failures which may lead to severe blackout.

2.6.1.2 Cyber-attack resilient PMU placement

The vulnerable lines, as identified in Table 2.2, are incorporated in PMU placement strategy of Eq. (2.20) and Eq. (2.21). Table 2.3 lists the optimal number of PMUs that are resistant to cyber-attacks. The results are also compared with the locations that arise when the typical unity cost function is used in place of the objective function in formulation Eq. (2.20), and vulnerable lines are eliminated from the system by altering the observability requirements in accordance with [201]. Although both the methods results in same number of PMU locations, as listed in Table 2.3. However, incorporating ξ_j in Eq. (2.21) as the merit of far-ness has increased the resiliency in the system as is demonstrated in the next subsection.

Table 2.3: Optimal PMU Locations For IEEE 14-Bus System Under Attack condition

Attack Resilient	Normal
4 at Bus-2, 7, 11, 13	4 at Bus-2, 7, 10, 13

2.6.1.3 Cyber-security resiliency assessment

To evaluate the impact of the proposed PMU placement strategy on system resilience, various test cases are conducted with varying levels of attack intensity and target bus locations for attack. The attack vector is modelled as a false data injection attack as described in [202] targeting non-linear state estimations running in the control centre. Across all test cases, a total of 79 measurements are available for SE algorithm execution. In Fig. 2.4(a), the IEEE 14-bus system is depicted, highlighting the placement of PMUs at four distinct locations: Bus-2, 7, 11, and 13, as determined by the proposed resilient PMU placement approach. With each PMU installed at a designated $bus - i$, the associated state variables V_i, θ_i and related power injections and flow measurements for that $bus - i$ are assumed to be secured (tamper-proof). Thus, Table 2.4 details all the secured measurements, resulted from the proposed resilient PMU placement. It can be observed from Table 2.4 that the total number of secured phasor measurements are found to be 36 in the IEEE 14-bus test system. These secured measurements are subsequently utilized in the Weighted Least Square AC state estimator, alongside conventional SCADA measurements, to ensure reliable state estimation and to detect any attacks tampering conventional measurements. Consequently, based on Eq. (2.24) detailed in Section 2.5 and the identified PMU locations from Section 2.4, the RI for this test system is calculated to be 45.56%. To illustrate the impact of a FDIA on state estimation performance, two

Table 2.4: Summary of secured measurements of the IEEE 14-bus system resulted from the attack-resilient PMU placement

IEEE 14-Bus System	
PMU Bus No	Secured Measurements
2	$V_2,\ P_2,\ Q_2,\ P_{1-2},\ P_{2-5},\ P_{2-4},\ P_{2-3},\ Q_{1-2},\ Q_{2-5},\ Q_{2-4},\ Q_{2-3}$
7	$V_7,\ P_7,\ Q_7,\ P_{4-7},\ P_{7-8},\ P_{7-9},\ Q_{4-7},\ Q_{7-8},\ Q_{7-9}$
11	$V_{11},\ P_{11},\ Q_{11},\ P_{6-11},\ P_{10-11},\ Q_{6-11},\ Q_{10-11}$
13	$V_{13},\ P_{13},\ Q_{13},\ P_{6-13},\ P_{12-13},\ P_{13-14},\ Q_{6-13},\ Q_{12-13},\ Q_{13-14}$
Total Secured Measurements = 36	

scenarios are considered. In Case-1), only conventional SCADA measurements are utilized in the state estimation process, while in Case-2), secured phasor measurements are also included. It is assumed that the attacker has access to multiple measurements, enabling them to inject false data ($\acute{a}$) into the original measurements set, (Z) via vulnerable SCADA communication channels. The attacker can thus compromise the measurement set as, $Z_a = Z + \acute{a}$, and can, thereby, deviate the original estimates of system states ($\hat{x}$) to some arbitrary bad states ($\hat{x}_{bad} = \hat{x} + \Psi$). This injection vector is generated based on the AC power flow model as $\acute{a} = \mathbf{H}(\hat{x}_{bad}) - \mathbf{H}(\hat{x})$ to circumvent conventional bad data detection mechanisms [202] such that $LNR_{bad} = LNR$, where $LNR_{bad} = ||Z_a - \mathbf{H}(\hat{x}_{bad})||$ and $LNR = ||Z - \mathbf{H}(\hat{x})||$. H is the Jacobian matrix relating available measurements with the state vector.

Case-1) When meters are not secured i.e., non-inclusion of PMU in the system: The effect of the FDI attack on the bus voltage magnitude of Bus-3, Bus-5, and Bus-9 is depicted in Fig. 2.5. FDIA targeted a fraction of attack buses, introduced bias to the magnitude of these attacked state variables, which causes incorrect estimation.

Based on the verification of residual checking, this attack remained undetected as shown in Fig. 2.6. This is because of the continuous adjustment of the associated

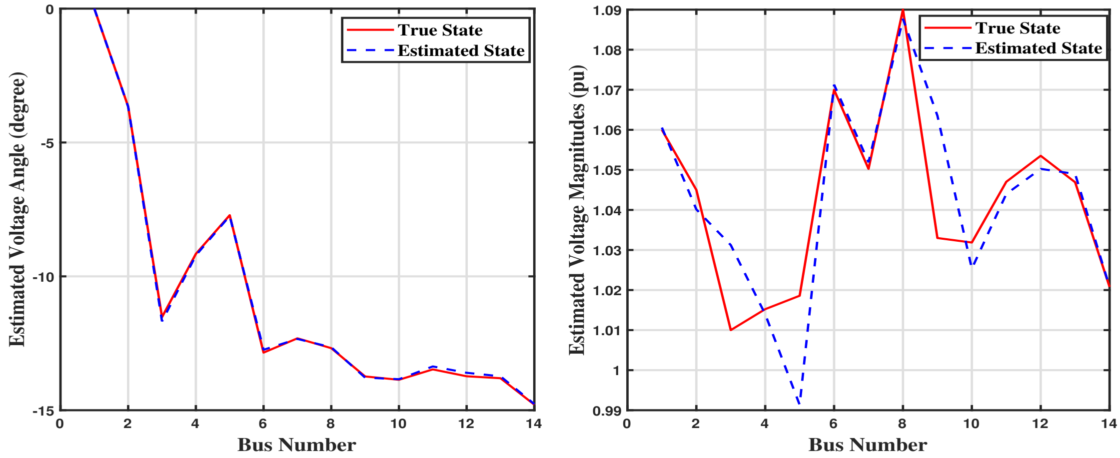


Figure 2.5: FDI attack on Bus-3, Bus-5 and Bus-9 with attack intensity (Ψ) of 0.03 pu

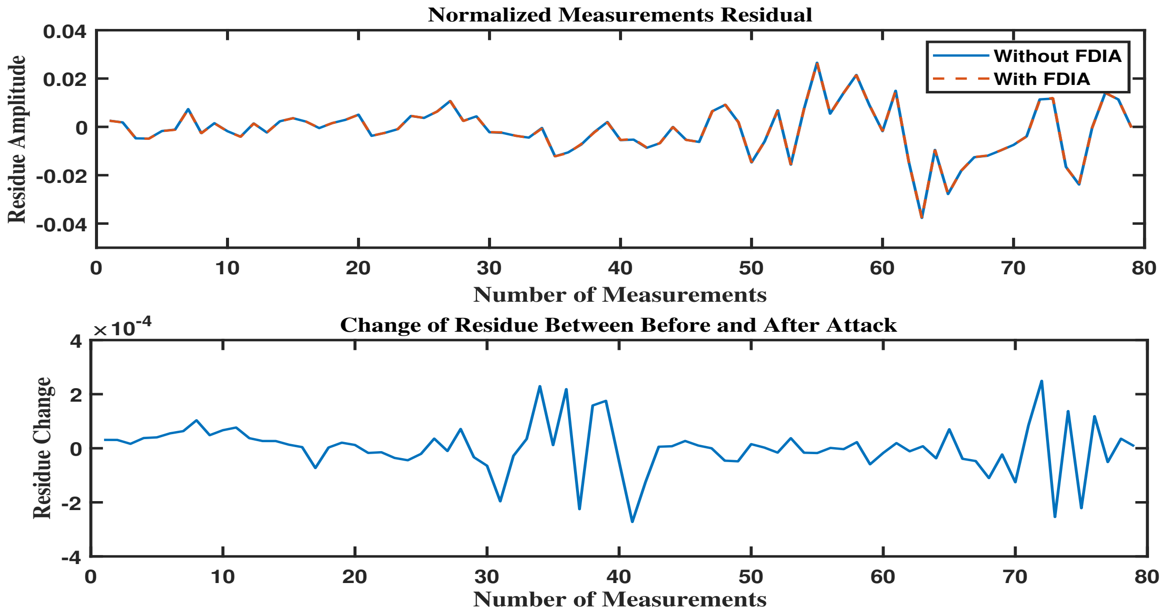


Figure 2.6: Performance of residue detector in presence of FDIA (Case-1)

measurements of those attack state variables. It is evident that the residue remained almost same for both type of cases i.e., with FDIA and without FDIA. The residual plot's threshold, as displayed below, is determined by using a chi-square distribution with a 95% confidence interval. Thus, the weighted sum of squares and L2 norm, or LNR, stayed below threshold to allow the attack to remain stealthy even in the presence of FDIA.

Case-2) When measurements are secured by optimal PMUs deployment: The performance of residual detector, in this case when the selected meters are inherently secured by the direct supervision of PMU is shown in Fig. 2.7. It can be observed that through the proposed PMU deployment process and full rank nature of the phasor measurement Jacobian matrix, most of the attacks get eliminated because their injection is hindered in the first place itself.

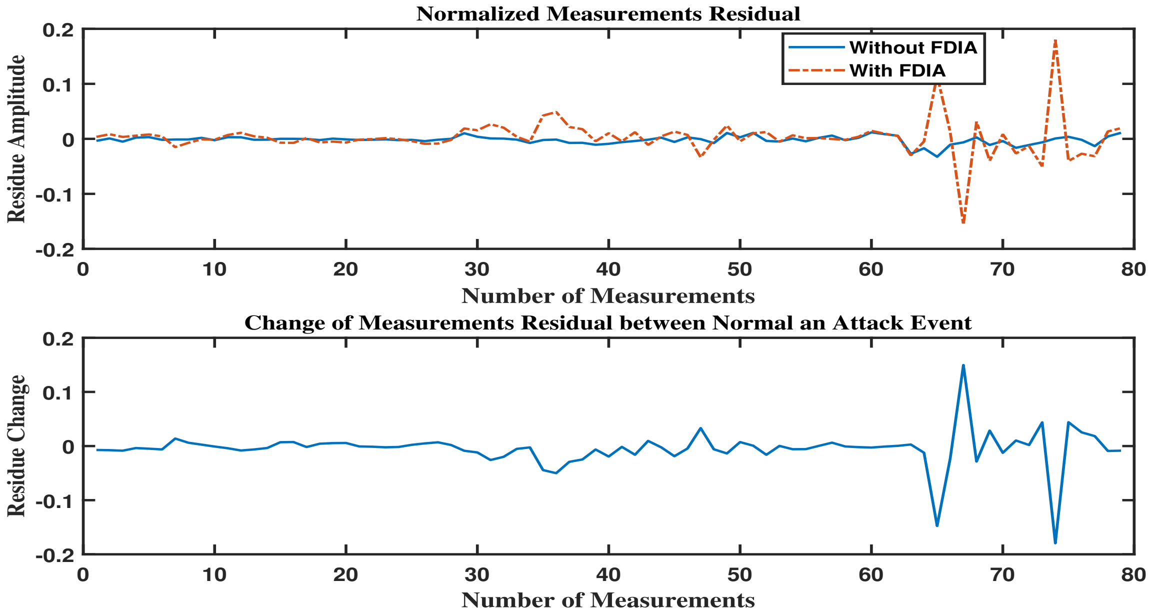


Figure 2.7: Performance of residue detector in presence of FDIA (Case-2)

2.6.2 New England 39-bus Test Systems

2.6.2.1 Vulnerable links identification

The New England 39-bus test system consists of 10 generators and 46 transmission lines are shown in Fig. 2.8(a). Table 2.5 shows the top-6 vulnerable lines identified by the proposed HBC-based attack strategy, and Fig. 2.8(b) depicts the drop of giant component size of the system when the attack lines are physically out from the system in subsequent order of this attack strategy. Here, the TBC- and EBC-based attack mechanisms demonstrate that, according to their respective techniques, the system giant component size reduction begins when one line goes out; nevertheless, the HBC-based attack mechanism experiences the same thing for a consecutive two lines outage. As shown in Fig. 2.8(b), the TBC-based attack method reduces the giant size to 69% after three consecutive line outages, and then it also stays constant for subsequent outages of other lines. In contrast, the EBC-based attack method drops the largest system size from 100% to 71% after one line outage, and then it remains almost constant for other remaining line outages. It is important to note that, despite its initial slight delay in system fragmentation compared to TBC and EBC, the HBC-based attack causes the network to be more torment as giant component size is drastically reduced below to 50%. Figure 2.8(c) explains that when top two lines are went out, the HBC-based method did not create any partition on the network, unlike the TBC and EBC-based attack mechanism had created same number of system components. However, after that the number of system partitions due to HBC keeps increasing gradually

Table 2.5: Ranking of Lines of IEEE 39-Bus System

Hybrid Betweenness Centrality		
Line No	Line Info	Value
4	Line 2-25	1.0
20	Line 15-16	0.7279
45	Line 29-38	0.7056
43	Line 25-37	0.7030
6	Line 3-18	0.6079
22	Line 16-19	0.5977

with increasing vulnerability measures, indicating more severe damage to the system. Another noticeable fact observed in EBC-based attack is that although the number of system segments keep increasing when line outage increases from three to five, the giant size has not shown any changes as shown by the HBC results. This is due to the fact that the nature of selection of vulnerable links for EBC is not so optimal and thus, in such period, it may pick up some unimportant links from some already created smaller partition, which basically has no contribution to the reduction of largest system component. That is why TBC and EBC depicts flat profile after certain line outages. So, it is hereby concluded that by developing such hybrid a betweenness centrality measure, one can anticipate a better understanding of the importance of nodes and lines of a specific complex network which assist operator in exploring the vulnerabilities more efficiently.

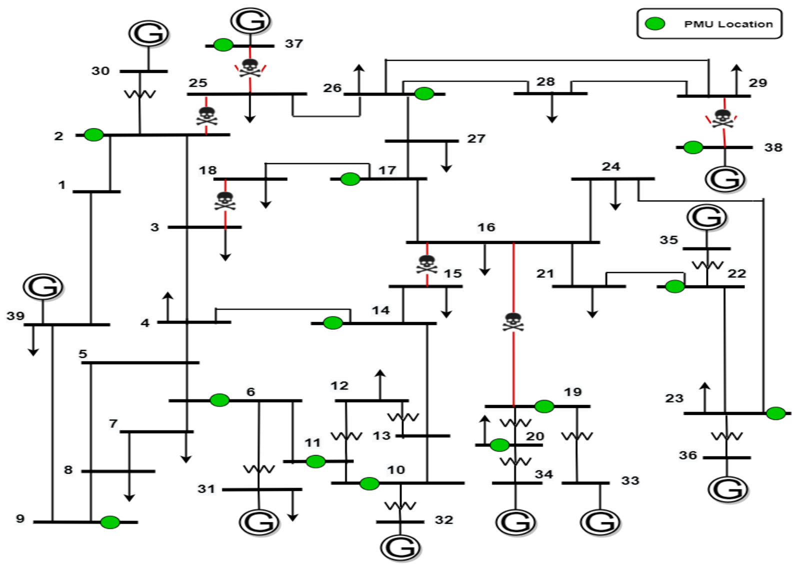
2.6.2.2 Cyber-attack resilient PMU placement

The top six vulnerable lines of the New-England Power System, as identified in Table 2.5, are now incorporated in proposed PMU placement scheme of Eq. (2.20) and Eq. (2.21). Table 2.6 shows the final number and locations of PMUs for both, the proposed placement scheme as well as for the normal PMU placement case. Once more, it is seen that although

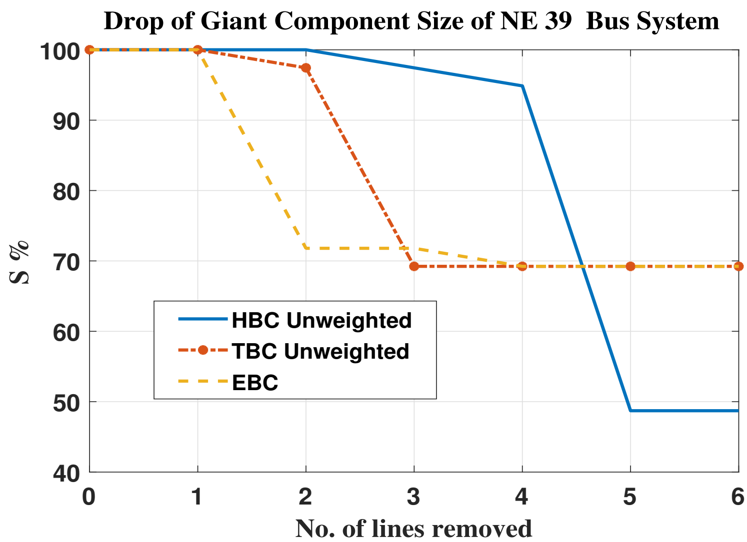
Table 2.6: Optimal PMU Locations For NE 39-Bus System Under Attack condition

Attack Resilient	Normal
14 at Bus-2, 6, 9, 10, 11, 14, 17, 19, 20, 22, 23, 26, 37, 38	14 at Bus-2, 6, 9, 12, 14, 17, 19, 20, 22, 23, 29, 32, 37, 38

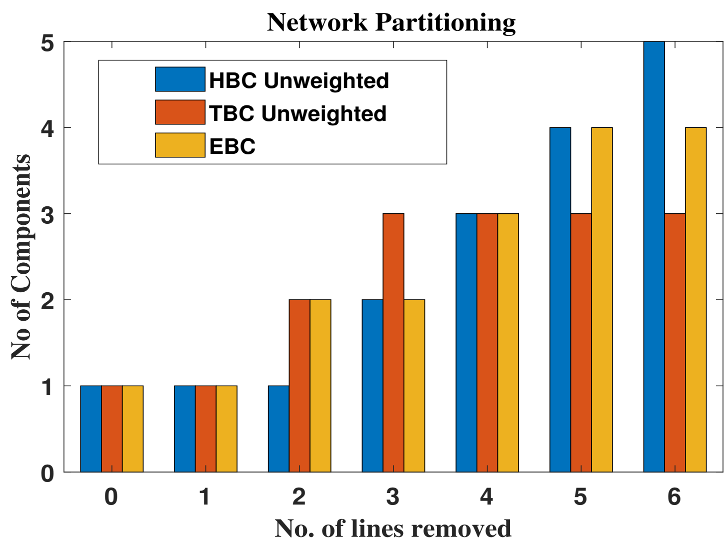
the number of PMU locations that are optimal for both the cyber-resilient and the normal procedures stays the same, the locations that are produced differ. The objective function in Eq. (2.21) traces the reduction of giant component at each step of attack execution strategy and based on that, the defense philosophy searches for new optimal locations of PMU based on available algebraic connectivity of the graph with shortest route features. It should be noted that at the end of the attack, the proposed PMU deployment method with some different PMU locations enhances the system resiliency while maintaining the full topological observability of the system throughout.



(a) Single Line Diagram



(b) Giant Component Size



(c) No. of Components

Figure 2.8: Effectiveness of HBC-based attacking strategy on NE 39-bus system

2.6.2.3 Cyber-security resiliency assessment

Similar to the FDIA test cases performed in the IEEE-14 bus system, same test cases are also carried out in NE 39-bus system. The total number of meters that are available in this system is 195. Figure 2.8(a) shows that the proposed OPP formulation strategy allocates fourteen specific bus locations dispersed geographically for the installation of PMUs. Due to the installations of PMUs at the those locations, a subset of measurements pertaining to PMU enabled bus injection meters and flow meters incident to the PMU equipped buses are being treated as secured and it is tabulated in Table 2.7. The table shows that

Table 2.7: Summary of secured measurements of NE 39-bus resulted from the attack-resilient PMU placement

NE 39-Bus System	
PMU Bus No	Secured Measurements
2	$V_2,\ P_2,\ Q_2,\ P_{1-2},\ P_{2-3},\ P_{2-4},\ P_{2-25},\ Q_{2-30},\ Q_{1-2},\ Q_{2-3},\ Q_{2-4},\ Q_{2-25},\ Q_{2-30}$
6	$V_6,\ P_6,\ Q_6,\ P_{5-6},\ P_{6-7},\ P_{6-11},\ P_{6-31},\ Q_{5-6},\ Q_{6-7},\ Q_{6-11},\ Q_{6-31}$
9	$V_9,\ P_9,\ Q_9,\ P_{8-9},\ P_{9-39},\ Q_{8-9},\ Q_{9-39}$
10	$V_{10},\ P_{10},\ Q_{10},\ P_{10-11},\ P_{10-13},\ P_{10-32},\ Q_{11-10},\ Q_{10-13},\ Q_{10-32}$
11	$V_{11},\ P_{11},\ Q_{11},\ P_{11-12},\ Q_{11-12}$
14	$V_{14},\ P_{14},\ Q_{14},\ P_{13-14},\ P_{4-14},\ P_{14-15},\ Q_{13-14},\ Q_{4-14},\ Q_{14-15}$
17	$V_{17},\ P_{17},\ Q_{17},\ P_{16-17},\ P_{17-18},\ P_{17-27},\ Q_{16-17},\ Q_{17-18},\ Q_{17-27}$
19	$V_{19},\ P_{19},\ Q_{19},\ P_{16-19},\ P_{19-20},\ P_{19-33},\ Q_{16-19},\ Q_{19-20},\ Q_{19-33}$
20	$V_{20},\ P_{20},\ Q_{20},\ P_{20-34},\ Q_{20-34}$
22	$V_{22},\ P_{22},\ Q_{22},\ P_{21-22},\ P_{22-23},\ P_{22-35},\ Q_{21-22},\ Q_{22-23},\ Q_{22-35}$
23	$V_{23},\ P_{23},\ Q_{23},\ P_{23-24},\ P_{23-36},\ Q_{23-24},\ Q_{23-36}$
26	$V_{26},\ P_{26},\ Q_{26},\ P_{25-26},\ P_{26-27},\ P_{26-28},\ P_{26-29},\ Q_{25-26},\ Q_{26-27},\ Q_{26-28},\ Q_{26-29}$
37	$V_{37},\ P_{37},\ Q_{37},\ P_{25-37},\ Q_{25-37}$
38	$V_{38},\ P_{38},\ Q_{38},\ P_{29-38},\ Q_{29-38}$
Total Secured Measurements = 112	

there have been total 112 secured phasor measurements found for this test case. It can be observed that through the proposed PMU deployment process and full rank nature of the phasor measurement Jacobian matrix, most of the attacks get eliminated because their injection is hindered in the first place itself due to presence of 112 secured measurements. Thus, calculated value of the proposed RI for this test case is as 57%.

2.7 Conclusions

In addition to natural random failures, modern power systems face increasing threats from malicious attacks, where adversaries aim to target influential nodes or critical lines to disrupt the system's functionality. To mitigate the risks posed by cyber-attacks, this chapter proposes a novel PMU-assisted framework designed to enhance the resilience of power systems. The primary goal of this framework is to strategically deploy a minimal number of PMUs to ensure system observability, even in the event of a data integrity attack such as FDIA in some top vulnerable lines. The proposed scheme is developed in three stages viz., (1) Identification of vulnerable links in the network, (2) Development of cyber-attack resilient PMU equipped network, and (3) Assessing the cyber-attack resiliency via proposed Resiliency Index.

- The proposed hybrid between-ness centrality index is found to be proficient attack strategy to identify group of transmission lines whose sequential outages may severely affect the system performance due to major structural breakdown.
- To prioritize the full system topological observability with higher resiliency in the presence of HBC-based attack, the novel development of PMU deployment framework works well and also defend the system from any data integrity types of attacks.
- The newly introduced resiliency index provides a quantitative measure of obtaining resiliency limit for the system operator to defend the system against typical ranges of stealth attacks in terms of meters found to be inherently secured.
- The proposed PMU placement strategy results in 46% and 57% secured measurements in the IEEE 14-bus and NE 39-bus system, respectively.

Chapter 3

A Novel Replay Attack Detection and Mitigation Framework for State Estimation

3.1 Introduction

In Chapter-2, a cyber-attack-immune metering framework was developed, wherein a vulnerability assessment in the transmission-level networks was carried out first. Thereafter, the system was made completely observable with minimum number of temper-proof PMUs to ensure system observability. Nonetheless, the attackers can still falsify the legacy SCADA measurements received from the Remote Terminal Unit (RTUs) via Replay Attacks (RAs), in order to hamper the critical state estimation application. This chapter, therefore, tries to develop a simple yet effective replay attack detection and mitigation framework in order to make Power System State Estimation (PSSE) application more resilient. To this end,

1. From attacker's viewpoint, Stage-1 of the proposed scheme identifies a minimal set of SCADA measurements from power flow, power injection, and voltage sensors which, if compromised, would result maximum error in the State Estimation (SE) results.

2. Stage-2 models different RAs which the attacker may inject into the measurement vector of the SE application.

3. Finally, Stage-3 presents a simple yet effective scheme for the detection and correction of RAs.

The rest of the chapter is organized as follows: In order to identify the highly sensitive active power flows, voltages, and injection sensor data that are vulnerable to RAs and could seriously disrupt PSSE estimates, a unique scheme is presented in Section 3.2. Section 3.3 represents the modeling of various RAs based on recording and replaying attack techniques with those vulnerable measurements. Section 3.4 depicts the flowchart for the SE, based on the mixed vulnerable SCADA and partial synchrophasors measurements, coherently detecting and correcting the RAs, if any. Section 3.5 details the validation results of the proposed algorithm on the two standard IEEE test systems i.e the IEEE-14 and IEEE 39-bus system, modelled in the Real-Time Digital Simulator (RTDS). Finally, Section 3.6

summarizes the major findings made from the work as is also illustrated in Fig. 3.1 below.

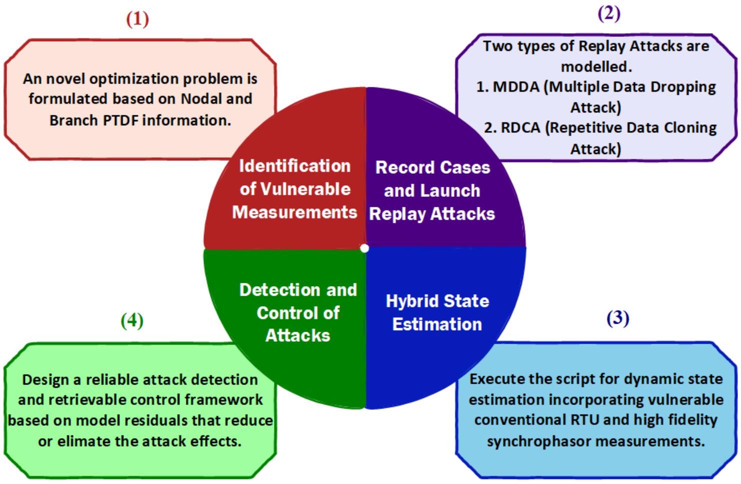


Figure 3.1: Distribution of contributions: A Quadrant Perspective Visualization

3.2 Stage-1: Identification of the Vulnerable SCADA Measurements

This is an offline stage, wherein a minimal number of power flows, voltages and injection sensors are identified, from an attacker's perspective, which can be compromised by RAs resulting serious disruption in performance as well as obtaining actual operational states of the power system. In pursuit of it, at the core of this scheme, a very commonly used factor known as Power Transfer Distribution Factor (PTDF) is exploited, which is usually defined as the incremental change in active flows of the line caused by the incremental change in active power injection at some nodes of the system [203]. Firstly, for finding minimum number of active power flow meters, an optimization problem is formed based on Branch PTDF (BR-PTDF) matrix which analyzes sensitivities of each lines towards change in 1 p.u. power injection and withdrawal at both the edge bus of each available transmission lines of the system, respectively. Finally, the top few sensitive lines are extracted, potentially resulting in significant vulnerabilities such as line overloads, load shedding, cascading failures, and blackouts if they are externally tampered with. Similarly, for minimum active power injection meters, a criteria is defined based on Nodal PTDF (N-PTDF) matrix information. The detailed descriptions of both the proposed schemes are elucidated as follows:

3.2.1 Selecting Critical Active and Reactive Power Flow Meters using Branch Power Transfer Distribution Factor (BR-PTDF)

BR-PTDF, defined as a factor that gives the fraction of power that is sent into the network at bus-s (source bus) to the bus-r (receiving bus) which flows over line l from bus-i to bus-j can be expressed as follows:

$$PTDF_{s\to r,ij} = \frac{1}{x_{ij}}\Big[(\overset{*}{X}_{is} - \overset{*}{X}_{ir}) - (\overset{*}{X}_{js} - \overset{*}{X}_{jr})\Big] \tag{3.1}$$

where, x_{ij} is the l^{th} line reactance, connecting bus-i and bus-j and $\overset{*}{X}_{ik}$ denotes the elements of bus reactance matrix $\overset{*}{\mathbf{X}}$ of size $(N \times N)$ pertaining to i^{th} row and k^{th} column. Alternatively, in matrix notation, the same can also be written as:

$$\mathbf{PTDF} = \mathbf{B_x} \times \overset{*}{\mathbf{A}} \times \widetilde{\mathbf{X}} \times (\widetilde{\mathbf{A}})^T \tag{3.2}$$

where, $\mathbf{B_x}$ and $\overset{*}{\mathbf{A}}$ are the branch network suspentance matrix of order $M \times M$ and line incidence matrix of order $M \times N$ respectively. $\widetilde{\mathbf{A}}$ is the modified version of $\overset{*}{\mathbf{A}}$ by setting the entry of the corresponding radial bus for any radial lines to be zero. Similarly, $\widetilde{\mathbf{X}}$ is also be modified to $\overset{*}{\mathbf{X}}$ by setting zero entries in the row corresponding to the reference bus and also for the rows corresponding to radial bus except its diagonal element to be placed as one.

At the end, each rows of the computed **PTDF** matrix denotes the line flow pairs $i \leftrightarrow j$ and columns denotes the injected/withdrawal pairs $s \leftrightarrow r$. This $s \leftrightarrow r$ can actually be any bus pairs of the network but for the simplicity, in this chapter, the chosen $s \leftrightarrow r$ pairs can be considered as only the end pairs of each lines. Now with the aim of maximizing the impact of RAs in terms of having highest flow changes in each line pairs by considering the cumulative effect of all $s \leftrightarrow r$ pairs into the account and also simultaneously search for the minimum number of meters of those highest sensible lines, the following novel optimization function with the linear constraint are formulated which is solved using usual Integer Linear Programming Model as shown below.

$$\begin{aligned}
\underset{x}{\text{Minimize}} \quad & f(x) = \sum_{j=1}^{n_l} \psi_j^T x_j \qquad x_j \in \mathbb{R}, \mathbb{Z} \\
\text{subject to} \quad & \Theta^T X \leq 1 \\
& x_j = (0/1), \text{a binary variable} \\
\text{where,} \quad & \psi(i) = \frac{1}{n_l} \times \sum_{j=1}^{n_l} PTDF(i,j) \\
& \Theta(j) = \sum_{j=1}^{n_l} PTDF(j,i)
\end{aligned} \tag{3.3}$$

Equation 3.3 introduces the binary variable x_j which serves as a selector index for active and reactive power flow sensors or measurements targeted for the attack. Specifically, when the decision variable, $x_j = 1$, it signifies that the i^{th} sensor is targeted and compromised, while a value of $x_j = 0$, signifies that the i^{th} sensor readings are plausible. The constraint outlined in Eq. (3.3) plays a crucial role in establishing a lower limit on the number of line flow sensors chosen for tampering by potential replay attackers. In order to obtained minimal sensors output, the design criteria is formed such that while 1 p.u power is injected to any one source bus "$\acute{S}$" and extracted from any one withdrawal bus "$\acute{R}$", the cumulative effect of all lines flows for that transaction are bounded by 1 p.u. Resulting it constitutes a right balance between selection of sensors for tampering while maximizing the overall vulnerability index to unity.

3.2.2 Selecting Active Power Injections and Voltage Meters using Nodal Power Transfer Distribution Factor (N-PTDF)

The N-PTDF matrix characterizes the influence of power injections at each node on a given individual line. In this matrix, there is a designated reference node i.e the slack bus through which all power transactions i.e withdrawal takes place between each injection node. By assuming a reference node (slack bus) available in the network, the N-PTDF are limited to only nodal bus injections instead of all combinations of transactions between each pair of busses. Thus, if m is designated as slack bus for corresponding power withdrawal and n be the bus where 1 p.u power is injected, in such case the distribution of power change on each line can be calculated from N-PTDF as follows:

$$\text{N-PTDF}_{m \to n} = PTDF_n - PTDF_m \tag{3.4}$$

Hence, the above N-PTDF matrix of size $n_l \times n_b$ can alternatively be written as following matrix notation:

$$\mathbf{N\text{-}PTDF} = \mathbf{B_x} \times \overset{*}{\mathbf{A}} \times \overset{*}{\mathbf{X}} \tag{3.5}$$

The above **N-PTDF** matrix is exploited in this chapter as mentioned in the following sequential steps to find the index of vulnerable injection meters. Same index of voltage meters can also be targeted by the attackers to maximize the impact of attack.

Algorithm 2: Selection of injections/voltage meters

Step-1: Calculate the Nodal PTDF (N-PTDF) matrix using Eq. (3.5).
Step-2: Compute a row vector containing the absolute sum of each column. This vector gives total absolute change for all the lines corresponding to each nodal point.
Step-3: Normalized and ranked the vector obtained from the Step-2 in descending sequence and then select top 25% of the total bus injection meters from it as target meters.

3.3 Stage-2: Modelling of Replay Attacks

Replay attacks involve unauthorized access by an intruder to secretly record the sensing data, which is later delayed or replayed with fraudulent manipulation to the control center (CC) during a sabotaging activities on the physical system without being noticed by the operator. In this process of data sniffing, the adversary analyzes the captured dataset to identify periods of disturbance and kept it separate from ambient data. The pre-recorded disturbance data is then subsequently replayed from sensor terminals, deceiving CC operators into taking falsified actions that may potentially favour the adversary. As these recorded packets are the duplicate copies of some original measured readings and adhere to the same protocol, conventional anomaly detection methods usually cannot recognize them as anomalies.

The above-mentioned intrusion strategy primarily consists of followings two steps:

3.3.1 Recording Window Phase:

This is the initial phase when the intruder infiltrates the system by breaking conventional IT network security gateways and discreetly records observed sensor values without altering any data for sufficiently long finite-time interval in a buffer. As the adversary has usage limitation of attack resources associated with this phase, it is reasonable to assume that they can only store n data packets of i^{th} sensor and the set of the capture data is denoted by $\mathcal{Y}_i(t) = \{y_i(t_0), y_i(t_1), y_i(t_2), ..., y_i(t_n)\}$, where $t_0 \neq t_1 \neq t_2 \neq ... \neq t_n$ be the discrete time instants of whole time period $\mathcal{T}$. Therefore, if the adversary starts eavesdropping at time instant t_0 (Let's denote it as, $\mathcal{T}_s$) to gather knowledge of sensor's data and continues to do so till time instant say $\mathcal{T}_e$, then the recorded interval is designated as $t_{rec} = \{t \in \mathbb{N} : t \in [\mathcal{T}_s, \mathcal{T}_e] = [t_0, t_0 + \acute{l} - 1]\}$, where, $\acute{l} \in \mathbb{N}$ denotes the window size of the attacker's recording phase and the recorded output follows the below mathematical equation:

$$\mathcal{Y}_i^r(t) = \mathcal{Y}_i(t), \quad t \in t_{rec} \tag{3.6}$$

3.3.2 Replaying Window Phase:

In the second phase, the attacker commences the tampering of pre-determined sensors identified in Section 3.2. This involves manipulating current observed values of those sensor measurements by substituting them with any related/unrelated previously recorded values for some specific time duration. Lets say, the replaying functionalities are initiated by the adversary at $\mathcal{T}_e + h$ sec and continues to persist for the whole duration up to $\mathcal{T}$ sec, then the playback time interval is denoted as $t_{rep} = \{t \in \mathbb{N} : t \in [\acute{n}\mathcal{T}_e + h, (\acute{n}+1)\mathcal{T}_e - \mathcal{T}_s]\}$, where, $\acute{n} = 1, 2, ...$ and h accounts for the total number of replay attack sequence and sampling period respectively. Thus, finally at the end of this second stage the malicious sensor reading that fraudulently transmitted to CC can be written as:

$$\mathcal{Y}_i^a(t) = \begin{cases} \mathcal{Y}_i^r(t - \acute{\tau}), & \text{if } \digamma_i = 1 \text{ and } t \in t_{rep}, \\ \mathcal{Y}_i(t), & \text{otherwise} \end{cases} \tag{3.7}$$

where, $\acute{\tau}$ is the time elapse between the onset of above two phases and $\digamma$ is the binary variable whose value is unity when RA is triggered and zero otherwise. Also, it is to be noted that in Eq. (3.6) and Eq. (3.7) the subscripts ^{r} and ^{a} denotes the sensor output during the record and replay window phase respectively.

3.3.3 Different RA Models Influencing PSSE:

Accurately estimating the state in real-time is crucial, especially in power systems, as it relies on sensor data from RTUs for monitoring and control. However, the integrity of these results is jeopardized by the occurrence of replay attacks, potentially leading to security breaches and power system instability. In pursuit of showing its impact on PSSE according to the execution plan mentioned above, two different RA models that aim to interfere with trend-based power system applications and cause possible failures, are discussed in this subsection below.

- **Multiple Data Dropping Attack (MDDA)**: In this MDDA-type replay attacks, attacker aims to capture the actual dynamic trends from the normal RTU sensor readings at distinct and convenient time intervals of attacker's choice in its recording window phase. During replaying window phase, the attacker substitutes the current stream of RTU readings with an interpolated data sequence derived from the previously recorded data as described in the following Algorithm. 3.

Algorithm 3: MDDA with interpolated data sequence

Step-1: Start eavesdrop and record the i^{th} RTU readings upto t_n^{th} sec with a gap of k sec intervals such that $\mathcal{Y}_i^r(t) = \{y_i(t_0), y_i(t_{0+k}), y_i(t_{0+2k}), ..., y_i(t_{0+\acute{l}k})\}$, where $t_{\acute{l}k} \leq t_n$ is satisfied.

Step-2: The incomplete recorded dataset due to dropout of multiple data at each $(\acute{l}k - 1)^{th}$ interval is then filled up by the computed steady state interpolated trends from the existing available i^{th} RTU dataset.

Step-3: In the replaying window phase, the original RTU measurements data from t_{n+1}^{th} sec to t_m^{th} sec are being replayed by the interpolated and recorded data provided by Step-2.

- **Repetitive Data Cloning Attack (RDCA)**: Replay attack can be launched via impersonating as a natural disturbance or an equipment fault. Therefore, the primary objective of RDCA here is to replace the current trends in multiple sensors with the cloning of pre-saved historical measurement data across repetitive time instances. With the aim of manipulating the dynamic signatures, in RDCA the attackers replaces the original dynamic data trends of the various RTU sensors with repeated sequences of pre-saved unrelated high disturbance or faulted signature trends. The following two algorithms i.e Algorithms. 4 and Algorithms. 5 outlines the modeling steps of different RDCAs.

3.4 Stage-3: PMU Sensor-Assisted RA Detection and Correction

In SCADA system, the traditional state estimation (SE) relies on non-linear weighted least square (WLS) algorithm for finding the best fit of the of system states, utilizing sensor measurements from RTUs across a wide geographical area. However, the susceptibility

Algorithm 4: RDCA with cloning of high disturbance event historical data (LT-RDCA)

Step-1: Choose a pre-saved unrelated historical huge load disturbance dataset from a historical database archives.
Step-2: Start recording of a k^{th} sec window from the above disturbance dataset such that $\mathcal{Y}_i^r(t) = \{y_i(t_0), y_i(t_1), y_i(t_2), ..., y_i(t_k)\}$, where $t_k \leq \mathcal{T}$.
Step-3: Now in the playback period, the current RTU data streams are being replaced by the recorded dataset and the same being replayed periodically from t_{n+1}^{th} sec onwards.

Algorithm 5: RDCA with cloning of Faulted event historical data (ST-RDCA)

Step-1: Choose a pre-saved unrelated historical Faulted dataset from a historical database archives.
Step-2: Record the fault signature from the chosen dataset from t_k sec to t_{k+4} sec duration.
Step-3: Now replace the original RTU measurements of a non-fault case by the same duration recorded data and replay it between t_k sec to t_{k+4} sec to misinterpret the non-fault case as fault case.

to high dB noise or replay attacks in the transmitted data between RTUs and other network components due to the absence of robust security mechanisms in telemetry poses a substantial threat. This lack of security measures leads to notable discrepancies in estimation accuracy and may also introduces potential catastrophic vulnerabilities in the physical system. However, the state-of-the-art PMU sensors, on the other hand, are considered to be resilient and cyber-secured due to their advanced cybersecurity features and robust design, providing accurate and time-stamped real-time data of voltage and current phasors [204, 65]. These devices play a crucial role in enhancing the grid's situational awareness and ensuring the integrity of the system against cyber threats. Nevertheless, the large scale deployment of these devices are quite expensive due to its requirement of communication infrastructure and maintenance cost and thus complete replacement of RTU device with PMU is not possible in near future. Therefore, the coexistence and complimentary co-operative support of both the sensors i.e., RTUs and limited PMUs is the only viable option to jointly estimates the power system states as well as ensuring global network observability and robustness against implausible cyber threats. Thus, in order to detect and correct the RAs in SE, the methodology proposed in this section is based on the improved synchrophasor-assisted hybrid state estimator (HYB-SE) and a model based residual check technique, as discussed in the following subsections.

3.4.1 Hybrid SE (HYB-SE) Model

In the WLS-assisted conventional SE, the measurement vectors (Z_{RTU}), comprising of $\mathcal{M}$ RTU readings (voltages, power injections, and power flows), are linked to the $2N-1$

system states (X_{RTU}) of order N using the following non-linear relationships:

$$Z_{RTU} = h_{RTU}(X_{RTU}) + \varepsilon_{RTU} \tag{3.8}$$

where $h_{RTU}(\cdot)$ is the non-linear mapping function that relates the RTU measurements with its corresponding state variables and $\varepsilon_{RTU} \in \mathcal{N}(0, \tilde{\sigma}_{RTU})$ is the measurement Gaussian noise or errors associated with RTU sensors.

Then, the conventional estimates of system state vector are obtained if the objective function ($\mathfrak{J}(X_{RTU})$) defined as the weighted sum of squares of measurement errors can be minimized through WLS method as shown below:

$$\mathfrak{J}(X_{RTU}) = [Z_{RTU} - h_{RTU}(X_{RTU})]^T \mathbf{R_{RTU}^{-1}} [Z_{RTU} - h_{RTU}(X_{RTU})] \tag{3.9}$$

where $\mathbf{R_{RTU}}$ be the $\mathcal{M} \times \mathcal{M}$ order RTU measurements error covariance matrix defined as $\mathbf{R_{RTU}} = \mathbb{E}[\varepsilon_{RTU} \cdot \varepsilon_{RTU}^T]$ The reciprocal of this matrix is called as weight matrix ($\mathbf{W_{RTU}}$) whose each diagonal elements represents the weightage of the corresponding available RTU measurements. It depends on the accuracy level of RTU sensors. At the end, solving the above objective function through iterative approach, the final estimated value of the states can be obtained as: [205]

$$\Delta Z_{RTU}(\hat{X}_{RTU}^k) = Z_{RTU} - h_{RTU}(\hat{X}_{RTU}^k) \tag{3.10}$$

$$\Delta \hat{X}_{RTU}^k = [\mathbf{H_{RTU}^T R_{RTU}^{-1} H_{RTU}}]^{-1} \mathbf{H_{RTU}^T R_{RTU}^{-1}} \Delta Z_{RTU}(\hat{X}_{RTU}^k) \tag{3.11}$$

$$\hat{X}_{RTU}^{k+1} = \hat{X}_{RTU}^k + \Delta \hat{X}_{RTU}^k \tag{3.12}$$

where, $\mathbf{H_{RTU}}$ is measurement Jacobian matrix of size $\mathcal{M} \times N$, $\mathbf{H_{RTU}^T R^{-1} H_{RTU}}$ is called as gain matrix and $\hat{X}_{RTU}^k = [\hat{V}_{RTU}^k \;\; \hat{\theta}_{RTU}^k]^T$ be the conventional estimated states at k^{th} iteration.

Now, to enhance the estimation accuracy, capturing dynamic state trends, and mitigate uncertainties, the PMU-derived bus voltage and current phasor measurements, in conjunction with the latest WLS-assisted conventional SE estimates, are utilised to form a new measurement vector (Z_{HYB}). Since voltage and current exhibit linear relationships with the relevant state variables, the new measurement vector is expressed in rectangular coordinates, forming a hybrid linear estimator model as shown below.

$$Z_{HYB} = \begin{bmatrix} \begin{bmatrix} \hat{V}_{Re} \\ \hat{V}_{Im} \end{bmatrix}_{RTU} \\ \begin{bmatrix} V_{Re} \\ V_{Im} \end{bmatrix}_{PMU} \\ \begin{bmatrix} I_{Re} \\ I_{Im} \end{bmatrix}_{PMU} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbb{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbb{C}_2 \\ \mathbb{C}_3 & \mathbb{C}_4 \\ \mathbb{C}_5 & \mathbb{C}_6 \end{bmatrix} \times \begin{bmatrix} V_{Re} \\ V_{Im} \end{bmatrix} + \begin{bmatrix} \varepsilon_{RTU}^{V_{Re}} \\ \varepsilon_{RTU}^{V_{Im}} \\ \varepsilon_{PMU}^{V_{Re}} \\ \varepsilon_{PMU}^{V_{Im}} \\ \varepsilon_{PMU}^{I_{Re}} \\ \varepsilon_{PMU}^{I_{Im}} \end{bmatrix} \tag{3.13}$$

$$Z_{HYB} = \mathbf{H_{HYB}} \; X_{HYB} + \varepsilon_{HYB} \tag{3.14}$$

The subscripts 'Re' and 'Im' in the above equation denotes the real and imaginary components of the measurements and hybrid states (X_{HYB}). Symbols $\mathbf{I}$ and $\mathbf{0}$ represents an unit matrix and zero matrix respectively. $\mathbb{C}_1$ and $\mathbb{C}_2$ are the matrices where each row corresponds to a specific PMU location, with ones placed in the columns corresponding to the measured voltage phasors by the respective PMU's row index. The elements of matrices $\mathbb{C}_3$ to $\mathbb{C}_6$ are represented in linear combinations consisting of line conductance and susceptance for the lines where current phasors are available. Finally, the WLS solution ($\hat{X}_{HYB}$) for the above linear model can be obtained in non-iterative manner and expressed in the same form as stated in Eq. (3.11) and Eq. (3.12).

$$\hat{X}_{HYB} = \left(\mathbf{H}_{\mathbf{HYB}}^{T}\mathbf{W}_{\mathbf{HYB}}\mathbf{H}_{\mathbf{HYB}}\right)^{-1}\mathbf{H}_{\mathbf{HYB}}^{T}\,\mathbf{W}_{\mathbf{HYB}}\,Z_{HYB} \tag{3.15}$$

where, $\mathbf{W_{HYB}}$ be the reciprocal of covariance matrix of hybrid estimator ($\mathbf{R_{HYB}}$) consist of error covariance matrices of conventional SE, PMU voltage phasors and PMU current phasors converted into rectangular format based on error propagation theory [206] of measurement transformation.

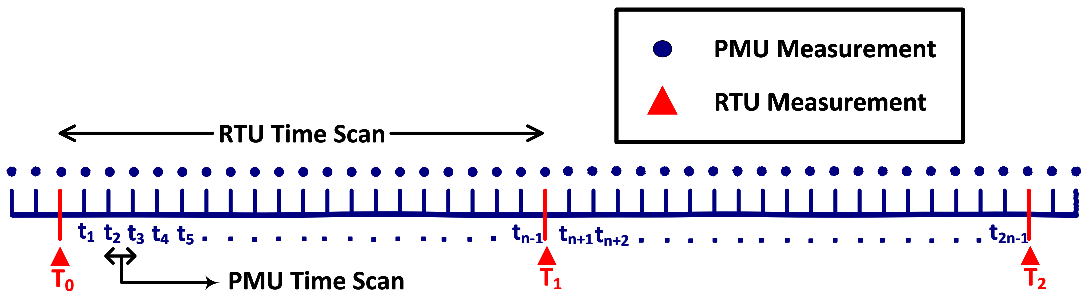


Figure 3.2: Time scale of arrival of RTU and PMU measurements

3.4.2 Proposed Phasor Measurement Based RA Detection and Correction

The detection and correction of the RAs on PSSE can be effectively achieved through the use of above described 2-stage hybrid state estimation model, considering real-time measurements collected using both RTUs and PMUs with different time scan periods. Usually, the measurement set updated by the RTU is within a time scan of few seconds and for PMU the reporting rate is quite high i.e., typically up to 60 frames per seconds. The aim here is to effectively integrate information from both the devices operating at different sampling rates, within the dynamic HYB-SE framework such that any manipulations in RTUs can be noticed easily. In this context, this chapter assumes that HYB-SE is executed whenever a new set of PMU data arrives, with a refresh rate of 20 frames per second for a 60 Hz system. Simultaneously, the conventional non-linear SE is executed at each one-second interval upon the arrival of a new RTU measurement dataset. This implies that between two consecutive RTU measurements, 20 PMU measurements are reported. The whole timeline diagram, including the arrival of the complete set of RTU and PMU measurements both and the HYB SE's execution cycle, is shown in Fig. 3.2. During periods when only PMU measurements are available, the latest states estimated solely by the conventional non-linear SE are incorporated to fill the incomplete measurement set and execute the HYB-SE. This is due to the fact that existing power system is observable by the available RTU measurements, but the inclusion of additional PMUs serves the purpose of transforming any critical existing measurements into redundant ones, thereby improving the overall accuracy of the measurement reconstruction process. This enhancement also aids in the identification of compromised RTU sensor readings and distinguishing the moment when a RA is initiated. After the **Stage-1** and **Stage-2**, steps have been followed in block (1)-(3) in Fig. 3.3, and a few critical RTU measurements are compromised, the RAs may be initiated by the attacker at any opportune time. The steb-by-step procedure for detecting and correcting the RAs in SE, are described as follows.

(1) In block-(4), the WLS-assisted SE algorithms estimate system states ($\hat{X}_{RTU}$) once every second. The linearized HYB-SE also runs through blocks 5–6 in parallel. The dynamic estimates $\hat{X}_{HYB}$ are produced recursively in HYB-SE using the estimates of the WLS estimator until new RTU data is received.

(2) Because the secured PMU measurements are quite accurate, ($\hat{X}_{HYB}$) can be used, using standard power flow equations, to reconstruct the estimation of the original RTU readings at each second interval following the arrival of Z_{RTU} in block-(7).

(3) In block-(8), a measurement residue (Λ) is computed using the original RTU sensor readings ($\tilde{Z}_{RTU}$) (which may or may not be compromised) and the reconstructed measurements ($\hat{Z}_{RTU}^{HYB}$).

(4) If the residue lies within an upper ($+\delta^{UB}$) and lower bound ($-\delta^{LB}$) in block-(9), the RTU measurements are labeled as Normal with a Flag=0 in block-(10), compromised otherwise. The measurements that have Flag=1 assigned to them are considered vulnerable and are kept in a variable called Z_v in block (11).

(5) To fix the bugged measurements, Z_v, block-(12) replaces the malicious RTU sensor readings with the reconstructed measurements by hybrid states and the toggling switch, S_w is set to position 2 as shown in Fig. 3.3.

(6) The conventional SE stage in block-(4) executes in loop, thereafter, to estimate the attack-free states with the corrected measurements through blocks (5)-(12).

The overall flowchart of the proposed scheme with various intermediate blocks is presented in Fig. 3.3.

3.5 Real-Time Digital Simulation (RTDS) Results

The proposed RA detection and correction framework is tested and validated on the IEEE 14-bus and New England (NE) 39-bus systems using the RSCAD software in the RTDS. The RTDS modeling closely mimics real field scenarios, providing accurate results.

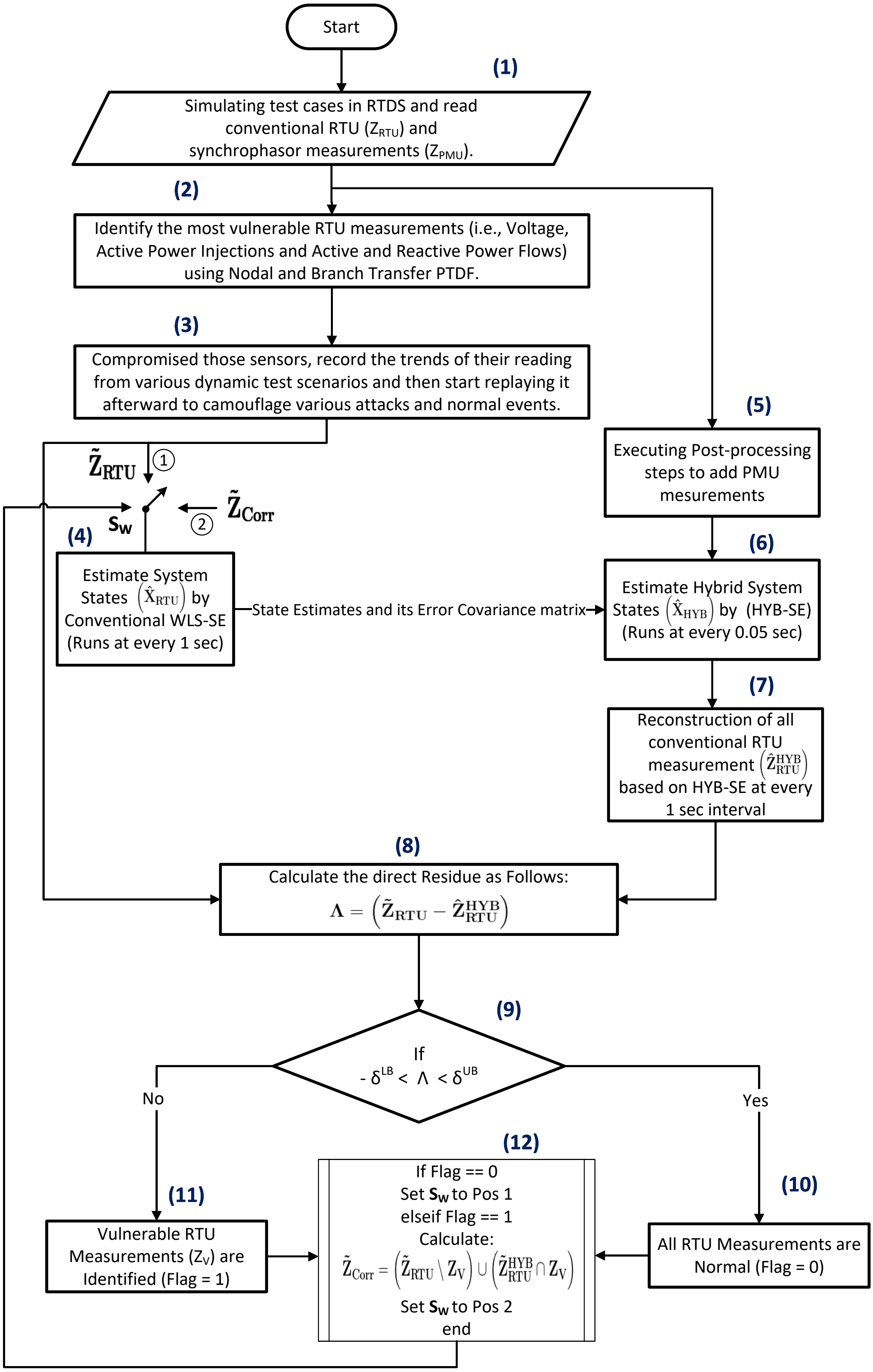


Figure 3.3: Flowchart of the proposed RA detection and correction scheme

Dynamic conditions of the grid are emulated by varying loads from ±10% to ±30% with user-defined logic sequences in RSCAD. Gaussian noise is introduced to both RTU and PMU readings to simulate measurement uncertainties. PMUs are installed at the same bus locations as specified in [207]. A MATLAB script implementing HYB-SE is then executed for randomly generated RA test scenarios. The framework's performance is evaluated based on factors such as correct identification of vulnerable measurements, detection time for launched RAs, and key metrics including True Negative (TN), False Negative (FN), False Positive (FP), and True Positive (TP), along with other derived evaluation indices.

3.5.1 IEEE 14 Bus Test System:

To assess the accuracy of the proposed method for detecting and correcting replay attacks, the IEEE 14-bus test system is initially selected. The system is tested against two attack strategies: MDDA and the RDCA algorithm (4), as discussed in Section 3.3. The details associated with various RTU, PMU and vulnerable measurements along with their standard deviation (SD) for the IEEE 14-bus system is tabulated in Table 3.1. Additionally, PMU sensors are installed at Bus-2, Bus-7, Bus-10, and Bus-13 to obtain voltage and current phasors for the execution of HYB-SE, which is used to generate refined reconstructed measurements. For this test system, the upper ($+\delta^{UB}$) and lower threshold ($-\delta^{LB}$) values are fixed at 0.03 p.u. through out the total simulation duration.

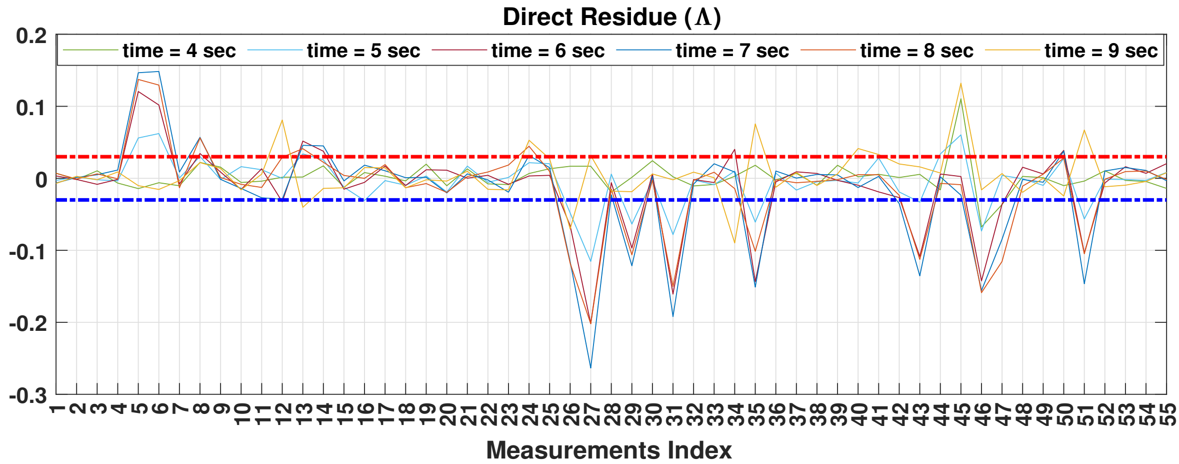
Table 3.1: Conventional, Compromised and Synchrophasor Measurements for IEEE 14-Bus Test System

IEEE 14-bus System					
Parameters	**Bus Location for Voltages (Bi)**	**Bus Location for Active Power Injections (Bi)**	**Bus Location for Reactive Power Injections (Bi)**	**Line Location for Active and Reactive Power Flows (Li-j)**	**No. of RTU Sensors**
	SD: 0.006 p.u.	**SD: 0.01 p.u.**	**SD: 0.01 p.u.**	**SD: 0.01 p.u.**	
Z_{RTU}	B1, B2, B6, B9, B10, B12, B14	B1, B2, B4, B6, B8, B10, B12, B14	B1, B4, B6, B7, B9, B11, B12, B13	L1-2, L1-5, L2-3, L2-4, L4-5, L4-7, L4-9, L5-6, L6-11, L6-12, L7-8, L7-9, L9-10, L9-14, L12-13, L13-14	7+8+8 +16+16 = 55
Z_v	B10, B12	B8, B10, B12		L2-3, L2-4, L4-7, L5-6, L7-8, L7-9, L13-14	2+3+7+7 = 19
Parameters	**Bus Location for Voltages Phasors**		**Line Location for Current Phasors**		**No. of PMU Sensors**
	SD: 1.0e-05 p.u (Mag), 0.001 (Ang)		**SD: 1.0e-05 p.u (Mag), 0.001 (Ang)**		
Z_{PMU}	B2, B7, B10, B13		L2-1, L2-3, L2-4, L2-5, L7-4, L7-8, L7-9, L10-9, L10-11, L13-6, L13-12, L13-14		4+12 = 16

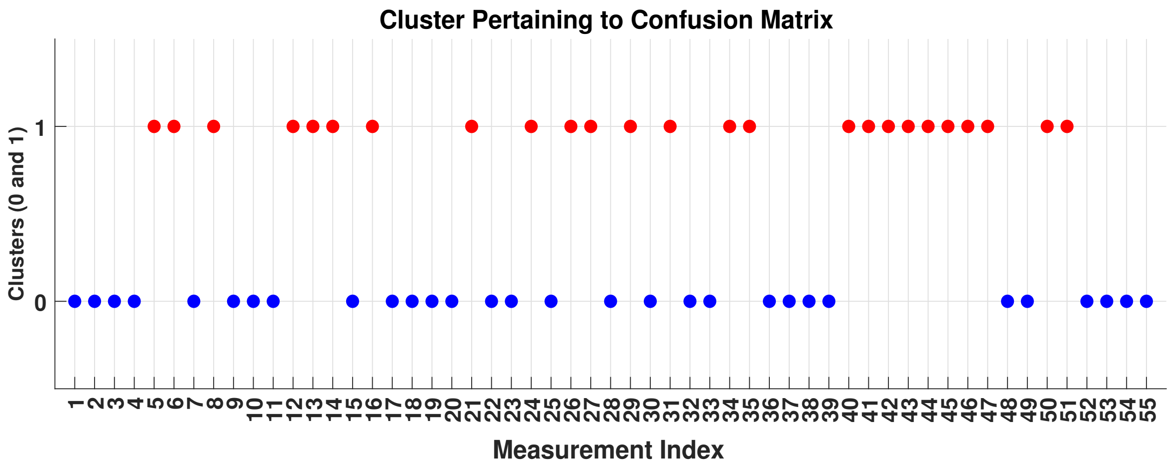
3.5.1.1 Multiple Data Dropping Attack (MDDA) Detection

In this study, before the attack initiation, the model is running in steady state with dynamic load variations are performed at the load bus 4, bus 5 and bus 9 to bus 13. The variations are simulated chronologically, with the model running at its base load for the initial 0 to 2 sec. Then, the load is increased by 30% for the next 3 sec, followed by another increment of 30% of the previous load for the next 3 sec. Finally, after 8 seconds, the load is reduced to its base load again. Throughout the first 5 sec simulation duration, in MDDA, the attacker eavesdrops and sniffs data packets to capture dynamic snapshots of the current simulation case at three distinct intervals: 0 sec, 2 sec, and 4 sec based on resource availability. Now to generate a complete set of replay attack vector for a 5 sec duration, interpolated steady-state values are computed for 1 sec and 3 sec based on the captured data set. Following this, the attacker is prepared to replace the current streaming of RTU sensor data with the pre-recorded data for a playback time of 5 sec, which is then transmitted to the control center from 5 sec onward.

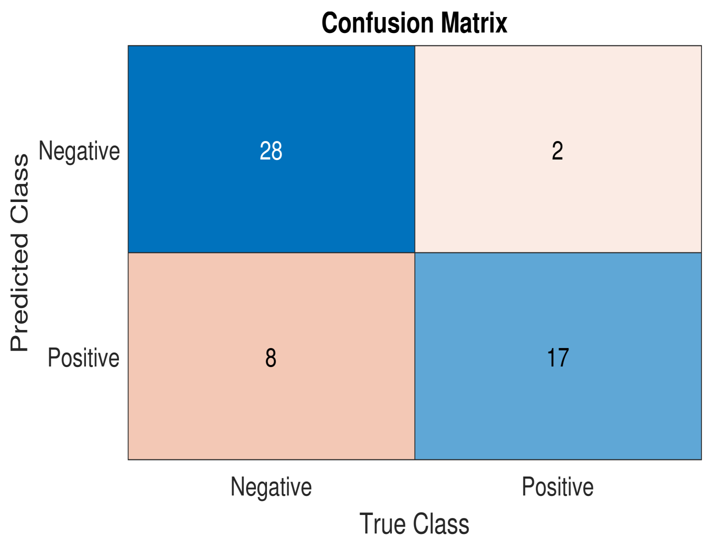
As per the flowchart of Fig. 3.3, with the onset of the attack, direct residue (Λ) in Fig. 3.4(a) exhibits significant changes in the model residual of some of the measurements crossing the thresholds limit denoted with red and the blue horizontal lines. At the end, the residual matrix is converted to a binary matrix by placing one while the observed residuals crossing upper and lower bound thresholds and place 0 otherwise. Then by calculating the total number of zeros and non-zeros at the end of simulation duration pertaining to individual measurements, a plot is generated as shown in Fig. 3.4(b) which represents the cluster of the measurements Flagged as normal (blue dots) and the ones flagged as vulnerable (Flag=1, with red dots). The same information is quantitatively assessed through confusion matrix as shown in Fig. 3.4(c) to count the number of correctly identified and labeled attacked and non-attacked RTU sensor measurements based on TP, TN, FP, and FN. These information also help to calculate some other crucial metrics such as precision, recall, and F1 score to gain profound quantitative insights into the efficacy of proposed attack detection method. It is noticed that the proposed method has relatively higher true positive rate (89.47%), accuracy (81.85%) with a little compromise in true negative rate (77.79%), F1 score (77.27%) and precision (68%). This is due to the presence of bad data in the measurement vector which is not attacked. Moreover, in order to visualize the timing instants of replay attack and to potentially identify the uncorrelated attacked measurements or outliers, various statistical study in terms of computing correlation coefficients via correlation matrix and box plot are conducted based on residual features matrix as shown in Fig. 3.4(d) and 3.4(e). The correlation matrix shown in Fig. 3.4(d) measures the relationships between different measurements over time for discerning distinct behavioral patterns during attack and non-attack periods. Thus, during 5 sec attack duration of MDDA, the coefficients values get surprisingly very high as compared to the same computed in other timings as visualized by the heatmap. The same can also be perceived by the boxplot as shown in Fig. 3.4(e) which facilitated the identification of outliers or potentially corrupted data using the symbol '+' during



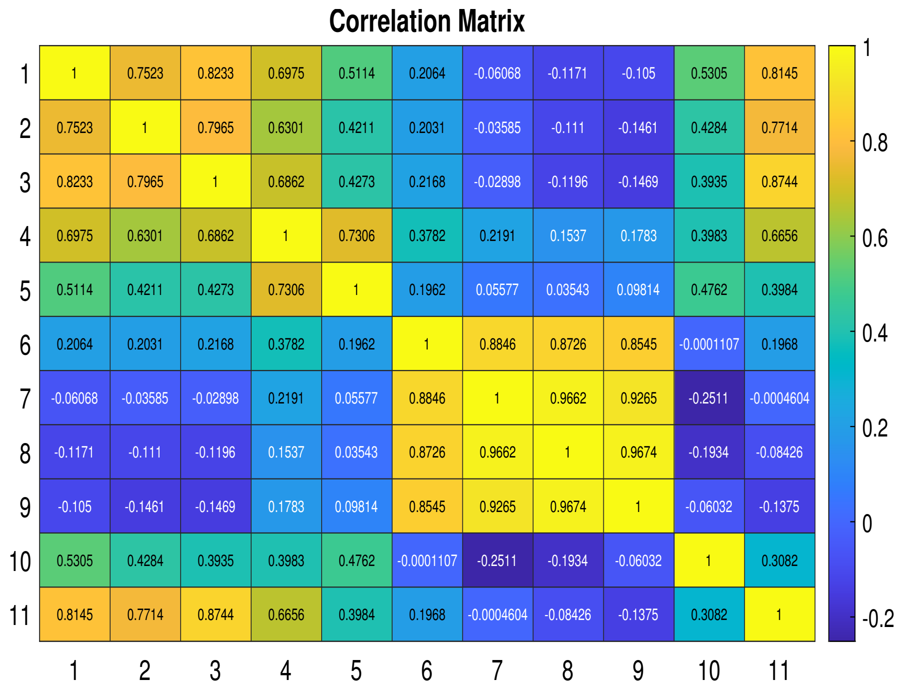
(a) Direct Residue



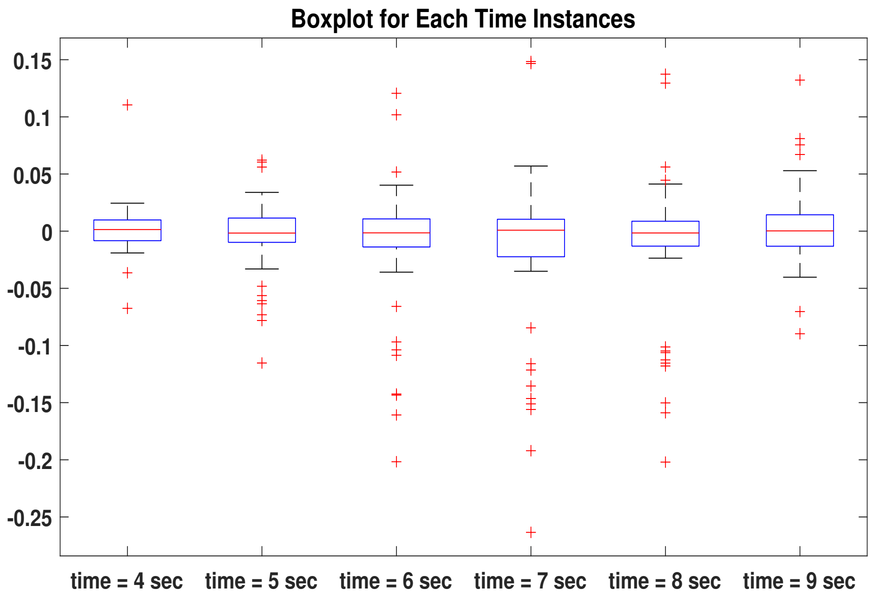
(b) Cluster of Measurements (Red-Attacked, Blue-Non-attacked)



(c) Confusion Matrix



(d) Correlation Matrix



(e) Box Plot

Figure 3.4: Detection phase of MDDA for IEEE 14-bus test system

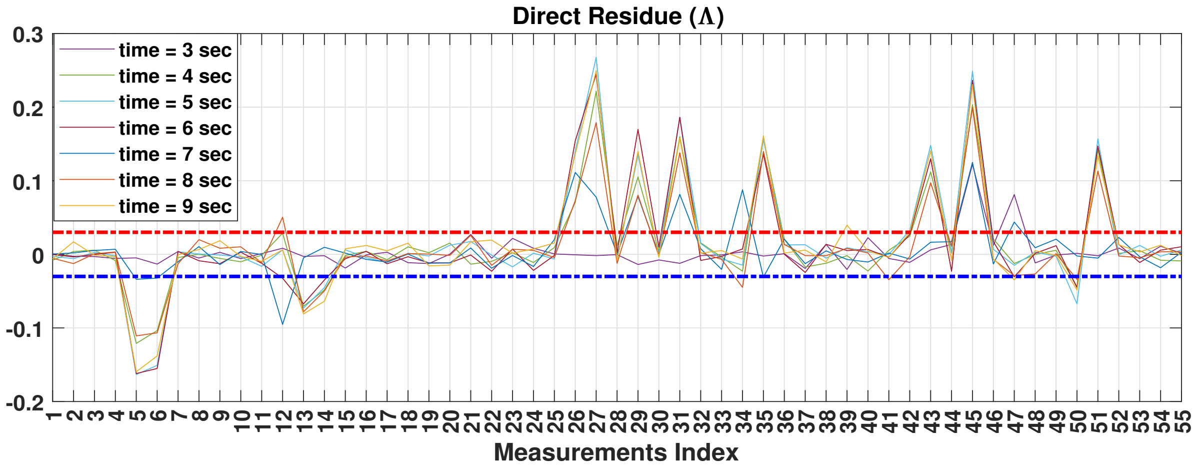
the attack periods (5 sec to 9 sec), as many data points significantly deviating from the expected distribution and lying outside of the interquartile range of the box.

3.5.1.2 Repetitive Data Cloning Attack (RDCA) Detection

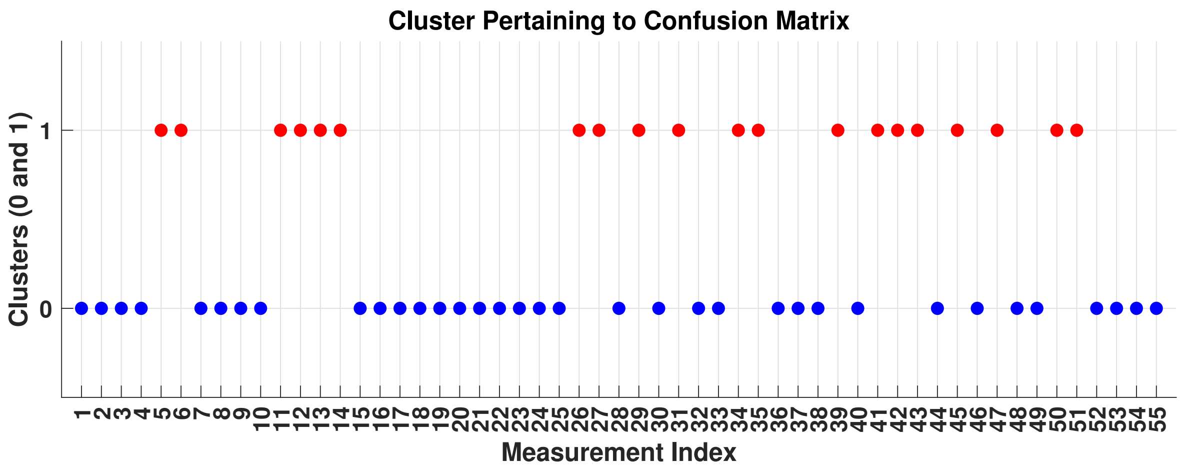
In contrast to the previous case, now the chronological load variation is done with low step load changes in such a way that the model is allowed to run under normal conditions for the initial 0 to 2 sec followed by 12% decrement for next 3 sec. In subsequent next 3 sec, the load is increased by 17% from its previous load value and finally the load settles at 10% below the previous value after 8 sec onwards. As per the RDCA algorithm 4, it is assumed that the attacker already got control access into the operator's network before and recorded an unrelated historical load disturbance data for 4 sec window span. Now in this case study, attackers start replaying those recorded data packets periodically in the replacement of original RTU data from 4 sec onwards. The direct residual plots, as shown in Fig. 3.5(a) reveals the presence of RDCA in some of the measurements. Figure 3.5(b) also successfully identified most of the vulnerable measurements marked as red dot while it crossing the set threshold limits. It is also noticed from the confusion matrix of Fig. 3.5(c) that the number of FN and FP are 1 and 2 respectively which is significantly improved then previous case study. Moreover, improvement can also be observed in other derived metrices such as true positive rate and true negative rate is almost 94.5% with precision, accuracy and F1 score are 90%, 94.6% and 92.3% respectively. To analyze temporal patterns in the context of potential replay attacks on those attack time instants, correlation matrix is also shown in Fig. 3.5(d). Stronger associations were shown by bright hues on the heatmap, which provided a visual representation of this increased correlation. This observation implies that replay attacks occurring at those specific time instance can have a noticeable impact on the temporal correlations between observations, resulting in patterns that can be identified in the correlation matrix. As the attack being start replayed from 4 sec by the adversary, the box plot depicted in Fig. 3.5(e) also shows a sudden increase in corrupted uncorrelated measurement readings which clearly discern irregularities and variations which corresponded with the occurrence of cyber attacks.

3.5.1.3 Attack Correction

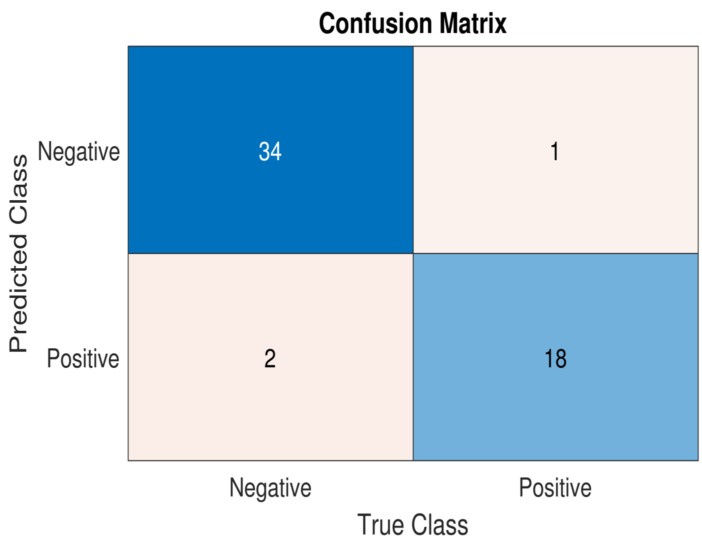
The correction phase of the proposed methodology against these two above mentioned attacks i.e MDDA and RDCA algorithm 4 is shown in Fig. 3.6. The vulnerable measurements identified by the above detection phase is now corrected with the reconstructed RTU measurements using the refined estimates of hybrid states incorporating secured PMU measurements. The re-execution of SE process thereafter takes place as per the algorithm explained in Section 3.4.2. As a result it can be evident from the direct residue plots of Fig. 3.6(a) and Fig. 3.6(b), that now the residuals are confined within the range selected by upper and lower level thresholds and thus none of the measurements are now treated as attacked or compromised. It clearly reveals that the error under compromised condition, i.e., before RA correction was maximum in the range of 0.2 pu, whereas after the correction as per the proposed scheme, the residual drops to a maximum value of about 0.02 pu. On the similar line, the correlation matrix depicted in Fig. 3.6(c) also demonstrates the strong correlation coefficients with its diagonal entity only which implies that the measurements received at different time instances are consistent and correlated with their respective time of arrival and there was existence of no anomalous readings. Also by exploring the temporal and spatial relationships between time, measurements, and attack labels, the 3-D scatter plot shown in Fig. 3.6(d) indicates that there is no measurements in the corrected RTU vector set



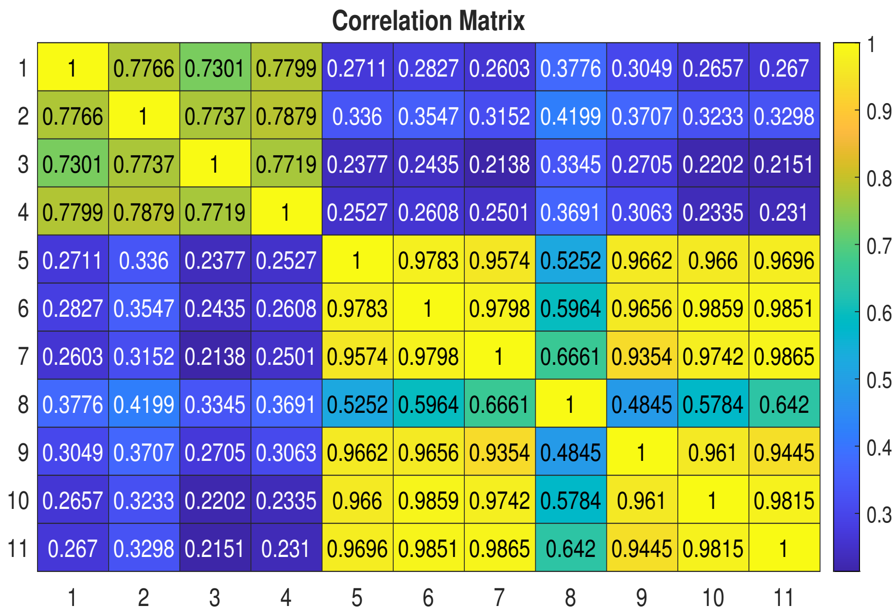
(a) Direct Residue



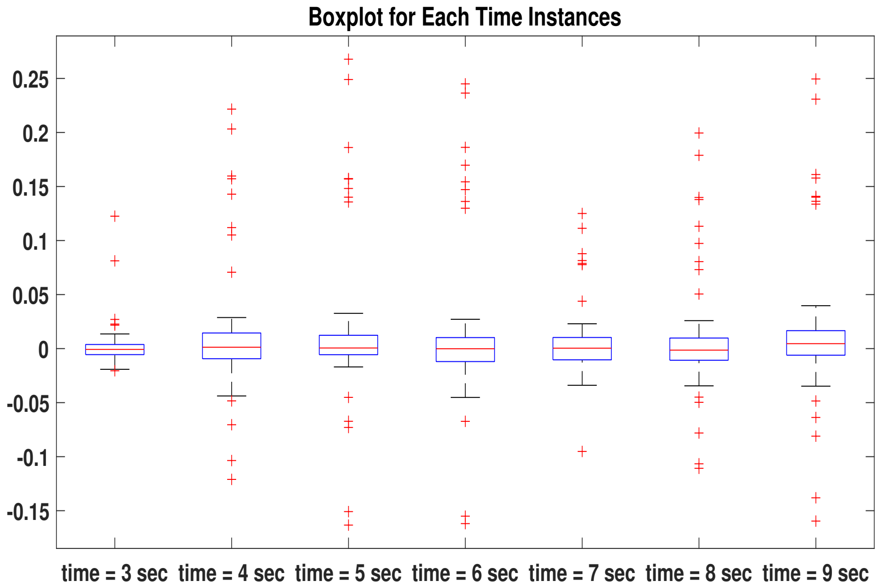
(b) Cluster of Measurements (Red-Attacked, Blue-Non-attacked)



(c) Confusion Matrix



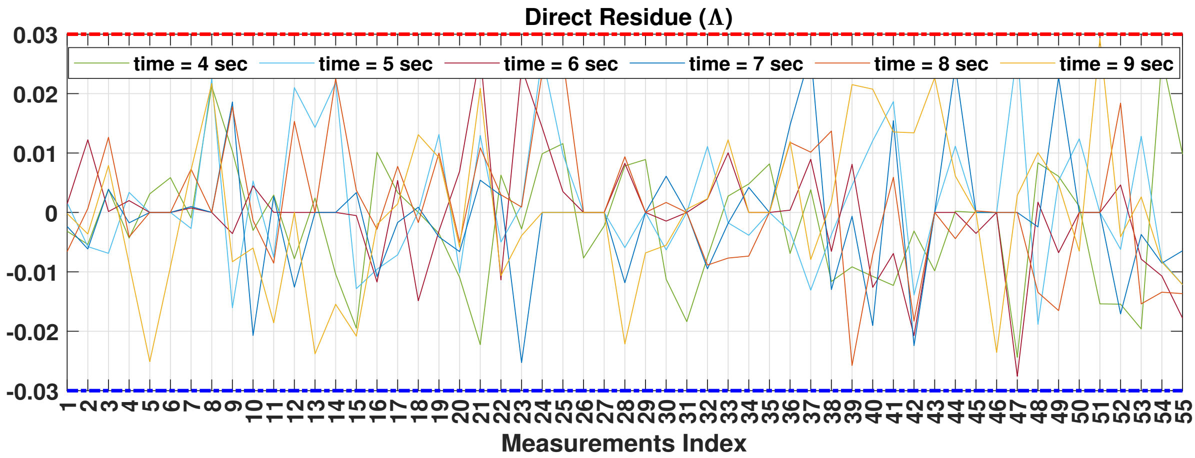
(d) Correlation Matrix



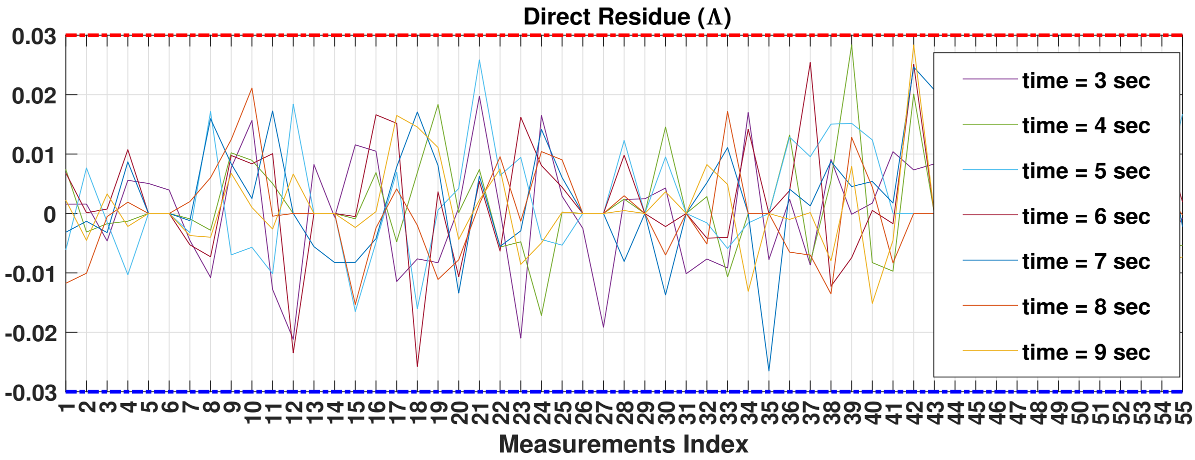
(e) Box Plot

Figure 3.5: Detection phase of RDCA Algorithm 4 for IEEE 14-bus test system

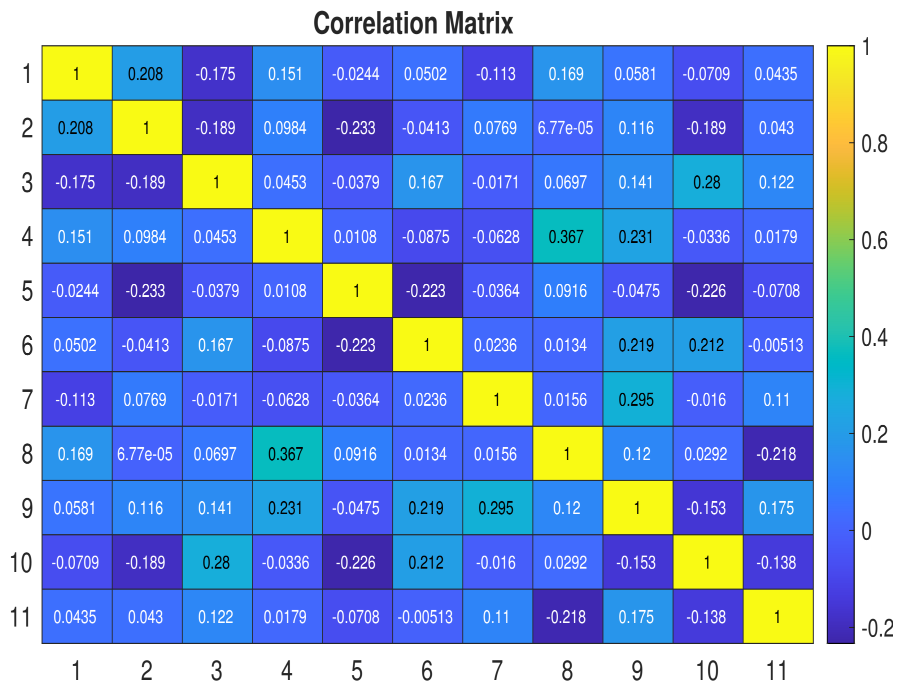
that are corrupted by any of the aforementioned attacks. Furthermore, in the correction phase of the MDDA, the maximum reduction in residue from the attacked state to the corrected state ranges between 83.33% (upper bound) and 89.73% (Lower bound). The highest root mean square error of the hybrid estimated states, post-correction for MDDA, is recorded at 0.4%. Similarly, during the correction phase of the RDCA, the maximum decrease in residue from the attacked state to the corrected state falls within the range of 89.51% (upper bound) and 83.65% (Lower bound). The maximum root mean square error of the hybrid estimated states, after correcting for RDCA, is observed to be 0.6%. Notably, in both case studies, following the correction process, the true negative rate reach 100%. The consolidated results derived from all the above post-correction plots



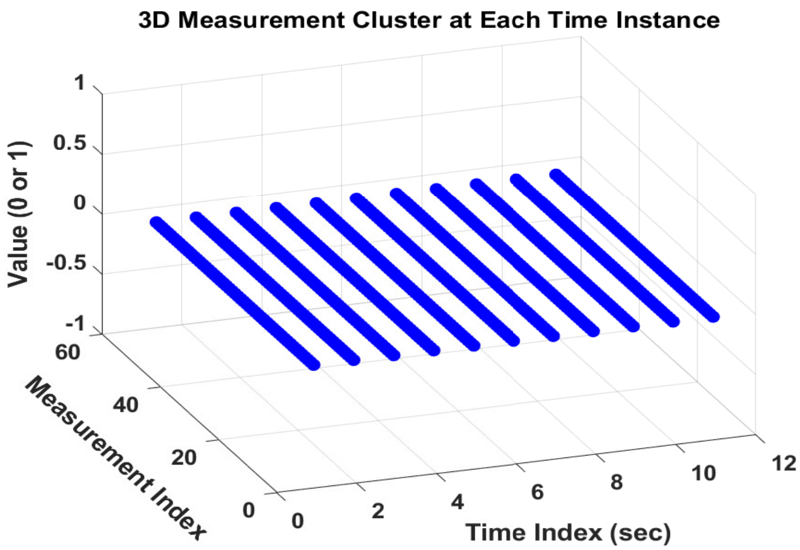
(a) Direct Residue of MDDA



(b) Direct Residue of RDCA



(c) Correlation Matrix



(d) 3-D Scatter Plot

Figure 3.6: Correction phase of proposed scheme for IEEE 14-bus test system

provide compelling evidence of the attack mitigation. Thus, the efficacy of proposed attack correction strategy has been proven to evident through its ability to restore the integrity of attacked measurements.

3.5.2 New England (NE) 39-Bus Test System:

To assess the effectiveness of the proposed attack detection and control approach on larger system, two different variants of RDCA as mentioned in algorithm 4 and algorithm 5 are now employed in NE 39-Bus test system. The details pertaining to various measurements have been summarized in Table 3.2. As the threshold selection is system specific, the upper ($+\delta^{UB}$) and lower ($-\delta^{LB}$) threshold level are set at 0.05 pu, corresponding to this test system.

3.5.2.1 Long Term Repetitive Data Cloning Attack (LT-RDCA):

This case study is similar to the RDCA Algorithm 4 as described in IEEE 14 test system where the adversary recorded the unrelated high disturbance data for replaying compromised RTU measurements periodically from 4 sec onward. Dynamic load variations are sequentially applied to 10 different bus locations: bus-3, bus-4, bus-8, bus-15, bus-16, bus-20, bus-24, bus-26, bus-29, and bus-39.The impact of these adversarial actions is evident in the direct residual (Λ) plot, illustrated in Fig. 3.7(a), which shows the clear evidence of adversarial action by effectively flagging out any potential corrupted sensor readings resulting from the attack. Figure 3.7(b) provides a granular visualization in pinpointing specific measurements affected during whole attack duration. Figure 3.7(c) shows the confusion matrix that provides a comprehensive picture of attack detection performance with very lower value of FP and FN, specifically 3 and 6, respectively. This leads to improved model's ability to successfully detect attacks while minimizing false

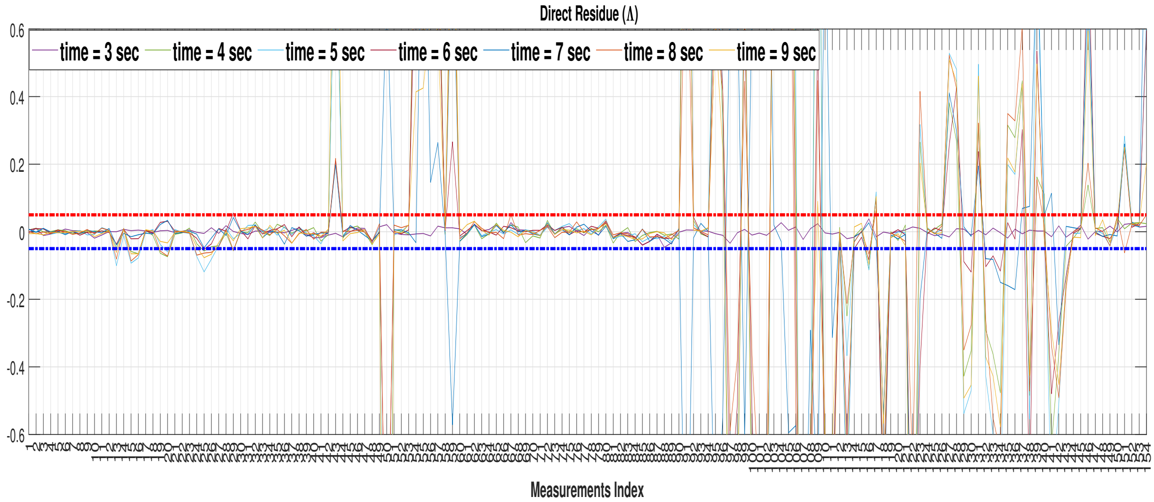
Table 3.2: Conventional, Compromised and Synchrophasor Measurements for NE 39-Bus Test System

NE 39-bus System					
Parameters	**Bus Location for Voltages (Bi)**	**Bus Location for Active Power Injections (Bi)**	**Bus Location for Reactive Power Injections (Bi)**	**Line Location for Active and Reactive Power Flows (Li-j)**	**No. of RTU Sensors**
	SD: 0.006 p.u.	**SD: 0.01 p.u.**	**SD: 0.01 p.u.**	**SD: 0.01 p.u.**	
Z_{RTU}	B1, B2, B3, B5, B6, B7, B8, B10, B12, B15, B16, B18, B20, B21, B22, B23, B25, B26, B28, B29, B30, B31, B32, B33, B34, B35, B36, B37, B38, B39	B1, B2, B3, B5, B6, B7, B8, B10, B12, B15, B16, B18, B20, B21, B22, B23, B25, B26, B28, B29, B30, B31, B32, B33, B34, B35, B36, B37, B38, B39	B1, B2, B3, B4, B7, B8, B10, B12, B15, B16, B18, B20, B21, B23, B24, B25, B26, B27, B28, B29, B30, B31, B32, B33, B34, B35, B36, B37, B38, B39,	L1-2, L1-39, L2-3, L2-25, L3-18, L4-14, L5-8, L6-7, L6-11, L7-8, L8-9, L9-39, L10-13, L13-14, L14-15, L15-16, L16-19, L16-21, L16-24, L17-18, L21-22, L22-23, L26-27, L26-29, L28-29, L12-11, L12-13, L10-32, L23-36, L25-37, L29-38, L19-20	30+30+30+ 32+32 = 154
Z_v	B20, B22, B23, B28, B29, B33, B34, B35, B36, B38	B20, B22, B23, B28, B29, B33, B34, B35, B36, B38		L1-2, L3-18, L4-14, L5-8, L6-7, L6-11, L7-8, L8-9, L9-39, L10-13, L13-14, L14-15, L15-16, L16-19, L16-21, L16-24, L17-18, L21-22, L26-27, L12-11, L12-13, L10-32, L19-20	10+10+23 +23 = 66
Parameters	**Bus Location for Voltages Phasors**		**Line Location for Current Phasors**		**No. of PMU Sensors**
	SD: 1.0e-05 p.u (Mag), 0.001 (Ang)		**SD: 1.0e-05 p.u (Mag), 0.001 (Ang)**		
Z_{PMU}	B2, B6, B9, B10, B11, B14, B17, B19, B20, B22, B23, B26, B37, B38		L2-1, L2-3, L2-25, L2-30, L6-5, L6-7, L6-11, L6-31, L9-8, L9-39, L10-11, L10-13, L10-32, L11-6, L11-10, L11-12, L14-4, L14-13, L14-15, L17-16, L17-18, L17-27, L19-16, L19-20, L19-33, L20-19, L20-34, L22-21, L22-23, L22-35, L23-22, L23-24, L23-36, L26-25, L26-27, L26-28, L26-29, L37-25, L38-29		14+39 = 53

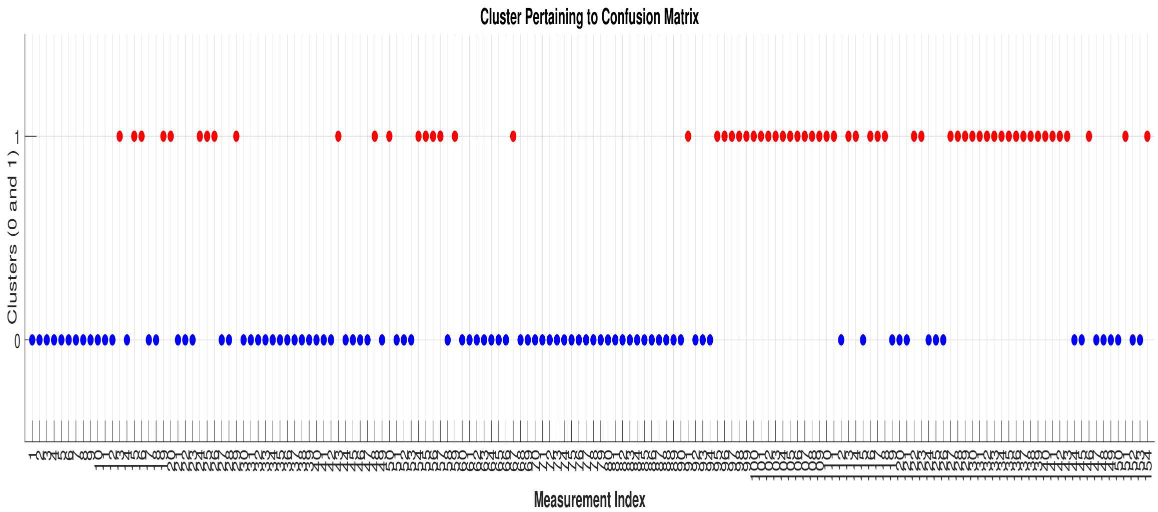
alarms. The resulted performance metrics includes: true positive rate of 91%, true negative rate of 97%, precision of 95.2%, accuracy of 94.2% and F1 score of 93.02%. To identify timing instances of potentially uncorrelated and compromised RTU measurement's arrival, based on columns features of residual matrix, the correlation coefficients are calculated as shown in correlation matrix of Fig. 3.7(d). By leveraging the heatmap's color intensity associated with elevated correlation coefficients, the timing of replaying historical recorded measurements are now clearly discernible.

3.5.2.2 Short Term Repetitive Data Cloning Attack (ST-RDCA)

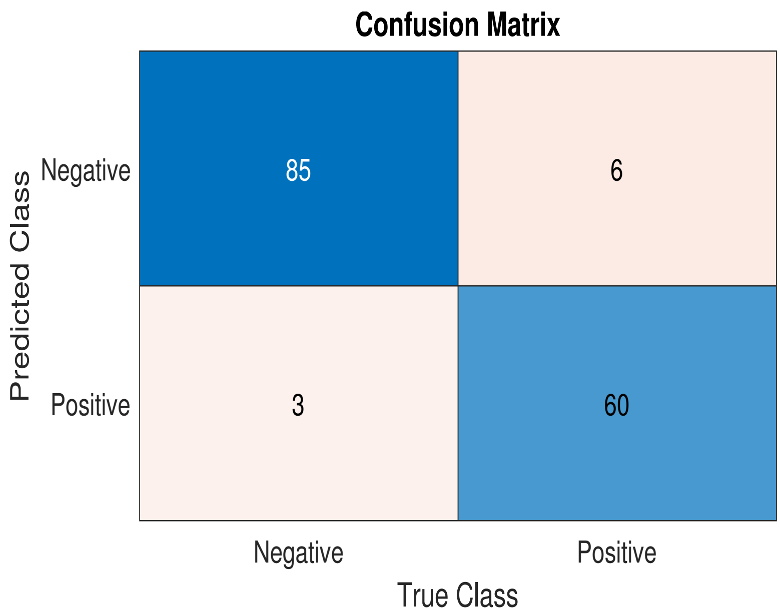
In the second variant of RDCA, the attacker initially recorded a 2 second windowed RTU data of a single phase to ground fault from any historical database and then these faulted RTU readings are being replaced with the normal RTU readings at 3 sec as per the Algorithm 5 to impersonate a normal event as a faulty one. Thus unlike previous attack cases, this attack features with limited duration but seemingly more hazardous as it could lead to false tripping of relays even if the system is in healthy condition. The proposed detection method effectively identifies such replay attacks, as evident from the



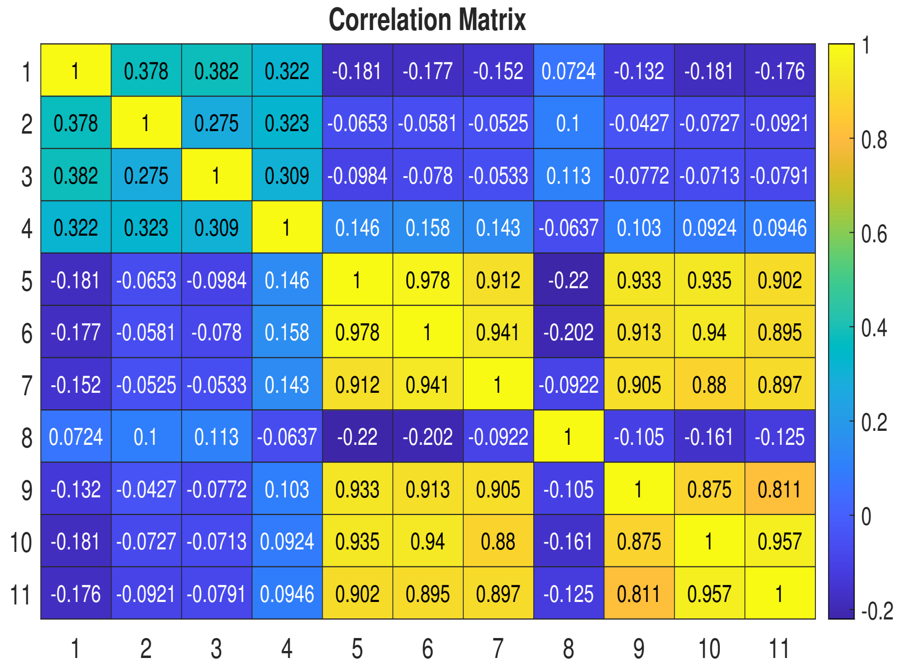
(a) Direct Residue of MDDA



(b) Cluster of Measurements



(c) Confusion Matrix



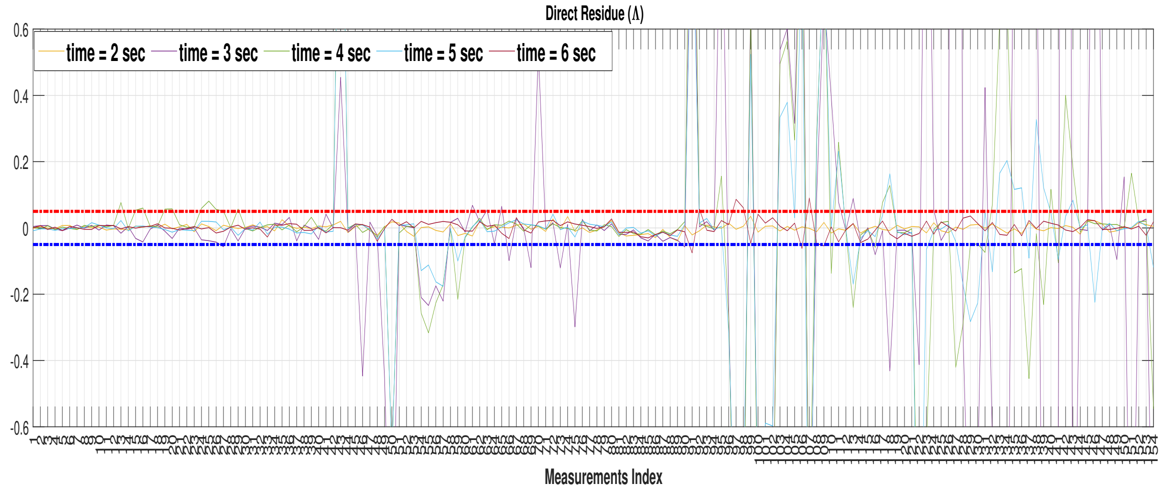
(d) Correlation Matrix

Figure 3.7: Detection phase of LT-RDCA (Algorithm 4) for NE 39-bus test system

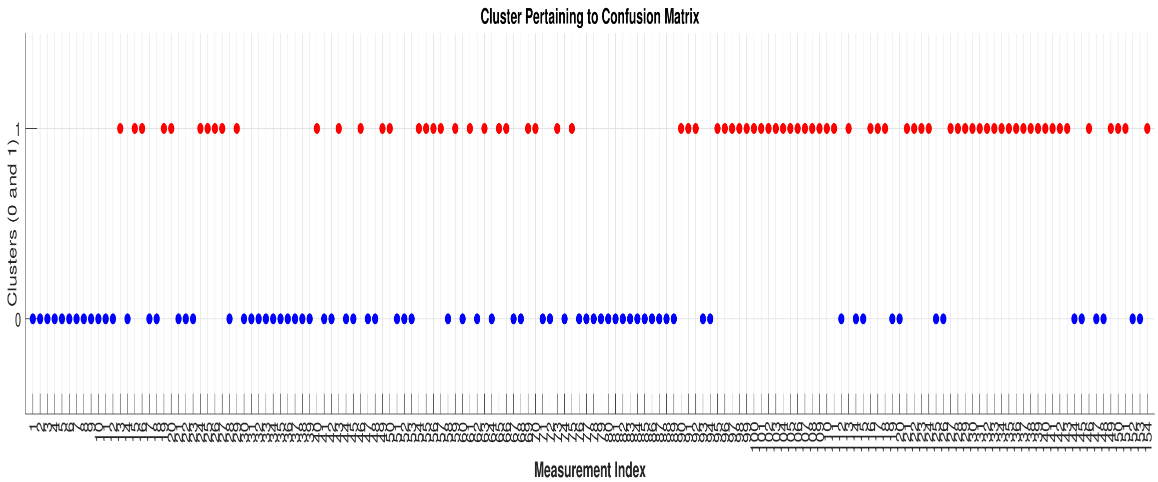
residue plot in Fig. 3.8(a). For the clear distinction of replay attack occurrence, in the above plot the pre-attack, during attack and post attack time legends are only shown. The distinct categorization between compromised RTU sensors data from the normal sensor data, offering a holistic view of measurement cluster during both routine and adversarial scenarios as depicted in Fig. 3.8(b). The confusion matrix in Fig. 3.8(c) reveals that application of proposed detection techniques over this test case results in 13 FPs and 1 FN in total. By analyzing the distribution of TP, TN, FP, and FN in the confusion matrix, a comprehensive evaluation of the model's success rates in detecting attacks can be accessed through the calculated values of other derived metrics such as true positive rate (98.49%), true negative rate (85.23%), precision (83.33%), accuracy (91%) and F1 score (90.3%). Lastly, the box plot in Fig. 3.8(d) offers detailed insights into the temporal dynamics of meter corruption, showcasing whiskers and individual data points. This visualization reveals high variability and skewed distributions in RTU sensor data during the initiation of replay attacks from 3 sec to 5 sec, highlighting anomalous behavior in detected meters.

3.5.2.3 Attack Correction

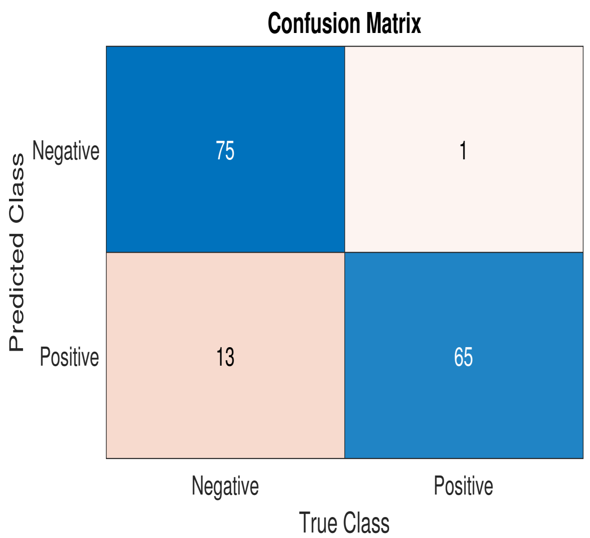
The correction phase of the proposed methodology against the aforementioned two attack variants of RDCA is depicted in Fig. 3.9. The vulnerable measurements identified in the detection phase are now corrected using the reconstructed RTU measurements, incorporating refined estimates of hybrid states with secured PMU measurements. The resulted residue plot of LT-RDCA, depicted in Fig. 3.9(a) is updated after incorporating corrected RTU measurements into the state estimation process. Notably, it is observed that the corrected residues consistently maintained a profile within the predefined range



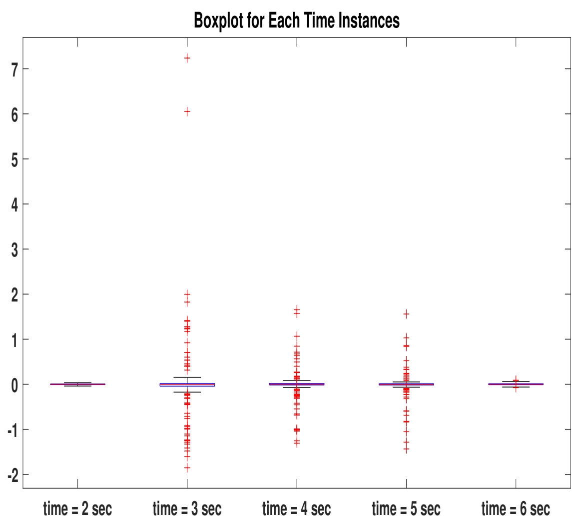
(a) Direct Residue of MDDA



(b) Cluster of Measurements

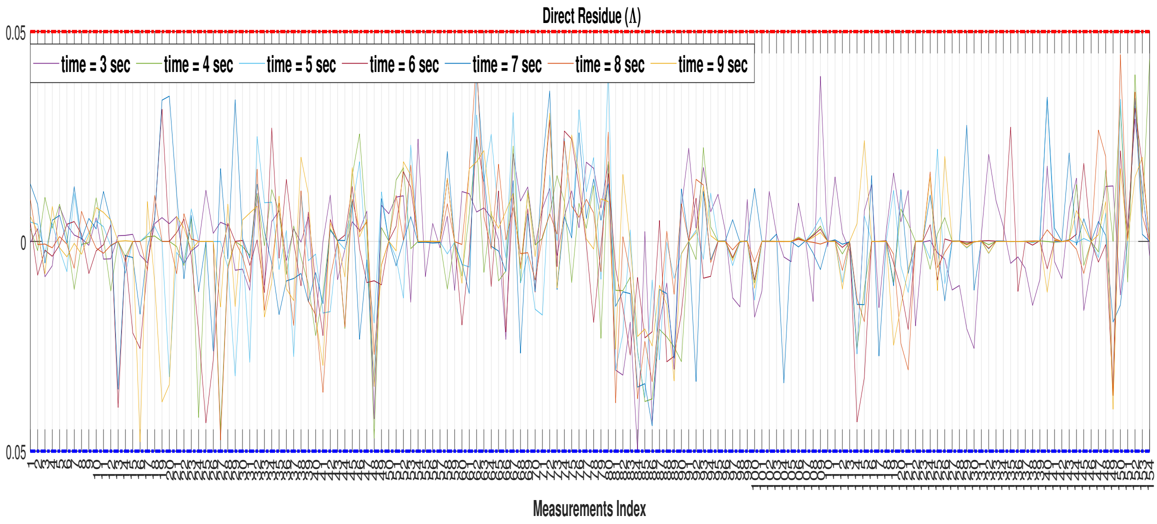


(c) Confusion Matrix

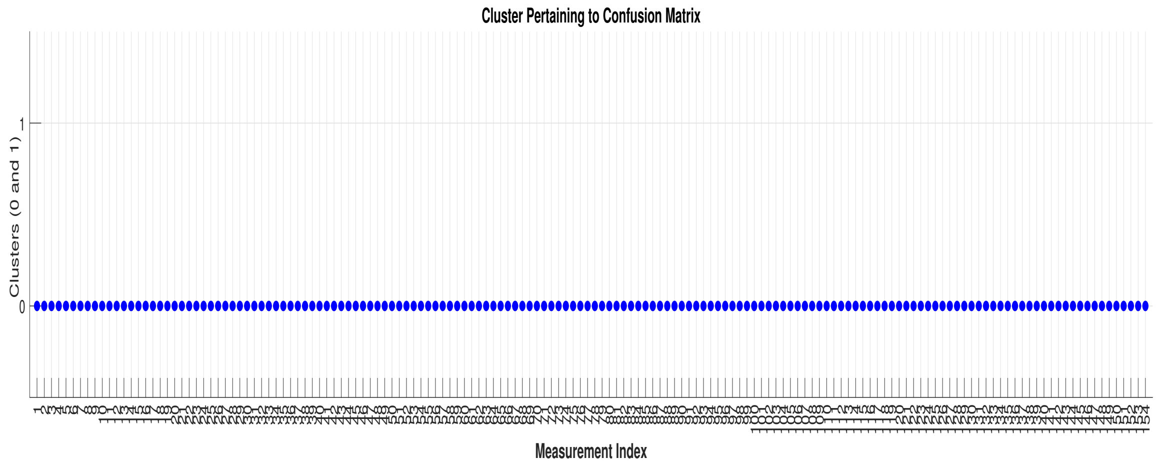


(d) Box Plot

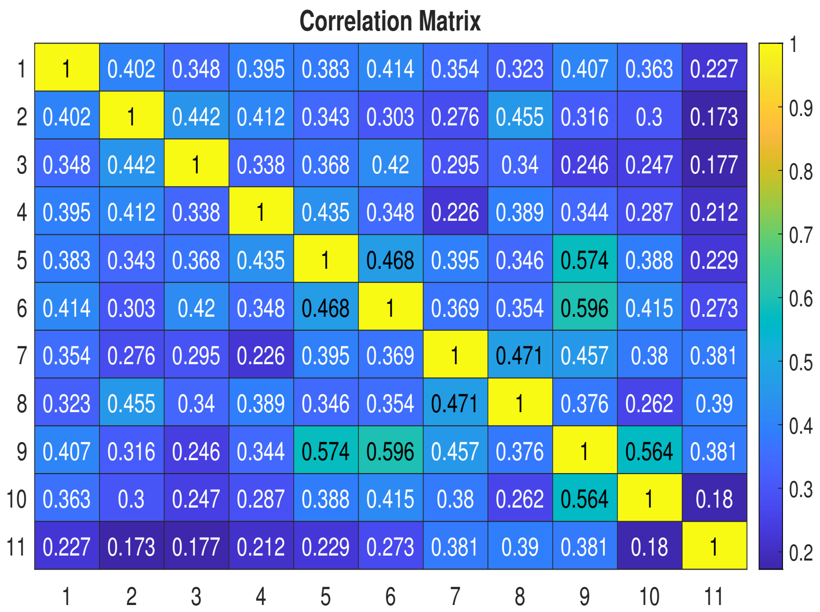
Figure 3.8: Detection phase of ST-RDCA (Algorithm 5) for NE 39-bus test system



(a) Direct Residue of LT-RDCA



(b) Cluster of Measurements



(c) Correlation Matrix for ST-RDCA

Figure 3.9: Correction phase of proposed scheme for NE 39-bus test system

delineated by the upper and lower level thresholds. This observation signified a successful mitigation of the detected attacks, as the corrected measurements no longer surpassed these established boundaries. As a result none of the RTU meters reading is now suspected as compromised which is clearly illustrated by all the measurements marked as blue dots in Fig. 3.9(b). Correlation matrix analysis of ST-RDCA is depicted in Fig. 3.9(c) reveals strong correlation coefficients only along the diagonal, indicating consistent and non-anomalous measurements at various time instances. Furthermore, for LT-RDCA correction, the maximum residue drop is approximately 96.5%, with a post-correction root mean square error of 0.4%. In ST-RDCA correction, the residue drop ranges between 97.81% (upper bound) and 82.87% (lower bound), and the root mean square error is 0.5%. The true negative rate for both the attack variant reach at 100% after the end of correction steps.

3.6 Conclusions

The aim of detecting and correcting replay-attacked measurements in the power system dynamic state estimator is achieved in this chapter based on developing three sequential stages. In Stage-1, based on nodal and branch power transfer distribution factors vulnerable RTU measurements are first identified, followed by corrupting those measurement values based on designing two novel RA models in Stage-2. Finally, in Stage-3, with the utilization of secured PMU sensors readings in a hybrid state estimator model, an attack detection and correction algorithm is developed to counteract the RAs. Rigorous testing on two popular standard IEEE test systems, i.e., IEEE-14 and NE-39, modelled in Real time Digital Simulators and the computational results of many performance indices has validated the effectiveness and remediation of proposed attack detection and correction strategy which in turn fortify resilience of cyber-physical systems against adversarial interventions. At the end, this chapter comes up with following salient features of the proposed method:

- The average detection rate for RDCA and MDDA is determined to be 94.6% and 90%, respectively. In the IEEE 14-bus system, there are a total of 8 FPs and 2 FNs for the MDDA and 2 FPs and 1 FN for the RDCA.
- The accuracy estimation for both MDDA and RDCA is found to be 82% and 93.27% respectively.
- The RMSE of estimated states of the estimator under MDDA and RDCA becomes 0.4% and 0.45%, respectively, after applying the attack correction algorithm.
- For both attack variations, the true negative rate constantly approaches 100%. This indicates the high specificity of the proposed algorithm in correctly identifying instances that do not belong to the attack class after correction.
- The resilience of the PSSE is enabled by secured PMUs, placed at approximately $1/3^{\text{rd}}$ of the system buses.

Chapter 4

Detection, Classification and Localisation of Cyber Attacks in Islanded AC Microgrid

4.1 Introduction

While Chapter-2 and Chapter-3 highlights the vulnerability of T-systems to cyber-attacks, it is also crucial to acknowledge the vulnerabilities that constantly faced by Distributed Energy Resources (DERs) at active distribution level network as well, particularly within the control mechanisms of Microgrids (MGs). MGs rely in seamless interconnections with various DERs for exchanging their own local information to the controllers and also highly interdependence with diversified communication and networking architecture for the efficient monitoring and control process. Consequently, now-a-days cyber criminals are increasingly shifting their attention to D-system, especially in the complex environment of islanded AC MG systems where the DERs controllers and its communication links are purposely targeted to be compromised, posing challenges for voltage and frequency stability. Thus, in D-System domain, particularity in the realm of MG, this chapter envisages for timely detection, classification and localization of cyber attacks which is of paramount importance for successful isolation of corrupted DERs from the MG topology under worst case situation. Thereafter, this detection and locational information will be utilized in next chapter to develop an effective resilient defense mechanisms to neutralize the cyber threats.

This chapter attains the above-mentioned objectives in 3 sequential steps. At first, Maximum Mean Discrepancy (MMD) based two-sample statistical hypothesis test is employed for analyzing and comparing distribution of the actual and estimated local frequency neighborhood tracking errors to detect the presence of misbehaving DERs or its corrupted incoming communication links. In next step, following detection of an attack, two statistical inconsistency measures—Shannon energy and entropy—are computed and used in a novel rule-based attack classification method that is integrated with the same XGBoost classifier to categorize different kinds of injection attacks in the controllers of the DERs. Upon classifying the nature of cyber-attacks, in the last step, this chapter introduces a multi-class attack localization scheme that leverages additional statistical features to be integrated into the XGBoost classifier. This facilitates the easier identification and quick isolation of compromised i.e., targeted DERs units from the system in the worst-case scenario.

This chapter is organized in six sections. In Section 4.2, a brief background of conventional primary control with communication based distributed secondary control of MG system is reviewed. Next, three consecutive sections i.e Section 4.3, Section 4.4 and Section 4.5 provides detailed description of the proposed methodology along with the real time digital simulation results pertaining to MMD based attack detection, proposed rule-based attack XGBoost-enabled attack classification and multi label attack localization scheme respectively. Finally, the key inferences are drawn and highlighted in conclusion Section 4.6.

4.2 Modelling Preliminaries of Islanded AC Microgrid

4.2.1 Cyber Graph Theory Terminology

Being a multiagent system [208], the distributed communication network of a microgrid are usually represented by a directed cyber graph topology such as $\mathcal{G} = (\mathbb{V}, \mathcal{E}, \mathscr{A})$, where each DERs are treated as vertices and the communication links associated with it are viewed as edges having a directed adjacency matrix $\mathscr{A}$. In such communicative microgrid environment, if the DER of vertex-j, v_j transmitting its own information to an another DER of vertex-i, v_i then the entity, a_{ij} of adjacency matrix $\mathscr{A}$ is defined as $a_{ij} = 1$, if $(v_j, v_i) \in \mathcal{E}$ otherwise $a_{ij} = 0$. The immediate neighbours of DER of vertex-i are represented mathematically by the set as $\bar{N}_i = \{j|(v_j, v_i) \in \mathcal{E}\}$. Every vertex of $\mathcal{G}$ is associated with in-degree and out-degree based on the information it received and transmitted respectively. Thus, the diagonal in-degree matrix of entire $\mathcal{G}$ is denied as $\mathscr{D} = diag\{d_i\} \in R^{N\times N}$ with $d_i = \sum_{j\in\bar{N}_i} a_{ij}$. This in-degree matrix indirectly helps to analyze the convergence rate of the system dynamics via calculating the graph Laplacian matrix as $\mathscr{L} = \mathscr{D} - \mathscr{A}$ with the assumption that the graph should have a spanning tree. If this assumption holds true, then $\lambda_1 = 0$ will be the simplest eigenvalue of $\mathscr{L}$ and $\tilde{\omega} = c\mathbf{1}$ becomes the solution to $\mathscr{L}\tilde{\omega} = 0$, c is any constant, which guaranteed the voltage and frequency synchronization of distributed secondary consensus law.

In such network multiagent system, there must be atleast one DER, who has the knowledge of MG's voltage and frequency reference values are considered to be as leader and its corresponding edge weight in the distributed secondary control is taken as $g_i \geq 0$ which is called as pinning gain [131].

4.2.2 Droop-Characteristics Based Primary Control

In an islanded operation, the voltage and frequency instability issues arise in MG due to the mismatch in power consumption and generation. In such critical scenario, power converters are used to operate in grid-forming mode where the primary control takes the control action for voltage and frequency regulation, and also to facilitate a proper active and reactive power sharing among parallelly operating voltage controlled voltage source inverters (VCVSI) without the use of communication links. This control can be achieved by designing the active power versus frequency and reactive power versus voltage droop characteristics as follows.

$$\omega_i = \omega_{n_i}^* - m_{P_i} P_i \tag{4.1}$$

$$\mathscr{V}_{odi} = V_{n_i}^* - n_{Q_i} Q_i \tag{4.2}$$

where, $\omega_{n_i}^*$ and $V_{n_i}^*$ are primary control nominal frequency and d-axis voltage reference respectively, obtained from secondary level control. m_{P_i} and n_{Q_i} are the active and reactive power droop coefficients, respectively which is selected based on converter power ratings and allowable maximum f and v deviations. P_i and Q_i are the filtered active and reactive power of i^{th} VCVSI, respectively [208, 145]. The output from the droop control acts as reference points for the internal zero level control loops of the VCVSI which generate switching modulation pulses for the DER's operation.

4.2.3 Communication Based Distributed Secondary Control

As the primary control of VCVSI only utilizes locally measured variables, it often leads to significant deviations in the global parameters of the MG system, such as frequency and voltage, from their reference values (ω_{ref} and $\mathscr{V}ref$). To address this, cooperative distributed secondary control leverages information from neighboring DERs through a defined cyber graph topology, enabling proportionate load power sharing and stable MG operation. To do so, the distributed secondary control provides the frequency and voltage set points, $\omega_{n_i}^*$ and $V_{n_i}^*$ in Eq. (4.1) and Eq. (4.2) for each DER-i in such a way such that the global frequency and voltage tracking synchronizing error quickly converges to zero as shown below:

$$\lim_{t \to \infty} \|\omega_i(t) - \omega_{ref}\| = 0 \quad \forall i \tag{4.3}$$

$$\lim_{t \to \infty} \|\mathscr{V}_{odi}(t) - \mathscr{V}_{ref}\| = 0 \;\; \forall i \tag{4.4}$$

Therefore, the secondary control inputs for the frequency and voltage of a distributive multiagent system can be written by differentiating the frequency and voltage droop characteristics in Eq. (4.1) and Eq. (4.2) respectively.

$$\dot{\omega}_n^* = \dot{\omega}_i + m_{P_i} \dot{P}_i = \mathcal{U}_{\omega_i}, \quad i = 1, ..., N \tag{4.5}$$

$$\dot{V}_{n_i}^* = \dot{\mathscr{V}}_{odi} + m_{Q_i} \dot{Q}_i = \mathcal{U}_{\mathscr{V}_i}, \quad i = 1, ..., N \tag{4.6}$$

where, $\mathcal{U}_{\omega_i}$ and $\mathcal{U}_{\mathscr{V}_i}$ are the frequency and voltage auxiliary control inputs for DER-i respectively. At this stage, the value of those control inputs at which synchronization of VSVCIs can be attainable if all DERs communicate its own information with its neighbours through a prescribed communication digraph $\mathcal{G}$ as.

$$\mathcal{U}_{\omega_i} = -c_\omega \delta_{\omega_i} \tag{4.7}$$

$$\mathcal{U}_{\mathscr{V}_i} = -c_{\mathscr{V}} \delta_{\mathscr{V}_i} \tag{4.8}$$

where δ_{ω_i} and $\delta_{\mathscr{V}_i}$ are the local neighbourhood synchronization errors for distributed secondary frequency control (DSFC) and distributed secondary voltage control (DSVC) respectively. $c_\omega \in \mathbb{R}$ and $c_{\mathscr{V}} \in \mathbb{R}$ are the control gains of DSFC and DSVC respectively which are choosen as per the following condition:

$$c_\omega = c_{\mathscr{V}} \geq \frac{1}{2\lambda^+_{min}(\mathscr{L} + \mathscr{G})} \tag{4.9}$$

where, $\mathscr{G}$ is the pinning gain matrix associated with communication graph $\mathcal{G}$ and λ^+_{min} is the minimum positive eigenvalue of matrix $(\mathscr{L} + \mathscr{G})$. This auxiliary control inputs $(\mathcal{U}_{\omega_i}, \mathcal{U}_{\mathscr{V}_i})$ for any DER-i are governed by a single integrator dynamics based on the v, f and power information of its own and its immediate neighbour as follows [209, 210]:

$$\delta_{\omega_i} = \sum_{j \in \bar{N}_i} a_{ij}(\omega_i - \omega^i_j) + g_i(\omega_i - \omega_{ref}) + \sum_{j \in \bar{N}_i} a_{ij}(m_{P_i}P_i - m_{P_j}P^i_j) \tag{4.10}$$

$$\delta_{\mathscr{V}_i} = \sum_{j \in \bar{N}_i} a_{ij}(\mathscr{V}_{odi} - \mathscr{V}^i_{odj}) + g_i(\mathscr{V}_{odi} - \mathscr{V}_{ref}) + \sum_{j \in \bar{N}_i} a_{ij}(n_{Q_i}Q_i - n_{Q_j}Q^i_j) \tag{4.11}$$

where, $\omega^i_j, \mathscr{V}^i_{odj}, P^i_j$ and Q^i_j are the frequency, voltage, active power and reactive power of DER-j, respectively that are being communicated to DER-i through separate channels. It is assumed that the MG is operating in an islanded mode with balanced loading and feeder model.

4.3 Cyber Attack Modelling and Proposed Attack Detection Scheme

Attack modeling and maximum mean discrepancy based detection mechanism are presented in this section for the distributed secondary control of the microgrid. Let's first define two usual definitions, that will be extensively used throughout the chapter.

Definition 1 (Compromised DER). *A compromised DER is one that is under direct attack.*

Definition 2 (Intact DER). *An intact DER is one that is not under direct attack or compromised.*

4.3.1 Attack Modeling

It is assumed that the adversary launches the FDI attacks in sensors, controllers or any decision-making units of the DERs in order to disrupt its operation and transmit the corrupted data to the control unit, affecting the MG data integrity and, thereby, jeopardizing its overall functioning. Based on the knowledge of distributive communication networks topology and DER's local information, perpetrators can hijack secondary controllers/corrupt the communication links which in turn, results in the auxiliary control inputs of each DERs being converged to some arbitrary wrong non-zero values that leads to frequency and voltage instability resulting collapse of the grid. A direct attack on the sensors or controller of the DER-i is modelled as, $\zeta_i^{attack} = \zeta_i + \Upsilon_i f_i^a$, where, ζ_i is the actual local frequency or voltage signal recorded by the sensors that is to be used by DSFC or DSVC for generating primary droop control reference, ζ_i^{attack} is the resulted compromised output after the data manipulation with attacker injected input denoted as f_i^a. Υ_i is unity when attack is initiated otherwise zero.

In a same manner, if the communication channel for outgoing frequency information of DER-i to DER-j is tampered with FDIA, then the malicious signal received by DER-j can be modelled as $(\omega_i^j)^a = \omega_i^j + \Upsilon_i f_i^a$, where, $(\omega_i^j)^a$ is the final manipulated information that has been transmitted to DER-j. Table 4.1 summarizes different types of attacks which have been studied and detected in the present work.

Table 4.1: FDIA Details

FDIA Types	Attack Signal Model (f_i^a)	Parameters
Step Attack	$f_i^a = \alpha_{st}$	α_{st} is a constant.
Ramp Attack	$f_i^a = \alpha_{ra}.t$	α_{ra} is a varying slope.
Scaling Attack	$f_i^a = \alpha_{sc}.y_i$	α_{sc} is scaling gain. α_{st} is the original signal.
Pulse Attack	$f_i^a = \alpha_{pu}(t),\ t \in \tau_p$	α_{pu} is a constant. $\check{\tau}_p$ is the attack duration
Sine Attack	$f_i^a = \alpha \sin \omega t$	α is a constant. ω is the injected frequency

4.3.2 Proposed Attack Detection Scheme

4.3.2.1 Maximum Mean Discrepancy (MMD)

At heart, the proposed cyber attack detection scheme exploits a two-sample distance-based measure called the maximum mean discrepancy (MMD) in a distributed cooperative secondary control of islanded MG. The distance of two distributions is calculated on the space of their probability measure based on the mean embeddings of two samples mapped into a reproducing kernel hilbert space (RKHS). In this chapter, basically, the most commonly used Gaussian radial basis function (RBF) kernel with kernel width ν, $\mathcal{K}(X_i, X_j) = \exp(-\|X_i - X_j\|^2/2\nu^2)$, which represents as a feature vector in some input space, is first applied over the two sample distributions and then the unbiased estimate of MMD is computed. RBF is a strictly continuous positive definite function. So, in formal sense, if P and $\mathcal{K}$ be the given probability measure and real-valued kernel defined on a topological space χ respectively, then the embedded mean for the samples drawn from P map to hilbert space $\mathcal{H}$ can be expressed as follows:

$$\mu_{\mathrm{P}} = \int_{\chi} \mathcal{K}(x,.)\; d\mathrm{P}(x) \tag{4.12}$$

where, x be the observation samples of distribution X. The expression of MMD can be easily represented in a more compact form by introducing the functional evaluation reproducing properties of RKHS in the following definition.

Definition 3 (Reproducing Kernal Property)**.** *Let, $\mathcal{H}$ be a Hilbert space of real valued functions on topological set χ i.e, $(\acute{f} : \chi \to \mathbb{R})$. Then the kernal function $\mathcal{K} : \chi \times \chi \to \mathbb{R}$ is called to be reproducing kernal of $\mathcal{H}$, if the following conditions holds [211]:*

1. *$\forall_x \in \chi$, $\mathcal{K}(x,.) \in \mathcal{H}$ and*
2. *$\forall_x \in \chi, \forall \acute{f} \in \mathcal{H}$, there must be a valid feature map $\varphi(x)$ from χ that map $\acute{f} \in \mathcal{H}$ to $\acute{f}(x) \in \mathbb{R}$ such that $\acute{f}(x) =< \varphi(x), \acute{f}(.) >_{\mathcal{H}}$ where $\varphi(x) = \mathcal{K}(x,.)$ represents the canonical representation of feature mappings.*
3. *In particular, x and y be the two samples drawn from two distributions that belongs to the non-empty set χ, then $\mathcal{K}(x,y) =< \varphi(x), \varphi(y) >_{\mathcal{H}} =< \mathcal{K}(x,.), \mathcal{K}(y,.) >_{\mathcal{H}}$.*

Therefore, by computing means via linearity, μ_P can also be expressed as expectation of feature map $\varphi(x)$ as follows:

$$\mu_{\mathrm{P}} := \mathbb{E}_{x \sim \mathrm{P}(x)}[\varphi(x)] = \mathbb{E}_{x \sim \mathrm{P}(x)}[\mathcal{K}(x,.)] \tag{4.13}$$

MMD, being a similarity measure, is applied over wide variety of problems, ranging from bio-informatics, neuroscience, machine learning to any other engineering applications to verify whether the two test samples defined on domain χ under study are statistically indistinguishable or not. The empirical estimates of MMDs heavily rely on the RBF class $\mathscr{F}(\acute{f} : \chi \to \mathbb{R})$ to be the unit ball in the universal RKHS, should have quick convergence and cheap computation, i.e for each m and n given points from two distributions, the cost is quadratic in time i.e $O(m+n)^2$.

Definition 4 (Maximum Mean Discrepancy)**.** *Let $\mathscr{F}$ be a class of functions, $\acute{f} : \chi \to \mathbb{R}$. $\mathcal{P} = \{p_1, p_2, ..., p_m\}$ and $\mathcal{Q} = \{q_1, q_2, ..., q_n\}$ are the observations that are being drawn independently and identically distributed (iid) from distributions X and Y be defined in the domain χ. Then the maximum mean discrepancy (MMD) can be written as [212, 213]:*

$$MMD[\mathscr{F}, X, Y] := \sup_{\acute{f} \in \mathscr{F}} (\mathbb{E}_{p \sim X}[\acute{f}(p)] - \mathbb{E}_{q \sim Y}[\acute{f}(q)]) \tag{4.14}$$

Now, from the properties described in Definition 1, one can write $\mathbb{E}_{p \sim X}[\acute{f}(p)] =< \mu_X, \acute{f} >$ and $\mathbb{E}_{q \sim Y}[\acute{f}(q)] =< \mu_Y, \acute{f} >$. Therefore, applying it Eq. (4.14) yields,

$$MMD[\mathscr{F}, X, Y] := \sup_{\|\acute{f}\|_{\mathcal{H}} \leq 1} < \mu_X - \mu_Y, \acute{f} > = \|\mu_X - \mu_Y\|_{\mathcal{H}} \tag{4.15}$$

As in practice, it is hard to compute expectations of μ_X and μ_Y, an empirical estimate of MMD is obtained by replacing the population expectations with empirical expectations on the sample of $\mathcal{P}$ and $\mathcal{Q}$ as follows:

$$MMD[\mathscr{F}, \mathcal{P}, \mathcal{Q}] := \sup_{f \in \mathscr{F}} \left(\frac{1}{m} \sum_{i=1}^{m} \acute{f}(p_i) - \frac{1}{n} \sum_{i=1}^{n} \acute{f}(q_i) \right) \tag{4.16}$$

The empirical estimates of MMD can be defined in a framework of statistical hypothesis testing where the computed estimates, $MMD[\mathscr{F}, \mathcal{P}, \mathcal{Q}]$ are compared with a predefined threshold γ. If the two distributions are found to be similar, $MMD[\mathscr{F}, \mathcal{P}, \mathcal{Q}]$ will be evaluated as zero and the null hypothesis $\mathscr{H}_0$: $X = Y$ gets accepted. On the other hand, if the distribution deviates far apart and become statistically distinguishable with loosing homogeneity, $MMD[\mathscr{F}, \mathcal{P}, \mathcal{Q}]$ result in crossing the threshold limit, which essentially means the alternative hypothesis gets accepted i.e $\mathscr{H}_1$: $X \neq Y$.

4.3.2.2 MMD based Cyber Attack Detection

Since MMD acts as a similarity measure to verify whether the two test series, defined on a domain χ, under study are statistically indistinguishable or not, it is applied on the frequency and voltage auxiliary control input signals of each DER in order to determine whether they are correctly participating as per the distributed secondary consensus-based protocol or not. While the controller of any DER or any of its incoming communication links is subjected to attacks, their local neighbourhood synchronization errors also get corrupted, resulting in a change in the statistical properties of the auxiliary control variables of DSFC and DSVC. As an example, under the typical satisfactory performance of DSFC, the frequency auxiliary control output ($\mathcal{U}_{\omega_i}$) of any DER, say DER-i can be represented by Eq. (4.7). But, while attackers penetrate into the MG multiagent system through any security breaches and take control of the frequency input of DSFC via hijacking the controller, then the auxiliary control variables of DER-i will get modified as,

$$\mathcal{U}^a_{\omega_i} = -c_\omega \delta^a_{\omega_i} \tag{4.17}$$

where, the previously clean local neighborhood frequency synchronization error (δ_{ω_i}) is now corrupted by some exogenous input Δ_i. This had been injected adversely due to the adversary's action while compromising the local frequency information of DER-i, (ω^a_i) used as an input for secondary control of the VSVCI inverter.

$$\delta^a_{\omega_i} = \delta_{\omega_i} + \Delta_i \tag{4.18}$$

$$\delta^a_{\omega_i} = \delta_{\omega_i} + \left[\left(\sum_{j \in \bar{N}_i} a_{ij} + g_i \right) \omega^a_i - \sum_{j \in \bar{N}_i} a_{ij} \omega^i_j \right] \tag{4.19}$$

Similarly, for each incoming link related to a particular DER, the corrupted auxiliary control inputs are computed in order to identify compromised communication lines. In general, the corrupted frequency auxiliary control required for the MMD calculation under attack in the communication links can be written as,

$$\delta^a_{\omega_i^j} = \delta_{\omega_i} + \left[\left(\sum_{j\in\bar{N}_i} a_{ij} + g_i \right) \omega_i - \sum_{\substack{j\in\bar{N}_i \\ j\neq k}} a_{ij}\omega^i_j - a_{ik}(\omega^i_k)^a \right] \tag{4.20}$$

$$\mathcal{U}^a_{\omega_i^j} = -c_\omega \delta^a_{\omega_i^j} \tag{4.21}$$

It is to be noted that in the presence of an attack, based on the communication cyber graph topology and distributed secondary cooperative control framework, one can easily inspect the corrupted frequency of DERs and thus, based on this corrupted frequency, compromised auxiliary controls ($\mathcal{U}^a_{\omega_i}$ *or* $\mathcal{U}^a_{\omega_i^j}$) can also be observed [116]. But the overall exogenous input, (Δ_i *or* Δ_i^j) injected by the attacker, is not required to be known or not measurable explicitly. Thus based on the statistical properties inferred from the both $\mathcal{U}^a_{\omega_i}$ and $\mathcal{U}_{\omega_i}$, one can compute the unbiased empirical MMD estimates by taking squares of Eq. (4.15) and then applying reproducing kernel properties to detect compromised DER as follows.

$$\begin{aligned} \|\mu_X - \mu_Y\|^2_{\mathcal{H}} &= <\mu_X - \mu_Y, \mu_X - \mu_Y>_{\mathcal{H}} \\ \|\mu_X - \mu_Y\|^2_{\mathcal{H}} &= <\mu_X, \mu_X>_{\mathcal{H}} + <\mu_Y, \mu_Y>_{\mathcal{H}} -2\times <\mu_X, \mu_Y>_{\mathcal{H}} \\ \|\mu_X - \mu_Y\|^2_{\mathcal{H}} &= \mathbb{E}_{X,X} <\varphi(\mathcal{U}_{\omega_i}), \varphi(\mathcal{U}'_{\omega_i})>_{\mathcal{H}} + \mathbb{E}_{X,X} <\varphi(\mathcal{U}^a_{\omega_i}), \varphi(\mathcal{U}^{a\prime}_{\omega_i})>_{\mathcal{H}} \\ &\quad - 2\mathbb{E}_{X,Y} <\varphi(\mathcal{U}_{\omega_i}), \varphi(\mathcal{U}^a_{\omega_i}))>_{\mathcal{H}} \end{aligned} \tag{4.22}$$

Substituting empirical estimates of the features spaces based on samples from $\mathcal{U}_{\omega_i} = \{u_1, u_2, ..., u_m\}$ and $\mathcal{U}^a_{\omega_i} == \{u^a_1, u^a_2, ..., u^a_m\}$, the final expression would be written as,

$$\begin{aligned} MMD[\mathscr{F}, \mathcal{U}_{\omega_i}, \mathcal{U}^a_{\omega_i}] := \Bigg[&\frac{1}{m^2} \sum_{s,t=1}^{m} \mathcal{K}(u_s, u_t) - \frac{2}{m^2} \sum_{s,t=1}^{m} \mathcal{K}(u_s, u^a_t) \\ &+ \frac{1}{m^2} \sum_{s,t=1}^{m} \mathcal{K}(u^a_s, u^a_t) \Bigg]^{\frac{1}{2}} \end{aligned} \tag{4.23}$$

Likewise, the communication link attack can also be detected by replacing the $\mathcal{U}^a_{\omega_i}$ with $\mathcal{U}^a_{\omega_i^j}$. Finally, the effect of the attacks on either DSFC or DSVC can be easily detected when the empirical MMD estimates cross some predefined design threshold γ.

4.3.3 Real-Time Digital Simulation Results

The efficacy of the proposed cyber attack detection scheme using the RTDS under various types of attacks are listed in Table 4.1 of Section 4.3. The experiments were conducted on a modified IEEE 13-node distribution feeder system, as depicted in Fig. 4.1, with averaged line parameters [214]. The detailed description of the modified test system with line specifications can be found in Appendix A.1. The system operates at a nominal frequency of 60 Hz and a line-to-line voltage of 4.16 kV. Four DERs, each with equal active power (1 p.u) and voltage rating (1 p.u), are interconnected via a 1.0 MVA, 0.48/4.16 kV Yg-Yg transformer, supplying power to an islanded AC MG. DER-3 is designated as the leader with a pinning gain of $g_3 = 1$, and communication between DERs is facilitated by a specified communication digraph. Conventional secondary control for voltage and frequency, based on Eq. (4.10) and Eq. (4.11), is implemented with control gains $c_{\mathcal{V}}$ and c_ω taken as 10 and 20, respectively. Real and reactive power droop coefficients are set to 2×10^{-4} and 1×10^{-3} respectively. The converter filter resistance (R_f), inductance (L_f) and capacitance (C_f) is set to 0.02 ohms, 100 micro-henry and 50 micro-farad respectively. Figure 4.2 shows the schematic laboratory hardware setup where the auxiliary control inputs of each DER are collected from the RTDS front analogue output panel, and then the proposed MMD-based attack detection algorithm has been converted to C code via MATLAB Simulink® C Coder builder, which is next compiled and run into DS1104 R&D controller board to observe its output via oscilloscope's screen.

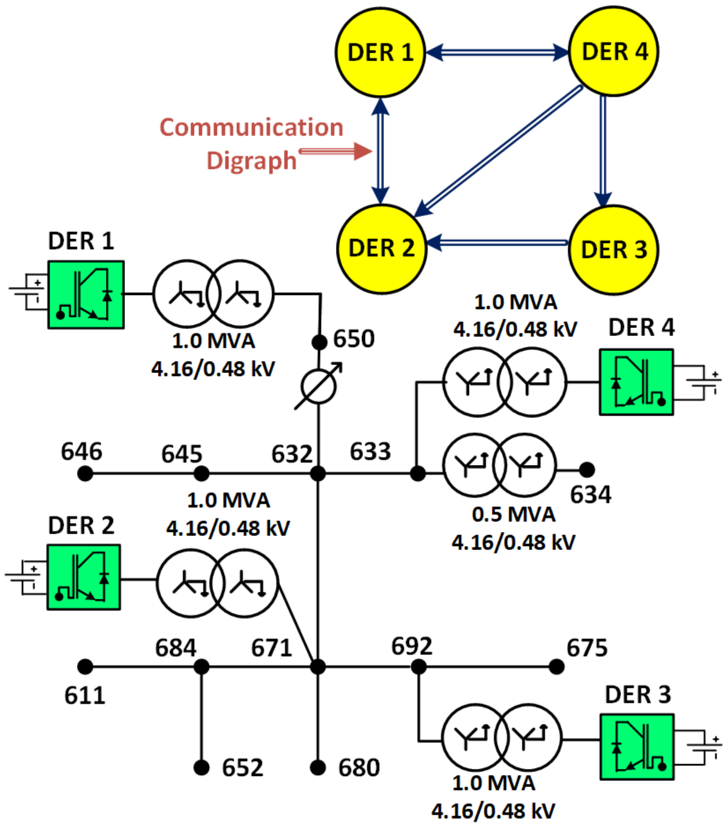


Figure 4.1: Single line diagram of the IEEE 13-node Microgrid test system

4.3.3.1 Single Attack Detection on Conventional DSFC

In the first case study *(Case A)*, a step-type attack is modeled to breach DER-1's DSFC as shown in Fig. 4.3. As seen in Fig. 4.3(a), a step attack with parameter $\alpha_{st} = 0.05$ pu is injected to the frequency input of the DSFC at about 3.4 seconds which causes a large variation in the local frequencies of all DERs below 220 rad/sec. Furthermore,

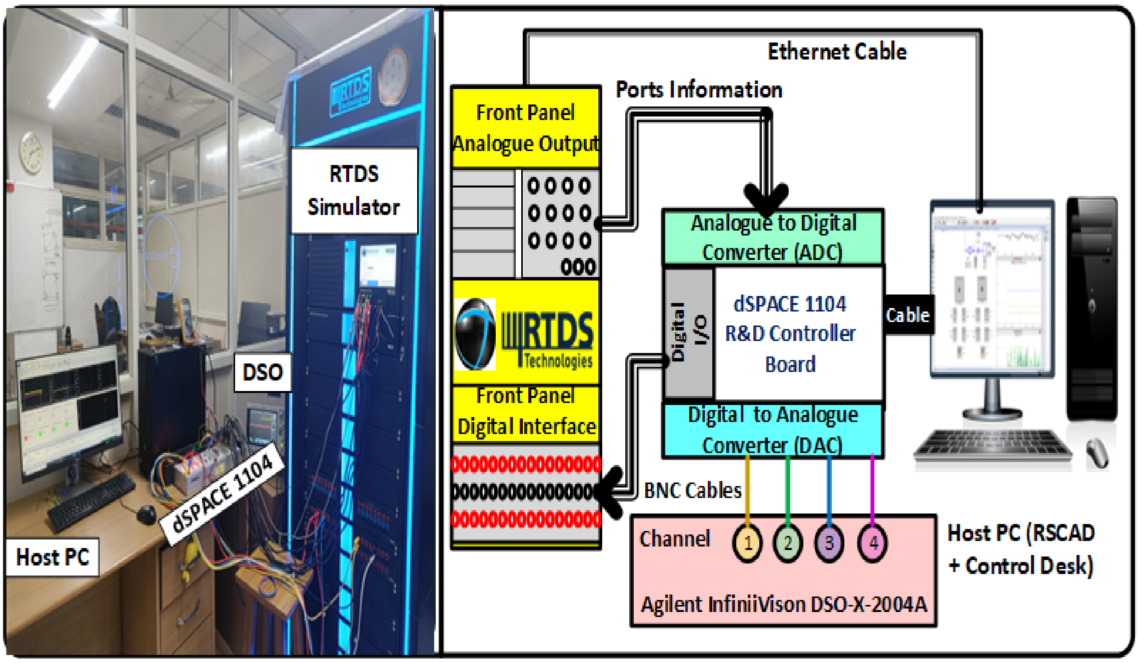
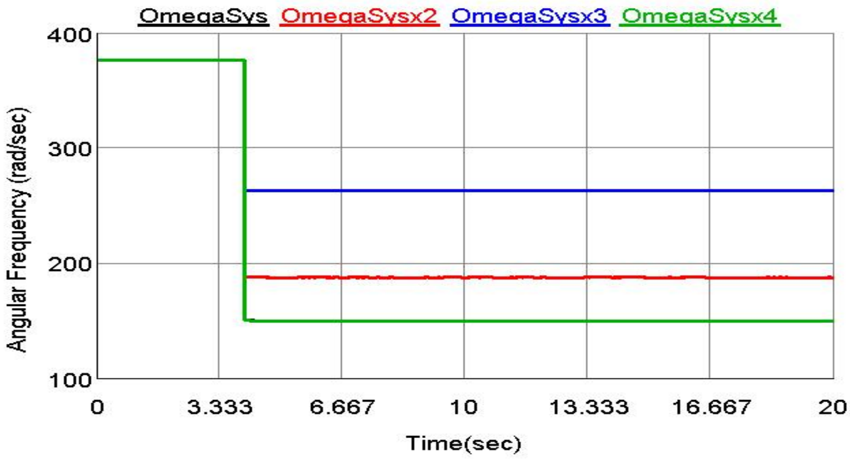


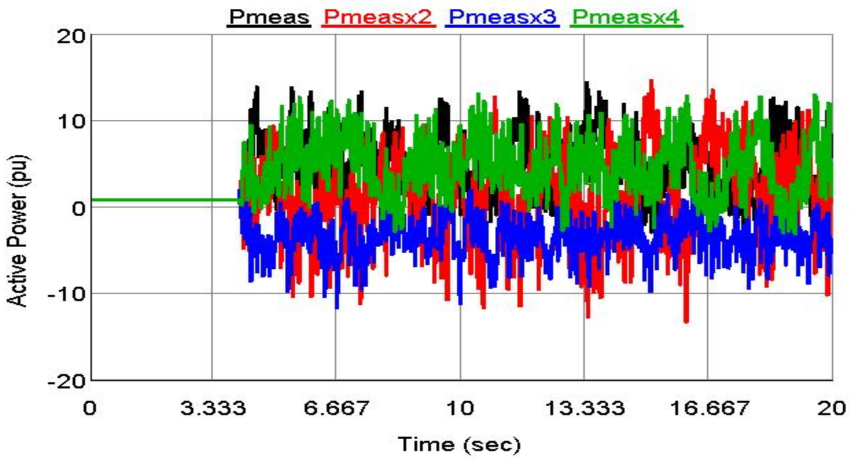
Figure 4.2: RTDS setup for HIL validation of the proposed scheme

Fig. 4.3(b) demonstrates the impact of this attack on the active power sharing among the individual inverters, revealing a consequential disruption in grid stability. Notably, DER-3, functioning as the leader DER with knowledge of the grid's reference voltage and frequency, experiences comparatively less impact than others. Nevertheless, as Figs. 4.3(a) and 4.3(b) show, it is difficult to identify compromised DERs because of the distributive cooperative consensus law that regulates DER interactions and the particular communication graph topology as depicted in Fig. 4.1. Additionally, Fig. 4.3(c) showcases the response of proposed MMD-based detector under attack condition, focusing on the irregularities in observed frequency auxiliary control input of DER-1. Since this attack is localized to DER-1, only the compromised inverter's MMD shows a notable rise, suggesting that local control variables differ from those of other DERs.

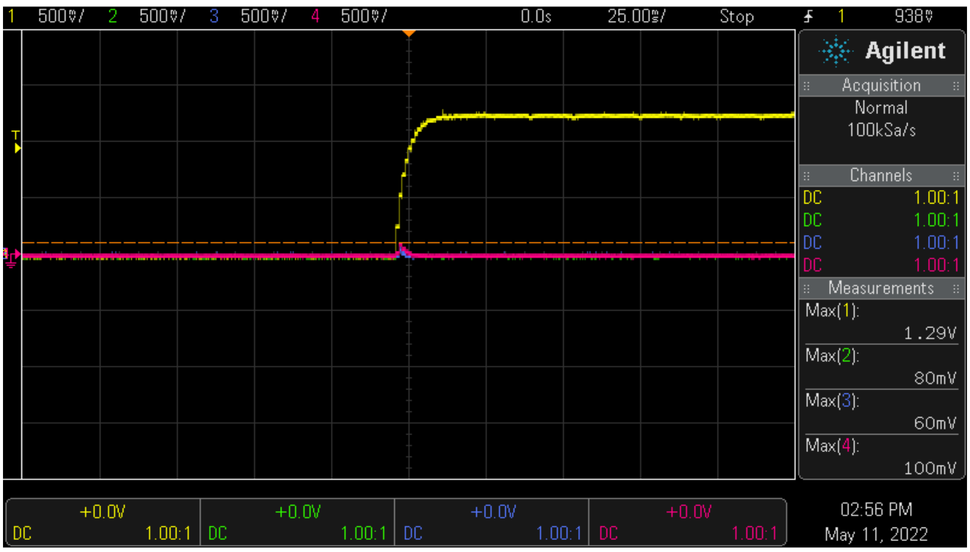
In the second case study *(Case B)*, a pulse type of attack with specific parameters of $\alpha_{pu} = 0.06$ p.u magnitude and $\check{\tau}_p = 0.5$ sec duration are being modeled and then injected to the controller of DSFC of leader DER-3 to compromise it frequency output as shown in Fig. 4.4. Similar to previous case, this manipulation results in significant deviation of local frequency of all DERs below 340 rad/sec as shown in Fig. 4.4(a). Figure 4.4(b) shows the impact of sharing active power among the inverters as a dire consequence of this periodic pulse attack, where the leader DER, i.e., DER-3 knows the overall reference voltage and frequency of the grid. Therefore, it drives down the frequency of all DERs in a coherent manner. But as all the DERs are also participating among themselves, obeying distributive cooperative consensus law for maintaining the grid stability through some specific predefined communication digraph, therefore, the direct attack impact of compromised DERs is inflicted differently to other DERs and thus correct identification of compromised DERs are also found to be very difficult as shown



(a) Frequency

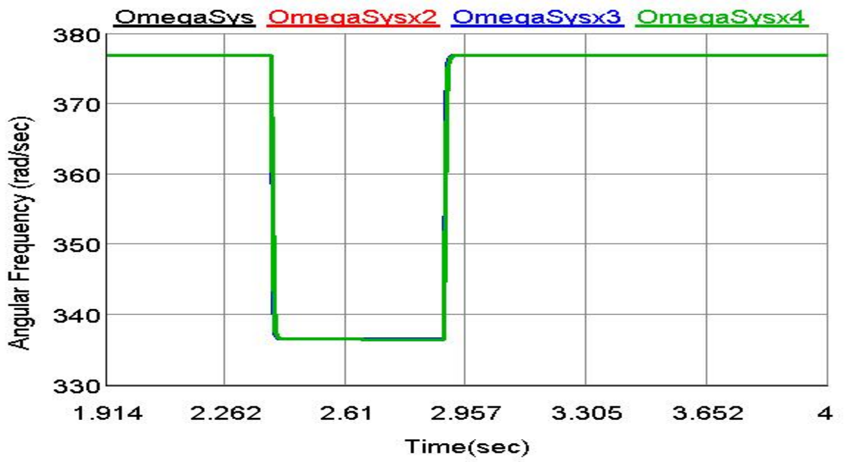


(b) Active Power Sharing

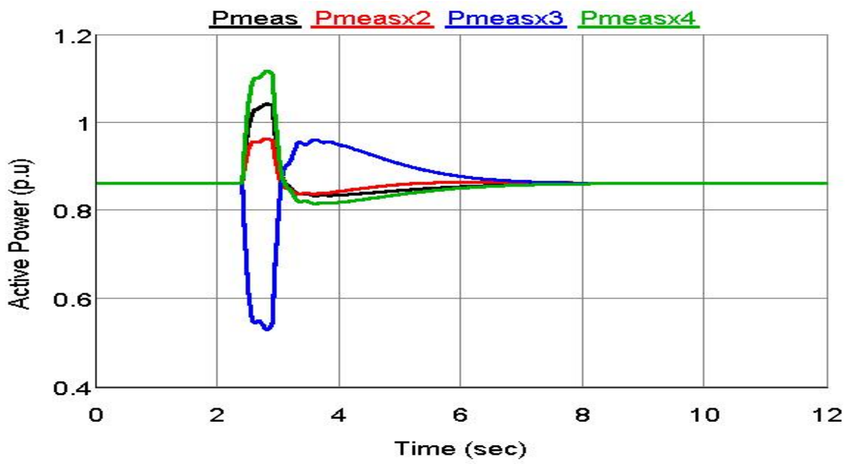


(c) Estimates of MMD Under Step Attack

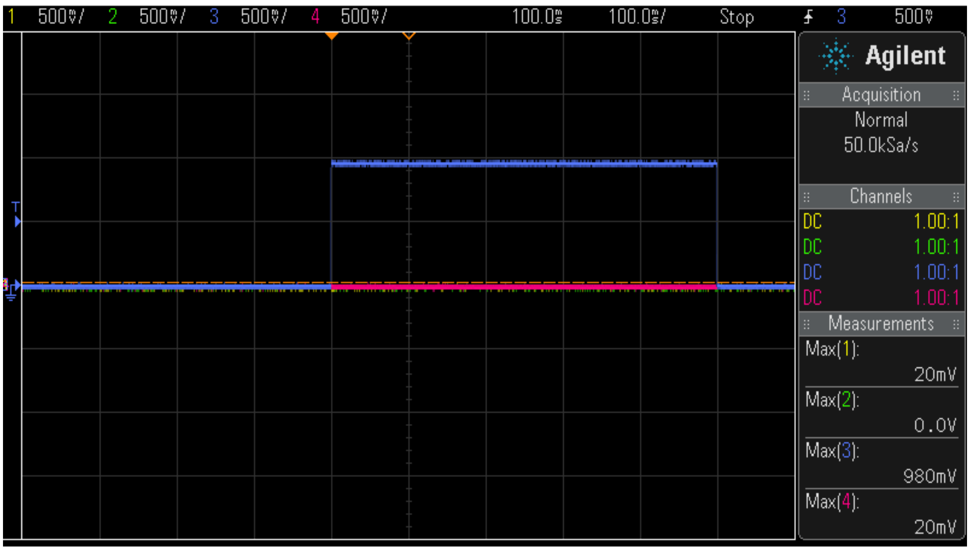
Figure 4.3: Case A: Effect of Step Attack on DSFC of DER-1 and MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.



(a) Frequency



(b) Active Power Sharing



(c) Estimates of MMD Under Pulse Attack

Figure 4.4: Case B: Effect of Pulse Attack on DSFC of DER-3 and MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.

in Fig. 4.4(a) and 4.4(b). Figure 4.4(c) exhibits the behaviour of MMD based detector under attack, targeting frequency auxiliary control input of DER-3. It is clearly revealed that the proposed MMD successfully captured the drift created between the actual and estimated local synchronization error after the attack launch which results in fulling the criteria of accepting null hypothesis as discussed in Section 4.3. As a results, MMD of the compromised inverter rises up to a significantly higher value due to disparity in local control variables, flagging out the alarm of cyber attack detection.

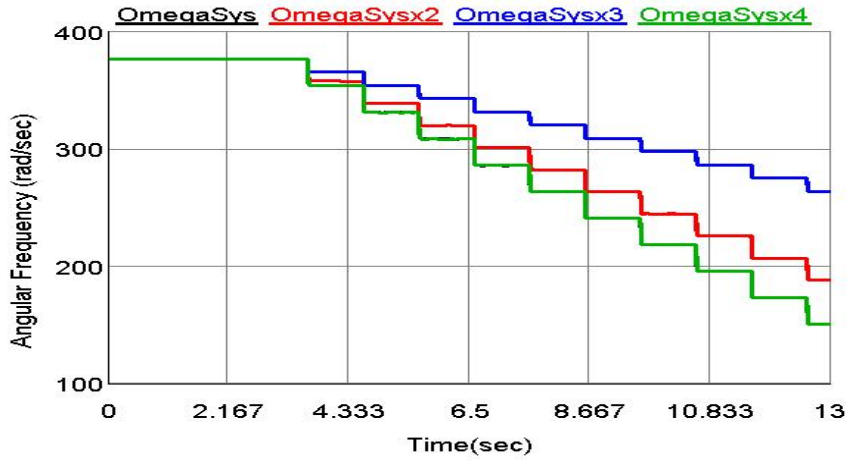
In the third case study *(Case C)*, a time-varying and non-linear sine attack signal ($f_i^a = \sin(0.01t)$) is injected into the DSFC of DER-1 around 3.5 seconds, aiming to disrupt the functioning of primary droop control techniques as shown in Fig. 4.5. Initially, all DERs' frequency and output power are regulated by conventional secondary control. However, upon injection of the malicious signal, consensus agreement is lost, leading to unstable frequency and active power exceeding acceptable limits across all DERs as shown in Fig. 4.5(a) and 4.5(b). This destabilization occurs as the victim DER begins to share falsified frequency, voltage, and power outputs with other DERs due to the addition of this time-dependent ambiguous input, disrupting their consensus protocols. As the aforementioned frequency attack model is updated with each second intervals, the simulation shows that the frequency and active power response exhibits a stair-case and oscillatory pattern, respectively. However, with the objective of correct and on-time identification of compromised units, the proposed detector successfully detect the attack instant based on observing the discrepancies in their embedding mean of theirs actual and observed auxiliary frequency control variables as shown in Fig. 4.5(c).

4.3.3.2 Single Attack Detection on Conventional DSVC

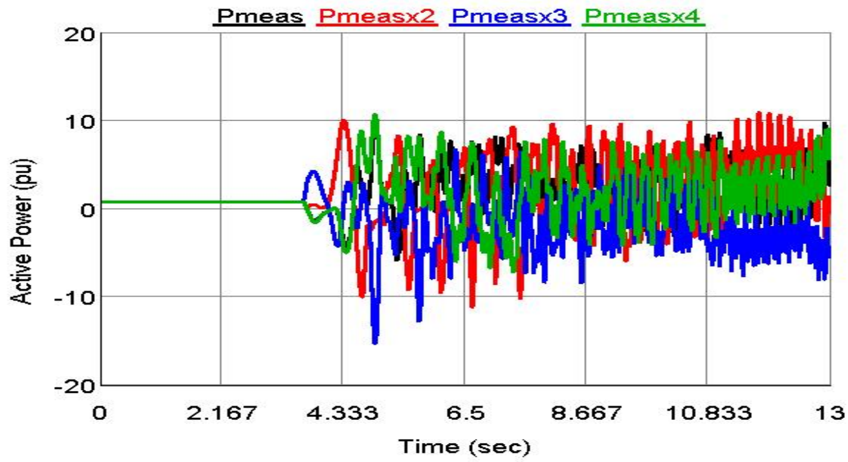
Similar to the previous case study, in this test scenario *(Case D)*, the attacker targets the secondary voltage controller of DER-2 as shown in Fig. 4.6. They're manipulating the DER's controller input signal with a scaling attack of attack parameter: $\alpha_{sc} = 1.05$ pu, on the DSVC. The on-set of the attack significantly impacts the voltage and reactive power profiles of the inverter, as depicted in Fig. 4.6(a) and Fig. 4.6(b). This manipulation leads to a erroneous voltage and reactive power response suggests all DERs lose their coordinated control, with global parameters deviating from normal due to the constant injection of false data. Failure to promptly address this issue and remove the compromised DER from the topology risks driving the microgrid towards instability. In this regard, it is observed from Fig. 4.6(c), that the proposed MMD algorithm demonstrates exceptional ability to identify such attacks and pinpoint the faulty DER.

4.3.3.3 Attack Detection for Simultaneous Attacks on Multiple DERS

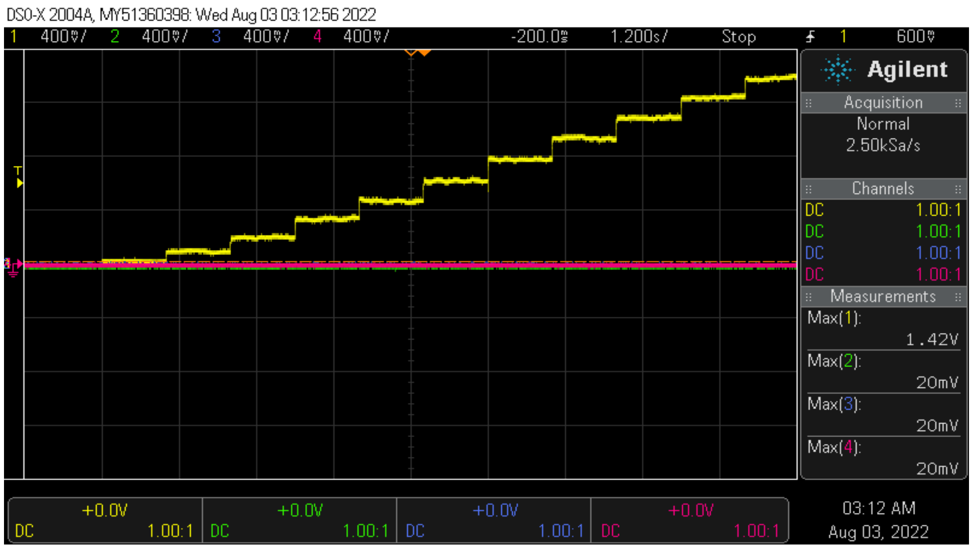
Similar to the case studies described earlier, in this test cases *(Case E)* and *(Case F)*, the DSFC of multiple DERs are now being targeted at the same time with a ramp type of signal first followed by a pulse attack as depicted in Fig. 4.7 and 4.8. In the first case, a slowly varying ramp attack signal, dynamically generated from a step signal with



(a) Frequency

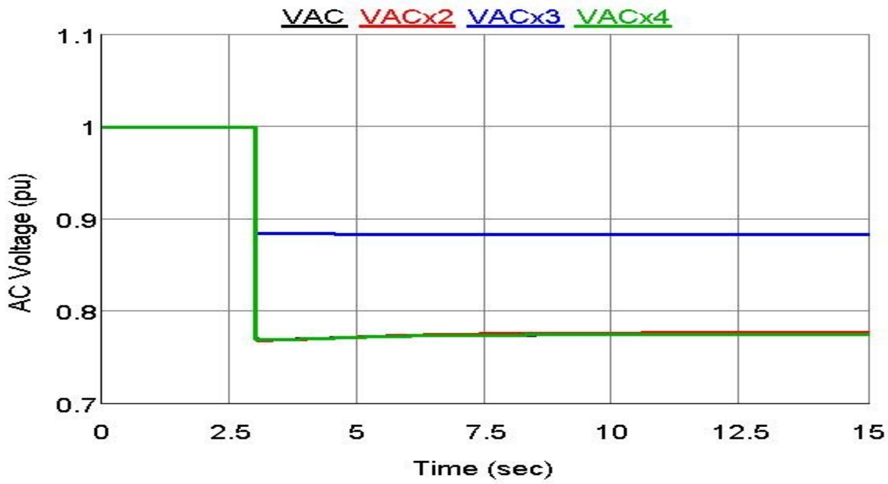


(b) Active Power Sharing

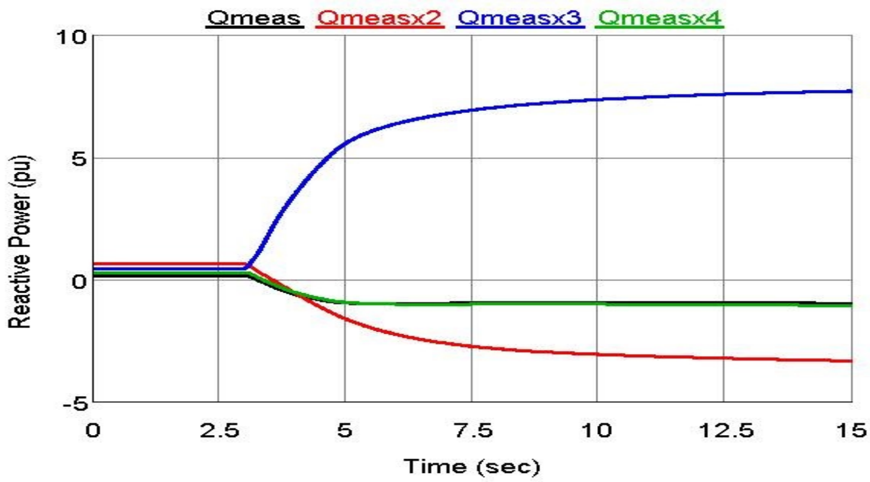


(c) Estimates of MMD Under Sine Attack

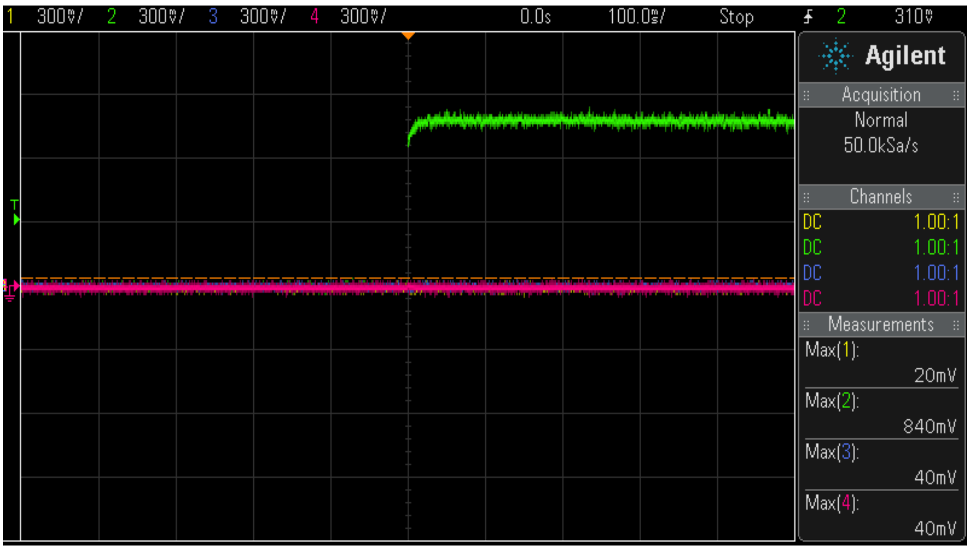
Figure 4.5: Case C: Effect of Sine Attack on DSFC of DER-1 and MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.



(a) DER Output Bus Voltage



(b) Reactive power sharing



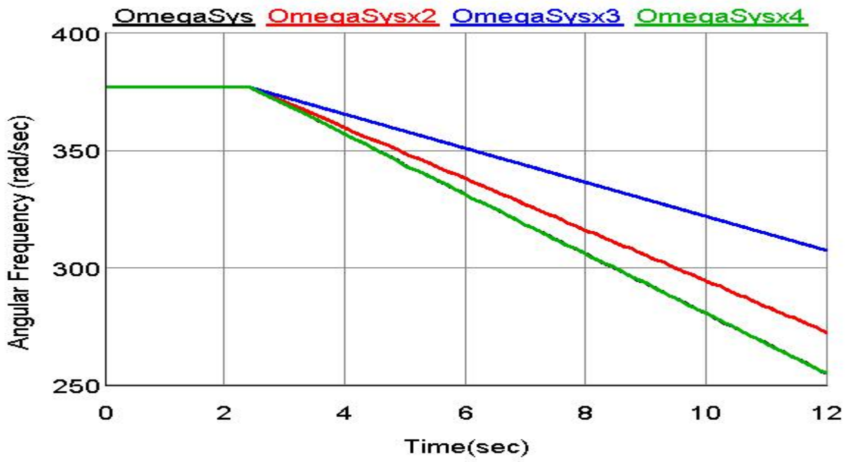
(c) Estimates of MMD Under Scaling Attack

Figure 4.6: Case D: Effect of Scaling Attack on DSVC of DER-2 and MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.

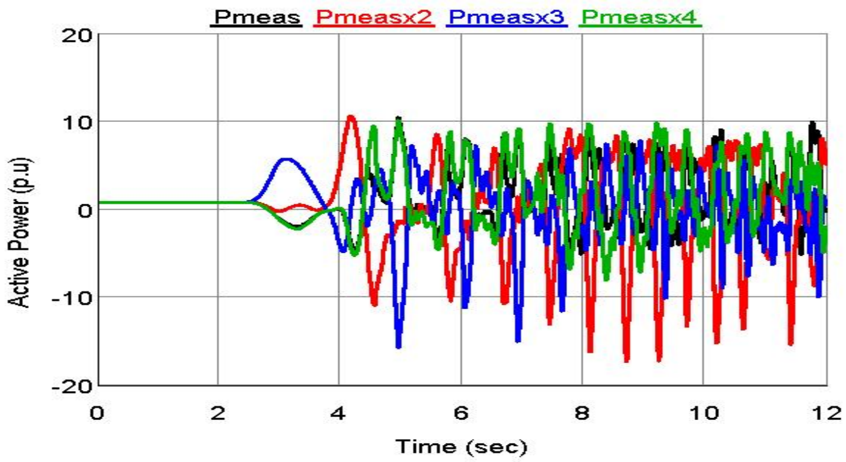
parameters $\alpha_{ra} = 1.2$ pu and a time constant of 500 sec, is injected into the frequency control of DER-1 and DER-3. This aims to disrupt the global parameters of all DERs and push them away from their desired consensus. And in the second case study, a pulse attack of magnitude $\alpha_{pu} = 0.04$ pu with duration of 0.5 sec are also injected to dishonestly alter the frequency setpoints for the primary control action of both the DERs affecting the MG's frequency and voltage stability. As shown in Fig. 4.7(a), 4.7(b) and Fig. 4.8(a), 4.8(b) the frequency and power of all DERs significantly disrupted due to such unbound attack effect. On the other hand, Fig. 4.7(c) and 4.8(c) shows the competence of the proposed MMD-based detector to correctly identify the DERs, victim of such severe attack. Here also, it is noticed that as this malignant attack is limited to DER-1 and DER-3, only the compromised inverter's MMD shows a notable rise, suggesting that local control variables differ from those of other DERs. Henceforth, this demonstrates the detector's effectiveness even in handling multi-DER attacks occurring simultaneously.

4.3.3.4 Attack Detection on Communication Links

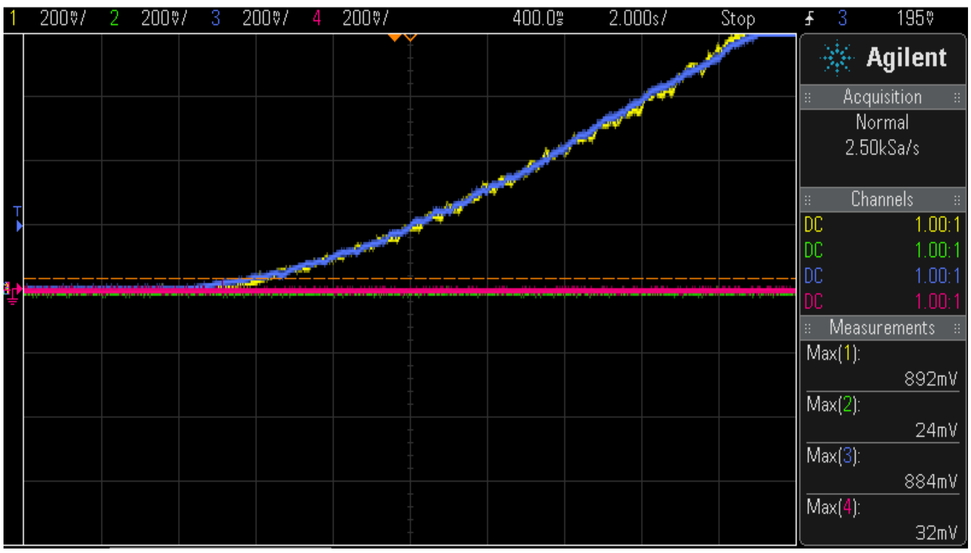
This case study *(Case G)* examines the impact of short-length pulse and slow-varying ramp frequency attacks on the communication links of an inverter, leading to deviations in DER parameters beyond acceptable ranges, challenging the detection of the compromised DER as illustrated in Fig. 4.9. The case studies is divided into two subparts: (a) Single communication link attack and (b) Multiple communication links attack. In the first subpart, a short-length pulse attack signal with a parameter of $\alpha_{pu} = 0.06$ pu, falsifies the frequency information of leader DER-3 while communicating with DER-2 for 0.5 sec, resulting in a significant frequency deviation, as depicted in Fig. 4.9(a). Since DER-3 serves as the reference for achieving consensus among all other DERs in a networked control MG system, compromising its outgoing information has a pronounced effect on the other DERs due to the rapid propagation of attack signals. Consequently, all DERs exhibit similar behavior to the corrupted DER. In this compromised scenario, the proposed detection scheme's performance is validated by calculating MMD for all the working links. Figure 4.9(b) illustrates that the MMD remains nearly zero for all intact communication links except the corrupted one, underscoring the proposed MMD's efficacy in identifying compromised links even when the leader DER information itself is compromised. On a similar line, another case study is being conducted in Fig. 4.9 where all the outgoing communication links of DER-1 are compromised with a slow varying ramp frequency attack signal of parameter $\alpha_{ra} = 0.06$ pu. With multiple attacks on DER-1's outgoing communication channels, the integrity of the communication graph topology is significantly compromised, resulting in the rapid and widespread dissemination of attack signals to other healthy DERs as depicted in Fig. 4.9(c). As the MMD calculation is now based on localized estimates of corrupted auxiliary control variables for each link, thus, as a next corrective step, the MMD calculation is carried out for all the available links. Clearly, Fig. 4.9(d) depicts that due to disparity of information received via any link with respect to the other neighboring links for each particular DER causes the respective link's MMD to



(a) Frequency

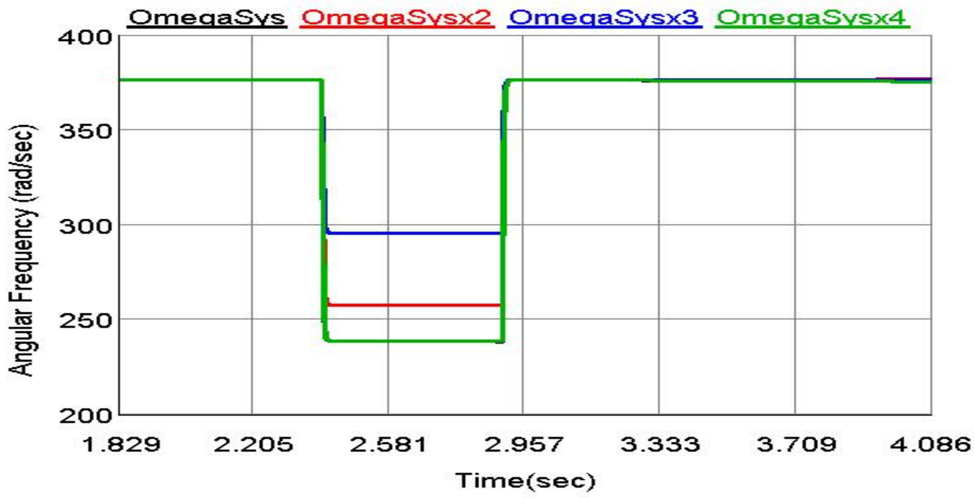


(b) Active Power Sharing

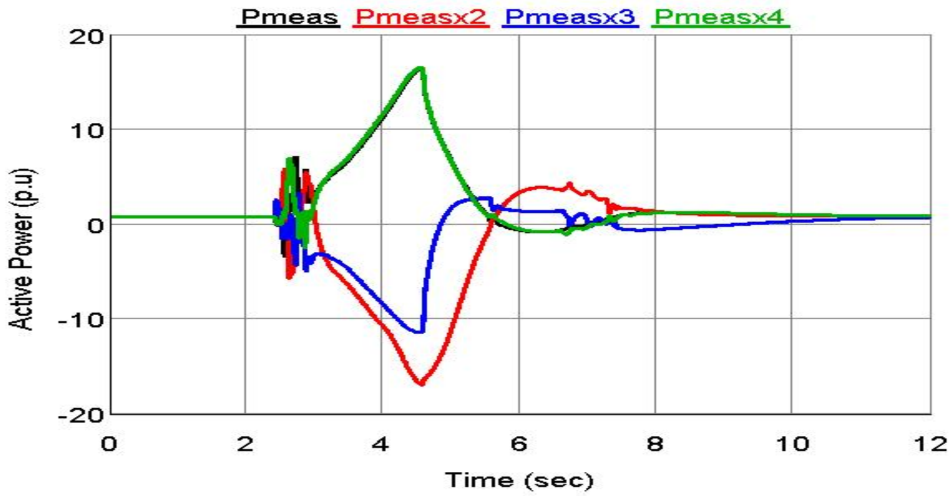


(c) Estimates of MMD Under Ramp Attack

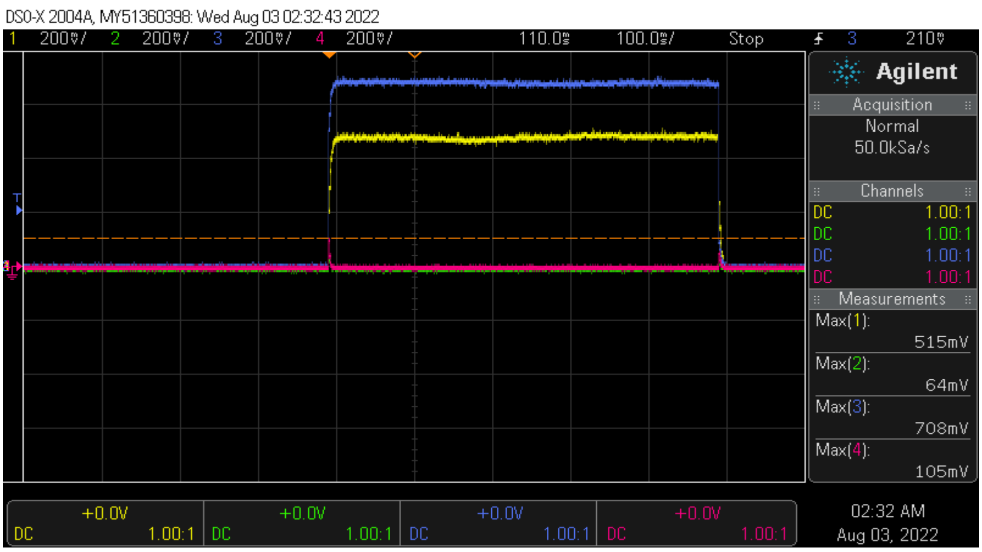
Figure 4.7: Case E: Effect of Ramp Attack on DSFC of DER-1 and DER-3 and their MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.



(a) Frequency

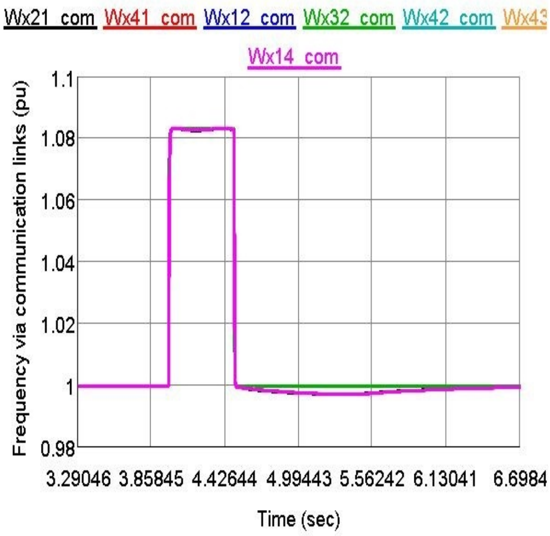


(b) Active Power Sharing

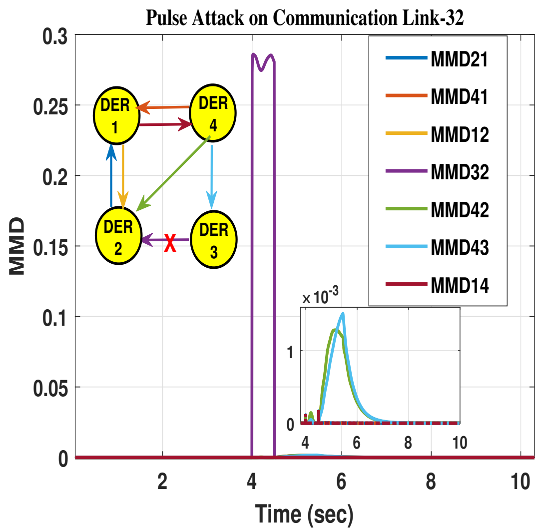


(c) Estimates of MMD Under Ramp Attack

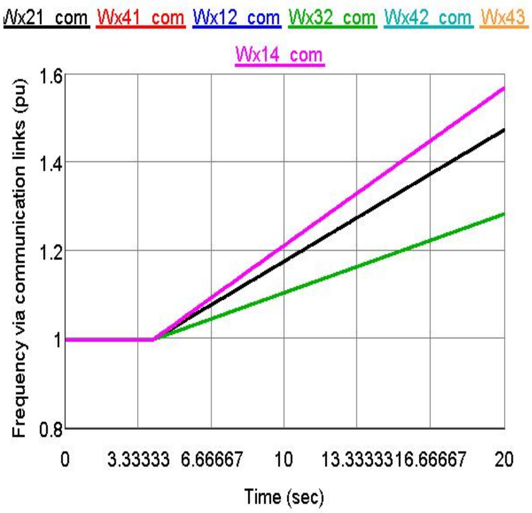
Figure 4.8: Case F: Effect of Pulse Attack on DSFC of DER-1 and DER-3 and their MMD estimates. In (a) and (b) the figure color labels black, red, blue and green represents frequency and active power of DER-1. DER-2, DER-3 and DER-4 respectively.



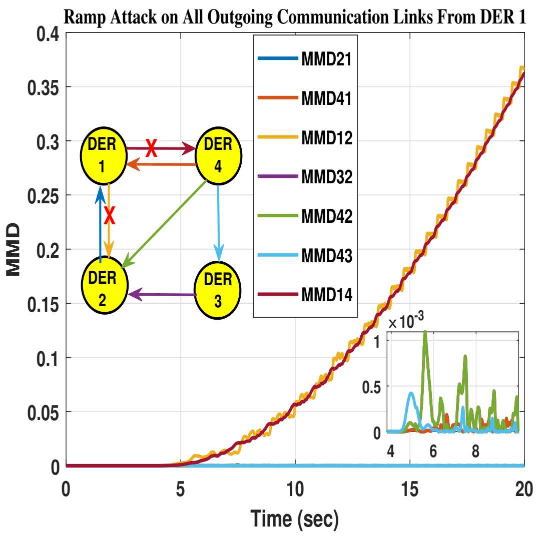
(a) Communicated Frequency



(b) Performance of MMD



(c) Communicated Frequency



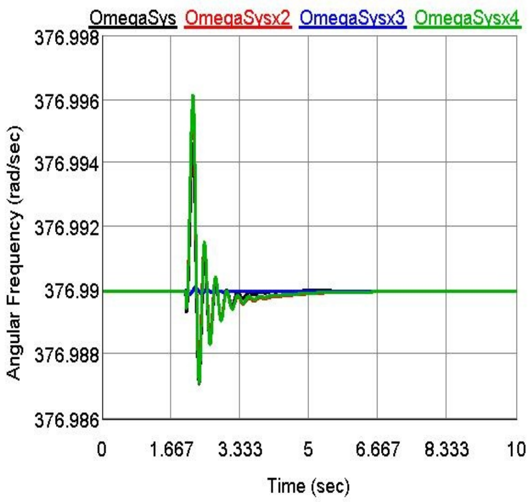
(d) Performance of MMD

Figure 4.9: Case G: Attack on communication link. (a),(b) Single line attack while transmitting ω_3^2 between DER-3 to DER-2. (c),(d) All outgoing communication links from DER-1 is compromised.

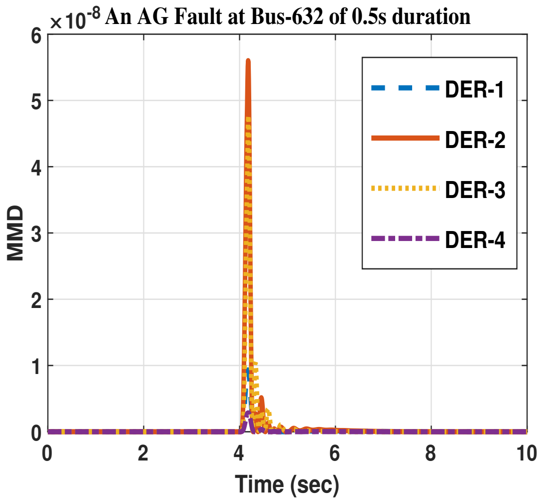
be remained almost zero for all the intact communication links except the corrupted one. This underscores the effectiveness of the proposed MMD-based detection scheme, even in scenarios where multiple communication lines are under attack.

4.3.3.5 Performance of the Proposed Attack Detector Against Natural Events

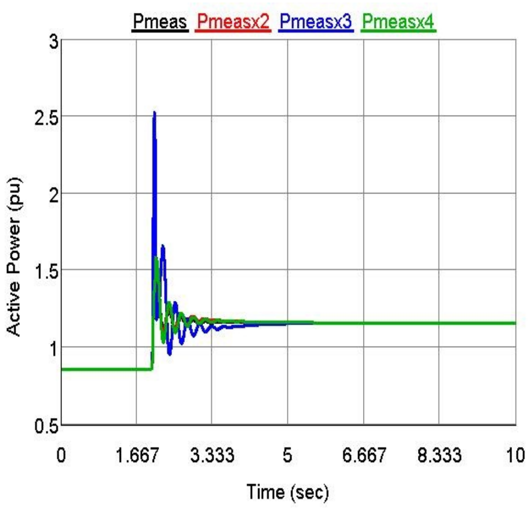
Power system undergoes certain changes in their operating states in the event of either any frequent natural disturbances or any unprecedented cyber-attacks. It is, therefore, crucial to differentiate between these two distinct events so that the operator should not initiate any adverse control action by misinterpreting a cyber attempt as a natural disturbance. To this end, in this test case *(Case H)*, a sudden single phase to ground (AG) fault of almost 10-cycles duration is created at Bus-632 of the IEEE 13 bus distribution test feeder as shown in Fig. 4.1. The frequency response and the MMD pattern are shown in Fig. 4.10(a) and 4.10(b). In contrast, another disturbance of load switching is performed in the system where a balanced load of $0.4MW$ active and $0.15MVAR$ reactive power are suddenly switched on at Bus-632. Figure 4.10(c) and 4.10(d) depict the change in active power sharing and the MMD detector performance, respectively, under such scenarios. It is well perceived from Fig. 4.10 that under both the events, MMD values come out to be significantly less than the threshold ($\gamma = 0.25V$) and henceforth the proposed detector is capable of accurately distinguishing cyber attack from fault and sudden load variations, ensuring that ***no false alarms*** are generated.



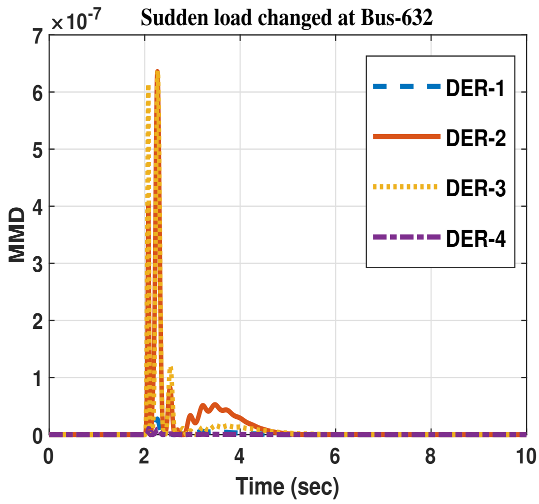
(a) Frequency



(b) Performance of MMD



(c) Active Power

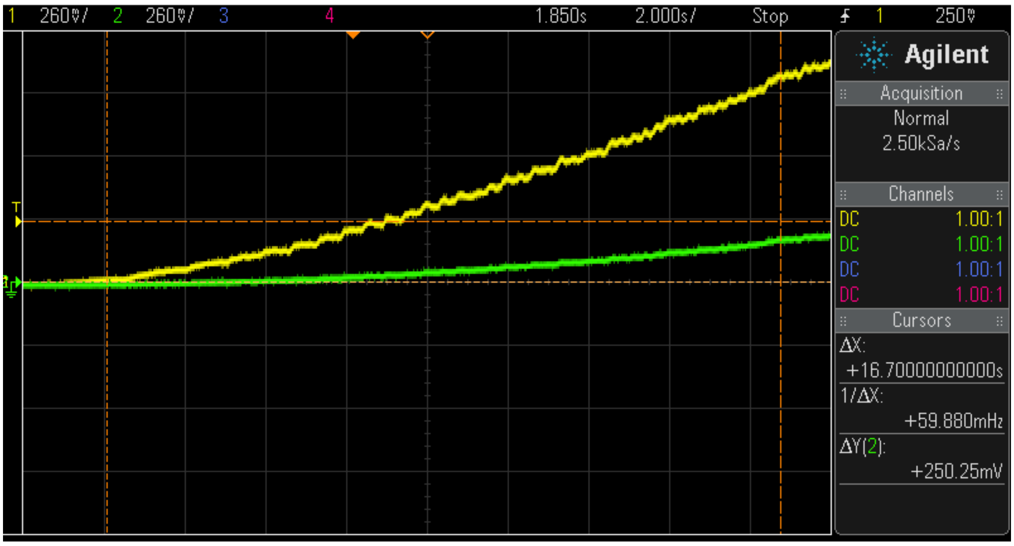


(d) Performance of MMD

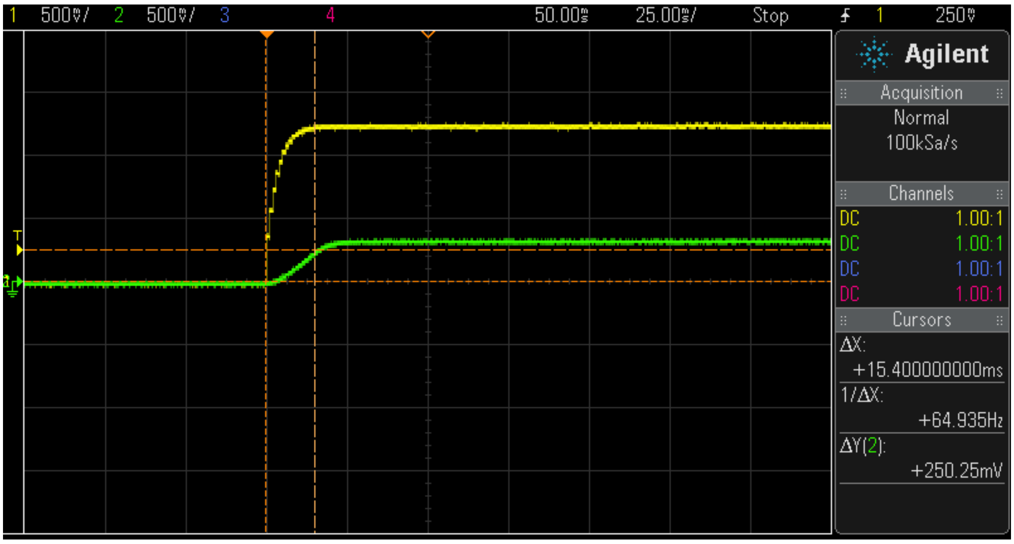
Figure 4.10: Case H: Performance of MMD against natural disturbances: (a),(b) Inception of a single-line-ground fault. (c),(d) Sudden switching of a balanced load.

4.3.3.6 Performance of the Proposed Attack on Comparative Assessment

The performance of the proposed MMD-based attack detection method has also been checked against a popular entropy-based attack strategy, i.e., Kullback-Leibler Divergence (KLD) [116]. To this end, Fig. 4.11(a) and 4.11(b) reveal the comparative behaviour of MMD and KLD in case of ramp and step type attack scenarios. The comparative assessment reveals that the response time of KLD is fairly sluggish as compared to that of the MMD. Also, the amplitude of KLD change is found to be very minimal with response to any attack event as it is not an exact measure of disparity and hence less sensitive, whereas, MMD is found to be simpler and more effective as compared to the entropy based method.



(a) Ramp-type FDI Attack



(b) Step-type FDI Attack

Figure 4.11: Case I: Performance between MMD and KLD: (a) Channel 1 (Yellow) – MMD, (b) Channel 2 (Blue) – KLD.

4.4 Rule-based EXtreme Gradient Boosting (XGBoost) Assisted Cyber Attack Classification

In the ever-evolving landscape of cyber security, accurately classifying cyber-attacks remains paramount importance for implementing effective defense strategies. With this as an aim, attack classification, in particular, plays a crucial role in identifying and categorizing different types of cyber threats on DERs controllers of MG system as mentioned in Table 4.1. To achieve this goal, a novel rule-based algorithm for feature extraction along with a renowned machine learning classifier i.e EXtreme Gradient Boosting (XGBoost) is utilized to classify those above detected attacks by the proposed detector in Section 4.2. In the domain of classification problem, XGBoost has emerged as a leading ensemble machine learning technique due to its exceptional performance across a multitude of tasks, especially in handling structured data with high dimensionality. Moreover, XGBoost's scalability and efficiency enable rapid training and deployment of attack classification models, crucial for real-time detection and response to cyber threats. XGBoost's core lies in its gradient boosting framework as shown in Fig. 4.12, which sequentially builds a set of weak learners, typically shallow decision trees, into a single, highly accurate "strong learner." Each new tree corrects the errors made by the previous ones through the minimization of the loss function, resulting in a model that can capture complex relationships within data. This iterative process allows XGBoost to continuously improve its predictive capabilities, achieving remarkable accuracy and generalization on various datasets. This makes XGBoost particularly well-suited for cyber attack classification tasks. Additionally, XGBoost's robustness to overfitting and its capacity to handle imbalanced datasets are particularly advantageous in the cyber security domain, where data may be scarce and class distributions uneven. Thus, by analyzing network topological features, system model and DER's control parameters XGBoost can learn to distinguish between normal network behavior and various attack types.

4.4.1 Introduction to Mathematical Operation of XGBoost Classifier

In boosting algorithm, in order to minimize the objective function, at first a base learner is chosen to fit it with the negative gradient of loss function at each iteration and thereafter the predicted outcome is added with the output from previous iteration value after being multiplied with some constant. In other sense it acts as performing gradient decent to the loss function and this negative gradients are usually termed as pseudo residuals. As it is known that XGBoost is nothing but the ensemble of several weak learners then the final prediction from a given dataset $D = (x_i, y_i)$ with n rows and m features is described as average weighted output of all the base learners as follows:

$$\hat{y}_i = \Phi(x_i) = \sum_{t=1}^{k} \grave{f}_t(x_i), \qquad \grave{f}_t \in \hat{F}. \tag{4.24}$$

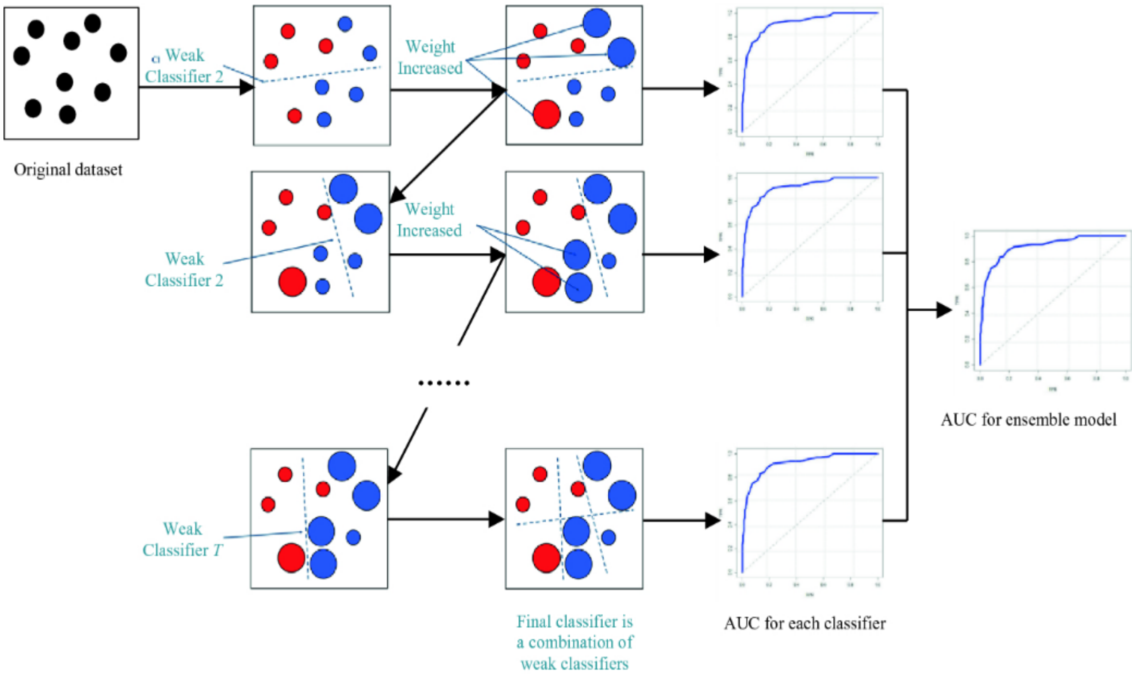


Figure 4.12: Visual representations of Gradient Boosting

where, $\hat{F} = \grave{f}(x) = \tilde{w}_{\grave{p}}(x)(\grave{p} : \mathbb{R}^m \to T_K, \tilde{w} \in \mathbb{R}^T{}_k)$ denotes the domain of regression trees, also referred to as CART (Classification and Regression Trees). Here, T_K, represents the total number of leaves in the tree and each $\grave{f}_t$ corresponds to a distinct tree configuration $\grave{p}$ and leaf weights $\tilde{w}$. This approach is based on approximating functions through the optimization of certain loss functions ($\mathcal{L}$) and the use of multiple regularization strategies as described below.

$$\begin{aligned} \mathcal{L}^{(t)} &= \sum_{i=1}^{n} \mathbb{L}(y_i, \hat{y}^{(t-1)} + \grave{f}_t(x_i)) + \vartheta(\grave{f}_t) \\ \text{where,} \quad \vartheta(\grave{f}_t) &= \beta T_K + \frac{1}{2}\eta \|\tilde{w}\|^2 \end{aligned} \tag{4.25}$$

Here, $\mathbb{L}$ is termed as differential convex loss function which calculates the difference between the target y_i and predicted outcome $\hat{y}_i$. The second term ϑ encounters the regularization concept, used to regulate the final weights to prevent over-fitting of the model. If this terms becomes zero, then it will become equivalent to conventional gradient boosting approach. The learning rate is represented by β; the larger the value of β, the simpler the tree. Another regularization term that lowers the step size in cumulative expansion is called shrinkage η. Now, to approximately calculate the value of the loss function for different possible base learners, taylor series expansion (up to second order derivative terms) must be applied to Eq. (4.25).

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \mathbb{L}(y_i, \hat{y}^{(t-1)}) + \underbrace{\frac{\partial^2 \mathbb{L}(y_i, \hat{y}^{(t-1)})}{\partial^2 \hat{y}^{(t-1)}}}_{\hat{\mathrm{g}}_i} \grave{f}_t(x_i) + \frac{1}{2} \underbrace{\frac{\partial \mathbb{L}(y_i, \hat{y}^{(t-1)})}{\partial \hat{y}^{(t-1)^2}}}_{\hat{\mathrm{h}}_i} \grave{f}_t^2(x_i) + \vartheta(\grave{f}_t) \tag{4.26}$$

As the first term is a constant and free from $\grave{f}_t$, Eq. (4.26) can be simplified as:

$$\hat{\mathcal{L}}^{(t)} = \sum_i \left[\hat{\mathrm{g}}_i \grave{f}_t(x_i) + \frac{1}{2}\hat{\mathrm{h}}_i \grave{f}_t^2(x_i) \right] + \vartheta(\grave{f}_t). \tag{4.27}$$

Let us define I_k be the set of instances belonging to leaf node $'k'$. Now, expanding ϑ from Eq. (4.25) in Eq. (4.27) yields as follows:

$$\hat{\mathcal{L}}^{(t)} = \sum_{k=1}^{T_K} \left[(\sum_{i \in I_k} \hat{\mathrm{g}}_i)\tilde{w}_k + \frac{1}{2}(\sum_{i \in I_k} \hat{\mathrm{h}}_i + \eta)\tilde{w}_k^2 \right] + \beta T_K. \tag{4.28}$$

Now, the optimal weight, $\tilde{w}_k^*$ for each leaf node-j, can be obtained by equating derivative of loss function with respect to each leaf node's weight i.e., $\frac{\partial \hat{\mathcal{L}}^{(t)}}{\partial \tilde{w}_k^*} = 0$, as follows:

$$\begin{aligned} 0 &= \sum_{i \in I_k} \hat{\mathrm{g}}_i + \frac{1}{2}(\sum_{i \in I_k} h_i + \eta) \times 2 \times \tilde{w}_k^* \\ \tilde{w}_k^* &= \frac{-\sum_{i \in I_k} \hat{\mathrm{g}}_i}{\sum_{i \in I_k} \hat{\mathrm{h}}_i + \eta} \end{aligned} \tag{4.29}$$

The obtained optimal weights is then substituted in Eq. (4.28) to get final optimal loss function value for a fixed tree structure $\grave{p}$ as shown below:

$$\hat{\mathcal{L}}^{(t)} = -\frac{1}{2}\sum_{k=1}^{T_K} \frac{(\sum_{i \in I_k} \hat{\mathrm{g}}_i)^2}{\sum_{i \in I_k} \hat{\mathrm{h}}_i + \eta} + \beta T_K \tag{4.30}$$

XGBoost employs several techniques to speed up training and reduce overfitting. It utilizes random subsampling of data and columns during tree building, introducing randomization. Additionally, XGBoost leverages a compressed, pre-sorted data structure that eliminates redundant sorting, significantly accelerating the search for optimal splits in decision trees. This allows for faster training and more efficient models [215].

4.4.2 Tuning of Hyper-parameters in XGBoost

The key features of XGBoost algorithm is its ability to handle missing and complex data, regularization and auto pruning techniques to prevent the model from being over/under-fitted and mostly excellent performance due to its parallelization feature concept to have efficient computation. However, such merits are obtained by carefully selecting and choosing values of some hyper-parameters through validation techniques like k-fold cross validation that drive the XGBoost model to enhance its performance and control its bias and variance in training and testing stages.

4.4.2.1 Learning Rate

The learning rate (β) parameter in XGBoost classifiers plays a crucial role in controlling how much each new tree in the ensemble contributes to the overall model. It is used to prevent the overfitting issues faced by the individual base learners but may increase training time. A smaller learning rate leads to more conservative updates, requiring more trees to achieve the desired accuracy. Conversely, a larger learning rate leads to more aggressive updates, potentially leading to faster training but also increasing the risk of overfitting. The optimal learning rate is achieved and set to the value 0.5 through experimentation or techniques like grid search.

4.4.2.2 Number of Estimator ($n\_estimator$)

In XGBoost, a critical hyper-parameter for achieving optimal performance is the number of estimators ($n\_estimator$), which refers to the quantity of decision trees built sequentially during the boosting process. While intuitively, a larger number of trees suggests a more intricate and potentially more accurate model, this isn't always the case. Increasing the number of estimators can lead to diminishing returns in terms of accuracy, and even introduce the phenomenon of overfitting and conversely, reducing the number of estimators can lead to underfitting and decreased accuracy. Therefore choosing a suitable count of n_estimator, necessitates a delicate balancing act between accuracy and model complexity. In this chapter, experimenting with a wide range of estimator values revealed 100 as the optimal choice, balancing accuracy and computational complexity.

4.4.2.3 Auto-pruning Hyper-parameter

XGBoost offers an automated approach to controlling tree complexity through the parameter denoted as ϱ, also known as the auto pruning parameter. This parameter establishes a minimum gain threshold, say 'max_depth' determining when splits are made during tree growth. Splits producing gains below this threshold are disregarded, preventing overfitting by focusing on informative splits that contribute meaningfully to the prediction accuracy. Similarly, if the 'max_depth is set to very low, underfitting situation may arise as the model is not able to be trained well to capture all the relevant patterns in the data. Thus While the 'max_depth' parameter sets an upper limit on tree depth, the auto pruning feature with the ϱ parameter allows for more granular control within that limit, leading to more efficient and potentially more accurate models. In this study, the max_depth is chosen to be 5 to balance between overfitting and underfitting.

4.4.2.4 Objective or Loss Function

The objective function in XGBoost is crucial for guiding the model's training process by evaluating its performance on the training data and calculating gradients for improvement. The choice of objective function depends on the task at hand; for binary classification, the binary logistic regression function is suitable, while for multi-class problems like the one addressed in this chapter, the 'softprob' function is recommended. This function is specifically designed to handle the complexities of multi-class classification tasks, ensuring that XGBoost learns effectively in these scenarios. Given that this chapter centers around a multi-class attack classification problem, the, softprob' objective function is opted in this study to calculate the probability of an observation belonging to a predicted class.

4.4.3 Dataset Preparation

For the generation of testing and training dataset for the chosen machine learning (ML) model, the same modified IEEE 13-bus islanded AC MG system incorporating 4 DERs with the pre-defined communication topology as shown in Fig. 4.13 is utilized for the classification problem. In this topology, the single head arrow represents one-way communication whereas the two head arrow represents two-way communication between the DERs. There are in total 5 different types of attacks are simulated: Pulse, Ramp, Random, Scaling and Sine. The attack is launched in the DERs by corrupting the frequency input of their respective secondary controllers as per the attack model summarized in Table 4.1. The generation of attacked dataset are prepared based on either increasing the attack impact through magnitude alteration or varying attack length through changing attack duration as described below and tabulated in Table 4.2 and 4.3. Finally, this comprehensive dataset would be used to train and evaluate the XGBoost model's effectiveness in detecting these attacks.

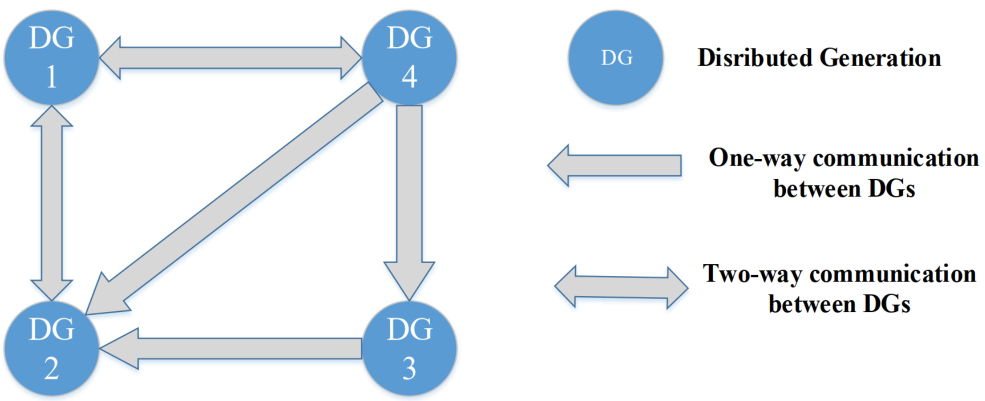


Figure 4.13: Communication topology of participating DERs in co-operative Distributed Secondary Control

In reference to dataset preparation, the first case study is dedicated for generation of such dataset which is later used to investigate the efficacy of the XGBoost model in pinpointing attacks that exhibited varying duration. To achieve this, they kept the attack magnitude (strength) constant while meticulously adjusting the time period during which the attack was launched. This manipulation allowed them to observe the system's response under different attack scenarios and identify the "sweet spot" from the attacker's perspective. A total window of 10 seconds was chosen to monitor the system behavior under various FDI attacks. Initially, the attack amplitude remains constant, while the attack duration increases incrementally for all distinct attack types to account for their inherent characteristics. For Pulse attacks, the time step is set to 0.01 seconds, while for other attack types, it is 0.05 seconds. The shorter duration for Pulse attacks is chosen to simulate the effects of very brief attacks, starting from 0.01 seconds and increasing up to 0.11 seconds. Similarly, the initial duration for other attacks is 0.5 seconds, increasing in steps of 0.05 seconds until reaching 1 second, as detailed in Table 4.2.

Table 4.2: First Case: Constant Attack Magnitude and Varying Attack Duration

Attack Types	**Magnitude (p.u)**	**Initial Attack Duration (sec)**	**Step Size (sec)**	**Final Attack Duration (sec)**
Pulse Attack	0.01	0.01	0.01	0.11
Ramp Attack	0.5	0.5	0.05	1.0
Random Attack	0.01	0.5	0.05	1.0
Scaling Attack	1.03	0.5	0.05	1.0
Sine Attack	0.3	0.5	0.05	1.0

In second scenario, the classifier ability is also need to be checked for such FDI attacks whose intensities (amplitudes) are varying while holding the attack duration constant. Here, the attack duration for the Pulse attack at its maximum value identified earlier (0.11 seconds) and for all other attack types at 1 second are fixed. However, the amplitude was now systematically increased in small increments (0.01 units) for a total of 10 steps. It's important to note that the attackers deliberately chose distinct initial amplitudes for each attack types. This strategic selection aimed to incorporate a broad spectrum of variations within the system parameters. In essence, they were creating a diverse dataset that reflected a wide range of attack intensities across different FDI attack types. A detailed breakdown of the initial amplitude values chosen for each attack type can be found in Table 4.3.

Table 4.3: Second Case: Constant Attack Duration and Varying Attack Magnitude

Attack Types	**Magnitude (p.u)**	**Initial Attack Duration (sec)**	**Step Size (sec)**	**Final Attack Duration (sec)**
Pulse Attack	0.11	0.01	0.01	0.1
Ramp Attack	1	0.5	0.01	0.59
Random Attack	1	0.01	0.01	0.1
Scaling Attack	1	1.03	0.01	1.12
Sine Attack	1	0.3	0.01	0.39

Therefore, in a nutshell for the training and evaluation of the XGBoost model, a comprehensive dataset is generated encompassing variation in both attack duration and intensity as discussed above. For each simulated attack scenario, attackers meticulously targeted a single DER at a time. Throughout the attack, four critical system parameters: frequency, active power, reactive power, and voltage are monitored and recorded. This data collection process encompassed all four DGs within the system, even though only one was under direct attack at any given moment. Therefore, by capturing all these parameters from all four DGs, under each attack scenario, 16 parameters are available for recording. Now referring back to Table 2.2, a total of 11 distinct attack scenarios were simulated. Considering the 16 parameters recorded for each attack on each DG, this translates to a total of 3520 attack instances. Similarly, Table 2.3, presented 10 unique attack scenarios. Following the same logic, this translates to an additional 3200 attack instances. Thus, by meticulously simulating attacks under various conditions , finally a total of 6720 attack instances (3520 instances from Table 4.2 + 3200 instances from Table 4.3) as summarized in Table. 4.4.

Table 4.4: Comprehensive Attack Dataset Generation

Description	**Values**
Number of DERs	4
Number of Attacks	5
Number of Recorded Parameters	16
Number of Attack Instances Generated from Table 4.2	3520
Number of Attack Instances Generated from Table 4.3	3200
Total Attack Instances	6720

4.4.4 Proposed Rule-based XGBoost Enabled Cyber Attack Classification Scheme

To effectively differentiate between statistically crafted cyber-attacks from the normal measurements received at the control centers, its crucial to identify and leverage distinctive patterns within the data. These patterns can significantly enhance the performance of XGBoost classifiers in detecting attacks. With this objective, this chapter delve into two key metrics: Entropy and Shannon Energy. These metrics are specifically chosen to help extract these crucial distinguishing patterns from the measured quantities.

4.4.4.1 Entropy

In signal processing, entropy reflects the information content within a signal to predict the outcomes of a random process. Essentially, entropy measures the average surprise or unexpectedness associated with each event in the signal. Signals with a more even distribution (all samples are equally likely) have higher entropy compared to those with a skewed distribution (where some values are more frequent). In other words, it can stated as measures of the average amount of information per event, with low probability events containing more information than high probability ones. Maximum entropy occurs when all events are equally probable, while zero entropy indicates certainty in one event and no information uncertainty. The equation for defining entropy can be written as follows.

$$E_{en} = \sum_{j} -q_j \log_2 q_j \tag{4.31}$$

where, q_i is the the probability of each state for all the possible states.

Entropy calculation in signal processing often involves estimating the probability distribution of signal samples, typically through methods like histograms or density estimation. The entropy formula is then applied to this resulting distribution to arrive at a numerical value representing the signal's information complexity. In scenarios where the exact probabilities are unknown, the "hist" function command in MATLAB is used for this purpose. This function divides the data range into bins of equal size, with the number of bins either specified by the user or determined by a default value. Selecting an appropriate number of bins is crucial. If too many bins are chosen, then the histogram lacks sufficient detail to accurately capture the underlying data distribution. Conversely, opting for an excessive number of bins can lead to an overly granular representation, potentially obscuring the overall picture.

To strike a balance between detail and simplicity in the histogram, in this study a default value of 10 bins is selected. Once the bins have been defined, the function counts the number of data points that fall into each bin. The relative frequency for each bin is then calculated by dividing this count by the total number of data points. Finally, by dividing each value by the bin width, the relative frequencies can be translated into probabilities. As a result, the relative frequency of a bin divided by its width determines the likelihood that a data point will fall inside that bin. This ensures that the histogram adequately captures the underlying distribution while maintaining a clear and interpretable visualization. Figure 4.14 displays the frequency's entropy's minimum and maximum ranges under various FDIAs, as tabulated in Table 4.1.

In Fig. 4.14, it is clearly evident that the Pulse attack occupies a distinct range on the entropy spectrum, clearly separated from the other FDI attacks under investigation. This distinction is be made out by establishing entropy thresholds for attack identification. If the calculated entropy of an attack falls within the range of 0.29 to 0.35, it can be confidently classified as a Scaling attack. Similarly, an entropy value between 0.73 and 1.26 suggests a Random attack. These clear demarcations based on entropy effectively classify a significant portion of the attacks. However, the may be such scenarios, where the process of distinction becomes slightly more intricate for the remaining attack types. Their entropy ranges exhibit some degree of overlap, making it challenging to differentiate them solely based on this metric. To address this hurdle, the help from the another statistical tool, named Shannon Energy is introduced in the proposed rule-based algorithm to establish a more comprehensive classification framework that could effectively distinguish between all the various FDI attack types, even those with overlapping entropy ranges.

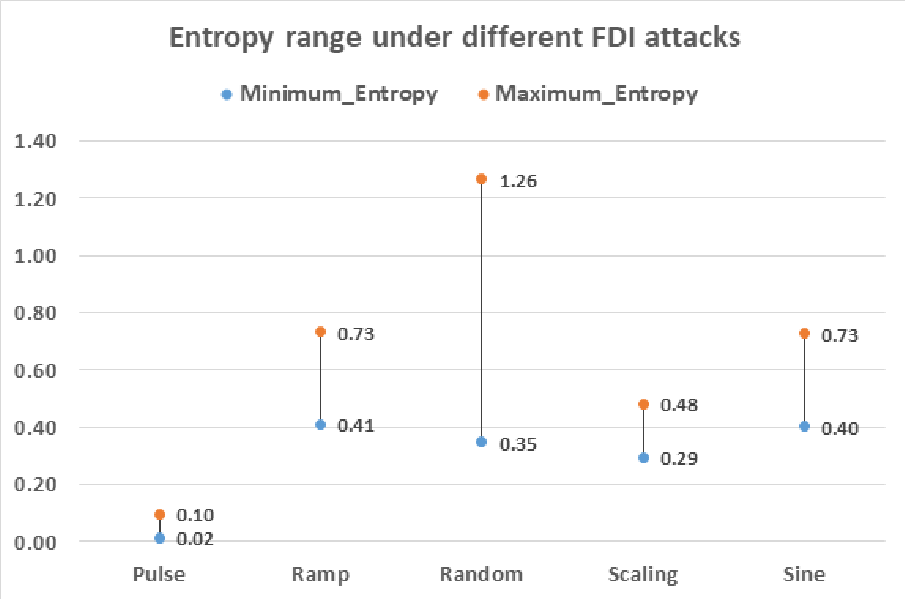


Figure 4.14: Ranges of Entropy values under different FDIAs

4.4.4.2 Shannon Energy

Shannon energy emerges as a significant concept, finding applications across diverse fields like audio and image processing, communication systems, and more. It calculates the signal's energy by analyzing its local spectrum at each individual sample. This local spectrum essentially breaks down the signal into its constituent frequency components, revealing how much energy is present at each frequency. By summing the energy contributions across all these frequencies, Shannon energy provides a comprehensive picture of the signal's overall energy distribution. It essentially quantifies the total energy embedded within a signal over a defined time interval. The equation used to calculate the Shannon energy is

$$E_{se}[n] = \sum_{j} -p^2[n] \log_2(p^2[n]) \tag{4.32}$$

where, $p[n]$ denotes the normalized signal.

Now, similar to the observations made with entropy, Fig. 4.15 also shows the distinct ranges of energy content information available between its usual minimum and maximum range values. This makes it particularly valuable in the context of attack classification of such sophisticated attacks that have entropy overlapped values. Thus by leveraging the combined power of entropy and Shannon energy thresholds, a significant portion of the attacks can now be differentiated. For instance, a Shannon energy value falling within the narrow band of 258.62 to 258.75 is treated as Random attack. Conversely, an energy level between 239.7 and 252.77 suggests a Scaling attack. At last, for all the remaining attacks with overlapping ranges, XGBoost classifier is exploited to classify the attacks. This classifier is then trained on the meticulously crafted dataset, possesses the ability to identify subtle patterns and relationships within the data that might not be readily apparent through basic thresholding. Thus, in the context of attack classification, employing entropy from section 4.4.4.1 alongside with the use of Shannon energy from section 4.4.4.2 in the XGBoost-enabled approach, a rule-based attack classification framework is developed for identifying and distinguishing between various types of FDIAs.

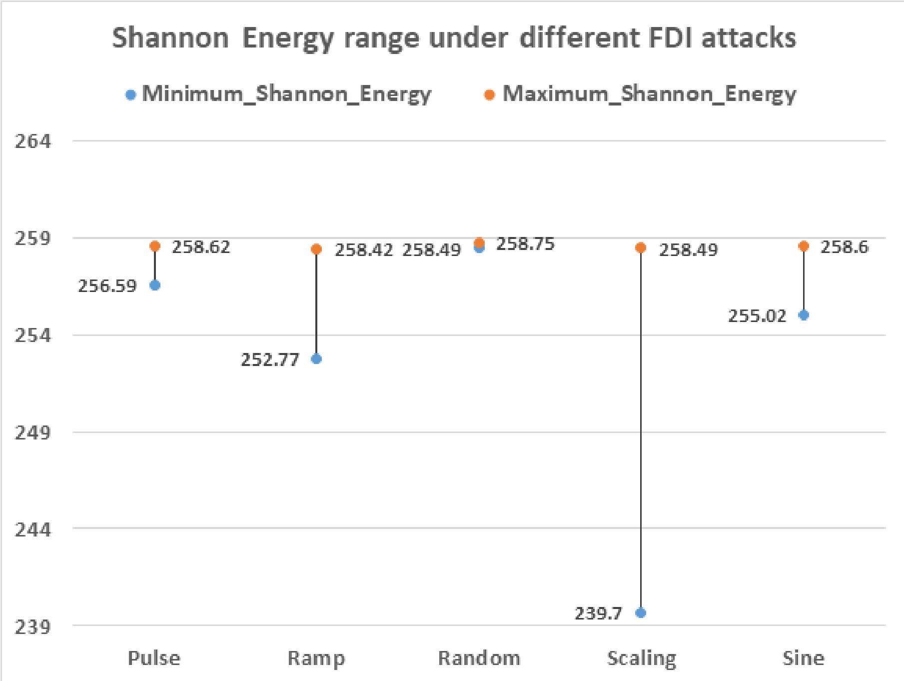


Figure 4.15: Ranges of Shannon Energy values under different FDIAs

4.4.4.3 Proposed Rule-based Flowchart

The proposed rule-based method for classifying different categories of FDIAs on a DSFC controller of islanded AC MG are summarized below:

1. **Data Collection:** The modified IEEE 13-bus islanded AC MG system was simulated in RSCAD software of NovaCor RTDS simulator with the time-step size of 50 μs. From the simulation, various monitoring parameters like frequency, power, and voltage of each DERs are collected from the RSCAD with an interval of 6.4 ms and the total simulation time window is chosen to be 10 seconds resulting in total 1562 data point for each monitoring parameters at the end.

2. **Attack Detection:** The proposed classification scheme can only be initiated when the proposed MMD-based detector discussed in section 4.3.2 detect the cyber attack in DER's controllers. While the metric exceeds the predefined threshold γ, the captured system parameters are then carry forwarded to the next for classification task.

3. **Entropy-based Classification:** If an attack is detected, entropy of the frequency is then calculated. By leveraging the observations discussed in the respective section

of entropy, Pulse, Scaling, and Random attacks are classified based on distinct non-overlapping entropy ranges.

4. **Shannon Energy for Overlapping:** Remaining attacks with overlapping entropy ranges are further analyzed using Shannon energy. Similar to entropy, non-overlapping ranges are used to classify Random and Scaling attacks. These above two stages creates a formation of rule-based approach to classify the attacks as an initial screening test.

5. **XGBoost for Remaining Cases:** Being trained on a meticulously prepared dataset, subtle patterns and relationships can be identified within the data that might be difficult to capture using basic thresholding techniques. Therefore, in cases where both entropy and shannon energy exhibit overlapping ranges, signifying a more intricate attack scenario, the rule-based method resorts to a machine learning model named XGBoost for attack classification. In this classification process, selecting the best possible feature subset from original set is very delicate task in order to have a highest predictive power for simplifying the analysis and enhance the model performance. In this work, pearson correlation coefficient is used as the feature selection technique. This approach is chosen because highly correlated variables tend to have a strong relationship with the target variable. If two variables are highly correlated, one variable can effectively predict the other, reducing redundancy in the model and saving computational resources. A threshold, such as 0.6 in this case, is set, and if the correlation between two variables exceeds this threshold, the variable with lower correlation with the target is dropped. Following this approach, entropy and shannon energy of the frequency were identified as the most correlated features with the target variable and were consequently chosen to form the basis of the classification rules.

Figure 4.16 presents a flowchart that visually summarizes this entire XGBoost-assisted rule-based FDI attack classification process.

4.4.5 Simulation Results Along with its Comparative Performance with Other ML Classifiers

This section provides the detailed insights of the performance of XGBoost classifier and compares it with three other popular ML classifiers: Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB). These models are trained and tested on a sizable dataset, split into a 70:30 ratio for training and testing, respectively. To mitigate any bias within the dataset, standard data normalization is performed using the *"StandardScaler()"* function from the sklearn library. Each classifier's performance relies on various regularization parameters as discussed previously like learning rate, number of estimators, maximum tree depth, and objective function are fine-tuned to get the best result out of the model. Performance evaluation is conducted using the Confusion Matrix (CM), which compares predicted and actual class labels to assess the classifier's accuracy

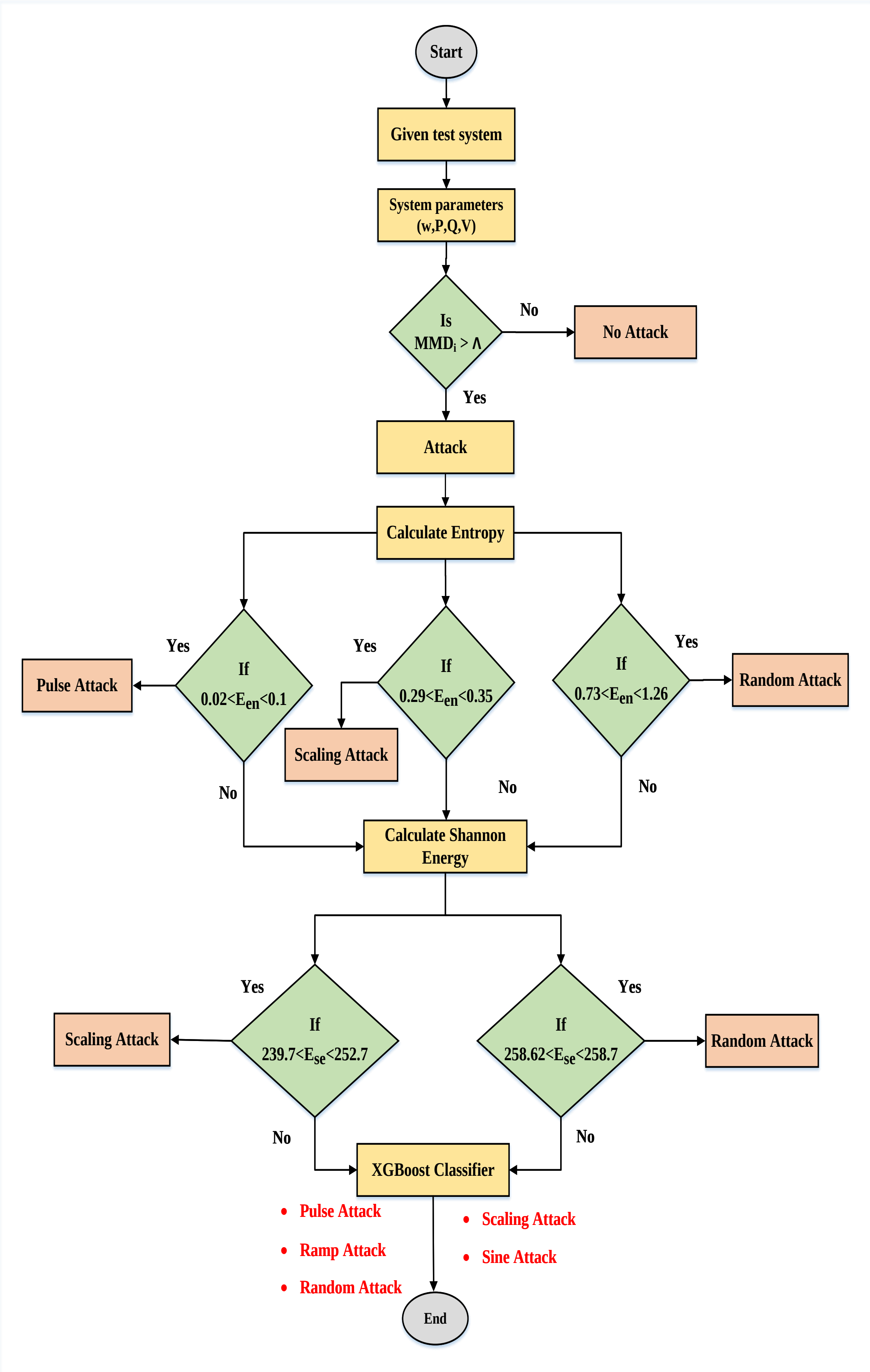


Figure 4.16: Flowchart of a rule-based XGBoost-enabled attack classification scheme

and identify misclassification patterns. Derived from the CM, metrics such as accuracy, precision, recall, and F1 score provide further insights into the classifiers' performance.

1. **True Positives (TP):** The number of instances correctly predicted as belonging to the positive class.

2. **True Negatives (TN):** The number of instances correctly predicted as belonging to the negative class.

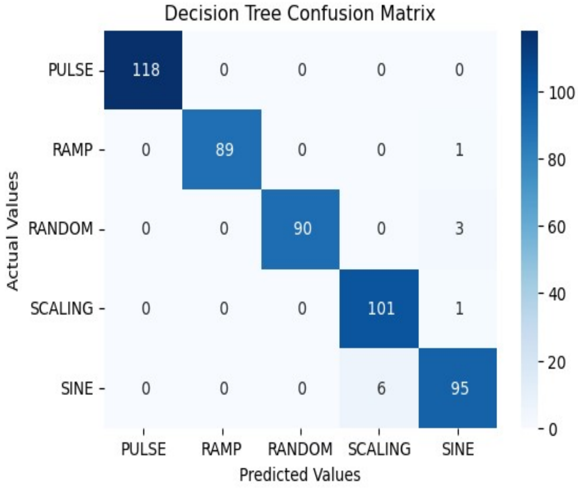
3. **False Positives (FP):** Also known as Type I errors, these are instances predicted as positive but are actually negative.

4. **False Negatives (FN):** Also known as Type II errors, these are instances predicted as negative but are actually positive.

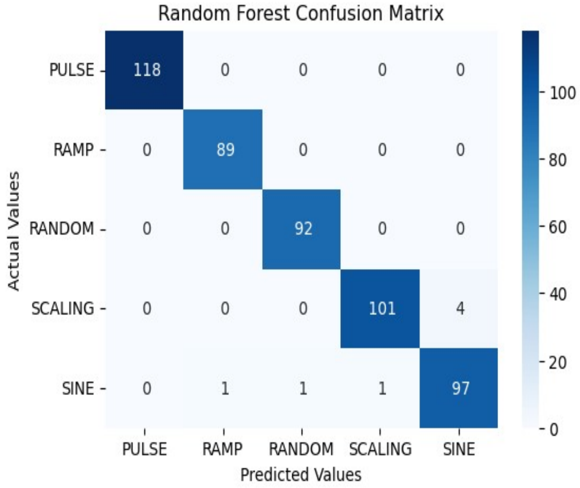
Figure 4.17, presents a confusion matrix summarizing the performance of various classifiers during comparative evaluation. Analyzing the rate of TP, TN, FP and FN, it can be seen that the DT classifier exhibits the weakest performance among all contenders. The RF and GB classifiers show improvement over DT, with XGBoost achieving the most remarkable accuracy. The CM also reveals that Scaling attacks pose the most significant challenge for classification. Their entropy and energy ranges often overlap with other attack types, making them difficult to distinguish using simpler models. However owing to advantages in speed and accuracy of the XGBoost classifier that attacks can be easily detected.

To comprehensively evaluate the model's performance, in this study k-fold cross validation, a robust technique widely used in machine learning is adopted. Selecting the optimal value for 'k' hinges on two crucial factors: dataset size and available computational resources. A common practice involves using 'k' values of 5 or 10. This choice strikes a balance between bias and variance, which are inherent trade-offs when selecting 'k.' Research has empirically shown that 'k' values of 5 or 10 often yield the most reliable results [216]. Considering these factors, in this chapter, 'k' is set to 10. Table 4.5 summarizes the overall accuracy and standard deviation of the model's accuracy obtained through this k-fold cross-validation. This basically illustrates correctly classified instances out of the total number of instances. Notably, the proposed method stands out as the most accurate classification technique among those evaluated by other ML classifiers, highlighting its effectiveness in identifying FDI attacks within the islanded AC MG.

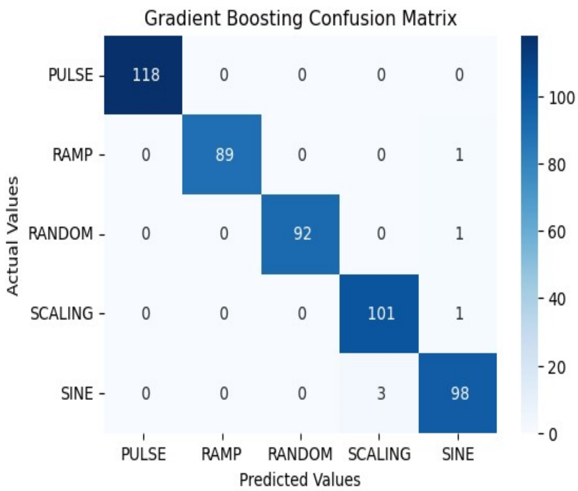
Now, leveraging these CM's information as shown in Fig. 4.17, Table 4.6 provides the other important comprehensive performance evaluation parameters such as precision, recall and F1 score for different classifiers under each categories of FDIAs. It is observed from Table 4.6 that as expected, the pulse attack, being the simplest to execute, is effectively detected by all classifiers with comparable performance. However, the efficacy of the models diminishes slightly for more intricate attacks like sine and scaling attacks. Nonetheless, the proposed rule-based XGBoost classifier consistently outperforms the other machine learning models in terms of precision, recall, and F1-score. This suggests that the XGBoost classifier achieves a better balance between correctly identifying attacks and minimizing misclassifications, even for the more challenging attack types.



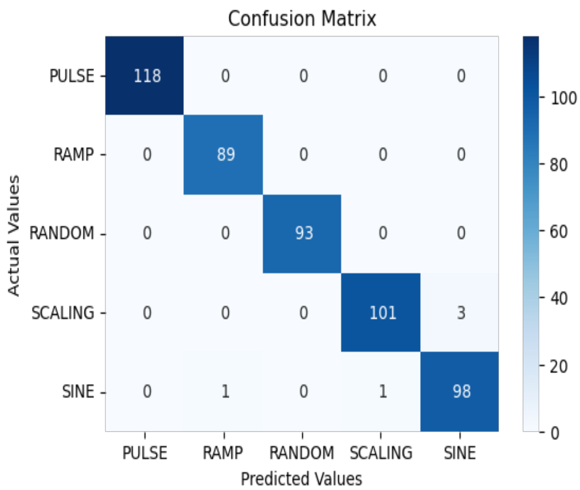
(a) Decision Tree



(b) Random Forest



(c) Gradient Boosting



(d) Proposed XGBoost

Figure 4.17: Confusion matrix for different machine leaning classifier.

Table 4.5: Comparative Assessment on Overall Accuracy of Different ML Classifiers

Classifiers	**Accuracy (%)**	**Standard deviation in Model's Accuracy (%)**
Decision Tree	98.47	1.31
Random Forest	98.98	0.74
Gradient Boosting	99.32	0.51
Proposed rule-based XGBoost	99.49	0.68

Table 4.6: Performance Metrics for Different FDI Attacks under Various ML Techniques

Types of FDIAs	**Decision Tree**			**Random Forest**			**Gradient Boosting**			**Proposed XGBoost rule-based**		
	Precision	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**
Pulse Attack	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Ramp Attack	1.0	0.99	0.99	0.99	1.0	0.99	1.0	0.99	0.99	0.99	1.0	0.99
Random Attack	1.0	0.97	0.98	0.99	1.0	0.99	1.0	0.99	0.99	1.0	1.0	1.0
Scaling Attack	0.95	0.99	0.97	0.99	0.96	0.97	0.97	0.99	0.98	0.99	0.97	0.98
Sine Attack	0.95	0.94	0.94	0.96	0.97	0.96	0.97	0.97	0.97	0.97	0.98	0.97

4.5 XGBoost Enabled Multi-Label Cyber Attack Localization Scheme for the Compromised DER Unit

4.5.1 Proposed Cyber Attack Localization Scheme using XGBoost with Extracting Additional Feature Inputs

Having successfully identified and classified the various attack types targeting the MG in Section 3 and 4, Section 5 delves into a critical aspect: attack localization. Precisely pinpointing the location of an attack within the MG equips system operators with invaluable information for formulating effective protection and defense strategies. However, pinpointing the attack source presents a significant challenge due to following two main factors.

1. Firstly, DERs within the MG possess unique characteristics and intricate interdependencies. These inherent complexities within the DER network make attack localization a non-trivial task.

2. Secondly, the MG's network topology itself adds another layer of difficulty. The interconnected nature of the MG's components creates a complex web, making it challenging to isolate the origin of an attack.

To address these challenges, this chapter proposes an extension of the XGBoost classifier specifically tailored for multi-label classification. Unlike single-label problems, multi-label classification predicts a set of binary values, each indicating the presence or absence of an attack on a particular DER within the MG. To accommodate this requirement of assigning multiple labels simultaneously, the model leverages techniques like sigmoid activation. This essentially treats each DER as an independent label during the classification process. However, during the simulation for generating test cases for this localization scheme, only one DER is considered to be attacked at a particular instant of time. This makes the localization task simpler. Moreover, to enhance localization accuracy, additional statistical features beyond Entropy and Shannon energy are explored, including Mean or median absolute deviation, standard deviation, kurtosis, skewness and crest factor. These features provide valuable insights into the variability, spread, and distribution of data points associated with each DER. Significant deviations from normal behavior in any of these statistical measures can serve as flags, potentially indicating the presence of an attack at a specific DER. Since the study focuses on four parameters of the DERs, i.e., frequency, voltage, active and reactive power - a total of 28 input feature variables are generated and saved in a .csv file for training the multi-label classification model. Given the imbalanced nature of real-time data, where certain attack types may be more prevalent than others, this chapter used 1000 estimators during training to ensure effective learning from minority classes within the dataset. Moreover, all the other hyperparameters such as ratio of training and testing data sets, learning rate, maximum depth, pre-processing and objective functions are kept similar to the previous classification problem. This

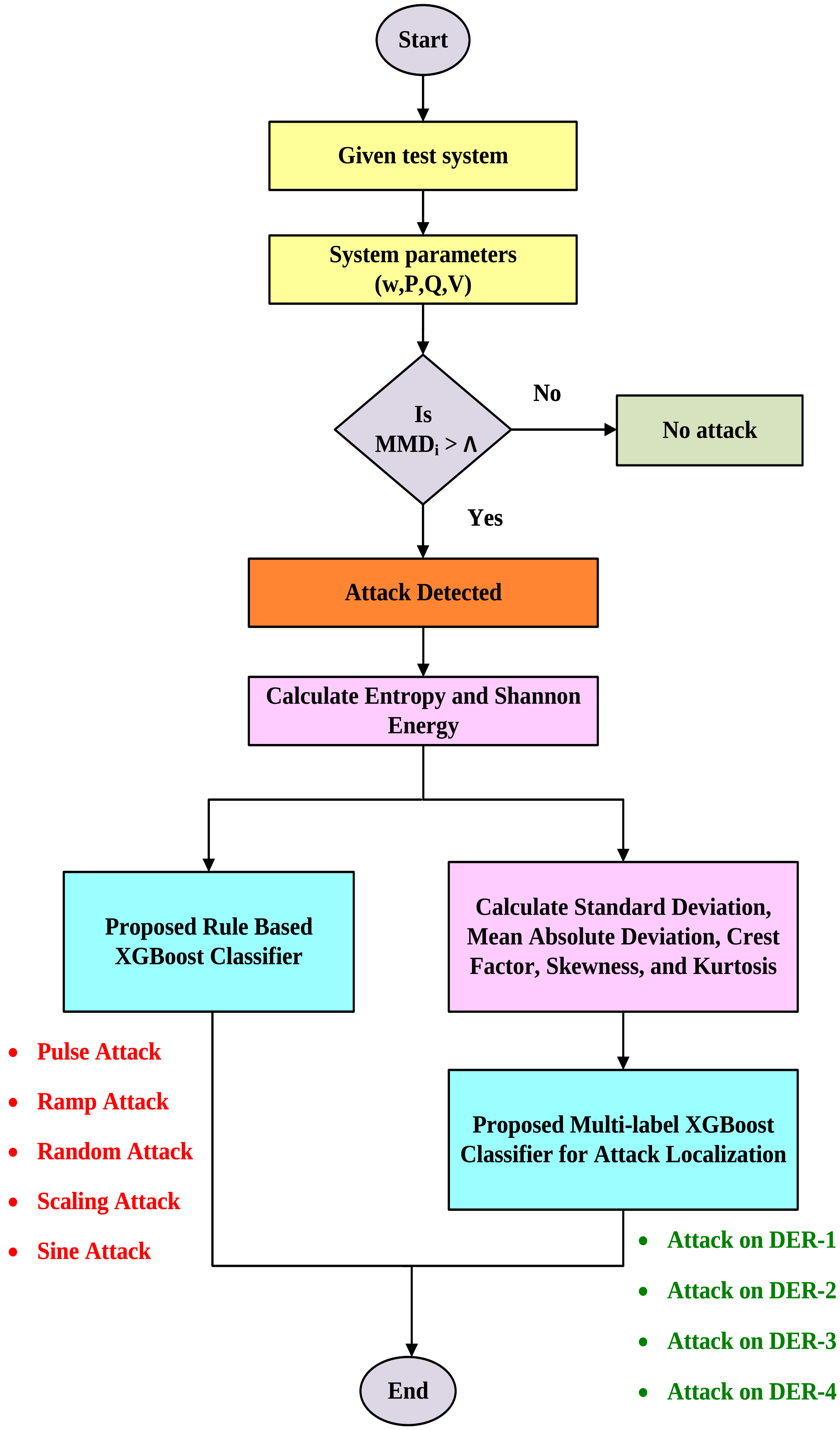


Figure 4.18: Flowchart of proposed XGBoost-enabled attack localization scheme

comprehensive feature set empowers the model to achieve a more accurate and robust attack localization within the MG. Once trained, the XGBoost classifier is equipped to predict the attack location within the test system and it is also used to compare its performance against established ensemble machine learning techniques for assessing its effectiveness. A comprehensive flowchart depicting the entire XGBoost-enabled multi-label classification process is presented in Fig. 4.18.

4.5.2 Performance Metrics for Multi-Label Classification

Apart from the performance metrics discussed in Section 4.4, a multi-label classification model's requires specific metrics for accessing its performance tailored to handle the complexity of multiple labels per instance. Those metrics are discussed below.

1. **ROC Curve:** In the realm of machine learning, particularly for problems with multi-label classifications the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) are popular metrics for evaluating model performance. The ROC curve visually depicts a classifier's ability to distinguish between positive and negative instances. It plots the True Positive Rate (TPR) on the y-axis against the False Positive Rate (FPR) on the x-axis. To generate the ROC curve, the classification threshold is progressively adjusted from 0 to 1. At each threshold, the TPR and FPR are calculated, forming a series of data points. The closer the ROC curve sits to the top-left corner of the graph, the better the classifier's performance. A random classifier, lacking any discriminatory power, would be represented by a diagonal line stretching from the bottom-left corner to the top-right corner as shown in Fig. 4.19. Once the ROC is plotted, the area under this ROC is called as AUC. It condenses a classifier's performance across all possible classification thresholds into a single, numerical metric whose values ranging from 0 (worst) to 1 (perfect). An AUC of 1 signifies a flawless classifier that flawlessly differentiates between positive and negative examples, while an AUC of 0.5 indicates a mere random guess.

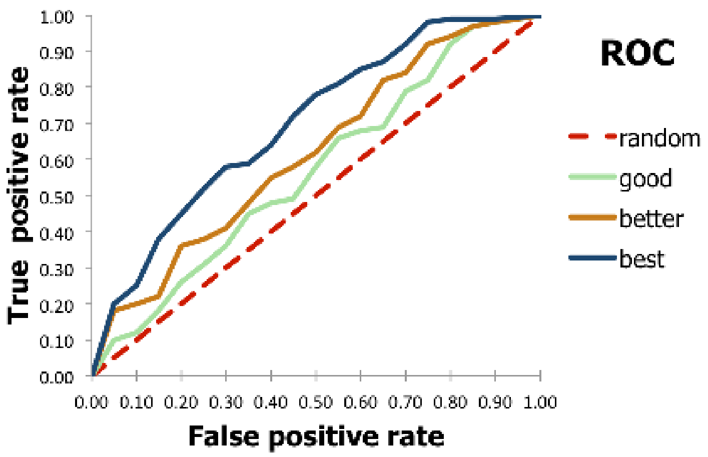


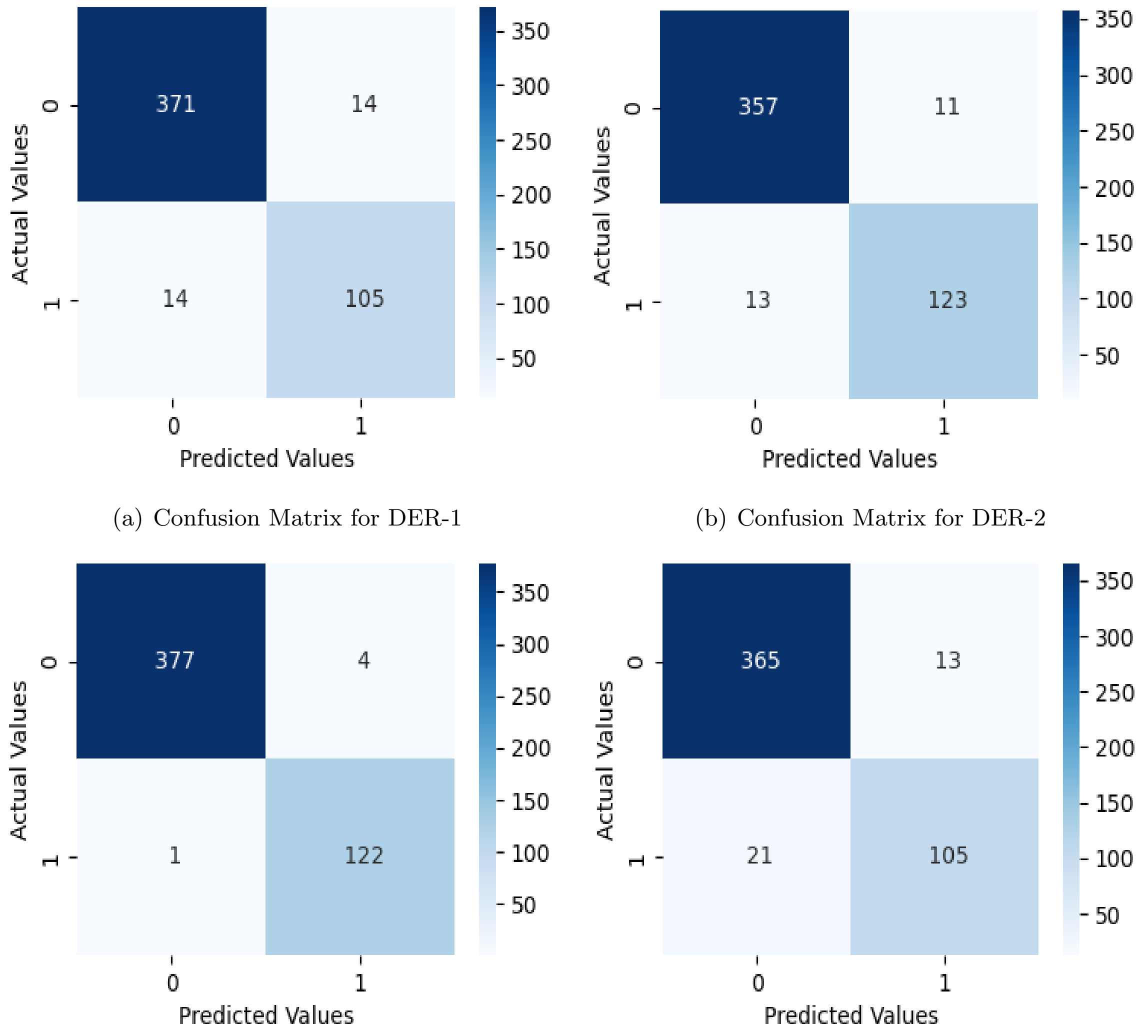
Figure 4.19: Visual representations of ROC and AUC

2. **Hamming Loss:** Alongside accuracy, *Hamming Loss* is also considered to assess model performance. It is the metric which calculates the fraction of erroneously predicted labels and mathematically represented as below:

$$\text{Hamming Loss} = \frac{\sum_j y(j) \neq \hat{y}(j)}{N} \tag{4.33}$$

where, N is the total number of labels, $y(j)$ is the j^{th} true label, and $\hat{y}(j)$ is the j^{th} predicted label.

Hamming Loss offers a thorough evaluation of the model's overall label accuracy performance. A lower Hamming loss indicates better performance, as it means that the classifier is able to predict more labels correctly.



(a) Confusion Matrix for DER-1

(b) Confusion Matrix for DER-2

(c) Confusion Matrix for DER-3

(d) Confusion Matrix for DER-4

Figure 4.20: Confusion matrix for different ML classifier.

4.5.3 Test Results

The performance of the proposed attack localization scheme is evaluated using CM as shown in Fig. 4.20 for each DERs within the microgrid system. Each CM visualizes the model's ability to distinguish between two classes viz., DER not attacked (0), and DER attacked (1).

As discussed, in this case, each DERs is considered at a time for the attack and TP, FP, TN and FN are calculated out of total 504 instances. Analysis of the CMs reveals that DER-1 has the highest False Positive Rate (FPR) at 0.04%, followed by DER-2 and DER-3 with a FPR of 0.03%. Conversely, DER-3 exhibits the lowest FPR at 0.01%. On the other hand, the False Negative Rate (FNR) is highest for DER-4 and DER-1, reaching 0.17% and 0.12% respectively. DER-2 follows with a FNR of 0.09%, and DER-3 has the lowest FNR at a mere 0.008%.

Using these information above, the performance parameters, namely, Precision, Recall, Specificity, F1-score and Accuracy, for the proposed FDI attack localization scheme is calculated and tabulated in the Table 4.7. Examining the table, it is observed that the

Table 4.7: Performance Parameters for the Proposed XGBoost Enabled Attack Localization Scheme

Attack Location	Precision	Recall	Specificity	F1-Score	Accuracy (%)
DER-1	0.88	0.88	0.96	0.88	94.44
DER-2	0.92	0.90	0.97	0.91	95.23
DER-3	0.97	0.99	0.99	0.98	99
DER-4	0.89	0.83	0.96	0.86	93.25

model achieves the highest precision and recall for DER-3. This translates to a very low chance of the model making a false positive or a false negative. This superior performance can be attributed due to the unique role played by DER-3 as "leader one" within the microgrid system. Leader DER always possesses knowledge of voltage and frequency set-points, making it a more predictable to detect compared to other DERs. Consequently, attacks launched on DER-3 are easier for the model to detect and classify accurately. Alongside the precision and recall, F1 score is also evaluated for all the DERs. An F1-score close to 1 signifies a well-balanced model that excels in both identifying true positives (attacks) and avoiding false positives. DER-3 shows the highest F1-score, indicating its well-balanced performance. Conversely, DER-4 has the lowest F1-score, suggesting the model struggles to accurately predict the attack status for DER-4 compared to others. This difficulty translates to a higher likelihood of both false positives and false negatives for DER-4. Finally, the accuracy of the model is shown which also follows the similar trend i.e highest accuracy is obtained by DER-3 and lowest for DER-4.

Figure 4.21, shows a graphical representation of true positive rate (TPR) and the

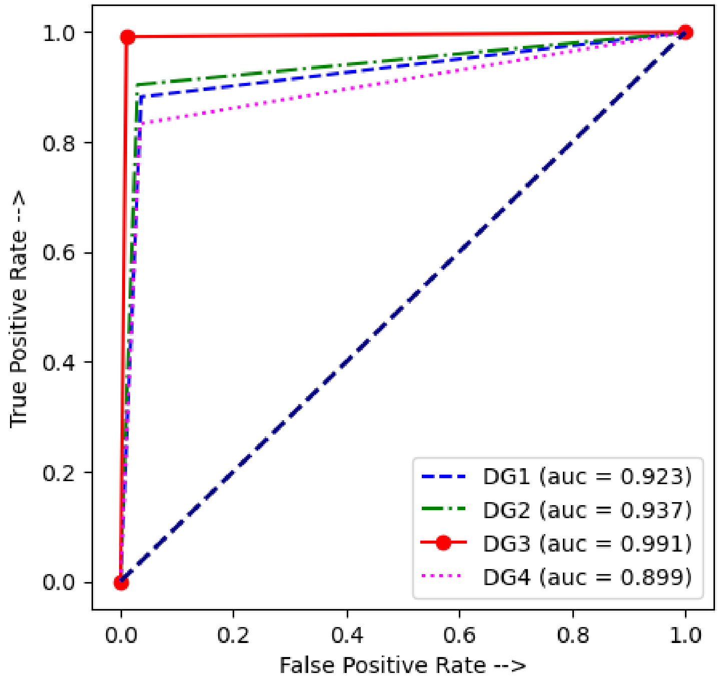


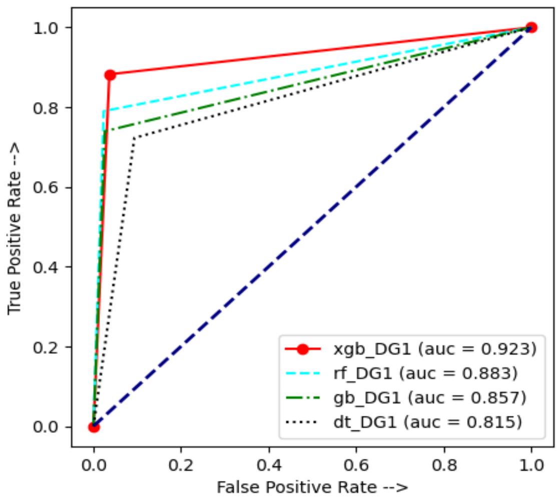
Figure 4.21: ROC curve for different DERs under attack condition

false positive rate (FPR) across different classification thresholds via ROC curve for each individual DERs. It can be seen that FPR values for DER-1 and DER-2 are closely aligned due to their similar roles (Degree score) in the prescribed communication topology, resulting in comparable detectability rates with similar TPR and FPR. The same interpretation can also be observed by computing the area under the ROC represented as AUC score. Higher is the AUC score, closer is classifier to the perfect classifier. In terms of the best attack localization performance, the DERs can be ranked from lowest to highest as follows: DER-3 > DER-2 > DER-1 > DER-4.

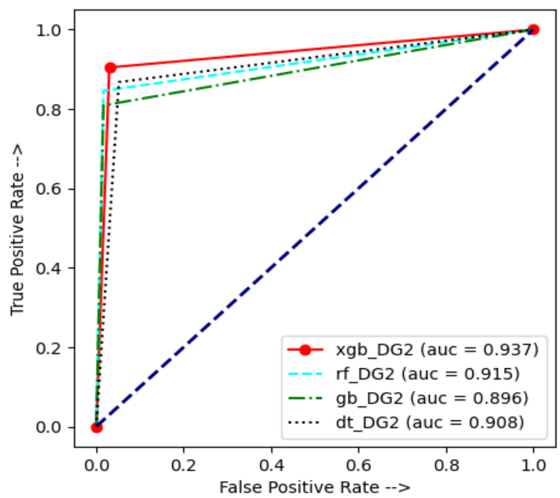
Table 4.8: Performance Comparison Among Different Classification Methods for Localization of Attacks

Classifier	**Accuracy (%)**	**Hamming Loss (%)**
Decision Tree	73.01	8.3
Random Forest	85.11	4.7
Gradient Boosting	81.34	5.8
Proposed Attack Localization Scheme	87.50	4.5

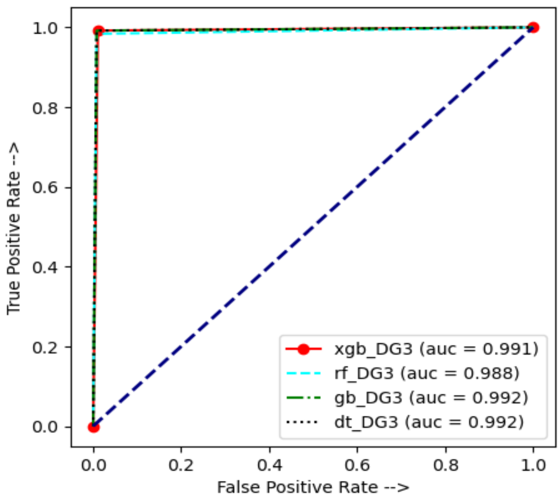
Now the performance of the proposed attack localization scheme with XGBoost as a classifier is compared with the other well known ML classifier as shown in Table 4.8. It is observed that among all the other classifiers such as RF, DT and GB, the accuracy of the proposed XGBoost assisted classifier is the highest. This is also worth noticing from



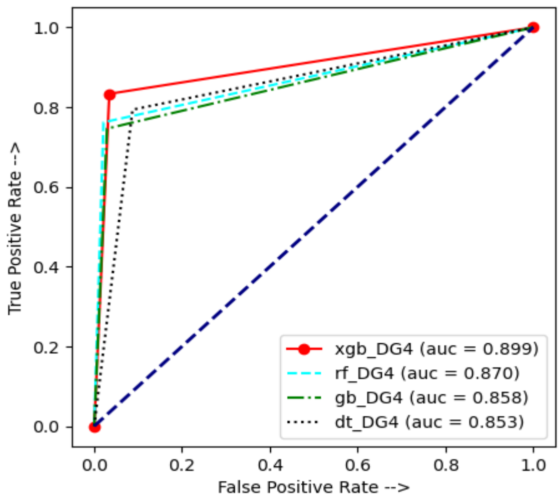
(a) ROC for DER-1 Under Different ML Techniques



(b) ROC for DER-2 Under Different ML Techniques



(c) ROC for DER-3 Under Different ML Techniques



(d) ROC for DER-4 Under Different ML Techniques

Figure 4.22: ROC curve for different DERs under different ML classifiers.

Table 4.7 and Table 4.8 that the accuracy of localizing the attack on individual DER is higher than the average accuracy of localizing the attack. The possible reason for the higher accuracy of individual DER can be the highly imbalanced dataset. Suppose, there is an attack on DER-4 i.e. the actual target value is [0 0 0 1], but the model predicts the attack on DER-3 i.e. [0 0 1 0]. Even though the model predicts incorrectly, the individual prediction for DER-1 and DER-2 is still correct resulting in higher individual accuracy for DERs. Thus, in multi-label classification, where each instance can be associated with multiple labels, accuracy alone may not be the most appropriate metric for evaluating the model performance. This is because accuracy measures the proportion of correctly predicted labels out of the total number of labels, considering each label prediction independently. This metric doesn't differentiate between minor and major classification errors. In scenarios where labels are imbalanced or where some labels are more prevalent than others, accuracy can be misleading as discussed previously. Therefore, in parallel to accuracy, hamming loss is also calculated by considering both false positives and false negatives across all classes, offering a more informative assessment of the model's ability to accurately predict the presence or absence of multiple labels for each instances. It would calculate the average number of incorrect labels across all four classes. It can be seen from the table that the proposed FDI attack localization scheme has the lowest hamming loss as compared to other considered ensemble ML techniques.

The effectiveness of the proposed attack localization scheme is also evaluated and compared to other ensemble machine learning techniques using ROC curve in Fig. 4.22. The AUC metric is employed to quantify the performance for each DER. The proposed scheme demonstrates that DER-1, DER-2 and DER-4 has significantly higher AUC values compared to the other ML techniques in the ROC curves. A higher AUC signifies a stronger ability to differentiate between attack and non-attack instances. For DER-3, the competition becomes tighter. Here, the AUC values for all the considered techniques, including the proposed scheme, are relatively close. This suggests that all methods perform at a similar level in identifying attacks on DER-3. In summary, even though the proposed approach may not be the best option for every DER, its outstanding results on DER-1, DER-2 and DER-4 demonstrate its overall effectiveness in attack localization within the microgrid system.

4.6 Conclusions

With the aim of ***accurate and timely detection*** of cyber attacks in the islanded AC MG, comprising of 4 grid forming inverters sharing the mutual information among each other through a prescribed communication topology, this chapter first introduces a novel attack detection mechanism based on Maximum Mean Discrepancy (MMD) test statistic which can calculate the discrepancies in unbiased estimates of local voltage/frequency synchronizing tracking errors for each DERs from the samples of two distributions after applying kernel tricks. Under the compromised situation, this statistical estimate behave erroneous leading the MMD to cross predefined threshold and raised the flag of attack detection. Having the attack being detected, in next stage a machine learning classifier, specifically the XGBoost is utilized with two statistical inconsistency measure i.e. Entropy and Shannon energy to form ***a novel rule-based attack classification*** approach for classifying various types of injection attacks in the DER's controllers. After the classification task being completed, ***a multi-label attack localization scheme*** is performed after exploiting a few more statistical features to be incorporated in the previous XGBoost classifier, which aids in pinpointing the specific attacked DERs, streamlining the process of isolating compromised components from the system in worst-case scenarios. Thus, by combining statistical measures and as well as ML techniques, this chapter introduced a comprehensive strategy for detecting and localizing attacks in a modified IEEE 13 bus islanded AC microgrids systems modelled in RTDS environment. The salient contributions of this chapter are as follows:

For Maximum Mean Discrepancy Based Attack Detection:

1. Accurate detection of different types of FDIAs on the controller inputs and its incoming and outgoing communications links.

2. Accurately distinguishes fault/switching events from cyber attacks, leading to no false alarms.

3. Superior as compared to Kullback Leibler divergence (KLD) [116] in terms of detection delay and threshold selection problem under varieties of attack models.

4. Due to its lower rate of change detection, the KLD performs worse than MMD under slow changing attacks with a predetermined threshold and may occasionally evade the attack.

For Proposed ML-based XGBoost Classifier to Classify and Localize Attack:

1. The proposed novel rule-base XGBoost classifier classify the FDIAs with an accuracy of 99.49% which outperform the existing ensemble ML techniques.

2. Furthermore, in terms of Precision, Recall, and F1 Score, the proposed rule-based approach similar to other ML classifiers, achieves 100% for detecting the simplest attack, such as the pulse attack. Additionally, it demonstrates significant performance for more complex attacks, such as sine attacks with the values of 97%, 98%, and 97%, respectively.

3. Apart from the classification, the proposed XGBoost enabled attack localization scheme also shows superior performance in terms of accuracy of 87.5% with a hamming loss of 4.5% which is significantly better than the existing ML classifier like Decision Tree, Random Forest, Gradient Boosting.

4. Pertaining to Precision, Recall, Specificity and F1 Score, the proposed localization scheme achieves superior performance for DER-3 followed by DER-2, DER-1 and DER-4 respectively. The same proposition is also validated from the plotting of ROC curve which results in highest AUC for each DERs with values of 92.3%, 93.7%, 99.1% and 89.9% for DER-1, DER-2, DER-3 and DER-4 respectively. The higher AUC for DER-3 signifies its superior ability to distinguish between attack and non-attack instances compared to the other DERs.

Chapter 5

Unknown Input Observer and Back-stepping Integrated Sliding Mode Control based Cyber Attack Mitigation Framework

5.1 Introduction

Having successfully detect the attack, followed by the precise attack classification and localization in Chapter-4, this chapter proposes a novel scheme to accurately estimate and mitigate cyber-attacks like unauthorized data manipulation attacks on DER's controller to maintain MG's voltage and frequency stability. This control scheme primarily comprises of two steps. The goal of the first step is to get a rough estimate of attacked DERs and the injected amount of attack bias by the perpetrator by utilizing the output of MMD obtained from Chapter-4 with an Unknown Input Observer (UIO) based control approach. In second step, the coarse estimated bias so obtained is then utilized in a Backstepping based Sliding Mode Controller (BSMC) design to generate a suitable control law that enforces the injected attack to be compensated by finer adjustments of the compensation signal due to anti-attack signal generation. Hence, from cyber attack detection in Chapter 4 to mitigation in Chapter 5, the development of the entire attack-resilient framework is outlined across four principal stages: Identification, Reconstruction, Mitigation, and Update, as depicted in Fig. 5.1. The efficacy of the proposed cyber-attack detection and mitigation scheme is tested under various types of cyber attacks on the IEEE-13 bus distribution test feeder operated in an islanded mode, modelled in RSCAD and is validated with RTDS. Moreover, the performance and superiority of the proposed detection scheme is compared with exiting ones through Hardware-in-the-Loop (HIL) simulation control environment.

The rest of the chapter is organized in four sections. Section 4.2 presents the proposed cyber attack resilient framework with details description of UIO and BSMC design methodology. The real time implementation results are analysed in Section 4.3 and finally Section 4.4 concludes the chapter.

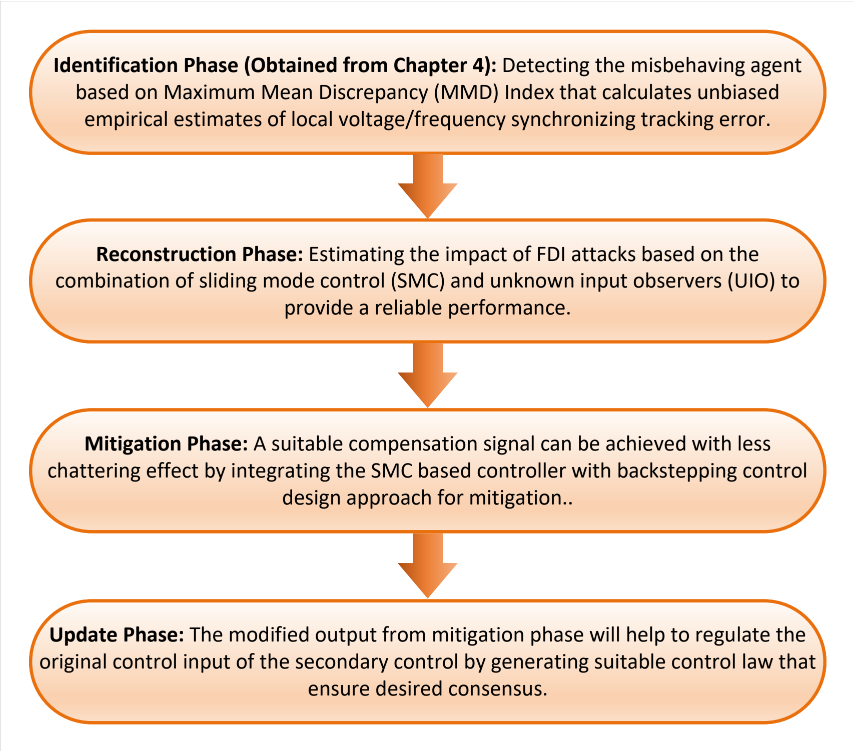


Figure 5.1: Illustration of the four key stages in the attack-resilient framework: identification, reconstruction, mitigation, and update.

5.2 Proposed Cyber Attack Resilient Framework

The overall proposed cyber-attack resilient framework is shown in Fig. 5.2 wherein a feedforward compensation method over the existing conventional secondary control system is proposed to make the existing consensus-based control scheme attack resilient. The design method initially needs an unknown input observer to monitor the system's global parameters affected by cyber-attacks and then make a rough estimation of exogenous false data injections (FDI) based on the initial observed states trajectory and designed state matrices. The estimated value of FDI is then used in the next stage, where a robust controller is designed based on sliding mode control concept and to get rid of the chattering issues and getting smooth performance, it is integrated with a very popular back-stepping control design approach so that a suitable counter control law can be generated that can compensate the effect of FDI attacks in real time on the local controllers of DERs and maintain the system stability throughout the process. This anti-attack signal generated from the control law is added with output of conventional secondary control signal to enforce the attack to be mitigated in secondary control operation of DERs. The joint utilization of this UIO and BSMC design illustrates excellent tracking and unknown injection compensation capabilities even under the compromised condition with no external hardware modifications of the existing MG distributed secondary controller and also not requiring any additional layers of communication channels.

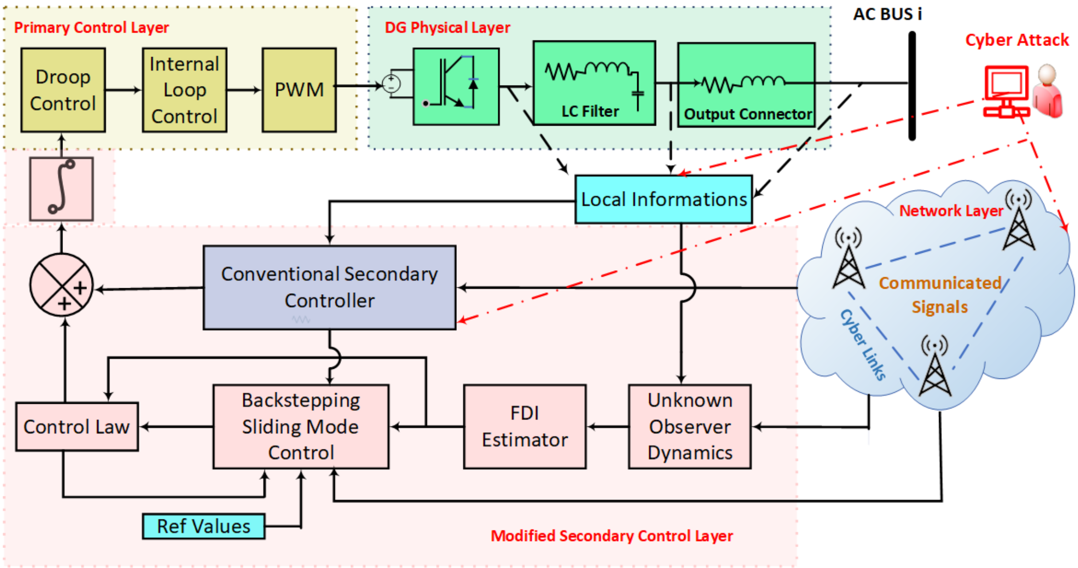


Figure 5.2: Proposed cyber attack resilient framework for MG's Distributed Secondary Control

5.2.1 UIO Design For DER's Secondary Control Layer

To start with the design process, the MG's distributed secondary control for each DERs is required to be represented in the state space domain with its networked system dynamics equations as shown below:

$$\begin{aligned} \dot{x}(t) &= \mathbf{A}x(t) + \mathbf{B}u(t) + \mathbf{E}f(t) \\ y_i &= \mathbf{C}_i x_i(t) \end{aligned} \tag{5.1}$$

where, $x(t) \in \mathbb{R}^N$ and $u(t) \in \mathbb{R}^P$ be the global states (frequency/voltage) and known input vector containing all the DER's states and $y_i(t) \in \mathbb{R}^{M_i}$ be the output vector available to DER-i. $f(t) \in \mathbb{R}^N$ be the unknown exogenous attack input scalar injected by the attacker in any respective state, and E is its associated full column rank vector defined by the output from $MMD[\mathscr{F}, \mathfrak{U}_{\omega_i}, \mathfrak{U}^a_{\omega_i}]$, obtained from Chapter-4. **A**,**B**,**C** and **E** are all known network system matrices with appropriate dimensions. Additionally, it should be noted that each agent i.e DER-i have access to its own and received measurements of its neighboring states.

Definition 5 (Distributed Estimation)**.** *In order to develop the proposed attack estimation and mitigation strategy, let's assume that each DER-i in the system Eq. (5.1) has a topological model of the MG secondary layer control systems and that a local set of measurements, $y(t)$, is accessible. Additionally, if the collection of DERs auxiliary control inputs respective to its local neighborhood synchronization errors detect the attack and locates the compromised unit—that is, locates the nonzero elements of $\boldsymbol{E}$ pertaining to an injection of $f(t)$—an attack estimation is said to have been acquired using observer's structure.*

Definition 6 (Unknown Input Observer (UIO))**.** *Considering the system dynamics in Eq. (5.1), an observer is defined as an unknown input observer, if error in estimation of states approaches to zero asymptotically i.e,* $\lim_{t\to+\infty} e(t) = \lim_{t\to+\infty}\|x(t) - \hat{x}(t)\| = 0$ *regardless of the presence of any exogenous input* $f(t)$ *in the system. Thus, the structure of the full-order UIO is given by the following system of equation [217]:*

$$\begin{aligned} \dot{z}_i^k(t) &= \boldsymbol{F}_i^k z_i^k(t) + \boldsymbol{J}_i^k \boldsymbol{B} u(t) + \boldsymbol{L}_i^k y_i(t) \\ \hat{x}_i^k &= z_i^k(t) + \boldsymbol{K}_i^k y_i(t) \end{aligned} \tag{5.2}$$

Now, if the observer equations Eq. (5.2) is applied to the MG network system dynamics equation Eq. (5.1), then the governing equation of estimation error of system states can be written in expandable form as shown below.

$$\begin{aligned} \dot{e}_i^k(t) = (\mathbf{A} - \mathbf{K}_i^k \mathbf{C}_i \mathbf{A} - \acute{\mathbf{L}}_i^k \mathbf{C}_i) e_i^k(t) + [\mathbf{F}_i^k - (\mathbf{A} - \mathbf{K}_i^k \mathbf{C}_i \mathbf{A} - \acute{\mathbf{L}}_i^k \mathbf{C}_i)] z_i^k(t) + \\ [\acute{\acute{\mathbf{L}}}_i^k - (\mathbf{A} - \mathbf{K}_i^k \mathbf{C}\mathbf{A} - \acute{\mathbf{L}}_i^k \mathbf{C}_i)\mathbf{K}_i^k] y_i(t) + [\mathbf{J}_i^k - (\mathbf{I} - \mathbf{K}_i^k \mathbf{C}_i)] \mathbf{B} u_i(t) \\ + (\mathbf{K}_i^k \mathbf{C}_i - \mathbf{I}) \mathbf{E} f_i(t) \end{aligned} \tag{5.3}$$

where k^{th} is the target node/DER compromised by cyber-attacks. Also, $\dot{z}_i^k(t) \in \mathbb{R}^N$ is the states of observer and $\hat{x}_i^k \in \mathbb{R}^N$ is the estimated states decoupled from compromised node-k and calculated by node-i. It is assumed that each DER is equipped with a UIO, which only needs its own state and local measurement (y_i) information of its neighbor. The matrices, $\mathbf{F}_i^k$, $\mathbf{J}_i^k$, $\mathbf{L}_i^k$ and $\mathbf{K}_i^k$ are the design consideration of proper dimensions which must satisfy the following relations.

$$\begin{aligned} \mathbf{F}_i^k &= (\mathbf{A} - \mathbf{K}_i^k C_i \mathbf{A} - \acute{\mathbf{L}}_i^k \mathbf{C}_i) \\ \mathbf{K}_i^k &= \mathbf{E}_i [(\mathbf{C}_i \mathbf{E}_i)^T \mathbf{C}_i \mathbf{E}_i]^{-1} (\mathbf{C}_i \mathbf{E}_i)^T \\ \mathbf{J}_i^k &= (\mathbf{I} - \mathbf{K}_i^k \mathbf{C}_i) \\ \acute{\acute{\mathbf{L}}}_i^k &= \mathbf{F}_i^k \mathbf{K}_i^k \\ \mathbf{L} &= \acute{\mathbf{L}}_i^k + \acute{\acute{\mathbf{L}}}_i^k \\ (\mathbf{K}_i^k \mathbf{C}_i - \mathbf{I}) &= 0 \end{aligned} \tag{5.4}$$

Henceforth, the state estimation error will be:

$$\dot{e}_i^k(t) = \mathbf{F}_i^k e_i^k(t) - \mathbf{J}_i^k \sum_{n \in \bar{N}_i \setminus \{k\}} \mathbf{E}_n f_n(t) \tag{5.5}$$

Here, $\mathbf{F}_i^k$ is a Hurwitz matrix i.e all the eigenvalues of $\mathbf{F}$ are stable, and thus asymptotically convergence of state estimation errors, $e(t)$ towards zero is guaranteed. Moreover, the unknown input term is also effectively decoupled from the observed states and gradually $\hat{x}$ converges to x. It is evident that the error dynamics Eq. (5.5) are stable and independent of $f_k(t)$, which aligns with Definition 2. Here, the observer matrices are designed based on pole-placement techniques [218]. Now, for the existence of an UIO and to solve Eq. (5.4), the necessary and sufficient conditions that need to be checked are as follow:

1. $\text{rank}(\mathbf{C}_i\mathbf{E}_k) = \text{rank}(\mathbf{E}_k)$

   $$\begin{bmatrix} sI - \mathbf{A}+ & (\mathbf{E}_k) \\ \mathbf{C}_i & 0 \end{bmatrix} = n + \text{rank}(\mathbf{E}_k)$$

2. $(\mathbf{C}, \mathbf{A}_1)$ is an observable pair, where $\mathbf{A}_1 = \mathbf{A} - \mathbf{E}[(\mathbf{CE})^T\mathbf{CE}]^{-1}\mathbf{CE})^T\mathbf{CA} = \mathbf{A} - \mathbf{AKC}$.

Thus, the estimate of unknown input injection can be obtained as follows:

$$\hat{f} = (\mathbf{C}_i\mathbf{E}_i)^{\dagger}[\dot{y}_i - \mathbf{CA}\hat{x}_i^k - \mathbf{CB}u] \tag{5.6}$$

In order to prove the observer global stability and asymptotic convergence of estimation of unknown exogenous input, lets the error is defined as

$$\tilde{f} = f - \hat{f} \tag{5.7}$$

Let, the Lyapunov function be taken as

$$V = \frac{1}{2}\tilde{f}^T\tilde{f} \tag{5.8}$$

Combining Eq. (5.1) and Eq. (5.6), the first derivative of V can be written as:

$$\begin{aligned}
\dot{V} &= \tilde{f}^T \ \dot{\tilde{f}}^T \\
&= \dot{\tilde{f}}^T[f - \hat{f}]^T \\
&= \dot{\tilde{f}}^T[f - (\mathbf{C}_i\mathbf{E}_i)^{\dagger}(\dot{y}_i - \mathbf{CA}\hat{x}_i^k - \mathbf{CB}u)]^T \\
&= \dot{\tilde{f}}^T[f - (\mathbf{C}_i\mathbf{E}_i)^{\dagger}(\mathbf{C}_i\mathbf{A}x_i^k + \mathbf{C}_i\mathbf{B}u + \mathbf{C}_i\mathbf{E}_i f - \mathbf{C}_i\mathbf{A}\hat{x}_i^k - \mathbf{C}_i\mathbf{B}u)]^T \\
&= \dot{\tilde{f}}^T[f - f - (\mathbf{C}_i\mathbf{E}_i)^{\dagger}\mathbf{C}_i\mathbf{A}e(t)]^T \\
&= -[(\mathbf{C}_i\mathbf{E}_i)^{\dagger}\mathbf{C}_i\mathbf{A}e(t)]^T\dot{\tilde{f}}^T \leq 0
\end{aligned} \tag{5.9}$$

Therefore, given microgrid system Eq. (5.1) and its initial condition $x(t_0) = x_0$, along with the UIO estimator Eq. (5.2) proposed for this system, if the parameters of the estimator in Eq. (5.2) meet the above conditions specified in Eq. (5.4) and remains constant, then as $t \to \infty$, $z \to Jx$ and $\hat{f} \to f$. Thus, it is observed that with proper selections of observer design matrices and following the necessary conditions for UIO existence, the unknown estimation error approaches to zero asymptotically with the estimation error of observer states. Figure 5.3 shows the block diagram of designing a UIO dynamic system for MG network, as described in Eq. (5.2).

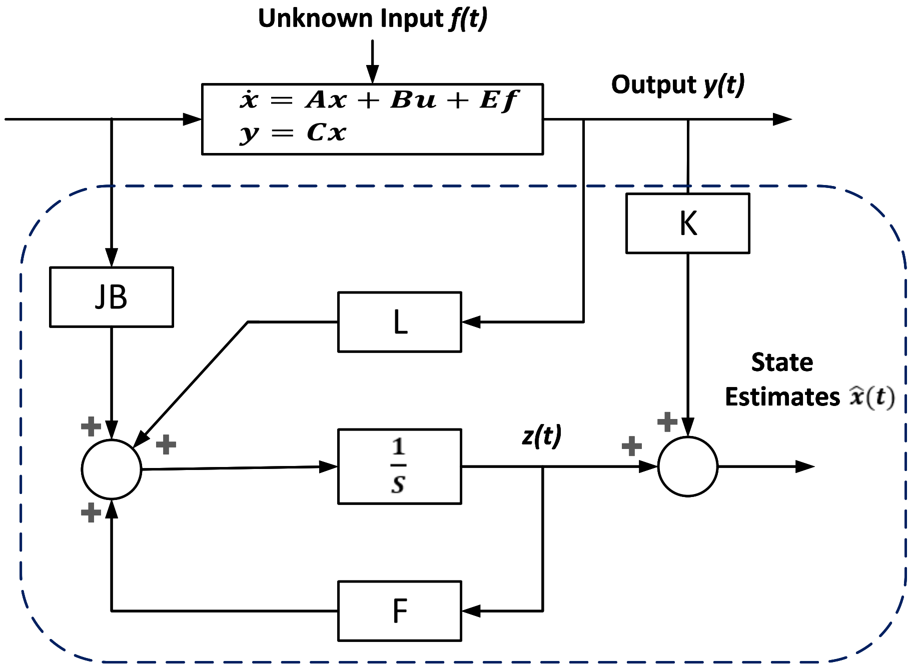


Figure 5.3: Schematic structure of a full order Unknown Input Observer

5.2.2 Back-stepping Integrated Sliding Mode Controller

This is the final stage of the proposed algorithm to design a compelling feedback control law under a fixed communication graph topology. This involves employing a back-stepping approach and sliding mode surface to counteract the impact of unknown FDI attacks on MG secondary controllers. Additionally, it includes the approximate estimation of injections using the above-mentioned UIO, contributing to the improvement of system stability and robustness. In the context of non-linear control theory, sliding mode control (SMC) is a widely reviewed researched area due to its superiority in trajectory tracking problems in multi-agent system fast response, controlling of model uncertainties, disturbance, and unmodeled system dynamics, etc [219, 220]. The main philosophy of SMC is to design a suitable sliding surface such that even in the presence of an unknown attack or disturbance, the states of the system can reach to this surface and stay over there to achieve resilient control performance. However, finite-time consensus and chattering issues in its control input are the major concern of its practical applications. To get rid of these above issues and to purposefully extract and utilize the benefits of SMC, another very popular non-linear recursive control technique, i.e., back-stepping control, is integrated with it, which is formulated based on Lyapunov functions in each step and found out to be superior for ensuring global stability of strict feedback systems. Initially, the entire system is splitted into several reduced-order subsystems, and some error or regulatory variables are introduced to redefine the networked system dynamics based on Lyapunov functions. Thereafter using the time derivative of those variables at each design step, a fictitious or virtual control input law is defined that acts as a stabilizing control for its previous states, and next stability is ensured by making the time derivative of Lyapunov functions to be negative definite as briefly described below.

Let, the MG's distributed secondary frequency control (DSFC) Eq. (5.1) for the topology shown in Fig. 4.13 of Chapter-4 is modified as follows:

$$\begin{aligned} \dot{X}_\omega &= \mathbf{A}X_\omega + \mathbf{B}_p U_p + \mathbf{B}_u U_{s\omega} + \mathbf{B}_f \hat{f}_\omega \\ y_\omega &= \mathbf{C}X_\omega \end{aligned} \tag{5.10}$$

where, X_ω and U_P be the DER's frequency and its injected active power input respectively. $\hat{f}_\omega$ is the rough estimation of unknown frequency bias output obtained from UIO and $U_{s\omega}$ be the desired control law that makes the DSFC cyber attack resilient.

$$X_\omega = [\omega_1, \omega_2, \omega_3, \omega_4]^T, \qquad U_p = [P_1, P_2, P_3, P_4]^T$$
$$U_{s\omega} = [u_{s\omega_1}, u_{s\omega_2}, u_{s\omega_3}, u_{s\omega_4}]^T, \qquad \hat{f}_\omega = [\hat{f}_{\omega_1}, \hat{f}_{\omega_2}, \hat{f}_{\omega_3}, \hat{f}_{\omega_4}]^T$$

$$\mathbf{A} = \begin{vmatrix} -2c_\omega & c_\omega & 0 & c_\omega \\ c_\omega & -3c_\omega & c_\omega & c_\omega \\ 0 & 0 & -2c_\omega & c_\omega \\ c_\omega & 0 & 0 & -c_\omega \end{vmatrix} \qquad \mathbf{B}_f = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

where, c_ω is the frequency control coupling gain value for MG's DSFC of voltage source voltage controlled inverters.

$$\mathbf{B}_p = \begin{vmatrix} -2c_\omega & c_\omega & 0 & c_\omega \\ c_\omega & -3c_\omega & c_\omega & c_\omega \\ 0 & 0 & -c_\omega & c_\omega \\ c_\omega & 0 & 0 & -c_\omega \end{vmatrix} \qquad \mathbf{B}_u = \begin{vmatrix} b_1 & 0 & 0 & 0 \\ 0 & b_2 & 0 & 0 \\ 0 & 0 & b_3 + c_\omega & 0 \\ 0 & 0 & 0 & b_4 \end{vmatrix}$$

Firstly, the tracking error of the DER-1's frequency is defined as:

$$\overline{\mathrm{w}}_{\omega_1} = \omega_{ref} - \omega_1 \tag{5.11}$$

The Lyapunov function is chosen as:

$$\mathrm{V}_1 = \frac{1}{2}\overline{\mathrm{w}}_{\omega_1}^2 \tag{5.12}$$

Combining Eq. (5.10) and Eq. (5.11), the first derivative of V_1 can be written as

$$\begin{aligned} \dot{\mathrm{V}}_1 &= \overline{\mathrm{w}}_{\omega_1}\dot{\overline{\mathrm{w}}}_{\omega_1} \\ \dot{\mathrm{V}}_1 &= \overline{\mathrm{w}}_{\omega_1}(\dot{\omega}_{ref} - \dot{\omega}_1) \end{aligned}$$

$$\dot{\mathrm{V}}_1 = \overline{\mathrm{w}}_{\omega_1}(\underbrace{\dot{\omega}_{ref} + 2c_\omega\omega_1 - c_\omega\omega_2 - c_\omega\omega_4 + 2c_\omega P_1 - c_\omega P_2 - c_\omega P_4 - b_1 u_{s\omega_1} - \hat{f}_{\omega_1}}_{-k_1\overline{\mathrm{w}}_{\omega_1}}) \tag{5.13}$$

According to the back-stepping control, the virtual control for DER-1 frequency, i.e, ω_1^* can be defined as

$$\omega_1^* = \frac{1}{2}(\omega_2 + \omega_4) - P_1 + \frac{1}{2}(P_1 + P_2) + \frac{1}{2c_\omega}(b_1 u_{s\omega_1} + \hat{f}_{\omega_1} - k_1\overline{\mathrm{w}}_{\omega_1} - \dot{\omega}_{ref}) \tag{5.14}$$

Using Eq. (5.13) and Eq. (5.14), the first derivative of V_1 is given by

$$\dot{\mathrm{V}}_1 = -k_1\overline{\mathrm{w}}_{\omega_1}^2 \leq 0. \tag{5.15}$$

In the similar fashion, the virtual controls of all other DERs such as ω_2^*, ω_3^* and ω_4^* and its respective time derivative of Lyapunov functions i.e $\dot{\mathrm{V}}_2 = -k_2\overline{\mathrm{w}}_{\omega_2}^2 \leq 0$, $\dot{\mathrm{V}}_3 = -k_3\overline{\mathrm{w}}_{\omega_3}^2 \leq 0$ and $\dot{\mathrm{V}}_4 = -k_4\overline{\mathrm{w}}_{\omega_4}^2 \leq 0$ are also calculated. Where k_1, k_2, k_3 and k_4 are feedback gains and should be greater than zero. In order to get the output from the controller, the sliding manifold for DER-1 is defined as follows:

$$S_{\omega_1} = \tilde{\omega}_1 - k_5\int_0^t \tilde{\omega}_1\, dt \tag{5.16}$$

where, $\tilde{\omega}_1 = \omega_1^* - \omega_1$. In the same way, sliding surfaces for all the other DERs are also calculated and the final Lyapunov function can be written as

$$\mathrm{V}_5 = \frac{1}{2}(S_{\omega_1}^2 + S_{\omega_2}^2 + S_{\omega_3}^2 + S_{\omega_4}^2) \tag{5.17}$$

Combining Eq. (5.10), Eq. (5.13) and Eq. (5.16) the time derivative of V_5 can be written as

$$\begin{aligned}\dot{\mathrm{V}}_5 = S_{\omega_1}[\dot{\omega}_1^* - \dot{\omega}_1 - k_5\tilde{\omega}_1] + S_{\omega_2}[\dot{\omega}_2^* - \dot{\omega}_2 - k_6\tilde{\omega}_2]\\ +S_{\omega_3}[\dot{\omega}_3^* - \dot{\omega}_3 - k_7\tilde{\omega}_3] + S_{\omega_4}[\dot{\omega}_4^* - \dot{\omega}_4 - k_8\tilde{\omega}_4]\end{aligned} \tag{5.18}$$

For brevity, only the first term of Eq. (5.18) is expanded as shown below which gives the suitable control law for DER-1

$$\begin{aligned}\dot{\mathrm{V}}_5(1^{st}\ term) = S_{\omega_1}[&\underbrace{\dot{\omega}_1^* + k_5\tilde{\omega}_1 + 2c_\omega\omega_1 - c_\omega\omega_2 - c_\omega\omega_4}\\ &+ \underbrace{2c_\omega P_1 - c_\omega P_2 - c_\omega P_4 - b_1 u_{s\omega_1} - \hat{f}_{\omega_1}}_{-k_9|S_{\omega_1}|^{\frac{1}{2}} sgn(S_{\omega_1}) - k_{10}|S_{\omega_1}|^{\frac{1}{2}} S_{\omega_1}}]\end{aligned} \tag{5.19}$$

Now, based on the back-stepping sliding mode principal, the output control law for DER-1 can be expressed as follows:

$$u_{s\omega_1} = \frac{1}{b_1}[\dot{\omega}_1^* + k_5\tilde{\omega}_1 + 2c_\omega\omega_1 - c_\omega\omega_2 - c_\omega\omega_4 + 2c_\omega P_1 - c_\omega P_2 - c_\omega P_4 - \hat{f}_{\omega_1} + k_9|S_{\omega_1}|^{\frac{1}{2}}sgn(S_{\omega_1}) + k_{10}|S_{\omega_1}|^{\frac{1}{2}}S_{\omega_1}] \quad (5.20)$$

Similar to the above approach, the control law for the other DERs can also be evaluated. Now according to Eq. (5.18), Eq. (5.19), Eq. (5.20) and the control laws of other DERs, the time derivative of the Lyapunov function of the entire controller can be written as

$$\dot{\mathrm{V}}_5 = -k_9|S_{\omega_1}|^{\frac{1}{2}}sgn(S_{\omega_1}) - k_{10}|S_{\omega_1}|^{\frac{1}{2}}S_{\omega_1} - k_{11}|S_{\omega_2}|^{\frac{1}{2}}sgn(S_{\omega_2}) - k_{12}|S_{\omega_2}|^{\frac{1}{2}}S_{\omega_2} - k_{13}|S_{\omega_3}|^{\frac{1}{2}}sgn(S_{\omega_3}) - k_{14}|S_{\omega_3}|^{\frac{1}{2}}S_{\omega_3} - k_{15}|S_{\omega_4}|^{\frac{1}{2}}sgn(S_{\omega_4}) - k_{16}|S_{\omega_4}|^{\frac{1}{2}}S_{\omega_4} \leq 0 \quad (5.21)$$

where all the parameters from k_9 to k_{16} are greater than 0 and $\dot{\mathrm{V}}_5$ will be zero if and only if all the sliding surfaces satisfy i.e., $S_{\omega_1} = S_{\omega_2} = S_{\omega_3} = S_{\omega_4} = 0$. Thus it is proved that the designed controller is asymptotically stable in a global sense based on Lyapunov functions. Finally, it can be concluded that in normal conditions, this computed control law will inject almost zero compensation, but in the presence of any cyber intrusion, a suitable counter value of the control law will be injected on the DSFC of the respective attacked DER identified by MMD to make it attack resilient.
The flowchart of the above mentioned unified end-to-end cyber attack detection and mitigation framework comprising of MMD, UIO and BSMC is depicted in Fig. 5.4.

5.3 Results and Discussion

This section demonstrates the effectiveness of the proposed cyber-attacks mitigation framework on a modified IEEE 13-node distribution feeder system, depicted in Fig. 4.1 of Chapter-4. This system incorporates four DERs connected through a 1.0 MVA, 0.48/4.16 kV Yg-Yg transformer, with DER-3 designated as the leader. The hardware-in-loop laboratory setup, illustrated in Fig. 4.2, collects auxiliary control inputs from each DER's RTDS front panel analogue output channels. Subsequently, the MMD-based attack detection algorithm outlined in Chapter-4 is translated into C code using MATLAB Simulink® C Coder builder and executed on the DS1104 R&D controller board. The digitized detection input is then transmitted to the RTDS runtime environment via the digital I/O panel interface to initiate the mitigation strategy.

5.3.1 Attack Mitigation on DSFC Against Scaling Attack

Figure 5.5 illustrates the applicability and robustness of the proposed mitigation strategy against a complex attack (State Dependent) i.e., scaling attacks. Here, the attackers use the attack parameter $\alpha_{sc} = 1.04$ pu to hijack the DSFC of DER-1 and falsely alter its frequency information. As soon as the attacked DER is identified by MMD, the column vector, "E" of UIO's governing equation, is updated to get a rough estimates of the

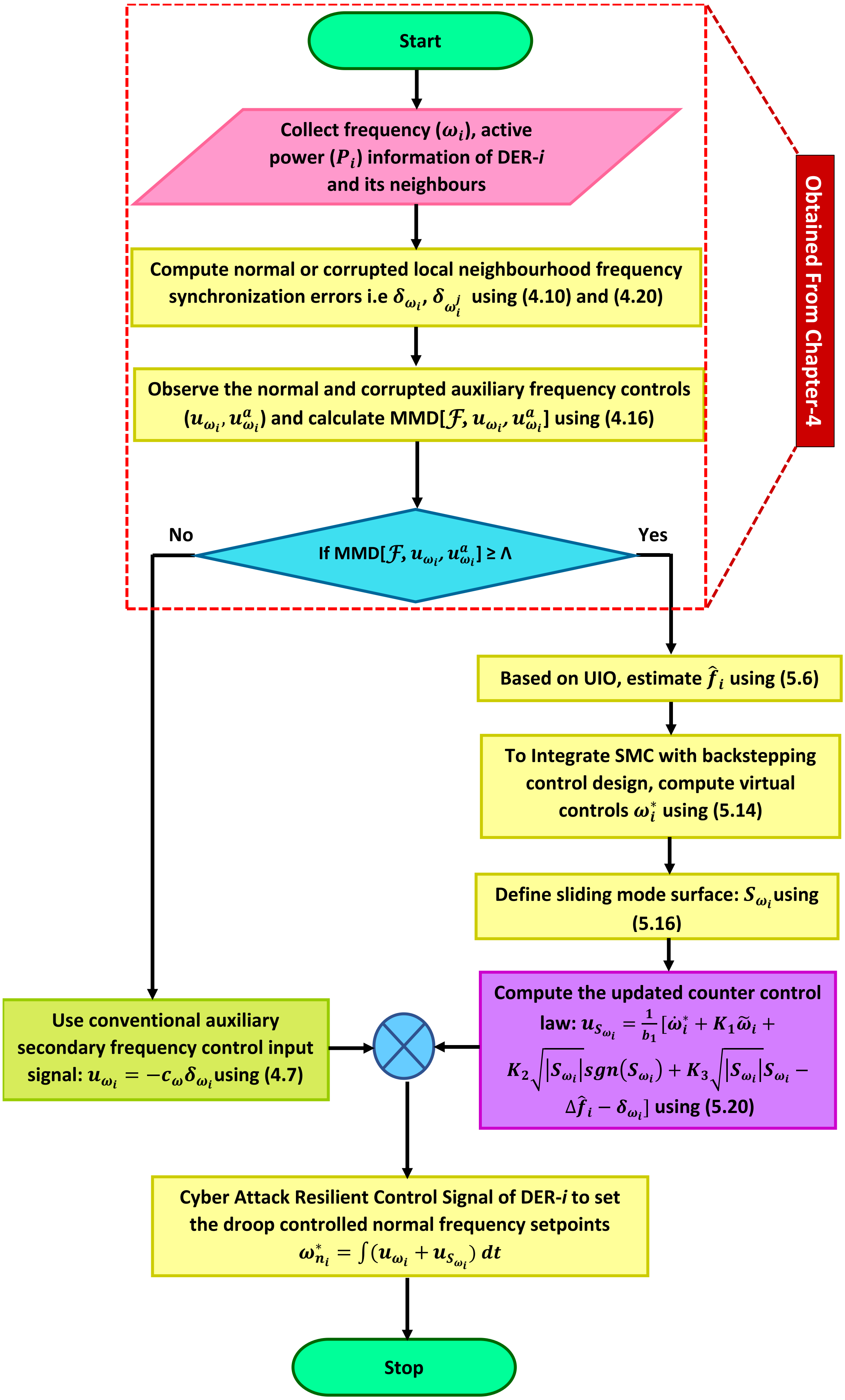


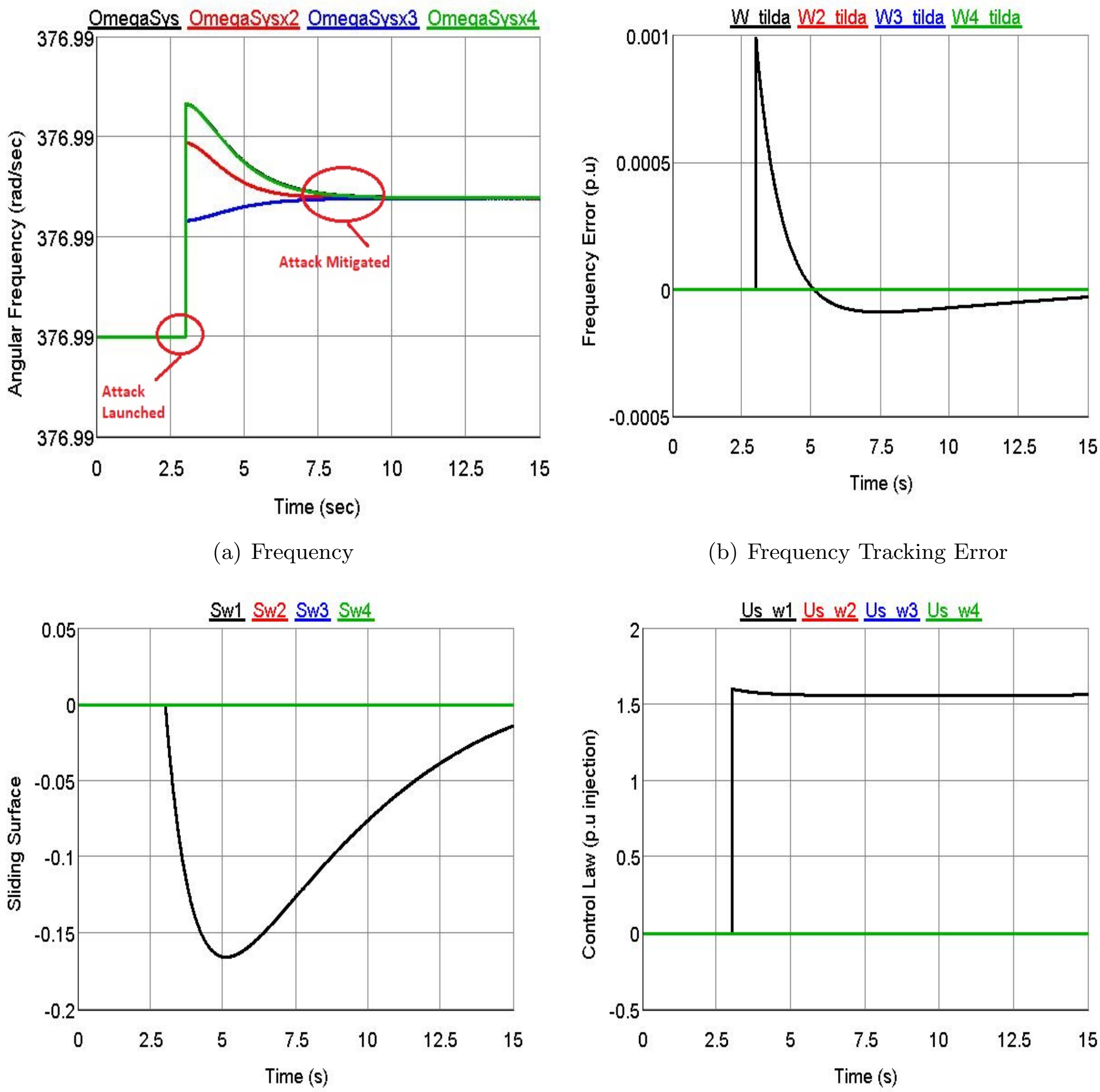
Figure 5.4: Flowchart of an unified cyber attack mitigation framework

unknown injection and the state trajectories. The proposed back-stepping assisted sliding mode-based resilient operation has activated instantly and the counter control law is being computed to restore the frequency back to normalcy. The controller and observer design requirements, which are less conservative, enable the control law to be updated in real time with the estimates of the UIO. In order to enforce robust behavior even in the event that the attack continues, it enforces the deviated state trajectories to follow back the initial sliding manifold. Figures 5.5(a) and 5.5(b) show that as soon as the attack is injected, the frequency begins to shoot out. However, the control law instantly corrects the attack's effect and restores the frequency, and within the due period, the frequency tracking error of DER-1 likewise reaches zero. The sliding manifold self-adjusted in accordance with changes in frequency deviation, as seen in Fig. 5.5(c). Next, the finer resolution of the injected control law and its endurance against persistent attack is also shown in Fig. 5.5(d). Thus it can be concluded that the combined effect of back-stepping and sliding mode control approach yields a suitable control law to maintain stability and performance of the system even in the presence of malicious disturbances, ensuring the system's resilience to cyber threats.

The estimation accuracy of the states, as resulted by the proposed UIO, is quantified in terms of mean squared error (MSE) and mean absolute error (MAE) of each DERs frequency deviation and unknown exogenous input estimation under FDIA, and is found to be — $(0.0020, 3.4959e^{-06}, 2.2979e^{-06}, 0.8132e^{-07}$ and $0.0020)$ and $(0.1004, 0.0027, 0.0019, 5.0747e^{-04}$ and $0.2079)$, respectively.

5.3.2 Attack Mitigation on DSFC Against Step Attack

Similar to the above case study, another type of commonly known attack i.e., step attack is encountered to replace the DER-1 frequency with the attack parameter of $\alpha_{st} = 0.05$ pu as modeled in [221]. As in the earlier case, the attack here is also launched at the same instant, and the proposed mitigation method is found to be working efficiently to damp out the attack's impact in DSFC, as shown in Fig. 5.6. At the onset of the step attack, Fig. 5.6(a) depicts the initial fall back of frequency from its nominal value with the correct tracking of frequency error terms for DER-1 only as identified from Fig. 5.6(b). This acts as a triggering instant for the quickest activation of the proposed BSMC assisted mitigation method, where the intermediate control signal is designed in each steps considering the system dynamics of each DERS to drive the error state towards a desired values, while also accounting for the previous control signals. This in turn introduces a resultant change in the sliding surface as shown in Fig. 5.6(c). This sliding surface henceforth generates a switching control law as depicted in Fig. 5.6(d) that guides the system trajectory towards a desired state despite the presence of cyber-attacks. The designed control law is then augmented with the governing equations of DSFC of each DERs of the MG system to include terms that actively compensate for the effects of the unknown attack inputs. In other words, this compensation mechanism dynamically adjusts the control inputs to counteract the disturbances caused by the attack, thereby minimizing its impact on the



(a) Frequency

(b) Frequency Tracking Error

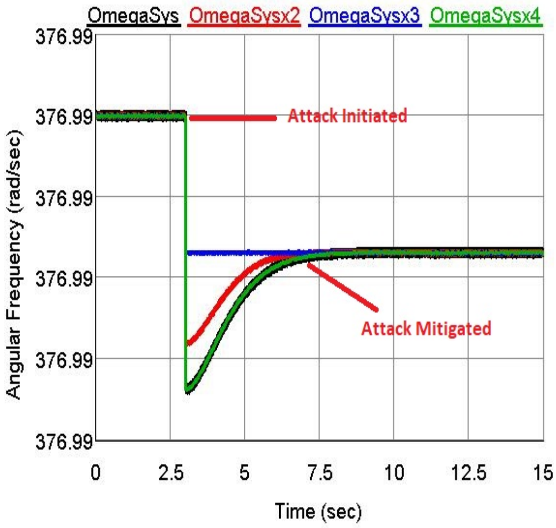
(c) Sliding Surface

(d) Control Law

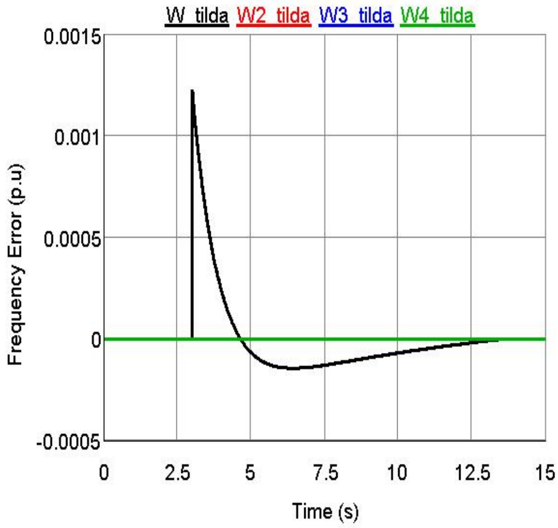
Figure 5.5: Performance of the proposed attack mitigation scheme for Scaling Attack on DSFC of DER-1. The figure color labels black, red, blue and green represents parameters for DER-1, DER-2, DER-3 and DER-4 respectively.

system.

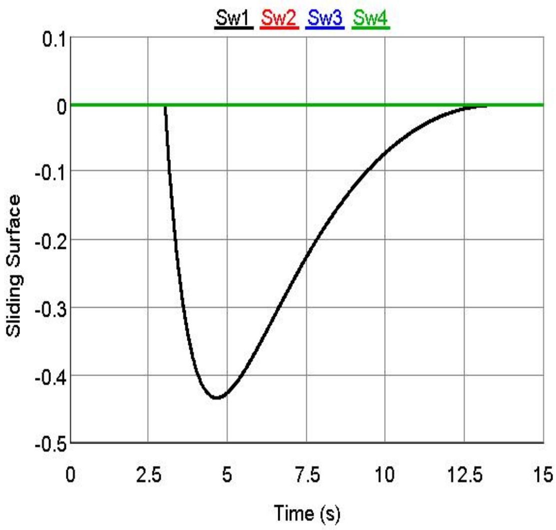
The MSE and MAE of all four DER's frequency and unknown exogenous input for this case are found to be $(0.0020, 3.4959e^{-06}, 2.2979e^{-06}, 0.8132e^{-07}$ and $0.0020)$ and $(0.1004, 0.0027, 0.0019, 5.0747e^{-04}$ and $0.2079)$ respectively. Overall, it is observed that the back-stepping approach helps account for the system's dynamics and the unknown attack's presence in addition to minimize the potential chattering issues in the control law, while the sliding mode component ensures the system's state reaches the desired behavior despite some level of disturbance in the unknown input estimation process.



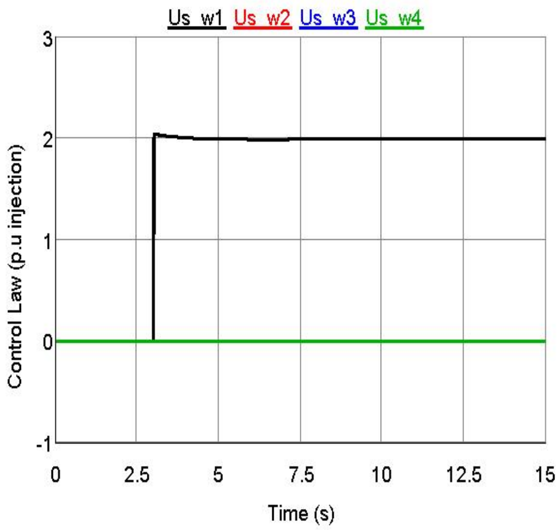
(a) Frequency



(b) Frequency Tracking Error



(c) Sliding Surface



(d) Control Law

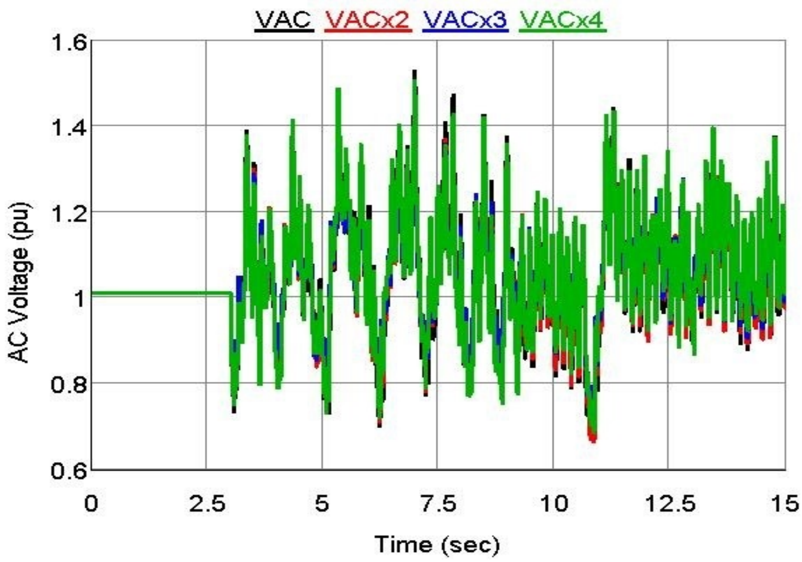
Figure 5.6: Performance of the proposed attack mitigation scheme for Step Attack on DSFC of DER-1. The figure color labels black, red, blue and green represents parameters for DER-1, DER-2, DER-3 and DER-4 respectively.

5.3.3 Impact on DER's Bus Voltage Profile by the Proposed Attack Mitigation Scheme

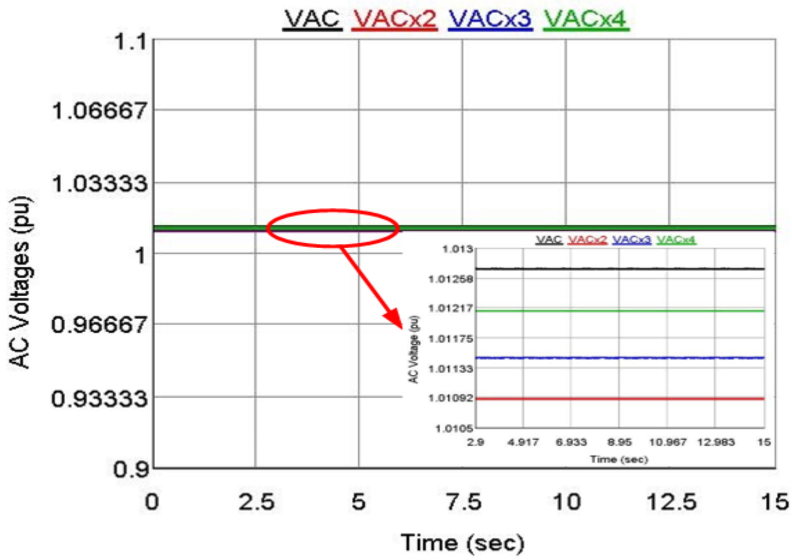
In reference to the above simulation scenario, Fig. 5.7 shows the explicit results of DER's bus voltage profile before and after activation of the proposed mitigation scheme. The overall stability profile of MG's bus voltage (Node 650) is compromised as a result of the attacker's malicious contamination of the victim DER's secondary controller's frequency input, as shown by the oscillatory behavior shown in Fig. 5.7(a). The remaining three DERs are connected to other nodes of the MG, and their voltage also deviates from the typical bus voltage range, as a result of the attack rapidly spreading to those nodes. Therefore, the stability of the entire system is compromised if the cyber-attack is not promptly detected and mitigated. As the actual voltage references are now lost due to such skillful manipulation of DER frequency parameter, the main objectives of modelling distributed cooperative secondary control in MGs i.e., *"All DERs should co-operate in maintaining consensus among themselves"* gets violated. The effectiveness of the proposed resilient control mechanism is demonstrated in Fig. 5.7(b), where the attack impact is adaptively mitigated without having a significant impact on the MG's DER bus voltage profile, which continues to remain within their acceptable voltage range, i.e., between 0.9 p.u. and 1.1 p.u., even in the event that the attack persists continuously. Additionally, it is clear from Fig. 5.7(c) that the MMD-based detection scheme proposed in Chapter-4 promptly identifies the attacked DER-1 (shown by the yellow curve). This, in turn, triggers the proposed attack mitigation scheme, in which the proposed back-stepping integrated sliding mode controller effectively reduces the impact of the attack by creating an adaptive control law in response to the attack scenario. Overall, this detection and mitigation stages are completed within a cycle of the nominal system frequency. Thus the real-time implementation of the proposed controller clearly establish the faster convergence speed and consensus agreement of MG's voltage and frequency parameters by the proposed mitigation scheme under bounded cyber intrusions. Also, the proposed controller is clearly very adaptive to the unknown bounded attack injections, and thus it does not need to isolate the infected DER's information; therefore, the system's resiliency and DER's utilization is increased.

5.4 Conclusions

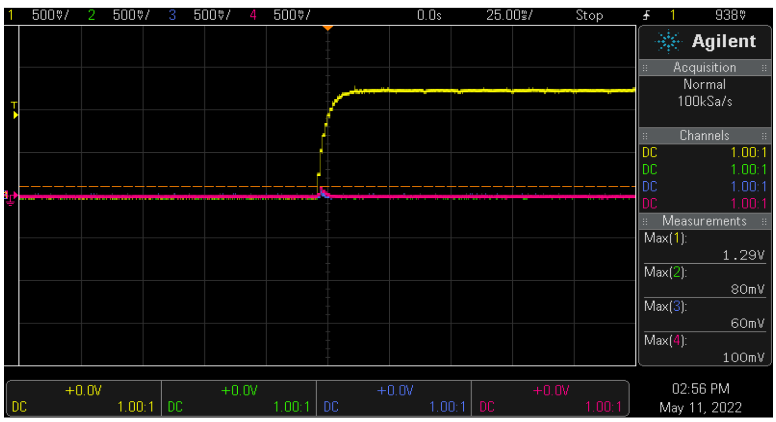
This chapter uses a sliding mode control-based attack mitigation method integrated with a back-stepping controller to help achieve consensus agreement even in a compromised situation. In its first stage, this framework needs the coarse estimation of the unknown input bias in the secondary frequency controller input obtained from a bank of unknown input observer (UIO). After that, designing an effective Back-stepping Integrated Sliding Mode Controller (BSMC) aids in nullifying the effect of the injected false data through the application of counter control law, pushing the system's behavior towards the desired trajectory despite the attack's influence. The efficacy and robustness of the proposed



(a) Before Application of Proposed Attack Mitigation Scheme



(b) After Application of Proposed Attack Mitigation Scheme



(c) Estimates of MMD Under Step Attack

Figure 5.7: Impact on DER's bus voltage profile before and after application of proposed attack mitigation scheme. The figure color labels black, red, blue and green represents AC terminal voltage of DER-1, DER-2, DER-3 and DER-4 respectively.

mitigation framework are validated by various Real-time digital simulations on the modified IEEE 13-bus system which led to the following key conclusions of the work.

- The proposed method does not impose any additional limitations on the proposed mitigation strategy, unlike [131, 116, 119], which assume that the leader DER must always be secured. Additionally, the number of role statuses (corrupted or healthy) of nearby DERs does not put any extra limitation over the proposed mitigation strategy.
- For both the attack cases, the Mean Squared Error (MSE) and Mean Absolute Error (MAE) values for the frequency of all four DERs and the unknown exogenous input are with in the acceptable tolerance which justify its accuracy of performance measure.
- The proposed mitigation method exhibits good robustness and faster convergence against different attacks and efficiently regulates DER's frequency and active power ratio.
- The proposed controller is very adaptive to the unknown bounded attack injections, and thus it does not need to isolate the infected DER's information; therefore, the system's resiliency and DER's utilization is increased.

Chapter 6

Synergistic Islanding and Cyber Attack Detection Scheme

6.1 Introduction

In the previous chapter, a cyber attack mitigation framework against the distributed consensus secondary control scheme for a MAS within a MG cyber-physical system is developed to enhance the resiliency and security of the D-Systems. Another pressing challenge that exists in D-Systems is to accurate detection of islanding scenario considering the threats of cyber-physical manipulations. Thus the goal of this chapter is to first propose a simple yet effective statistical parameters based passive islanding detection scheme (IDS) that relies only on the one phase voltage data, measured at the point of common coupling (PCC) followed by a signal processing based cyber attack detection method that aims to make existing islanding detection methods "attack-proof" or less susceptible to cyber attacks. To this end, at first an accurate islanding detection scheme is proposed, which is comprised of three main stages. In *Stage-1*, the scheme performs a quick analysis of the mean value of the PCC bus voltage under both balanced and unbalanced conditions to detect islanding in a coarse manner. *Stage-2* involves the computation of a Decaying DC Detector (DDCD) using statistical properties of the input signal. Finally, *Stage-3* introduces a Statistical Relay Digital Logic (SRDL) circuit based on output of *Stage-1* and *Stage-2* to differentiate between islanding and non-islanding events. Next, to prevent cyber attacks from manipulating the islanding decisions and misleading system operators, the proposed IDS is combined with a signal processing based Cyber Attack Detector (CAD) module for the detection of statistically crafted cyber-attacks. The proposed CAD detects a cyber-attack in the contaminated islanded data by computing a stochastic non-parametric correlation coefficient, i.e., Spearman's rank correlation in conjunction with a deterministic Cosine-Similarity measure.

The efficacy of the proposed method is rigorously tested and assessed under various circumstances on real-life Banshee industrial microgrid system modelled in the RTDS simulation environment, on the basis of the IEEE-1547, UL 1741 and IEC-62116 standards.

The rest of the chapter is organized as follows. The statistical analysis of each phase PCC voltage mean variation under various islanding and non-islanding events is presented in Section 6.2. The proposed statistical property-based passive islanding detection technique is explained in Section 6.3. Section 6.4 discusses the simulation results of the proposed

IDS. Next, Section 6.5 presents the proposed signal processing based cyber attack detection method. The attack detection results is then depicted in Section 6.6. Finally, Section 6.7 concludes the chapter by drawing the main findings of the overall framework.

6.2 Statistical Analysis of Various Islanding and Non-Islanding Scenarios

In order to statistically analyze the voltage signal received at the PCC under islanding or fault/switching scenarios, a test system is implemented in PSCAD/EMTDC as per the IEEE-1547 and UL 1741 standards listed in Table 6.1 which helps to devise a satisfactory relay logic in the forthcoming section. The system is shown in Fig. 6.1, where a single DG is connected to the grid at the PCC, along with a parallel connected RLC load. The constant current control mode of DG operation is used for the islanding studies, and unity power factor operation of the inverter is considered. The corresponding parameters details are also listed in Table 6.1. The sampling frequency of 7.68 kHz is used for 60Hz, which results in N=128 samples per cycle. Further, the actual voltage signal V^{PCC}, as measured from the PCC, is used to compute and analyze the absolute mean of the one cycle window data, as follows.

$$V^{PCC}_{mean} = \left| \frac{1}{N} \sum_{i=1}^{N} V^{PCC}(i) \right| \tag{6.1}$$

Table 6.1: Test System details as per IEEE 1547 and UL 1741 standards

S.No	Parameter	Details	S.No	Parameter	Details
System parameters			**Load parameters**		
1	Frequency	60 Hz	6	R (Ω)	2.304
2	Voltage (L-L)	480	7	L (H)	0.00611
3	DG power output	0.1 MW	8	C (F)	11.51.29
4	Input Voltage of DG	900 V	9	Quality Factor	1
5	Switching Frequency	8kHz			

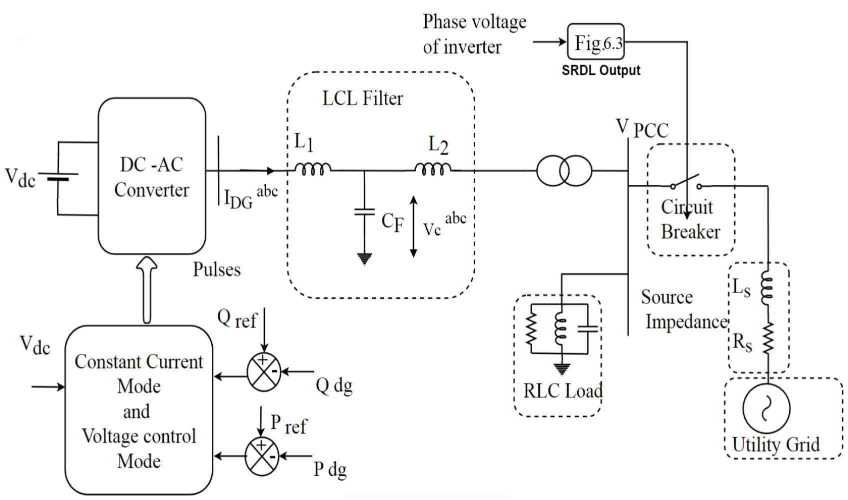


Figure 6.1: IEEE 1547 and UL 1741 Standard based Islanding test system

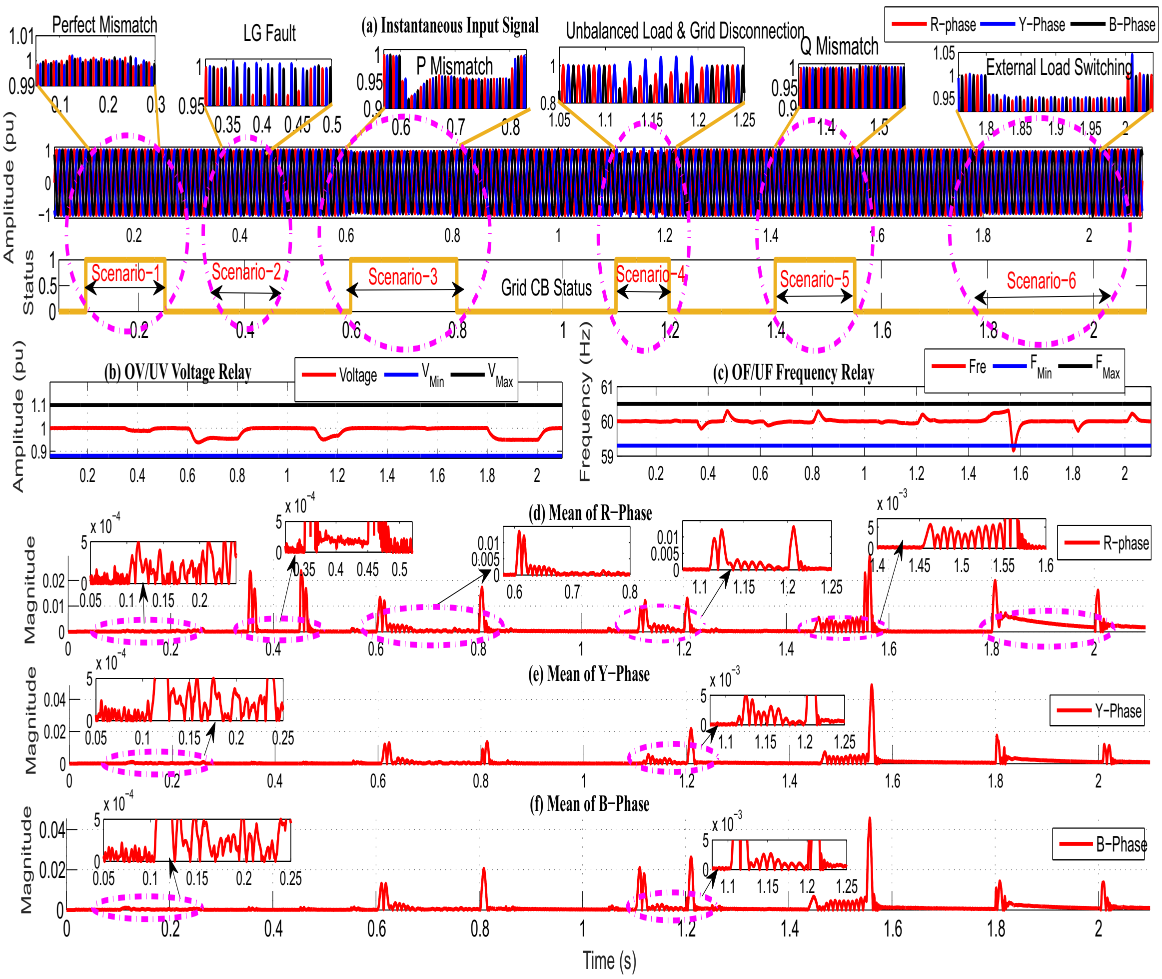


Figure 6.2: V_{mean}^{PCC} analysis of various Islanding and Non-Islanding scenarios

Figure 6.2(a) shows the instantaneous three phase voltage, below which is shown the grid Circuit Breaker (CB) status. Grid CB status = 0 indicates CB is CLOSED and grid CB status = 1 means CB is OPEN. The statistical analysis of the PCC voltage is done for six scenarios, as described below.

1. **Scenario-1:** In this scenario, an islanding event is simulated by opening the grid CB at 0.1s. It can be observed from Fig. 6.2(b-c) that the islanding condition does not affect the voltage and frequency due to perfect zero power mismatch of both active power and reactive power. Figure 6.2(d-f) shows the one cycle average as per Eq. (6.1) for each three phase signals respectively. From the zoomed portion of all the three phase one cycle averages of Eq. (6.1), it is observed that there is a significant change before and after the islanding scenario. A high frequency and low frequency nature can be observed before and after islanding, respectively in the calculated one cycle average of each phase.

2. **Scenario-2:** In this scenario, a non-islanding event is simulated with grid CB closed. A LG fault is applied at 0.35 s and cleared at 0.45 s in the R-phase with a fault resistance of 1Ω. Clearly, the voltage and frequency is significantly affected, as

shown in Fig. 6.2(b-c). One cycle V_{mean}^{PCC} for R-phase also reveals sudden changes in its value during the start and end of the fault. The other two healthy phases do not undergo any significant changes, as shown in Fig. 6.2(d-f). Likewise, under various fault conditions, calculated one cycle V_{mean}^{PCC} of each phase as per Eq. (6.1) has significant transient effect at the start and end of the fault.

3. **Scenario-3:** In this scenario, grid CB is opened at 0.55 s for the duration of 0.25 s to analyze the islanding scenario under P mismatch, as shown in Fig. 6.2(a). The voltage and frequency experience a significant effect under this condition, as can be observed from Fig. 6.2(b-c). One cycle V_{mean}^{PCC} for all phases is shown in Fig. 6.2(d-f) which also depict significant variation.

4. **Scenario-4:** In this scenario, the system is working under unbalanced conditions in all the three phases. Islanding scenario is then simulated by opening the grid CB at 1.1 s under perfect mismatch of P & Q, as shown in Fig. 6.2(a). Under this scenario, again the voltage and frequency experience significant change as shown in Fig. 6.2(b-c). One cycle V_{mean}^{PCC} for all phases is shown in Fig. 6.2(d-f). Due to the unbalanced operation, magnitude of one cycle average in each phase is different. Yet V_{mean}^{PCC} reveals the Islanding condition successfully.

5. **Scenario-5:** In this scenario, Q mismatch is simulated as an Islandng scenario by opening the grid CB at 1.4 s. Majorly, the voltage remains unaffected, and frequency has significant variations, as shown in Fig. 6.2(b-c). Figures 6.2(d-f) depict the significant changes in one cycle V_{mean}^{PCC} for all the three phases.

6. **Scenario-6:** An external load switching is simulated at 1.8 s as a Non-Islanding scenario, with grid CB as closed, as shown in Fig. 6.2(a). In this case, the external load is added at 1.8 s and removed at 2 s from the system. Interestingly, the one cycle V_{mean}^{PCC} of all phases, as shown in the Fig. 6.2(d-f), reveals a decaying DC effect.

The aforementioned simulated scenarios cover a wide variety of possible conditions in an MG system, and the signal behavior under these real-world scenarios has been extensively studied in order to produce a reliable passive islanding detection method.

6.3 Proposed Islanding Detection Method

It is expected that the low-voltage grid will have voltage and frequency relays that trip the generator in the event that at least one phase's voltage or frequency restrictions are exceeded. If the frequency and voltage are both within the allowed ranges, then a Non-Detection Zone (NDZ) is present. As was discovered in Section 6.2, the voltage mean V_{mean}^{PCC} shows a pattern that is constant across all three phase voltages under the various scenarios evaluated. Therefore, the instantaneous voltage of any one phase at the output of the DG can be utilized to detect NDZ using one-cycle information. This is accomplished by computing the V_{mean}^{PCC} according to Eq. (6.1) utilizing a 128-point buffer.

It is noteworthy that the statistical analysis carried out in Section 6.2 under different conditions demonstrates that, under Islanding conditions, a significant shift in V^{PCC}_{mean} for one cycle can be observed. However, depending solely on the voltage magnitude could lead to false alarms when it comes to Non-Islanding events like as failures, external load switching, etc. The proposed IDS technique thus uses the new idea of detecting the presence of decaying DC in the signal before making a judgement on the unintentional Islanding in addition to utilizing V^{PCC}_{mean} information. To this end, Fig. 6.3 depicts the overall framework of the Proposed Islanding Detection Method (PIDM). As illustrated in this figure, the PIDM is comprised of three stages, viz., 1) Mean-based coarse Islanding Detection (MID), 2) Decaying DC Detector (DDCD, and 3) finally, the Statistical Relay Digital Logic (SRDL). The description of each stage is discussed briefly in the following subsections.

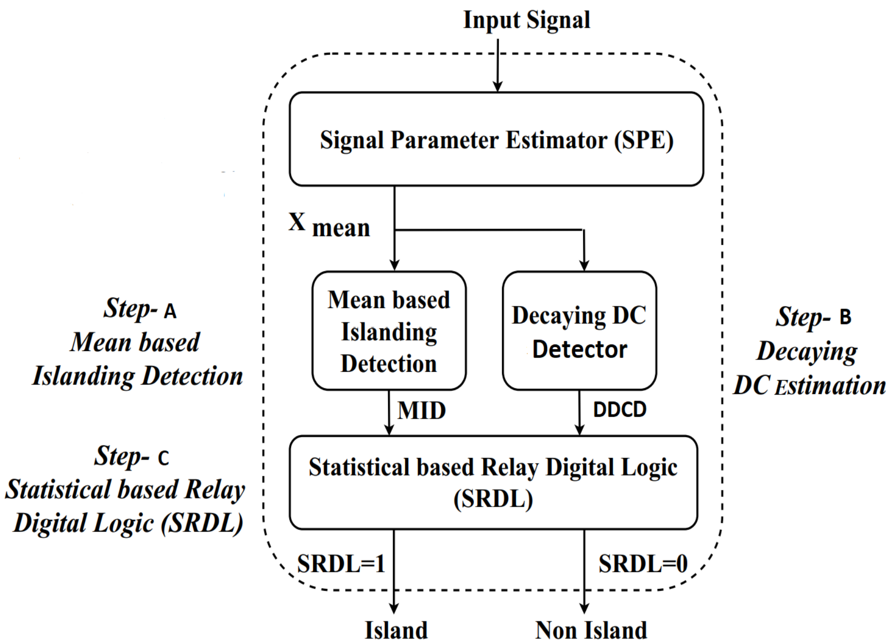


Figure 6.3: Block diagram of the proposed data driven Islanding detection scheme

6.3.1 Mean based Coarse Islanding Detection (MID)

A coarse estimation of of islanding detection is produced by this block. At this stage, islanding is just suspected and is finally confirmed in the last stage of the proposed SRDL. The input for this stage is V^{PCC}_{mean}. The following are the procedural steps involved in the calculation of MID.

Step-1: Create a 32-bit buffer of V^{PCC}_{mean} values, i.e., $\mathbf{V^{PCC}_{mean}} = V^{PCC}_{mean}[1:32]$, and find the minimum, V^{PCC}_{min}, in the buffer. The V^{PCC}_{min} is significantly influenced by the presence of fixed DC offset and lower level of decaying DC quantity which arises during faults, load switching, and harmonics.

Step-2: Via moving window concept, maintain a 32-bit buffer of $\mathbf{V^{PCC}_{min}}$ values, termed as V^{PCC}_{rem}, i.e., $\mathbf{V^{PCC}_{rem}} = V^{PCC}_{min}[1:32]$.

Step-3: Finally, the entropy in V_{rem}^{PCC} is calculated. The entropy value so obtained is termed as Mean based Islanding Detector. Inside the observation window of 32-bit buffer, the value of MID remains fairly constant (less than $1e^{-4}$ or zero) under non-islanding scenarios. On the other hand, an appreciable disorder or uncertainty is shown by MID (in the range of 0.1-4) during possible/suspected islanding events.

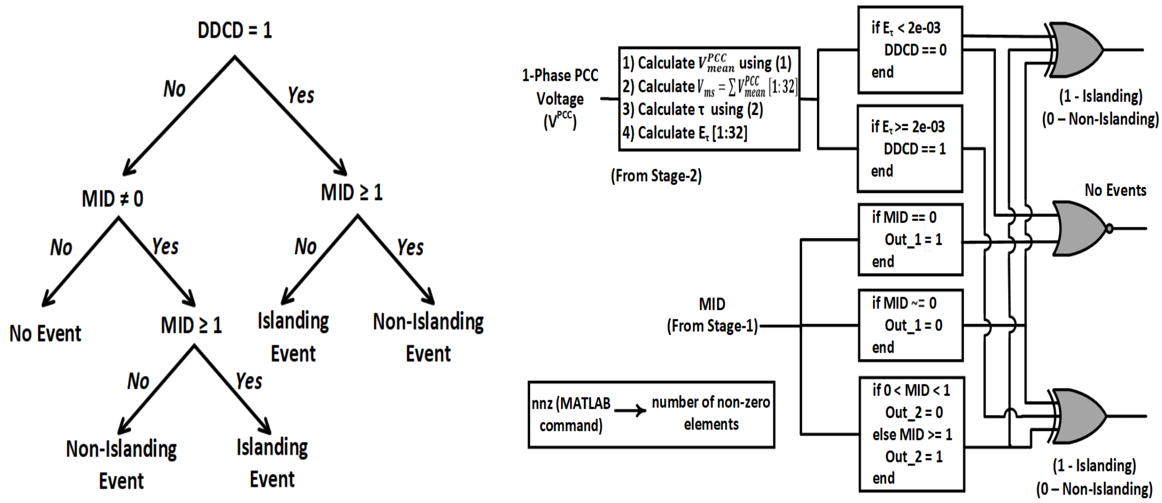
6.3.2 Decaying DC Detector (DDCD)

The DC decay time is not used as a direct indication of an islanding situation but as an indicator of transient behaviour within the MG. A MG's electrical topology and operating circumstances alter during an islanding event, resulting in transient phenomena including voltage and frequency aberrations. Under such circumstances, the DC offset present in the signal will decay over time due to the oscillating components present in the signal. The main objective of this stage is, therefore, to assess the presence of decaying DC component quantity, which arises due to faults and load switching events.

Estimating DC decaying time: The V_{mean}^{PCC} information is summed up using a 32-bit buffer data as $V_{ms} = \sum V_{mean}^{PCC}[1:32]$. The rate at which the magnitude of V_{ms} is decaying is calculated as :

$$\tau(i) = \text{abs}(-dt/(4\log(V_{ms}(i) - V_{ms}(i-1)))) \tag{6.2}$$

where, dt is the sampling time. τ, thus, represents the amount of time that the decaying DC quantity (if present) will take to decay to approximately 36.8% of its original amount. Finally, the entropy in $\tau[1:32]$, i.e., (E_τ) is computed in order to see the randomness in the computed decay time and for deciding if the decaying DC component is present or not in the signal. For the discretization purpose, the entropy of DC decaying quantity (E_τ) present in the signal is compared with a threshold of 2×10^{-3}. In case, the entropy is found to be more than the set threshold, DDCD will be set to 1 or 0 otherwise.



(a) Statistical based Relay Logic Tree (b) Statistical based Relay Digital Logic (SRDL) Diagram

Figure 6.4: Proposed Îslanding Detection Method (PIDM) logic

6.3.3 Statistical based Relay Digital Logic (SRDL)

As seen in Fig. 6.4(a), the output of the MID is utilized in tandem with the output of the DDCD in this last step to create the breaker logic for the operation. The binary output variable SRDL is set to 1 to indicate islanding and 0 for a non-islanding situation based on the following rules:

1. **Rule-1:** As mentioned in section II.B, if the studied signal is heavily influenced by the presence of decaying DC, then DDCD will be set 1 as a suspected abnormal events. In order to ensure that this arises due to event of a non-islanding scenario, non-zero MID value is then compared with unity. If the MID value is equal to or greater than 1, it is classified as a non-islanding event; otherwise, it is declared as an islanding event.

2. **Rule-2:** If the signal does not have the signature of decaying DC of significant tolerance, the DDCD will not be triggered and thus it should be set to state 0. According to Rule-2, in this case, an event is considered an islanding scenario only when $MID \geq 1$. This is due to the fact that ideally DDCD should not respond and always be obscured during most of the islanding scenarios. Thus, in such case, unlike the rule-1, now SRDL logic will be set to 1 indicating an islanding scenario when MID$\geq$1. On contrary, if the MID value falls within between open and closed interval range such as MID $\in (1 \times 10^{-4}, 1]$, the event is confirmed as low impacted Non-islanding event, and SRDL is set to 0 in that case.

3. **Rule-3:** This is the generalized case where these two main decision making components i.e DDCD and MID both are found to be zero. To avoid any misjudgement, MID values less than 1×10^{-4} are treated as 0 in this context, signifying that neither islanding nor non-islanding events have occurred, indicating a normal system state.

Based on the above rules, a Statistical Digital Relay Logic (SRDL) is designed to indicate the presence of islanding and non-islanding scenarios as illustrated in Fig. 6.4(b). At last, in order to avoid nuisance tripping of DGs due to some stringent scenarios and to take confirm decision over an event, the toggling outputs of the logic gates are constantly monitored for a cycle using sliding window concept. If the presence of number of ones i.e., non-zeros (NZ) is more than zeros (Z) for that period of time, then the final SRDL decision confirms the event as an islanding scenario otherwise it indicates ia as a non-islanding scenario.

6.4 RTDS Simulation Results

Test System: The efficacy of the PIDM is carried out via real time digital simulation on Banshee power system [222]. Banshee MG system is a real-life small scale industrial facility that receives power from three utility radial feeders, as shown in Fig. 6.5. This three radial feeders provide three zones with different levels of connectivity via normally open switches. Each area's mainstream feeder is connected to the utility grid via a different PCC circuit breaker. The distribution voltage level of this MG has a system voltage of 13.8 kV. Eighteen aggregated dynamic loads with a power factor of 0.9 lag are supplied by those feeders. Additionally, there are two 200 horsepower induction motors that have compressor loads. There is a 4 MVA diesel generator in area-1 and 3.5 MVA natural gas fired combined heat and power plant operating at 13.8 KV in area-3 and they usually have a controller operating on 4% linear voltage and frequency (V/F) droop. Area-2 formed by feeder 2 contains a 2.5 MVA battery energy storage system and 5 MVA photovoltaic array designed via average modelling with time-varying irradiance profile and temperature. Additionally, 3 more grid following PV units, i.e., DG1, DG2 and DG3, of equal 1.25 MW rating, are located at Bus-107, Bus-305 and Bus-209, respectively. More details about the Banshee MG system layout, configuration and various other source, line and load component details can be found out in Appendix A.2.

Test Scenario: Following test cases have been considered for establishing the efficacy of the PIDM.

- *Zero Power Mismatch:* Representing no or very small net imbalance between active and reactive power. A passive islanding detection scheme finds it difficult to discriminate between grid-connected and islanded scenarios.

- *Load and Capacitor Switching*: Sudden connection/disconnection of load of different power factor and switching of capacitor introduces transient in the systems which poses challenges to islanding detection methods (IDMs) to accurately discriminate them as non-islanding events.

- *Fault condition with different resistance*: Single-phase faults are very common in power systems, and their severity can vary based on fault resistance. To assess the IDM method's sensitivity to fault characteristics and its capacity to distinguish between faults and islanding occurrences, PIDM was tested under various fault resistance situations.

- *Loss of parallel feeder (LOPF):* Loss of Parallel feeder is a spurious non-islanding event that usually leads to a false triggering of islanding detection. It is important to consider such events because such events affect inverter frequency, and may lead to a cascaded failures in the system.

Apart from the above mentioned test case, the proposed method is rigorously verified on other events also such as linear and non-linear load switching, capacitor switching, induction motor starting and tripping of other DG than targeted DG as shown in Table 6.2.

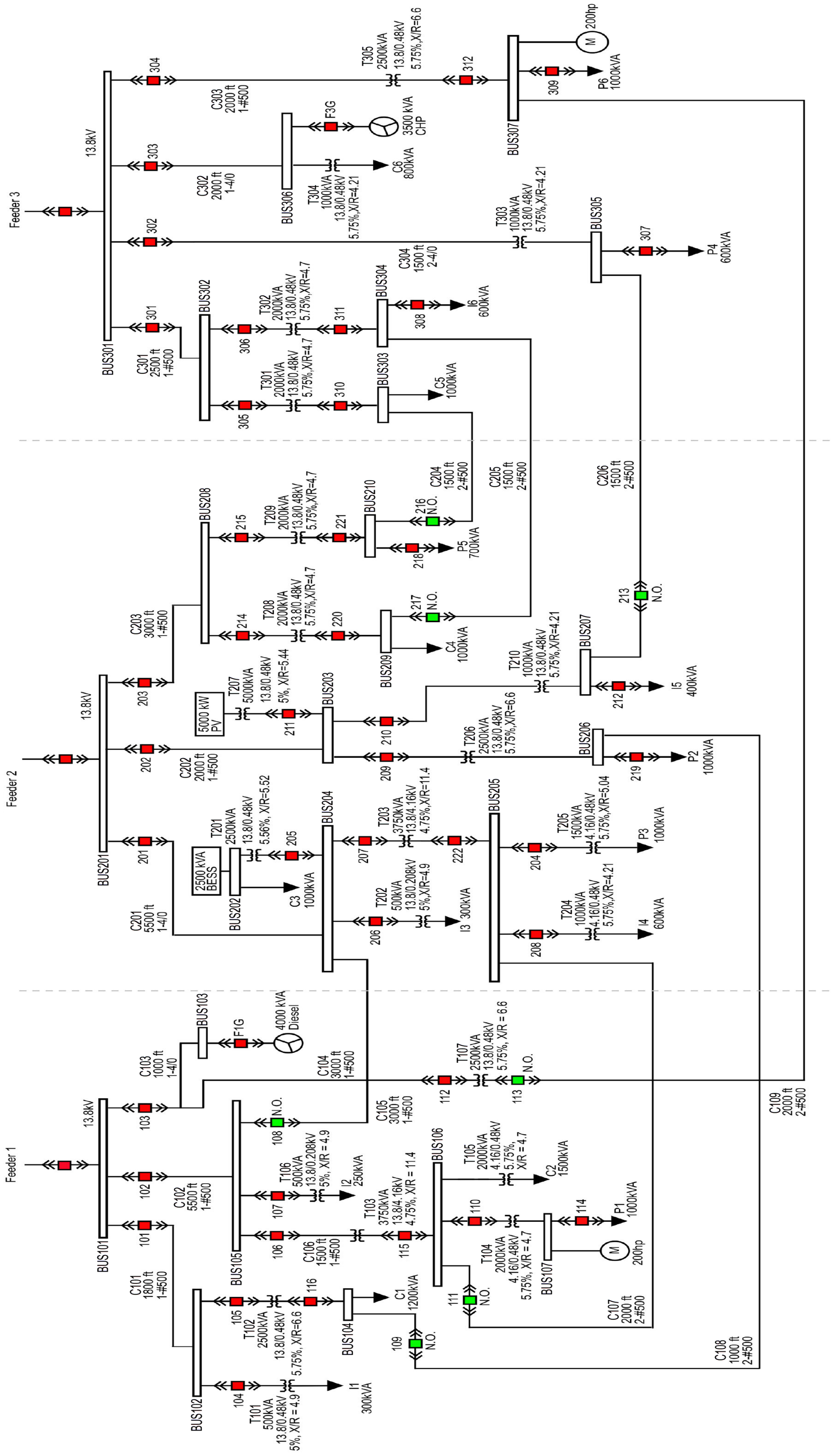


Figure 6.5: One-line diagram of Banshee Industrial Microgrid model

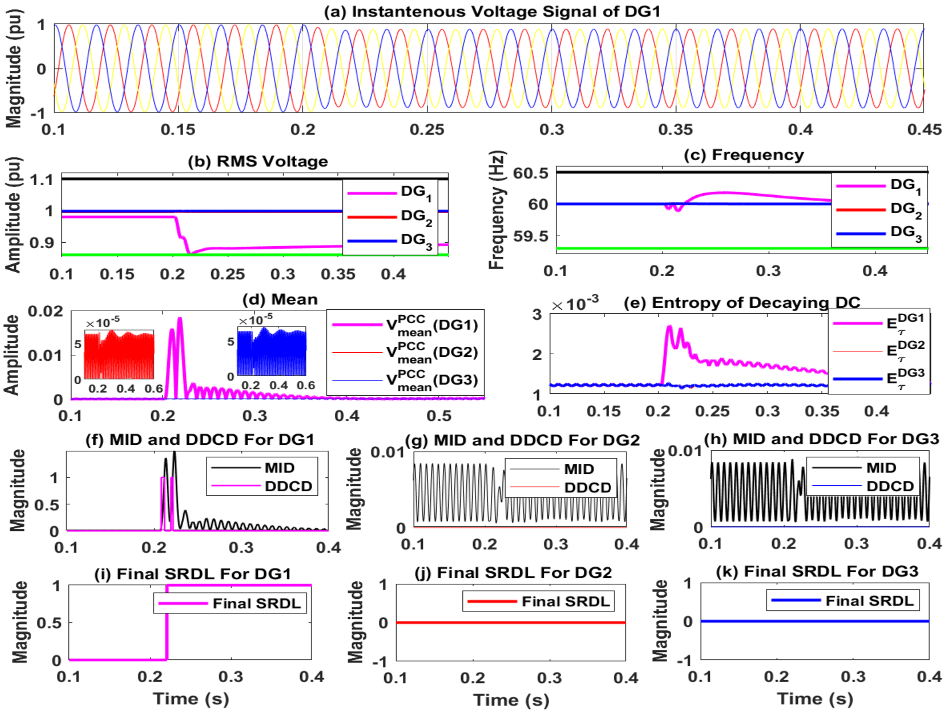
Table 6.2: Simulated Islanding and non-islanding scenarios.

Islanding Events	
Type of Islanding	**Details**
Active Power Mismatch	-50% to +50%
Recative Power Mismatch	-4% to +4%
Non-Islanding Events	
Type of Non-islanding	**Details**
Three Phase to Ground Faults	0.01Ω to 10Ω
Double Phase to Ground Faults	0.01Ω to 10Ω
Single Phase to Ground faults	0.01Ω to 10Ω
Capacitor Switching	500 kVA
Load Switching	125 kVA, 0.8 pf lag
Non-linear Load Switching	100 kW
Induction Motor Load Switching	200 HP, 0.48 kV
Tripping Other DGs Except Targeted DG	Trip a 5 MW DG
Loss of Parallel Feeder	Trip Feeder-2

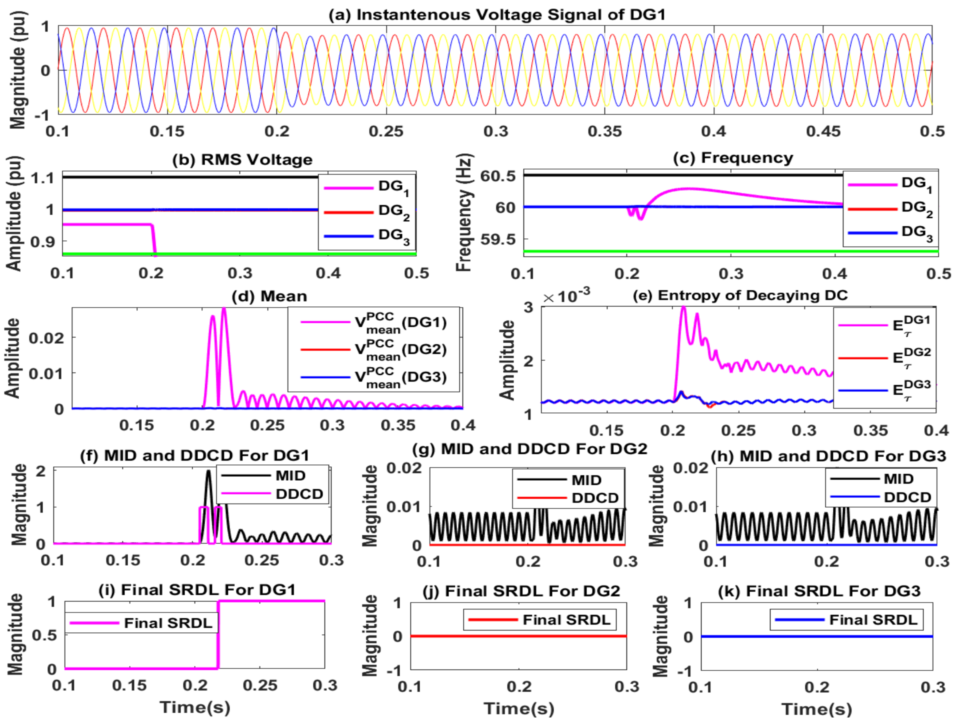
6.4.1 Validation of Proposed Method on Banshee Industrial Microgrid with High Penetration of Renewables

6.4.1.1 Islanding Scenarios

In order to analyse the performance of the proposed method, different cases of islanding situation based on active and reactive power imbalances are simulated. Figure 6.6(1) shows the performance result of the proposed method when the load active power and reactive power was adjusted to 105% and 100% **(Case 1)** of generated power, respectively. Figure 6.6(1)(a) shows the instantaneous voltage waveform collected from DG1 terminal and Fig. 6.6(1)(b-c) shows the RMS voltage and frequency respectively for all the three DGs. It can be perceived from the plots that DG1 is islanded, therefore there is some significant changes can be seen in PCC voltage mean and its entropy of decaying DC as shown in Fig. 6.6(1)(d-e). It is observed from the Fig. 6.6(1)(f) that at the point of time when entropy of decaying DC just crosses the threshold 2×10^{-3}, the MID value is lying below 1 which triggers the SRDL logic as a suspected event. Thereafter, with reducing the decaying DC, the MID value keeps on increasing for sometime within a 1-cycle window. Now as these two events hold on for sufficient amount of time in a 1-cycle waiting period, the number of ones count is found to be more than the zeros and hence the Final SRDL of DG1=1 is flagged out stating it as an Islanding event as shown in Fig. 6.6(1)(f). But, the same kind of pattern has not been captured by the SRDLs of non-targeted DGs i.e DG2 and DG3, thus it does not raise any flag which can be easily visualized from Fig. 6.6(1)(g-h) and (j-k). Likewise, **Case 2** depicts another UL 1741 standard islanding case studies where the load active and reactive are set to 125% and 100% respectively. The results of Fig. 6.6(2) also demonstrate the efficacy of the PIDM in accurately detecting the condition.



(1) Case-1 - 105% P and 100% Q



(2) Case-2 - 125% P and 100% Q

Figure 6.6: Power mismatches as Islanding scenario

6.4.1.2 Non-islanding Scenarios

Case 3 - LLG Fault with 2Ω resistance:- Figure 6.7 shows a non-islanding scenario where a LLG fault **(Case 3)** of 2Ω fault resistance took place at 0.2s and persist for 0.3s as seen from the mean information of Fig. 6.7(d). From Fig. 6.7(e-h), it is observed that the entropy of decaying DC for all the three DGs are lying below threshold and the non-zero MID values are also below 1. Therefore, the proposed method gives the final decision of this event as a non-islanding event for all the DGs based on the designed SRDL tree as shown in Fig. 6.2. The final SRDL waveforms for all three DGs are shown in Fig.6.7(i-k). The other variety of faults studies such as LG, LLLG with varying low to high fault resistance on DG sides are also tested by the proposed method and all the events are found out to be non-islanding scenarios successfully.

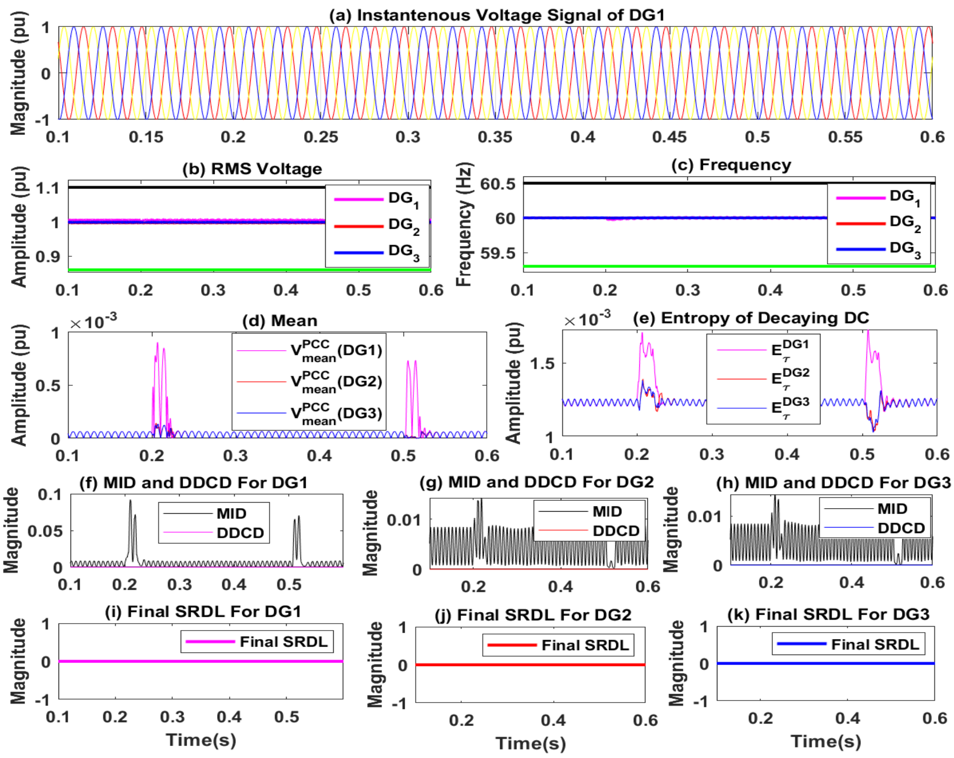


Figure 6.7: Case 3 - LLG Fault with 2Ω resistance

Case 4 - 500 kVA capacitor bank switching:- Figure 6.8 shows another event of non-islanding test case where a 500kVA capacitor bank **(Case 4)** is switched on at around 0.2s at Bus 107. Due to such insertion, there are slight changes in the RMS voltages as shown in Fig. 6.8(b) and consecutively mean value of PCC voltages face some small variation. Overall, for this event MID and entropy of decaying DC both lies below the limit for all of the DGS as shown in Fig. 6.8(d-h) and thus SRDL logics treated this event as non-islanding as shown Fig. 6.8(i-k).

Case 5 - 200HP induction motor switching:- The reliability of the proposed method is now demonstrated by switching a 200HP, 0.075 MWs/MVA inertia, induction motor

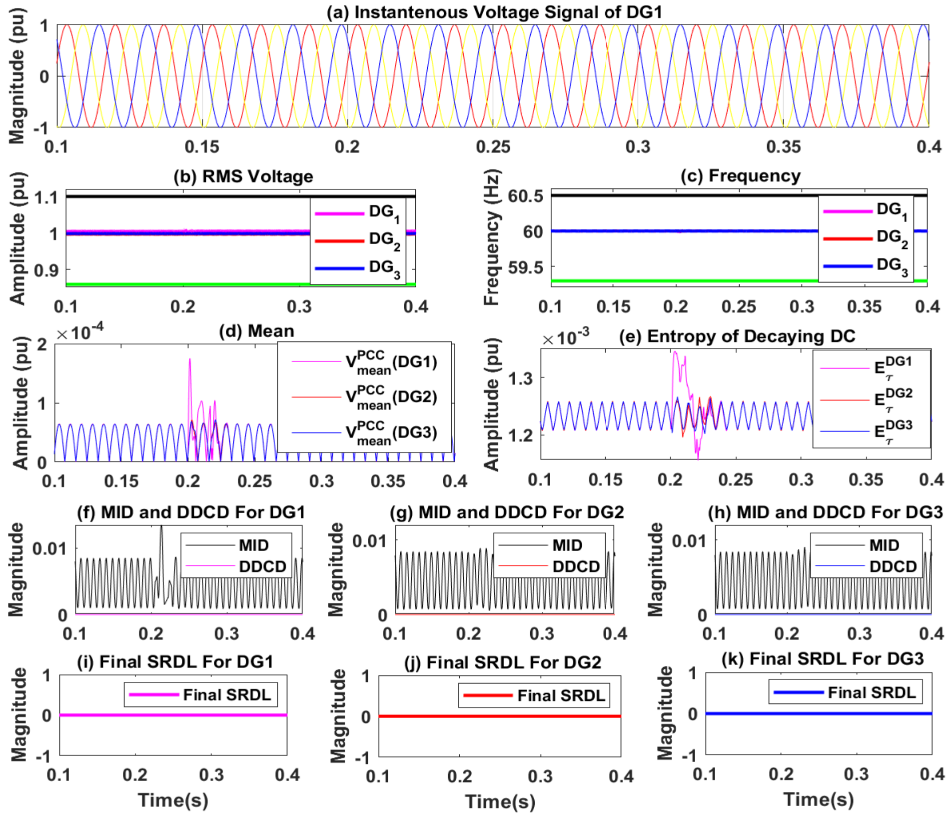


Figure 6.8: Case 4 - 500 kVA Capacitor bank switching

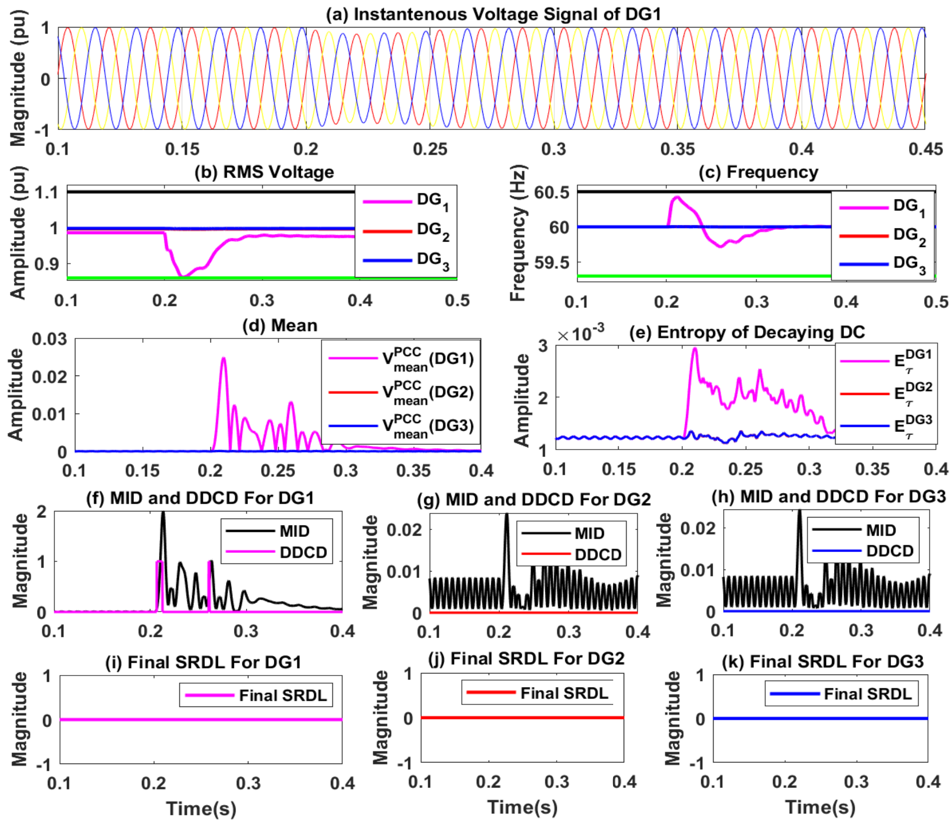


Figure 6.9: Case 5 - 200 HP Induction motor switching

(IM) load **(Case 5)** at the same bus location 107 operating at 0.48kV. The motor load is switched on at about 0.2s as shown by the RMS voltage and frequency waveforms in Fig. 6.9(b) and (c) respectively. During the start-up time, IM requires a large amount of reactive power for a period of time and this, in turn, causes some changes in voltage and current profile. These abrupt changes sometimes mimic the exact islanding signature and thus can cause most of the islanding detection methods to fail. But the proposed method deals this situation well. It can be seen in Fig 6.9(d)-(e) that during the switching action, the variation in mean and entropy (E_τ) takes place. It can be seen from Fig. 6.9(e)-(f) that while the entropy of decaying DC is crossing the thresholds at that instant MID value is not more than one resulting SRDL to be 1, which is an indication of false islanding scenario. Thereafter, MID value keeps increasing more than one for sufficient amount of time while the entropy of decaying DC decreasing and finally stayed below its threshold limit resulting SRDL to be 0 which indicates a non-islanding event. As the proposed Final SRDL logic take its final decision based on the waiting period of 1-cycle of the intermittent nature of logic gates output, thus it is found out at the end that the number of zeros are more than the ones. Therefore, based on majority voting in favour of non-islanding, the Final SRDL logic is set to 0 which confirms this event as non-islanding. The above description can also be easily discernible from the Fig. 6.9(f-h). Thus, it can be verified that the proposed method detects this non-islanding event even if the load is large induction motor type.

Case 6 - 100kW non-linear load switching:- The response of the proposed method was also assessed under the influence of nonlinear loads. In this investigation, a 100kW three-phase diode rectifier **(Case 6)** with a resistive load was introduced as a nonlinear load. The connection of this load at bus-107 was simulated at approximately 0.2 seconds. The observations from Fig. 6.10(f-h) reveal that the switching of the nonlinear load does not significantly impact either the MID or DDCD criteria for any DGs. Consequently, the SRDL component of the proposed method correctly classifies this event as non-islanding.

Case 7 - Tripping of other DG except of targeted DG:- Another non-islanding scenario is also simulated at the feeder-2 i.e., area-2 of the Banshee microgrid model as shown in Fig. 6.11, where DG2 is kept as the targeted DG and the existing large 5 MW Banshee PV DG are now suddenly tripped at 0.2s **(Case 7)**. It is clear from the Fig. 20 due to tripping of non-targeted DG causes some level of oscillation in the MID and entropy of decaying DC index but they are still lies within the threshold. Thus, it can be stated that tripping of other DG except the targeted DGs does not have any negative influence on the Final decision over SRDL of the targeted DGs of the proposed method.

Case 8 - Loss of parallel Feeder:- In this case, loss of a parallel feeder that usually results in nuisance tripping of relays caused by the misinterpretation of a non-island scenario as an island scenario is simulated. The efficacy of the proposed method is tested for a sudden disconnection of feeder 2 **(Case 8)** from the utility grid and the performance

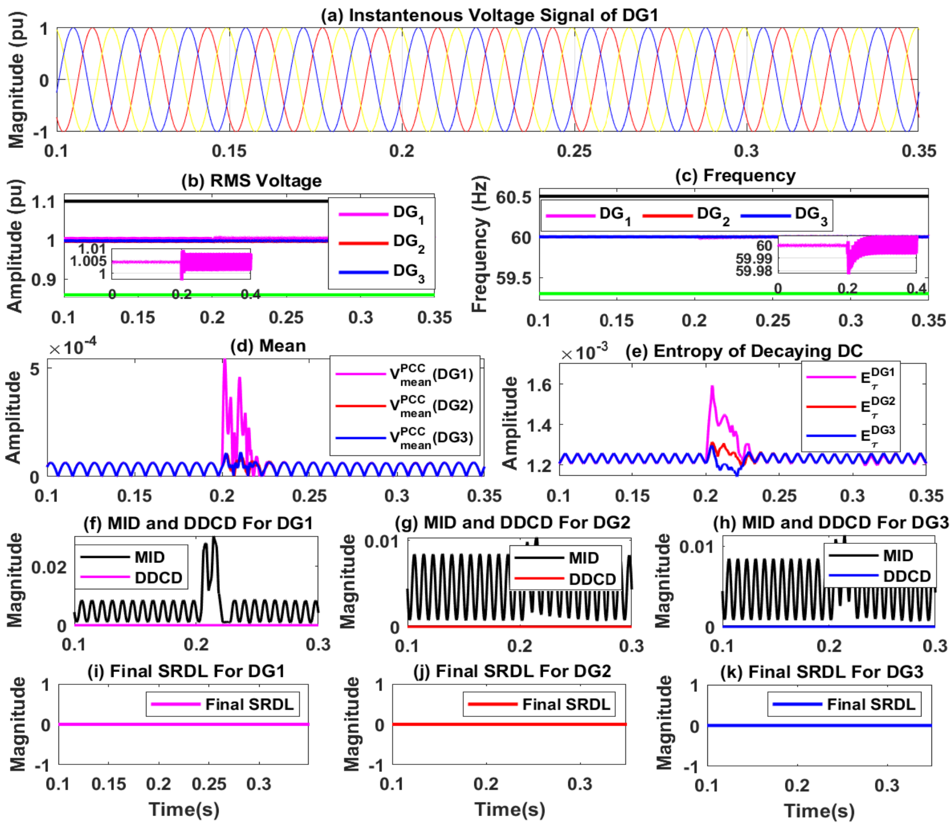


Figure 6.10: Case 6 - 100 kW Non-linear load switching

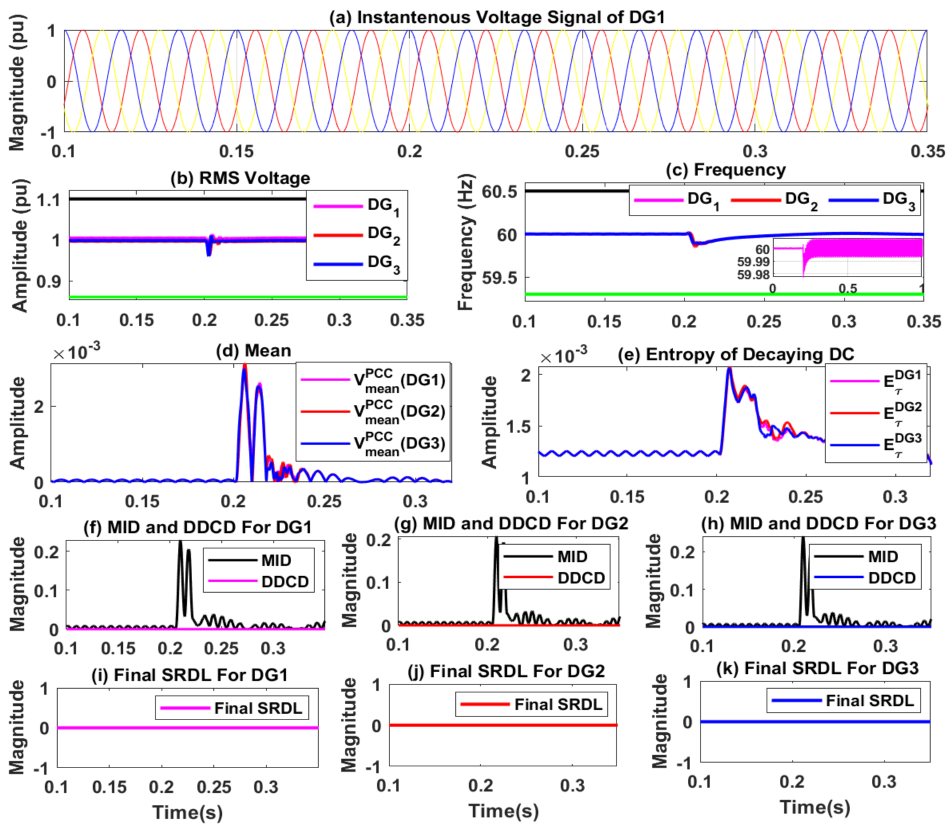


Figure 6.11: Case 7 - Tripping of other DG except of targeted DG

of the three DGs are then monitored as shown in Fig. 6.12. As DG3 is located at Bus 209 which falls under the area of feeder-2, this situation is an islanding from DG3's perspective but from the viewpoints of the other two DGs located at Bus 107 and Bus 305 it is a non-islanding situation. Furthermore, as a sizable 1.25 MW DG3 and a massive 5 MW PV array simultaneously gets eliminated from the utility grid, a significant imbalance in the local PCC voltage of the DG3 terminal occurs, as illustrated in Fig. 6.12(a). Since the disconnection of feeder 2 forms a huge island and all the DGs in that island are designed to operate in grid following mode, they have lost their voltage and frequency references which results in fluctuations of the DG3 rms voltage and frequency beyond UV/OV and UF/OF limits as shown in Fig. 6.12(b-c). But this situation does not create any negative impact for DG1's and DG2's performance as they remain connected to the upstream grid. Therefore, unlike DG3, MID and DDCD of DG1 and DG2 are lying within their safe limits and from their viewpoints, this event is successfully treated as non-islanding which can be seen in Fig. 6.12(f-g) and Fig. 6.12(i-j), respectively. Also, to be noted that the huge random oscillations in the MID and DDCD of DG3 triggers the SRDL logic intermittently to consider it as a suspected event. Moreover, this huge amount of power loss because of feeder disconnection leads to severe distortion in DG3's rms voltage and frequency beyond limits, resulted the MID and DDCD to cross their individual respective threshold at the same time for certain duration. As a consequence, the actual signature or confirmation details of islanding with respected to DG3 can't be satisfying periodically on time and thus it introduce a sufficient delay in responding DG3's SRDL. As the condition of getting number of non-zeros values more than zeros in a one cycle window is being achieve late, an inadvertent delay of 100 ms is observed in Fig. 6.12(k).

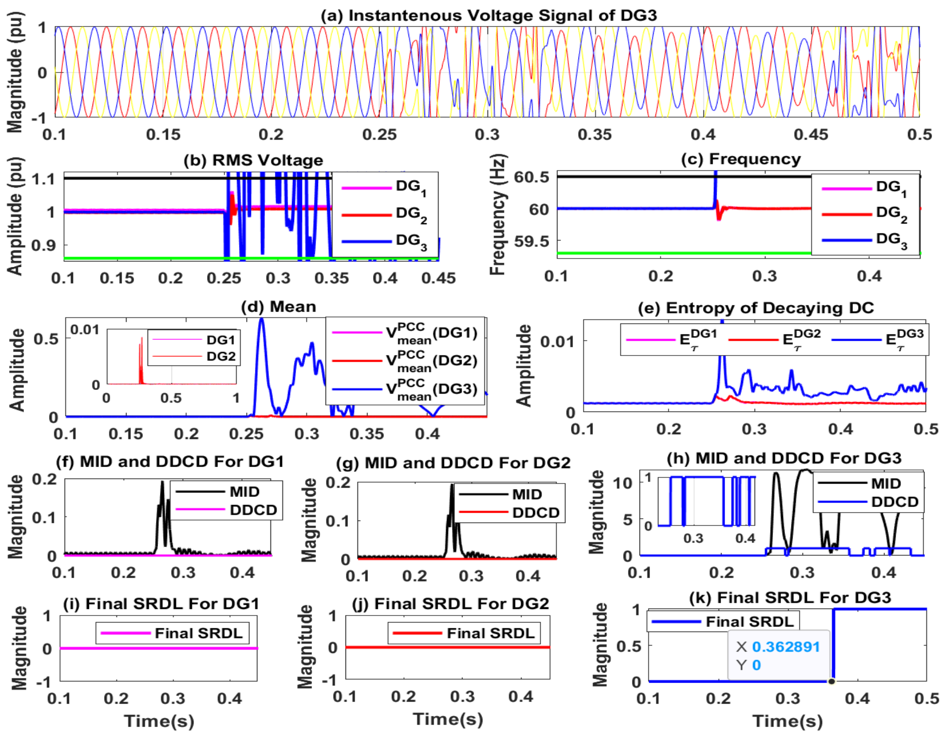


Figure 6.12: Case 8 - Loss of parallel feeder

6.4.2 Comparative Assessment with ROCOV

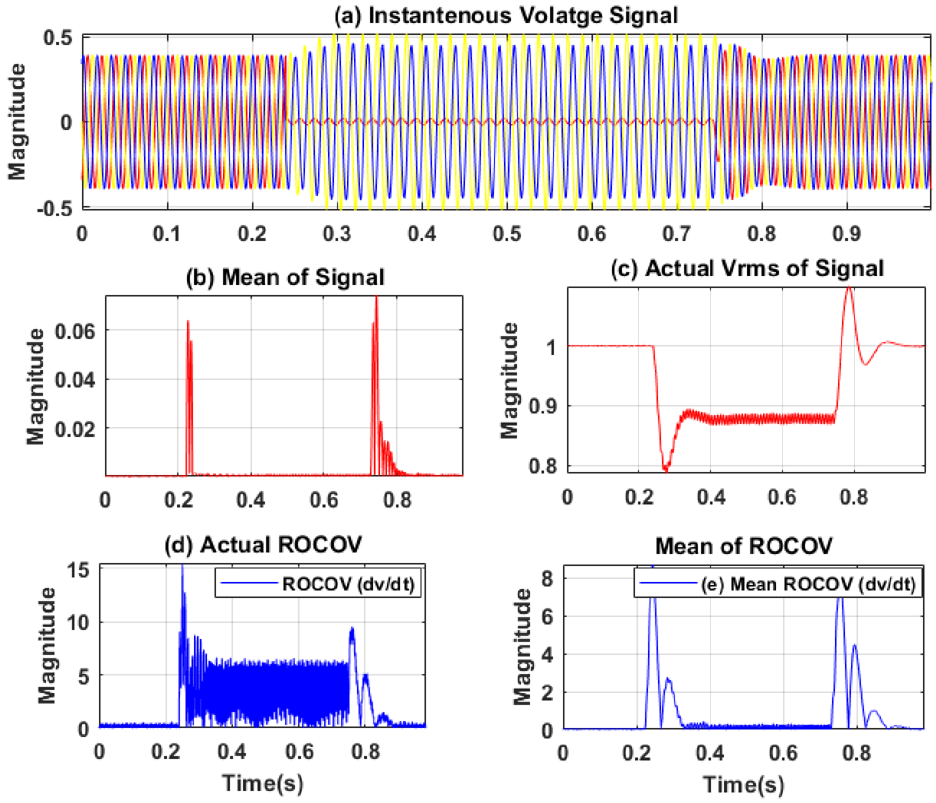
At the outset, ROCOV [223] methods looks quite similar to the PIDM. Nevertheless, when there is an imbalance in reactive power, the ROCOV faces significant challenges in effectively distinguishing islanding conditions. This is primarily because ROCOV relies on detecting variations in reactive power only as an indicator of islanding. For instance, unplugging a low power factor load from the network under normal circumstances can result in a large reactive power imbalance and voltage oscillations. Likewise, during switching of capacitors, motors or due to low resistance fault high inrush current will flow, which causes the voltage of PCC to decrease. As a result, the ROCOV value may surpass the predetermined threshold in this situation and malfunction. On the other hand, in instances of islanding when the reactive power imbalance is subject to minimal fluctuations, the ROCOV value might not rise over the predetermined threshold, thereby failing to identify the islanding condition. The above two scenarios are explained below through simulation in terms of the ROCOV and PIDM, where the mean of the 1-cycle single phase PCC voltage information are exploited.

6.4.2.1 Case 1

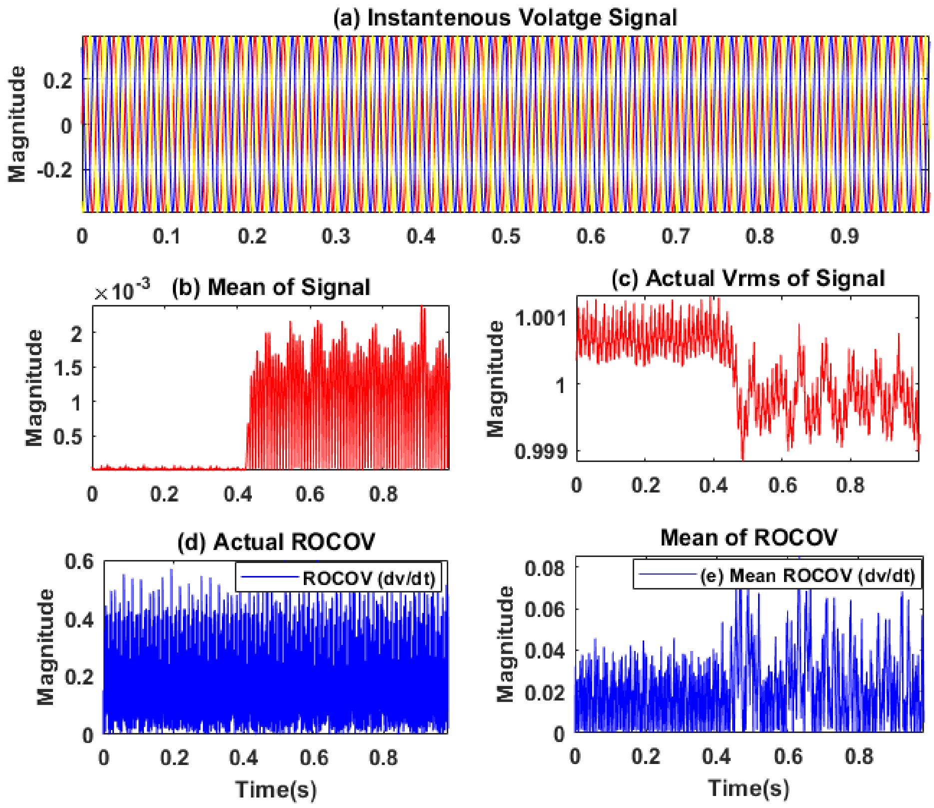
Figure 6.13(1) shows the Non-islanding case study where a temporary low resistance fault (1-phase to ground) event of 0.1Ω fault resistance takes place at 0.23s for 0.5s duration. The mean of the signal shown in Fig. 6.13(1)(b) clearly distinguishing this event and exhibits its low amplitude and low frequency components. Thus, it does not mal-operate by treating this event as islanding and do not trip the Over-frequency/Under-frequency (OF/UF) and Over-voltage/Under-voltage (OF/UF) relays. As the PIDM exploits the mean features of PCC voltage thus, it behaves in a similar manner, as shown in Fig. 6.13(1)(b). On the contrary, the ROCOV and mean of ROCOV shown in Fig. 6.13(1)(d) and (e) depicts significant oscillations with very heavy high frequency content along with high magnitude sharp spikes that creates trouble via nuisance tripping of relays by misinterpreting this event as an islanding event.

6.4.2.2 Case 2

Figure 6.13(2) shows the case study where an Islanding event is simulated with 0% active and 2% reactive power imbalance in the PCC voltage. As one can perceive from Fig. 6.13(2)(b), methods that based on only mean based features exploitation (like PIDM), there is a clear distinction in the high and low frequency content before and after Islanding scenario. But the frequency content is not distinctly visible whenever rms voltage information is used as shown in Fig. 6.13(2)(c). Moreover, with the implementation of ROCOV, the difficulty of differentiating Islanding and Non-Islanding situation under reactive power imbalance case is becoming more cumbersome as illustrated in Fig. 6.13(2)(d). Thus, in that respect the PIDM and ROCOV methods are not found out to be similar in performance.



(1) Performance Comparison Between ROCOV and PIDM During Non-Islanding Event



(2) Performance Comparison Between ROCOV and PIDM During Islanding Event

Figure 6.13: Comparative assessment between ROCOV and PIDM

Clearly, simplicity is an advantage of the ROCOV method, but there are situations in which a little complex approach can offer a superior performance. To this end, the PIDM balances this complexity with the aim of achieving better results in terms of detection accuracy and reliability. This is achieved by integration of the additional feature i.e., DDCD with MID, which, in turn, improves the performance in terms of accuracy and reliability as compared to traditional methods like ROCOV.

6.4.3 NDZ Analysis

Numerous islanding situations with various active and reactive power unbalance levels were simulated in order to determine the NDZ of the proposed method. Figure 6.14 provides a comparative illustration of the NDZ results between the proposed method and three other existing techniques [162, 224, 225]. The selection of these techniques was based on their relevant features.

Clearly, the proposed method exhibits a notably smaller NDZ compared in terms of active power imbalances and larger NDZ for reactive power imbalance to the other techniques as shown in Fig. 6.14. This reduction in the NDZ is attributed to its decreased sensitivity to non-islanding events. Furthermore, the proposed method is characterized by its simplicity, rapid response time, high reliability, and low computational complexity.

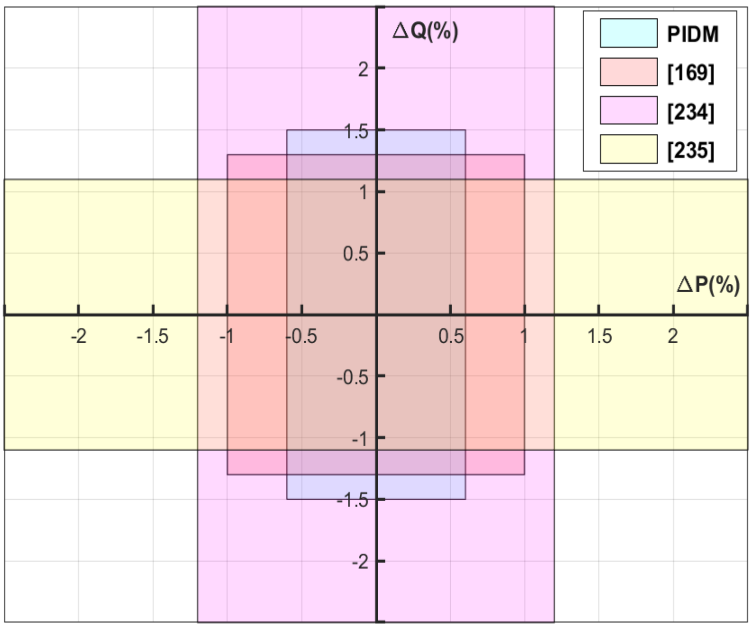


Figure 6.14: Non-Detection Zone of the proposed method

6.5 Development of Cyber Attack Immune Secured Islanding Detection Framework

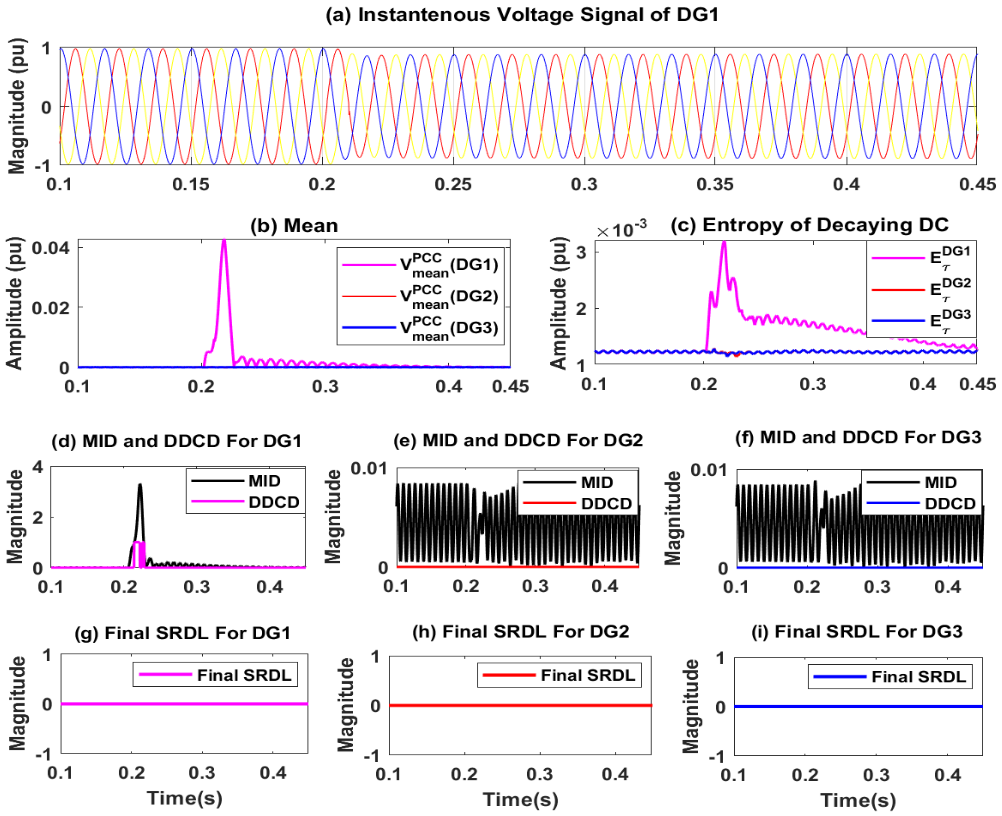
In the proposed islanding detection framework based on the SRDL output as mentioned above, the Distribution Management Operator (DMO) relies on the precise signals to discern between islanding and non-islanding events. However, the inherent vulnerability of SRDL to cyber attacks may pose a significant challenge. Attackers with prior knowledge of SRDL's functionality can take advantage of this vulnerability in a number of ways. They may manipulate the input voltage signals to mask the genuine islanding events, subtly altering the parameters to evade detection by SRDL. Alternatively, attackers could introduce spurious signals that imitate the features of real islanding instances in order to create fraudulent or fake islanding events. Such tactics can deceive the DMO, leading to erroneous decisions and potentially catastrophic consequences for grid operations. This necessitates to include a cyber attack detection module in conjunction with the SRDL based proposed IDM to strengthen its resilience against adversarial interference. As, it would be shown in the upcoming subsection of this chapter, that through the integration of a statistically crafted proposed cyber attack detector module it is possible to successfully detect and counteract the hostile attempts of the cyber attackers to disrupt the islanding detection process. This proactive strategy not only safeguards against potential cyber threats but also enhances overall cyber situational awareness, ensuring the integrity and reliability of islanding detection in the face of evolving security threats.

6.5.1 Vulnerabilities of the Proposed SRDL's Output based Islanding Detection

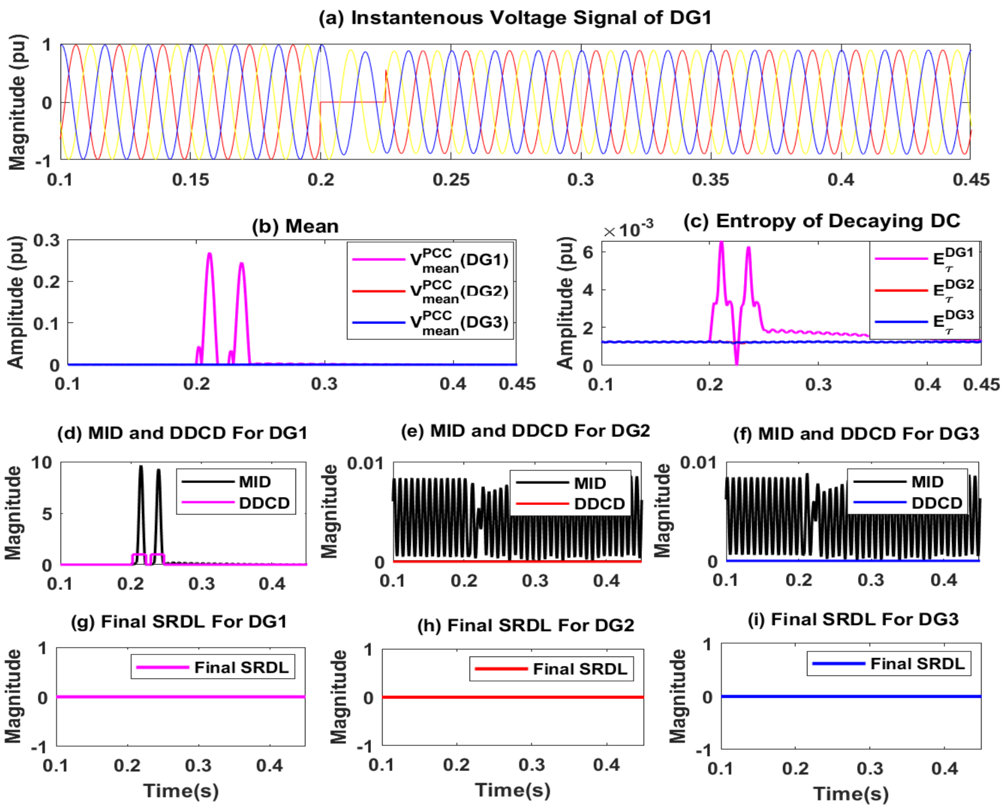
The vulnerabilities of exploiting the SRDL can result in either camouflaging the actual occurrence of an islanding event or imitating a non-islanding event occurrence with a fake islanding event as discussed below.

6.5.1.1 Masking of a Genuine Islanding Event

In this first case study, the performance of SRDL's output based proposed IDM is re-evaluated for the islanding test **Case 1**, shown in Fig. 6.6(1) i.e., at 105% active and 100% reactive power mismatch condition. In this test case, two types of cyber attacks are considered i.e FDIA and DoS to falsify the actual decision of SRDL as depicted in Fig. 6.15(1) and Fig. 6.15(2). As the SRDL takes its final decision based on observing its toggling outputs pairs i.e presence of ones and zeros of the first one cycle sliding window buffer, attacker takes the advantage of exploiting those one to three cycle islanding input signal information by either injecting synthetic false data or intermittently block the signal by launching Denial of Service (DoS) attack to mislead the DMO about actual islanding case. In Fig. 6.15(1), a FDIA attack is injected in the instantaneous voltage signal of DG-1 at the same time instant of islanding occurrence i.e., 0.205 sec, which results in change of



(1) FDIA Launched to Mask the Actual Islanding Event



(2) DoS Attack Launched to Mask the Actual Islanding Event

Figure 6.15: Masking the real Islanding event

the shape of mean and entropy of decaying DC as shown in Fig. 6.15(1)(b) and (c). It is observed from Fig. 6.15(1)(d) that at the onset of islanding and attack simultaneously, there is a delay in pick up of both DDCD and MID to cross its individual threshold and thus initially SRDL start treating it as non-islanding events. Thereafter, although the entropy of decaying DC start crossing its threshold 2×10^{-3} but the MID values lying below 1 for that duration is too small to raise its first flag of suspected islanding for significant time length. It is also noticed that due to injected FDIA for few cycle, the decaying DC and MID may persist for the longer time than usual (unlike Fig. 6.6(1)) which results in more number zeros than ones at the end of 1 cycle window of SRDL signal. This clearly indicates the hostile attempt of masking the actual islanding event as depicted in Fig. 6.15(1)(g).
The same attacking philosophy also holds equally good for the DoS attack as illustrated in Fig. 6.15(2), where it can be used to masquerade an islanding event as a non-islanding scenario. In this scenario, following an islanding event, the perpetrator interrupts the transmission of legitimate data sample for 40 ms as shown in Fig. 6.15(2)(a). Figure 6.15(2)(b),(c) and(d) demonstrate the significant change in the mean and entropy of decaying DC information over a longer duration as the aftermath of DoS attack which successfully hide the islanding scenario.

6.5.1.2 False Triggering of an Islanding Event

The second case study resembles to a very similar non-islanding test i.e., **Case 3** of Fig. 6.7, where a 0.75 ohm line to line fault took place at 0.2059 sec. At the same time instant, the attackers also launched a random FDIA to distort the islanding input signal vigorously to deceive the SRDL output as illustrated in Fig. 6.16. Here the attacker's aim is to cease the DG-1 generation by pretending a non-islanding event as an islanding event. Thus in that case, neither the attack effort needs to be so rigor nor the attack vector needs to be devised so stealthily. A simple and constant random injection for a sufficient time duration on the instantaneous one phase voltage waveform of the DG-1 is sufficient enough to alter the actual detector's output. As it is evident from the Fig. 6.16(a), (b) and (c), that due to such consistent random attack, the mean value of the signal significantly raised to a high erroneous value throughout the whole attack interval while decaying DC only persist for very small duration comparatively due to fault. This results in MID value for staying above 1 for prolong time which obviously introduces more number of ones than zeros in the SRDL signal over a cycle. This ultimately forces the SRDL detector to flag this event as an islanding event when actually it isn't.

6.5.2 Proposed Cyber Attack Detection Framework Using Kalman Filtering Technique

With the aim of securing the above SRDL's output based islanding detector, a novel cyber security framework is first established as demonstrated in Fig. 6.17, which usually works by representing the instantaneous input voltage waveform into its mathematical state space

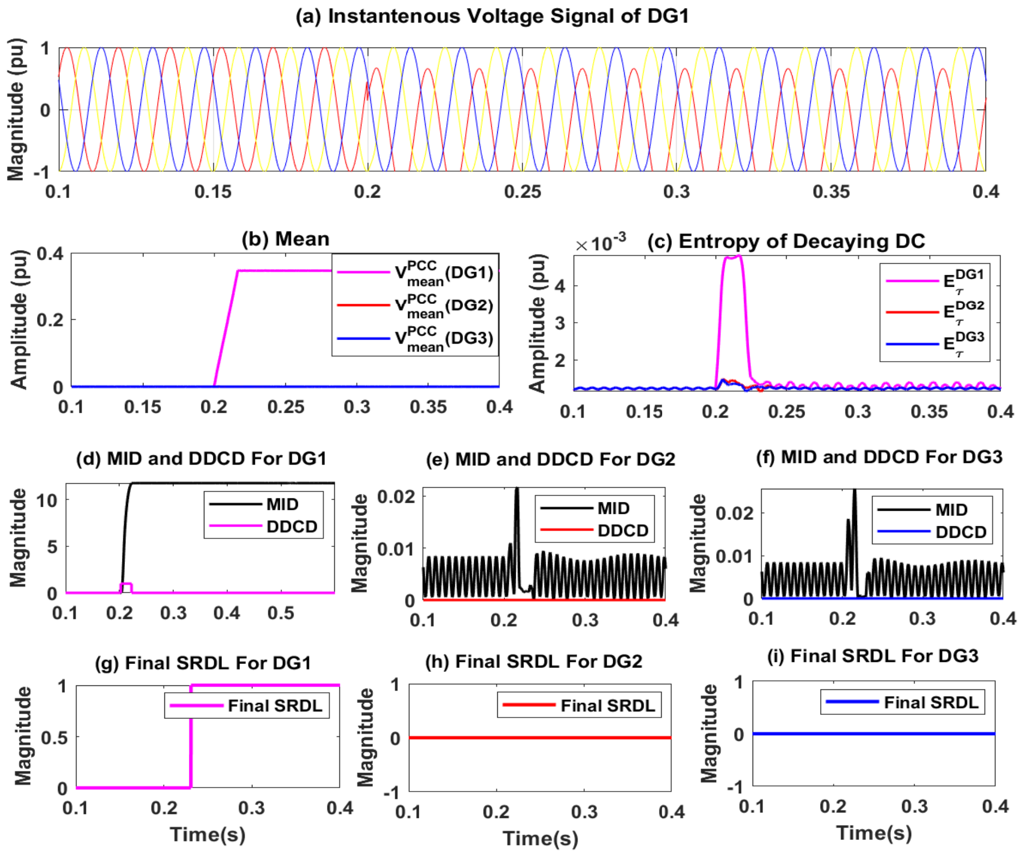


Figure 6.16: Triggering of a fake Islanding event

equations and thereafter applying Kalman Filter (KF) based state estimator method to estimate the purified states of the model even in the presence of measurement noise and cyber attacks. This estimated states at the end are reused to reconstruct the estimated input signal waveform and then two signal processing techniques are applied to find the discrepancies between the actual raw input and KF-assisted estimated output. This leads to the development of a novel CAD which comprises of a stochastic non-parametric correlation coefficient, i.e., Spearman's rank correlation in conjunction with a deterministic Cosine-Similarity measure. The proposed KF-based cyber secured islanding detection framework integrated with CAD can detect onset of any cyber attack within 2 to 3 cycles window while an unauthorized attempt is made to disrupt the actual DG's islanding or non-islanding voltage waveform signal by manipulating its voltage, frequency and phase information. The key advantages of the proposed detector is that it is simple, fast, threshold-free, and accurate against sophisticated FDIA even in the presence of white Gaussian measurement noise of 20 dB. The efficacy of the proposed detection mechanism is validated on the Banshee MG, modelled in RTDS.

According to the Fig. 6.17, the voltage phasors measurements are firstly collected at each time step Δt from RTDS runtime environment, and treated as real-time measurements (RTM). The RTMs are thereafter utilized in KF based state estimator and the CAD as described below.

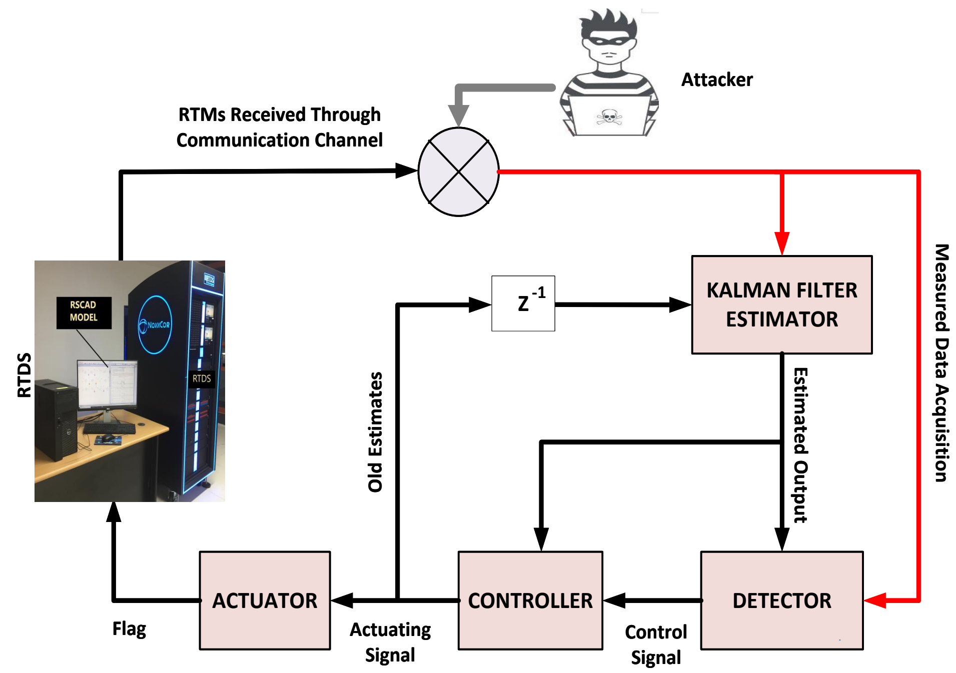


Figure 6.17: Block diagram of proposed generic KF-assisted cyber security framework

6.5.2.1 Kalman Filter and Its State Space Modelling

The RTMs received from a meter-m, at a time instant-t, can be modelled as a sinusoid, i.e, $X(t) = A_m \cos(\omega t + \phi_m)$, where A_m *and* ϕ_m are the magnitude and phase, respectively of the signal. The above equation can be further expanded as follows.

$$X(t) = A_m \cos(\omega t) \cos(\phi_m) - A_m \sin(\omega t) \sin(\phi_m)$$
$$X(t) = x_1 \cos(\omega t) - x_2 \sin(\omega t) \tag{6.3}$$

An equivalent state variable (one cosine projection and another sine projection) representation of Eq. (6.3) can then be used before it is being processed by KF algorithm, i.e.,

$$x(t+1) = \mathfrak{F}x(t) + \check{\nu}(t) \tag{6.4}$$
$$y(t) = \bar{\mathbf{H}}x(t) + \ell(t) \tag{6.5}$$

where, $\boldsymbol{x} = [x_1; x_2]$, with the KF state variables $x_1 = |A_{X_i}| \cos(\phi_{X_i})$, and $x_2 = |A_{X_i}| \sin(\phi_{X_i})$. y represents real time noisy measurement signal, fed to KF estimator at each $Deltat$ time interval. $\mathfrak{F}$ is the state transition Identity Matrix, $\check{\nu}$ is process noise with noise covariance matrix $\bar{\mathbf{Q}}_k$ i.e v_k $N(0, Q_k)$ and ℓ is measurement noise with covariance matrix $\mathbf{R}_k$ i.e ℓ_k $N(0, R_k)$ assumed to be white Gaussian and statistically independent of process noise, and row vector $\bar{\mathbf{H}}$ which is dynamic with respect to time is defined as, $\bar{\mathbf{H}} =$ [cos (2πft) -sin (2πft)]. In this paper the standard deviation of state noise pertaining to $\bar{\mathbf{Q}}_k$ is taken as 0.01 pu and $\mathbf{R}_k$ contains the scalar variance of measurement noise which is calculated from 20 db SNR. Finally, the KF based state estimation is obtained through the following predict and update steps.

$$\textbf{Prediction:} \qquad \hat{x}_i^- = \mathcal{F}_{i-1}x_{i-1}^+ \tag{6.6}$$

$$\bar{\mathbf{P}}_i^- = \mathcal{F}_{i-1}\bar{\mathbf{P}}_{i-1}^+\mathcal{F}_{i-1}^T + \bar{\mathbf{Q}}_{i-1} \tag{6.7}$$

$$\tag{6.8}$$

$$\textbf{Updation:} \qquad \bar{\mathbf{K}}_i = \bar{\mathbf{P}}_i^-\bar{\mathbf{H}}_i^T(\bar{\mathbf{H}}_i\bar{\mathbf{P}}_i^-\mathbf{H}_i^T + \mathbf{R}_i)^{-1} \tag{6.9}$$

$$\bar{\mathbf{P}}_i^+ = (\mathbf{I} - \bar{\mathbf{K}}_i\bar{\mathbf{H}}_i)\bar{\mathbf{P}}_i^- \tag{6.10}$$

$$\hat{x}_i^+ = \hat{x}_i^- + \bar{\mathbf{K}}_i(y_i - \bar{\mathbf{H}}_i\hat{x}_i^-) \tag{6.11}$$

where predicted (priori) and updated (posteriori) estimates are represented by the superscripts '-' and '+' respectively. It essentially means that based on the current input sensor measurement y at time t and previous estimated states reading $\hat{x}^-$ at time $(t-1)$, the estimator of the system produces estimated readings $\hat{x}^+$ at time t i.e every states are being updated at each Δt time interval of KF run. $\bar{\mathbf{P}}$ and $\bar{\mathbf{Q}}$ are the process covariance and model error covariance matrix, respectively. $\mathbf{R}$ is the sensor measurement noise matrix and $\bar{\mathbf{K}}$ represents Kalman gain which can quickly be converged in a few steps. After completion of predict and update operations at each time step, the estimation error in states is computed as , $\breve{e}_i = \hat{x}_i^+ - x$.

6.5.2.2 Proposed Cyber Attack Detector (CAD)

In order to detect the onset of any cyber attack on the voltage sensor measurements, the raw input measurements (X) and the reconstructed estimated output (Y) of the KF are utilized. Since various kinds of cyber attacks are usually crafted statistically in disguise, any significant dissimilarity between the both, i.e., the actual and estimated measurements, can enable the system operator to detect and raise a cyber attack flag for further preventive actions. Now in search for the appropriate statistical tool to measure this degree of dissimilarity, in this section two different similarity measures are utilized to propose a new comprehensive cyber attack detection techniques that involves the computation of stochastic non parametric correlation such as Spearman's rank correlation in conjunction with a deterministic Cosine-Similarity measure which is described below. The knowledge of this correlation coefficients and similarity measures are highly demanding for judging the robust parameter estimation in presence of high frequency noise and outlier data in the observation samples which causes instability issues in the system.

Cosine Similarity Coefficient (CSC): In context of deterministic similarity index, Cosine Similarity Coefficient is now discussed first. This is an interesting measure of resemblance between two non-null vectors of their inner product space that compute the cosine angle between them. Therefore, if majority of observe samples of a bivariate random variable (X,Y) differ of each other in a continuous data streaming process, the cosine angle is going to be increased, and therefore indicates lower similarity. Mathematically, CSC can be expressed as:

$$CSC(X,Y) = cos(\theta) = \frac{\sum_i \hat{x}_i^+ y_i}{\sqrt{\sum_i \hat{x}_i^{+2}}\sqrt{\sum_i y_i^2}} = \frac{< X.Y >}{||X||||Y||} \tag{6.12}$$

Where, $\hat{x}_i^+$s and y_is as the elements of vectors X and Y respectively. However, CSC suffers from a drawback that it is not a proper distance metric, and requires threshold value setting to identify any data manipulation in the measurement data. Further, the *sole* use of this detector tend to give false alarms in presence of noise in the measurement.

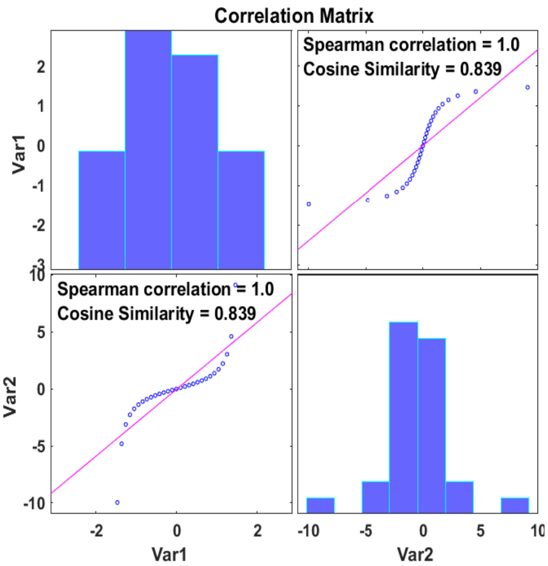
Spearman Rank's Correlation Coefficient ($\mathbf{r_s}$): To circumvent above stated drawbacks of CSC, a non-parametric correlation coefficient, i.e., **Spearman rank's correlation coefficient** (r_s) is utilized in this work in conjunction with CSC. The Spearman rank's correlation coefficient provides a non-parametric correlation that measures the strength and direction of two monotonically changing random variables based on assigning ranking on them, and is calculated as,

$$r_s(X,Y) = 1 - \frac{6\ \sum_{k=1}^{N}(u_k - v_k)^2}{N(N^2-1)} \tag{6.13}$$

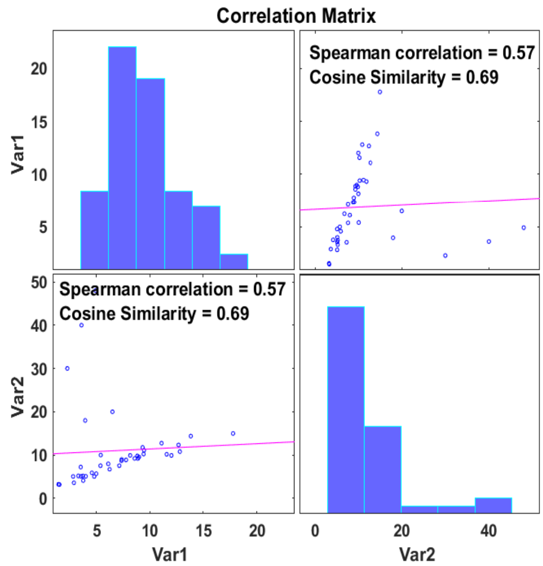
where, u_k and v_k are the corresponding ranks of two data vectors X and Y for k = 0...,N-1, and N be the total number observation pairs.

The performance of Spearman rank-order correlation and CSC is depicted in Figs. 6.18(a) and 6.18(b) under normal and attack scenario.

Figure 6.18(a) depicts that in case of monotonic relationship Cosine Similarity Coefficient (CSC) and Spearman rank correlation (r_s) both are higher with slight difference in value because CSC cannot recognize the exact association between the two variables unlike r_s does. Besides from Fig. 6.18(b), it is revealed that r_s helps to obtain a valid result as compared to the CSC for it is more robust in estimation process when data contains strong outliers or heavy tailed errors. CSC is not a direct measure of statistical association of two variables and therefore it is not invariant to shifting of data which makes is sensitive to outliers. While on the other hand, the unique feature gained by the spearman correlation (r_s) is just because of it determination based on ranking the real observation instead of direct using of raw measurements. Thus it can be stated, if any changes occur in original measurements that do not have significant impact on the earlier rank order, should not alter r_s unlike CSC. In context of cyber-attack detection, this chapter considers the use spearman correlation and cosine similarity jointly as a detection index when both diverge or converge to each other from a certain point. A detailed study of various test cases revealed that the CSC and r_s, although, may follow different pattern but at the *onset* of malicious data attack in measurements, their values coalesce. Based on this feature,



(a) Data with Monotonically Increasing Relationship



(b) Data Contain Heavy Tailed Outliers

Figure 6.18: Comparison of performance between Cosine Similarity and Spearman's Rank Correlation Coefficient

the proposed cyber attack detector (CAD) is defined as the absolute difference between Spearman rank-order correlation, r_s and CSC values, i.e.,

$$CAD = |r_s - CSC| \tag{6.14}$$

And, at the time of cyber attack in the measurement, $CAD = 0$. This implies that the flag raised by CAD during attack is 1.

The overall working procedure of CAD's operation is summarized in the flowchart of Fig. 6.19. Here, a small threshold on CAD less than 0.002 is chosen for practical purposes. Moreover, as the transients involved at the inception of any power system operational events is hardly lasting for 2 to 3 cycles, t_{wait} is fixed to 0.04 sec at the starting of the simulation. Moreover, in order to judge the efficacy of CAD's performance along with its KF-based estimator, a 20 dB white Gaussian noise is also added with the studied input voltage signal.

6.5.3 Proposed Cyber Attack Immune Islanding Detection Framework

The previous Section 6.5.1, reveals how different intent of a cyber attacker can raise serious concern about the cyber security issues over the SRDL's output based islanding detection to mislead the actual islanding and non-islanding scenarios and in Section 6.5.2, a KF-assisted attack detection framework incorporating two statistical similarity measure was demonstrated to identify various cyber attacks in the sensor measurements. It is evident from the above two that SRDL is very trustworthy in terms of pure islanding and non-islanding event identification. However, as it lacks anomaly detection awareness, it is

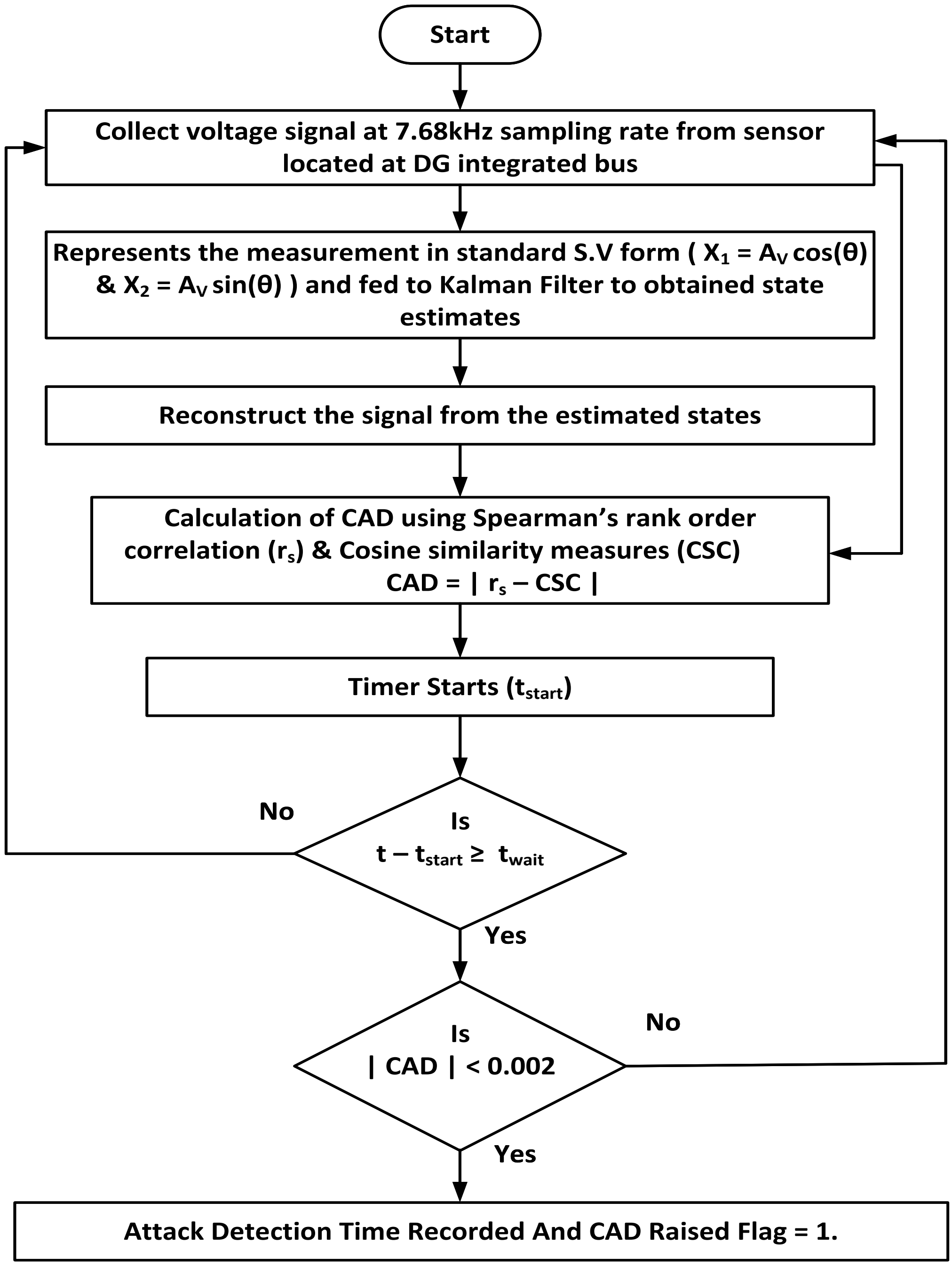


Figure 6.19: Proposed algorithm for Cyber Attack Detector (CAD)

expected to maloperate in any kind of cyber intrusion. CAD on the other hand, is faithful in differentiating between attacked and non-attacked state. This motivated to shift the focus in achieving a secured islanding detection operations in this subsection where an attempt is made from the defender perspective to rectify the manipulated SRDL decision in favor of the accurate assessment of the actual event that had occurred in the MG. In particular, this can be obtained by combining the proposed islanding detection scheme with that of proposed CAD. Such integration results in four possibilities of test scenarios which are illustrated in the truth table shown in Table 6.3. This table presents a variety of test cases that assess how well the cyber secured islanding detection scheme performs in various scenarios and against potential attacker actions. It is evident that the flag

Table 6.3: Truth Table for Cyber Immune Islanding Detection Logic

Possible Test Scenarios	**Attacker's Motive**	**Flag Generated by SRDL**	**Flag Generated by CAD**	**Final Relay Trip Logic (FRTL) Considering Cyber Intrusion**	**Actual Decision Made by the DMS Operator based on FRTL Output**
No Islanding + No Attack	No Motive	0	0	0 **(No Trip)**	**Normal State:** System is in healthy condition.
Islanding + Cyber Attack	Mask Actual Islanding Event	0	1	1 **(Send Trip Command)**	**Masked Islanded State:** System encounter a genuine islanding in real.
Non Islanding + Cyber Attack	Trigger Fake Islanding Event	1	1	0 **(No Trip)**	**Fake Islanded State:** System encounter a genuine non-islanding in real.
Islanding + No Attack	No Motive	1	0	1 **(Send Trip Command)**	**True lslanded State:** System face a genuine islanding in real.

produced by the CAD is always 0 in an islanding or non-islanding event provided there is no cyber interference on the system. Conversely, when the attacker attempt to manipulate the sensor measurements of the system via means of FDIA or DoS attack, CAD raise a flag of 1. This signifies that CAD is insensitive to islanding or non-islanding event but sensitive to the occurrence of cyber attacks. Thus, based on this behavioral attributes, a logical XOR operation can be applied over these two flags, generated by both SRDL and CAD in order to formulate a Final Relay Trip Logic (FRTL), which is secured than SRDL and considered to be as *attack-proof.* The simulation results of the next section will illustrate the compelling evidence of abilities of FRTL in enhancing the cyber situational awareness of islanding detection problem in presence of cyber intrusions and enabling the DMO to take proper decision and control action against the real occurring events in the grid. Therefore, following each possible instances as indicated in Table 6.3, the schematic architecture of cyber secured FRTL, merging the SRDL's and CAD's output is designed as shown in Fig. 6.20. In that reference, it is also important to highlight two crucial observations i.e., (1) Decision given by SRDL is relatively faster than CAD and (2) From

MG's safety point of view, masking of an actual islanding state is more hazardous or devastating for the grid functioning than the creation of fake islanding state. Thus, based on the level of importance in detecting masked and faked islanding state and impact of attack consequences, it is noteworthy to consider the following two design criteria for FRTL to get a reliable and safe islanding performance: **(a)** *While SRDL is 0, the FRTL will be instantly activated when CAD raised a Flag = 1.* This is because apart from *No Attack-No Islanding* state, SRDL can also be 0 in masked islanded state. In that context, fastest action in ceasing of islanded DG's generation is needed with atmost priority to maintain grid's voltage and frequency stability. **(b)** *While SRDL is 1, the FRTL will wait for maximum 2.5 cycles to seek for the Flag raised by the CAD. If the Flag is raised to 1 within the due waiting time, FRTL will be activated instantly as soon as CAD detect the attack.* This criteria is very important to prevent the FRTL in taking wrong decision (i.e., sending false trip command to circuit breaker) while SRDL is manipulated to 1 but the CAD is yet to reach its designated threshold to raise an attack detection flag due to its sluggish operation as compared to SRDL. From the simulation results, it would be clearly manifested that this maximum waiting cycle introduces a delicate balance between detecting a camouflaging islanding state and preventing a non-islanding state pretending to be islanding with an efficient manner.

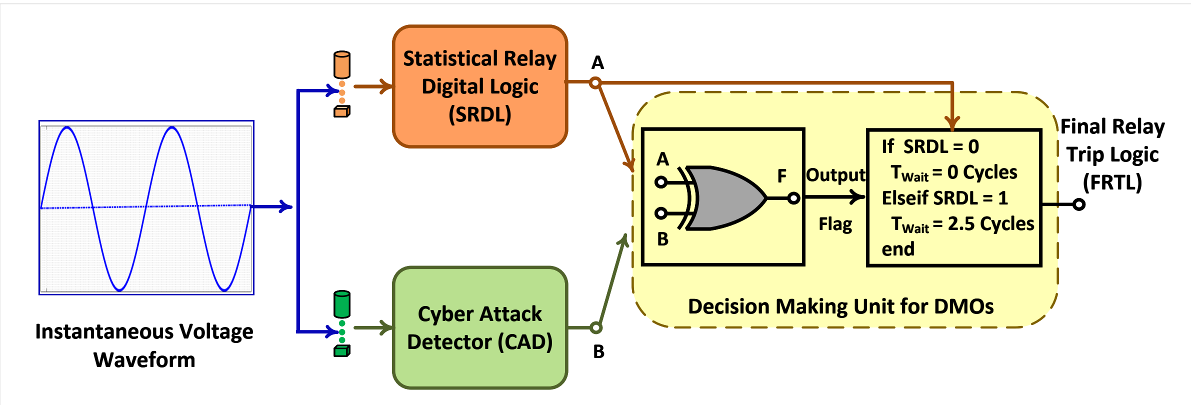


Figure 6.20: Schematic architecture of cyber immune Islanding detection scheme

6.6 Attack Detection Simulation Results

6.6.1 Islanding State With No Cyber Intrusion

This is the same islanding case study as shown in Fig. 6.6 where there is no intervention of the cyber attackers are considered initially. Figure 6.21 depicts the overall performance of the proposed cyber immune islanding detection scheme. In Fig. 6.21(a), three different versions of input instantaneous voltage waveform are shown. The green waveform represents the actual (true) voltage readings at the DG1 terminal. To emulate a realistic scenario and assess the detector's robustness, a 20 dB noise signal is added to produce a measured observation (blue waveform), which serves as the input to the KF-assisted estimator. The output of the KF estimated filter, which needs to be processed by the proposed CAD, is highlighted in the red dotted curve in Fig. 6.21(a). Figure 6.21(b) and (c) illustrate the performance of the proposed detector, i.e., CAD by observing the behavior of two similarity indices. In Fig. 6.21(b), it can be observed that in the absence of any attack, the two coefficients exhibit a non-interactive nature, and therefore, CAD never reaches a zero value or falls below the set threshold, even in the presence of noise, as depicted in Fig. 6.21(c). Thus as expected, the SRDL correctly identifies the islanding (flag = 1) situation as done previously, while the flag of CAD is set to 0 always. Consequently, based on the XOR operation between these two's as shown in Fig. 6.20(d), the FRTL issues a final relay trip command after 2.5 cycle of waiting and informed the DMS operator about the authenticate true islanded state of the systems.

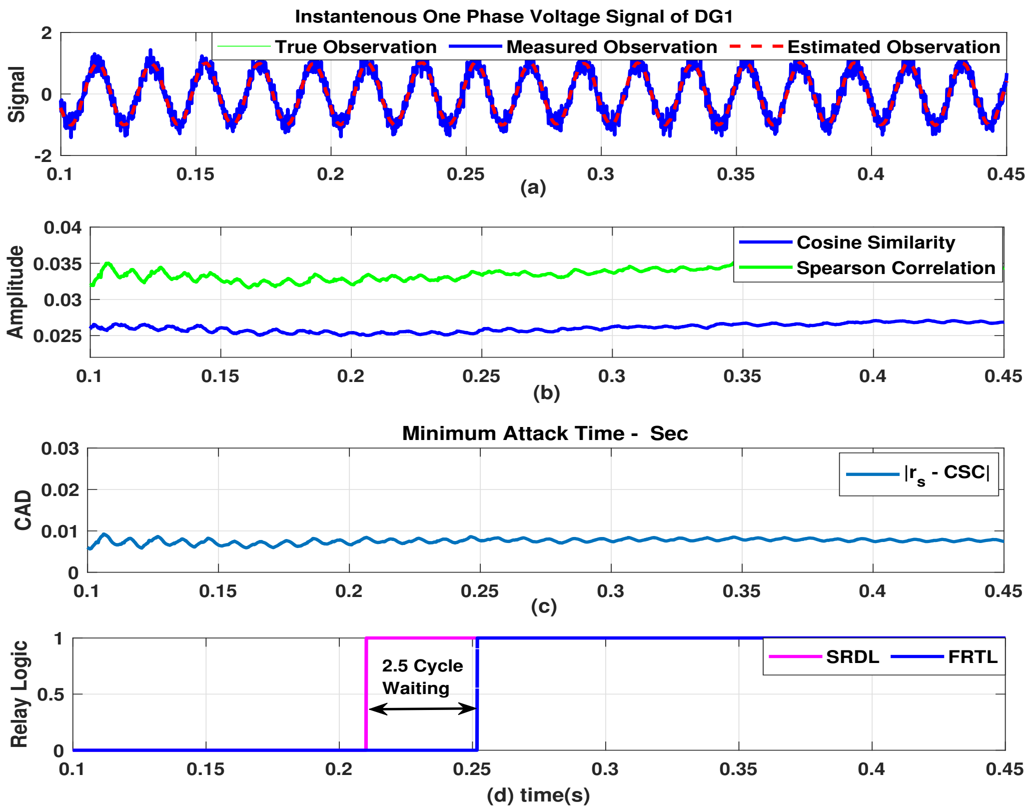


Figure 6.21: True Islanding condition with No cyber attack

6.6.2 Non-islanding State With Random Nature of Cyber Attack

In this case, the same non-islanding case study as depicted in Fig. 6.7 is taken but with considering the fact that now the cyber attacker has access to the DG1's control interface to manipulate the input instantaneous voltage signal. A LLG fault with low resistance of 0.75 ohm and a random kind of cyber attack are injected at the same instant of the time (0.2059 sec) to the input signal as shown in Fig. 6.22(a). Such attack can be generated at any time instant with randomly crafted mechanisms to introduce some arbitrary errors into the measurements and state variables of KF to mislead the operation. Random attack

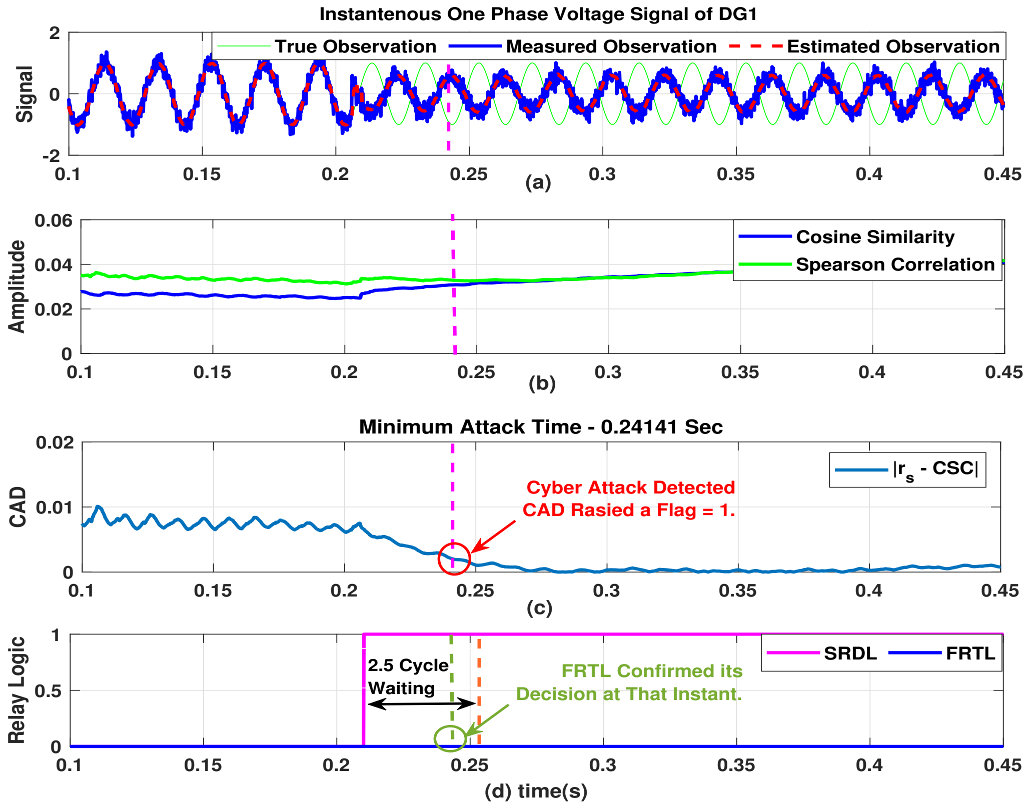


Figure 6.22: Fake Islanding condition with Random FDI attack

can be modelled as, $y^a(t) = C(t)x^a(t) + y_k^v(t)$, where, $y_k^v(t) = (1 - \beta_k)G_k\delta_k(t)$, $\delta_k \in \mathbb{R}^n$ be a manipulating signal parameter by the attackers, G_k is known system topological information i.e about system's states. β_k is an independent Bernouli distributed series whose values lies in between 0 and 1, deciding the strength of the injected attack. The first important thing here is to notice that SRDL is failed to detect the non-islanding event in this case due to the random attack intervention which basically alters the rule of its computational parameters and as a result it generated the trip signal which is nothing but the straightforward indication of fake islanding event as illustrated in Fig. 6.22(d). Thus, once the SRDL's flag is raised to 1, an alarming suspected islanding situation in the MG is created which initiate the task of 2.5 cycles observation window to monitor any compromised behavior exhibited by the CAD. Interestingly, it is observed from Fig. 6.22(b) that initially before the attack, when measurements are clean and normal both the coefficients (CSC and r_s) are behaving independently. But as soon as the malicious data randomly starts getting injected, the Spearman's rank correlation, i.e., r_s and the CSC very quickly start to trace the estimated change in the observed data pairs after the attack's initiation. Subsequently, the proposed CAD starts converging to zero and reached below threshold limit ($2e^{-3}$) at about 0.2414 sec with raising its Flag = 1. This successful detection of random attack is completed by the due time less than the waiting cycles as shown in Fig. 6.22(c). This small detection delay, involved in this detection process may due to their underlying performance measure which is inherently different.

Finally, it is evident from the Fig. 6.22(d) that this flag, output by the CAD acts as a correcting measure through the formation of actual FRTL to revert the erroneous trip signal generated by the SRDL previously due to misinterpretation of the non-islanding event as an islanding event.

6.6.3 Non-islanding State With Denial-of-Service (DoS) Attack

This test case is very similar to the case study conducted previously with the difference that this time attacker choose the DoS attack to manipulate the SRDL's decision. In context of disruption in islanding detection problem of power systems, a DoS attack usually be launched by interrupting the legitimate transmission of accurate and timely data between sensors and monitoring units for sufficient duration which causes the system to get congested or may overloaded as shown in Fig. 6.23(a). The lack of sensor data due

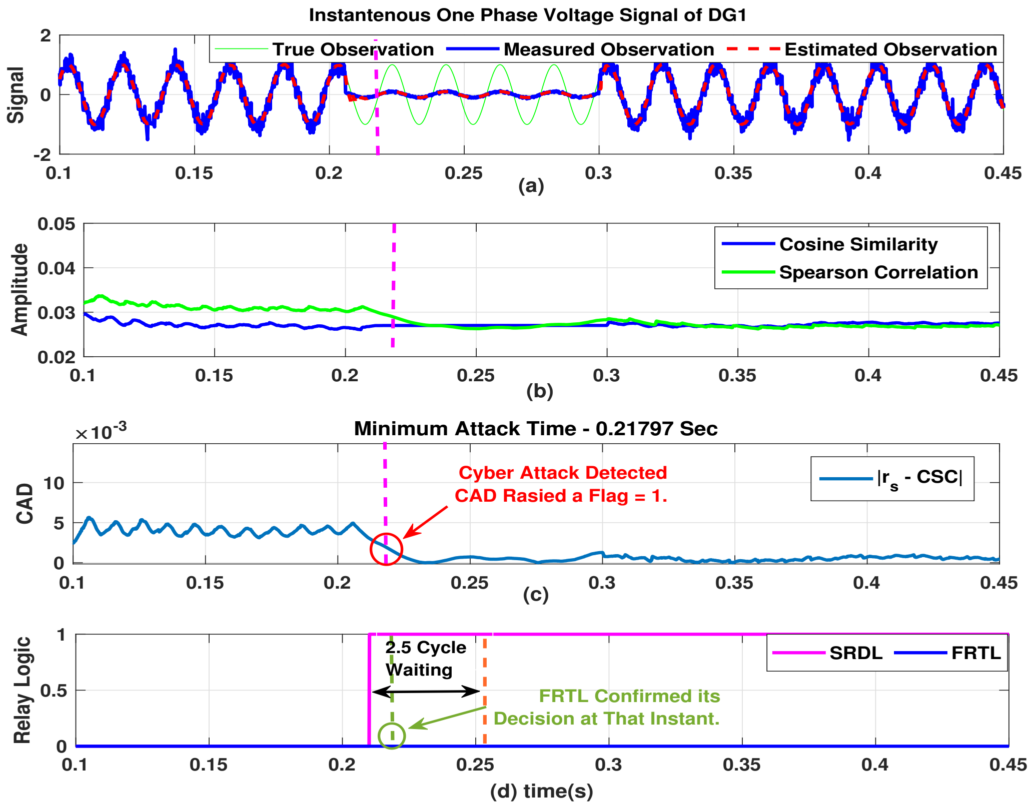


Figure 6.23: Fake Islanding condition with DoS attack

to such DoS attack posses challenges to SRDL's output based PIDM in terms of delaying detection of actual islanding event, potentially leading to blackouts and equipment damage or may make the input data missing or incomplete resulting in inaccurate and falsified islanding detection decision. This same observation is very well validated from Fig. 6.23(d), where SRDL's operation is found to be deceived in accurately distinguishing between genuine islanding data or non-islanding state. As a result of such misconception, SRDL inadvertently trigger false alarms against the actual non-islanding scenario, causing DMO to take unnecessary system actions that further complicate and compromised system reliability and security. Figure 6.23(b) and (c) shows the evidence of how the proposed detector, CAD assist in rectifying the vulnerable SRDL's control decision by tracing the independent behavior of r_s and CSC before and after attack. It can be seen that within a few cycles of SRDL's trip command, the proposed CAD reached the below threshold at 0.2178 sec which finally restricts the FRTL to raise Flag = 1 unlike SRDL and thus the fake islanding decision given by SRDL is successfully suspended.

6.6.4 Islanding State With False Data Injection (FDI) Attack

This case study represents an another motive of attacker's where the attempt is to mask a genuine islanding event through the strategic execution of FDI attack into the islanded voltage waveform data with the aim of jeopardizing the stability of the grid, leading to cascading failures and equipment damage. Here the islanding event has been taken place from 0.2059 sec and thereafter a small amplitude of false data are being injected very stealthily by the adversary in each positive and negative half cycles of the instantaneous voltage waveform for an attack duration of 2 cycles. From Fig. 6.24(d) it is revealed that, as this is a case of real islanding event, SRDL is expected to raise its flag to 1. But due to such synthetic injection of carefully constructed stealthy attack, SRDL is deviated from its actual expected decision by raising its Flag = 0 and treat this islanding event wrongly as a no-islanding case which may impose serious repercussions on DG's operation. Nevertheless, as the defender is equipped with an attack detector, CAD in parallel with

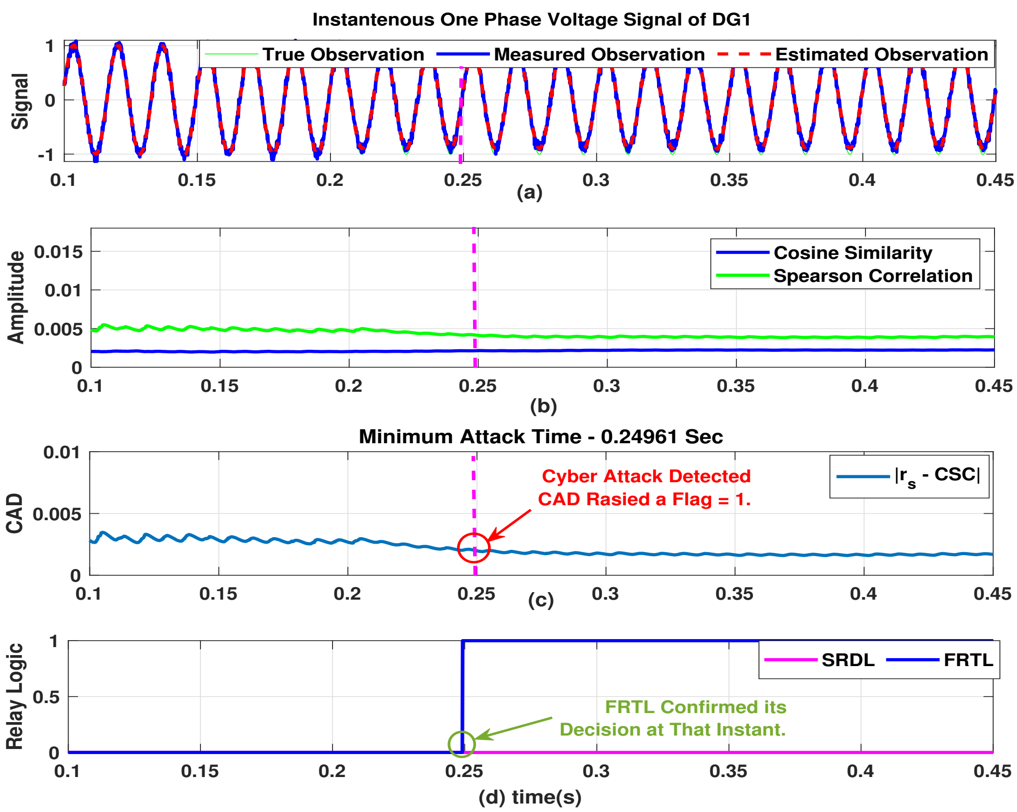


Figure 6.24: Masking of an Islanding condition with stealthy FDI attack

that of SRDL, it is observed from Fig. 6.24(b) and (c) that at around 0.249 sec, the r_s and CSC are converged to a desired value where the absolute difference between these two's lies below the threshold limit ($2e^{-3}$). Therefore CAD has raised its flag to 1 instantly at that timing instant. Now, as the SRDL was 0 previously, the instant change in CAD's output immediately raise the FRTL to change to 1 as illustrated in Fig. Fig. 6.24(d) and send the trip command to the circuit breaker by informing the DMO that an actual islanding state is being attempted to mask by the adversary and therefore a prompt action of ceasing DG1's generation is required.

6.7 Conclusions

In this chapter, primarily a data driven passive islanding detection technique is proposed which is later being integrated with an attack detection framework to keep it safe and secured against unprecedented cyber intervention. The proposed islanding detection technique first exploit the voltage mean value and the entropy information of any one phase to develop a Mean based Islanding Detector (MID) along with an entropy-based Decaying DC Detector (DDCD). The MID and DDCD information is finally utilized to design a Statistical Relay Digital Logic (SRDL) that accurately distinguishes the islanding and non-islanding events. The Proposed Islanding Detection Method (PIDM) is rigorously tested on Banshee industrial MG system, modelled in Real-Time Digital Simulation platform revealing the following notable conclusions.

1. Simple implementation, requiring only 254 samples/scan.
2. Requires only one phase voltage mean information.
3. A small non-detection zone (NDZ) of approx 0.25% is resulted.
4. Fast islanding detection within 2 cycles.
5. Integration of the additional feature i.e., DDCD with MID improves the performance in terms of accuracy and reliability as compared to traditional like ROCOV.
6. Loss of Parallel Feeder (LOPF) case establishes the capability to accurately locate exact point of disconnection.

In conclusion, the PIDM has been demonstrated to be fast, reliable, and capable of accurately detecting island formation at the point of DG interconnection. However, its vulnerability to cyber attacks, stemming from a lack of cyber situational awareness, presents a significant challenge to achieve a faithful islanding detection operation. These attacks could involve malevolent actions aimed at masking an islanding event or falsifying a non-islanding event as an islanding scenario. To overcome these challenges, this chapter also introduces a statistical similarity-based Cyber Attack Detector (CAD) within a Kalman filtering framework where two similarity coefficients, viz., Cosine Similarity Coefficient (CSC) and Spearsman's Rank order coefficient (r_s) are explored to identify potentially corrupted samples of the input islanding data. The CAD operates synergistically with the PIDM, providing a smooth and cyber-secured islanding detection experience. Variety of cases studies led to the following key conclusions about the proposed CAD.

1. Accurately detects the onset of cyber attacks with a maximum delay of two and half cycle.
2. Performance is fairly immune to noise.
3. Integrating CAD with the existing PIDM, the system gains enhanced resilience against cyber threats, ensuring more robust and accurate detection of islanding events while mitigating the risk of false alarms caused by malicious cyber activities.

Chapter 7

Conclusions and Future Scope

7.1 General

The integration of advanced sensing, computing, communication, internet and networking technologies in the power sector has transformed electrical grids, making them more flexible, reliable, and efficient. This evolution enables the grid to actively manage electricity flow, monitor consumption, optimize resources, and integrate renewable energy. However, due to extensive dependencies of grid over the communication and several layers of the cyber network, this transformation at both the transmission **(T-system)** as well as the distribution **(D-system)** level has exposed the power grid to unprecedented vulnerabilities, primarily stemming from the increasing threat of cyber-attacks.

In response to the emerging cyber-attack concern, this research aims to assess power system vulnerabilities at transmission and distribution level with focus on detection and mitigation of cyber threats across the spectrum of transmission to active distribution power networks, safeguarding the reliability and security of our vital energy management system applications. To begin with the *T-systems* first, the thesis developed a cyber-attack resilient secured metering infrastructure through optimal placement of PMUs (OPP) to ascertain full topological observability and global situational awareness against false injection attack on top ranked transmission lines that are structurally more vulnerable. Next, by exploiting those secured meters' information resulting from OPP, a reliable replay attack detection and correction framework is devised to safeguard one of the core instrument of power grid energy management system i.e Power System State Estimation (PSSE). While conventionally, cyber attacks have been perceived as primary threats to transmission systems due to the significant power flows associated with these networks, it is imperative to recognize the growing importance of safeguarding active distribution systems as well, specifically the microgrid (MG) in light of the heavy integration of distributed energy resources (DERs). DERs requires seamless communication among its neighboring units or master controller units to regulate power flows and maintain stability within MG, yet this interconnectedness makes them susceptible to cyber threats. Therefore, as the cyber security landscape is in the process of shifting from transmission to distribution systems, the next focus of this thesis is to explore the vulnerabilities in communication architectures and controllers of DERs of MG to develop an end-to-end attack resilient and control framework for *D-systems* that offers a multifaceted defense mechanism capable of detecting, classifying, isolating, and neutralizing cyber-attacks in distribution grid with unprecedented efficiency and efficacy. At the last, the thesis deals with an another looming challenge in the distribution side of MG i.e developing of a cyber-vigilant robust passive islanding detection technique for a cyber-physical smart grids. In particular, the major contributions of thesis can stated as follows,

- Assessment of the structural vulnerability of power grid network and thereby development of a cyber-attack resilient secured metering infrastructure for the **T-system** based on optimal placement of PMUs.
- Development of a novel replay attack detection and mitigation framework for power system state estimation by exploiting the limited secured measurements obtained from OPP locations.
- Accurate detection, classification and localization of cyber attacks tailored to **D-system** specifically islanded AC MG system.
- Following the attack detection and localization information, devise an Unknown Input Observer (UIO) and Back-stepping Integrated Sliding Mode Control (BSMC) based Cyber Attack Mitigation Framework for MG system.
- Development of a passive islanding detection scheme (IDS) in MG which is immune to maloperation caused by possible cyber attacks.

7.2 Summary of Contributions

The first crucial steps towards the development and implementation of the aforementioned **Cyber Attack Resilient Monitoring and Control Framework** for fortifying *T-system*'s resilience against cyber attacks involves performing structural vulnerability analysis to identify vulnerable points within the grid that require protection or reinforcement. This analysis emphasis on understanding the physical behavior of power systems with its topological structure which in turn serves as a proactive measure to identify and mitigate potential cyber-physical vulnerabilities within power systems. In this regards, **Chapter-2** firstly introduced a novel attack strategy, termed Hybrid Between-ness Centrality (HBC) from the perspective attackers to identify top-ranked transmission lines vulnerable to malicious tripping, thereby compromising system observability and situational awareness. In next, the framework strategically places Phasor Measurement Units (PMUs) through a unique objective function to protect those vulnerable lines against false data injection attacks (FDIAs), ensuring secure measurements and system integrity. This proactive approach inturn enhances system resilience and maintains observability even during data integrity attacks. The effectiveness of the proposed framework conducted on the IEEE 14-bus and New England (NE) 39-bus systems provide the following conclusions:

- The Hybrid Between-ness Centrality (HBC) index emerges as a proficient attack strategy, effectively identifying groups of transmission lines whose sequential outages top 20% lines could lead to significant structural breakdowns within the system. This results in a total of 4 vulnerable lines for IEEE 14-bus and 6 lines for NE 39-bus system respectively.

- In comparison of HBC with another two conventional attack strategies i.e., Topological Betweenness Centrality (TBC) and Electrical Betweenness Centrality (EBC), the HBC results in maximum decline in Giant component size (S^l). For IEEE-14 bus system, after all the top four vulnerable links are consecutively attacked, the TBC strategy results in no change in reduction of S^l, where as EBC has 87% and HBC has highest 57% reduction. Similarly for NE-39 bus system the reduction of S^l for TBC, EBC and HBC are 69%, 71% and below 50% respectively.

- The novel PMU deployment framework developed to prioritize full system topological observability demonstrates effectiveness in enhancing system resilience against HBC-based attacks and defending against data integrity threats. By strategically placing PMUs to the almost $\frac{1}{3}^{rd}$ of total busses, the framework ensures the availability of secure measurements, thereby safeguarding system integrity.

- Finally, based upon number of secured measurements attained through optimal PMU deployment, the resiliency score of IEEE 14-bus and NE-39 bus systems are found out to be 46% and 57% respectively.

Building upon a secured metering infrastructure through optimal PMU placement in the previous chapter, **Chapter-3** delves into the another critical aspect of safeguarding the heart of the Energy Management System (EMS), the Power System State Estimation (PSSE), against stealthy cyber threats, particularly Replay Attacks (RAs). To this end, this chapter firstly leverage the topographical information along with branch and nodal version of Power Transfer Distribution Factor (PTDF) to identify most sensible vulnerable Remote Terminal Unit (RTU) meters and thereafter these are exploited to launch two different variants of RAs (i.e., MDDA, RDCA) to disrupt the decision making operation of control center operator via compromising PSSE. Subsequently, a detection and correction approach is developed to safeguard the PSSE against RAs, utilizing secured phasor measurements from optimally placed PMUs in a Hybrid State Estimation (SE) algorithm. The effectiveness of the proposed framework, demonstrated on the previous two standard test system modelled in RSCAD software of Real-Time Digital Simulator (RTDS) leads to the following key conclusions.

- In this study, the proposed branch and nodal PTDF based attack strategy leads to identification of total 36% and 40% of vulnerable meters from the available RTU measurement set for IEEE 14-bus NE 39-bus system respectively.

- The likelihood of detection of the two proposed attack strategies, i.e., Repetitive Data Cloning Attack (RDCA) and Multiple Data Dropping Attack (MDDA), is evaluated, with average detection rates of 94.6% and 90%, respectively. Despite some False Positives (FPs) and False Negatives (FNs) in both algorithms, from being attacker perspective RDCA exhibits slightly inferior performance to MDDA.

- The accuracy estimation of the correction method against MDDA and RDCA is assessed, revealing accuracy rates of 82% for MDDA and 93.27% for RDCA. This indicates that more precise correction of RAs is achieved for RDCA.
- The Root Mean Square Error (RMSE) of estimated states of the PSSE is evaluated under both MDDA and RDCA, resulting in RMSE values of 0.4% and 0.45%, respectively, after applying the attack correction algorithm.
- The true negative rates for both attack types are approximately 100%, indicating the high specificity of the suggested detection and correction techniques. This indicates the robustness of the algorithms in correctly identifying instances that do not belong to the attack class after correction, thereby ensuring the integrity of system operations.

Besides addressing the vulnerabilities inherent to *T-system*, the next three chapters contributes significantly for enhancing the overall security and reliability of *D-system* as well particularly in realm of communicative MG environment via development of Cyber Attack Resilient Monitoring and Control Framework. To this end, **Chapter-4** firstly utilized a statistical two-sample hypothesis test called the Maximum Mean Discrepancy (MMD) for the attack detection process over DER's controller or its associated communication links. Having the attack detected, thereafter few more statistical properties are exploited to formulate a rule-based algorithmic flowchart integrated with a popular ML classifier i.e XGBoost for the efficient attack classification and precise attack localization inflicted to perverted DERs. The main findings of the proposed approach which were verified on a modified IEEE 13-bus islanded AC microgrid system modeled in the RTDS environment, are listed below.

For MMD Based Cyber Attack Detection in Distributed Secondary Frequency Control (DSFC) of MG

- The proposed non-parametric statistical test, MMD successfully able to detect FDIA targeted in either the secondary frequency controller's DERs or its incoming and outgoing communication links.
- The proposed detection strategy is not further limited by the number of role statuses (corrupted or healthy) of nearby DERs, nor is it subject to the strict premise that the leader DER must always be secured.
- It is successfully able to differentiate between an cyber events and normal physical events i.e fault/ switching events, leading to no *false alarms*.

- The proposed detection strategy shows its superiority over the existing Kullback Leibler divergence (KLD) based detection method in terms of evading delay in detectability and threshold selection problem under varieties of attack models.

For XGBoost Enabled Proposed Rule-based Precise Attack Classification and Localization Scheme

- The proposed novel rule-based XGBoost classifier exhibits exceptional performance in classifying FDIAs, achieving an accuracy of 99.49%. This outperforms existing ensemble machine learning (ML) techniques.

- The rule-based approach demonstrates remarkable precision, recall, and F1 Score for detecting various attack types. It achieves 100% in detecting simpler attacks like pulse attacks and maintains high performance for more complex attacks such as sine attacks, with precision, recall, and F1 Score values of 97%, 98%, and 97% respectively.

- The proposed XGBoost-enabled attack localization scheme showcases superior accuracy (87.5%) and a lower hamming loss (4.5%) compared to existing ML classifiers like Decision Tree, Random Forest, and Gradient Boosting.

- In terms of precision, recall, specificity, and F1 Score, the proposed localization scheme excels particularly for DER-3, followed by DER-2, DER-1, and DER-4. This observation is further supported by the Receiver Operating Characteristics (ROC) curve analysis, which demonstrates the Area Under the Curve (AUC) for each DER: i.e., 92.3% for DER-1, 93.7% for DER-2, 99.1% for DER-3, and 89.9% for DER-4. The higher AUC for DER-3 indicates that the proposed localization method perform better in differentiating between attack and non-attack instances for the attacked in DER-3 compared to other DERs.

After successful detection, classification and localization of attacks in Chapter-4, the next very crucial step is to nullify the effect of attack in compromised DER's secondary controller to bring back the MG system to normalcy. With this as an aim, **Chapter-5** develops an unified cyber attack resilient framework comprising of an Unknown Input Observer (UIO) and Back-stepping Integrated Sliding Mode Controller (BSMC). The UIO estimates attack bias injected into the controller, which is then used by adaptive BSMC to generate a counter control law that enforce the attack to be mitigated. The validation of the aforementioned detection and mitigation techniques in Chapter-4 and Chapter-5 is performed on an modified IEEE 13-bus distribution system through the hardware-in-loop testing environment which extract the following key concluding remarks of the proposed attack resilient framework.

- The main advantage of the proposed attack mitigation method is that it neither depends on the limitation of the number of indegree healthy DERs of the compromised unit nor the role-status of leader DER information to be secured.

- The proposed scheme does not demand any modification of the existing hardware of the DSFC or the inclusion of additional communication channels to achieve this resilient action. Therefore, it is simple, cost effective and less computationally expensive.

- The key highlight of the proposed resilient method is due to its faster convergence and good robustness against different attacks. Moreover, it efficiently regulate the power sharing between DERs in the MG even in the presence of attacks.

- The proposed controller is very adaptive to the unknown bounded attack injections and offers higher resilience and better utilization of DERs as the information of infected DERs no longer needs to be separated now from the existing communication topology to stop the spread of the attack effect.

Finally, the **Chapter-6** delves into an another challenge of accurately detecting islanding events in smart active distribution cyber-physical systems amidst emerging cyber threats. It introduces a novel passive islanding detection method (IDM) based on entropy information from decaying DC detector (DDCD) and mean-based coarse islanding detector (MID). Later, the decision of this proposed islanding scheme has been integrated with a novel kalman-filter based cyber attack detector (CAD) module in order to identify statistical inconsistency in the signal that confirms potential manipulation of islanding input. Testing on the Banshee industrial MG system in the Real-Time Digital Simulation (RTDS) platform validates the effectiveness of the proposed hybrid method with following concluding remarks.

- The statistical property inherited by IDM with kalman filter assisted attack detector makes the system operator well-informed from being taken any wrong decision over any suspected islanding event caused due to cyber attacks.

- With a maximum half-cycle delay, the proposed CAD reliably identifies the onset of the attacks.

- The CAD's performance is fairly noise-immune, and no threshold selection is required.

- Besides the reliable performance of CAD to detect attack from cyber manipulation perspective, the proposed IDM also exhibits its robustness in identifying various islanding cases like different active and reactive power mismatches and various non-islanding cases like fault, linear and non-linear load and capacitor switching etc.

- Unlike the most of the other islanding method, the proposed AID schme also shows its superior performance in terms of not falsely detecting the loss of parallel feeder and removal of other DERs except targeted DER as an islanding event.

- The proposed AID method is also simple as it only need one cycle (254 samples/scan) voltage data of any one phase for the implementation and resulted only a small non-detection zone (NDZ) of 0.25%.

- With the integration of DDCD with MID as an additional features enhances the performance accuracy and reliability of AID scheme compared to traditional ROCOV method.

7.3 Scope for Future Work

This thesis investigated the growing cyber security challenges faced by the power grid due to strong interdependence over the critical cyber infrastructure and heavy integration of communication and networking technologies within the transmission **(D-system)** and active distribution systems **(T-system)**. The research in this area can be further extended as follows.

- In order to develop an end-to-end attack detection and control framework from cyber-physical smart grid's perspective, there are still few research areas that are untouched and needs more defense-in-depth research studies. The examples for such future research avenues are: (1) *Generation Side Perspective:* Resilient defense against FDIA on measurements and control signals for the normal operation of Automatic Generation Control and Load Frequency Control etc, (2) *Transmission Side Perspective:* Defensive measures to resist the line current differential relays from being maliciously tripped against cyber attacks, (3) *Distribution Side Perspective:* Enhancing security and reliability of deregulated electricity market and its trading application through proposing an offer-breach detection process.

- While the focus of existing studies has predominantly centered on AC MGs, the growing prominence of hybrid AC-DC architectures underscores the need to extend and adapt proposed cyber resilience framework to encompass and address vulnerabilities of these systems as well. Therefore, testing and validating its effectiveness in a hybrid microgrid would be a valuable area of exploration which significantly contribute to the overall security posture of modern distribution grid.

- Moreover, as the cyber threat security landscape is ever evolving, the proposed cyber attack resilient control framework for the *D-Systems* needs further enhancement to deal with hybrid attack model and unbounded attack scenarios while guarantee the asymptotic convergence of the control response.

By addressing these above areas of future research, it is possible to advance the state-of-the-art in cyber-attack resilience for power systems, ultimately ensuring the reliability, security, and resilience of critical infrastructure in the face of unprecedented cyber threats.

References

[1] Ahmed S. Musleh, Guo Chen, and Zhao Yang Dong. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 11(3):2218–2234, 2020. doi: 10.1109/TSG.2019.2949998.

[2] Xiaoge Huang, Zhijun Qin, and Hui Liu. A survey on power grid cyber security: From component-wise vulnerability assessment to system-wide impact analysis. *IEEE Access*, 6:69023–69035, 2018. doi: 10.1109/ACCESS.2018.2879996.

[3] Chih-Che Sun, Chen-Ching Liu, and Jing Xie. Cyber-physical system security of a power grid: State-of-the-art. *Electronics*, 5(3), 2016. ISSN 2079-9292. doi: 10.3390/electronics5030040.

[4] Haftu Tasew Reda, Adnan Anwar, and Abdun Mahmood. Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts. *Renewable and Sustainable Energy Reviews*, 163:112423, 2022. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2022.112423.

[5] Ics-cert year in review 2016, Nov 2016. URL https://www.cisa.gov/sites/default/files/Annual_Reports/Year_in_Review_FY2016_Final_S508C.pdf.

[6] Threat landscape for industrial automation systems statistics for h2 2020, Jan 2023. URL https://ics-cert.kaspersky.com/publications/reports/2021/03/25/threat-landscape-for-industrial-automation-systems-statistics-for-h2-2020.

[7] Kevin E. Hemsley and Dr. Ronald E. Fisher. History of industrial control system cyber incidents. *Idaho National Lab. (INL), Idaho Falls, ID (United States)*, 12 2018. doi: 10.2172/1505628. URL https://www.osti.gov/biblio/1505628.

[8] Utpal Bhaskar. India's power industry comes under increasing cyberattacks from hackers, Sep 2019. URL https://www.livemint.com/industry/energy/how-cyber-attacks-are-increasing-in-india-s-power-sector-1568107532851.html.

[9] Cert-in cyber incident reporting guidelines, Feb 2024. URL https://www2.deloitte.com/in/en/pages/risk/articles/CERT-IN-direction-for-reporting-cyber-incidents.html.

[10] Dou An, Feiye Zhang, Feifei Cui, and Qingyu Yang. Toward data integrity attacks against distributed dynamic state estimation in smart grid. *IEEE Transactions on Automation Science and Engineering*, 21(1):881–894, 2024. doi: 10.1109/TASE.2023.3236102.

[11] Patrick Wlazlo, Abhijeet Sahu, Zeyu Mao, Hao Huang, Ana Goulart, Katherine Davis, and Saman Zonouz. Man-in-the-middle attacks and defence in a power system cyber-physical testbed. *IET Cyber-Physical Systems: Theory & Applications*, 6(3): 164–177, 2021. doi: https://doi.org/10.1049/cps2.12014.

[12] Matthew Porter, Pedro Hespanhol, Anil Aswani, Matthew Johnson-Roberson, and Ramanarayan Vasudevan. Detecting generalized replay attacks via time-varying dynamic watermarking. *IEEE Transactions on Automatic Control*, 66(8):3502–3517, 2021. doi: 10.1109/TAC.2020.3022756.

[13] Sara Siamak, Maryam Dehghani, and Mohsen Mohammadi. Dynamic gps spoofing attack detection, localization, and measurement correction exploiting pmu and scada. *IEEE Systems Journal*, 15(2):2531–2540, 2021. doi: 10.1109/JSYST.2020.3001016.

[14] Elif Ustundag Soykan, Mustafa Bagriyanik, and Gurkan Soykan. Disrupting the power grid via ev charging: The impact of the sms phishing attacks. *Sustainable Energy, Grids and Networks*, 26:100477, 2021. ISSN 2352-4677. doi: https://doi.org/10.1016/j.segan.2021.100477.

[15] Wei Chen, Derui Ding, Hongli Dong, and Guoliang Wei. Distributed resilient filtering for power systems subject to denial-of-service attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(8):1688–1697, 2019. doi: 10.1109/TSMC.2019.2905253.

[16] Fei Tao and Dan Ye. Secure state estimation against eavesdropping attacks based on time-varying coding and noise-adding. *IEEE Transactions on Network Science and Engineering*, 11(1):174–184, 2024. doi: 10.1109/TNSE.2023.3293106.

[17] Tùng T. Kim and H. Vincent Poor. Strategic protection against data injection attacks on power grids. *IEEE Transactions on Smart Grid*, 2(2):326–333, 2011. doi: 10.1109/TSG.2011.2119336.

[18] Xueping Li, Yaokun Wang, and Zhigang Lu. Graph-based detection for false data injection attacks in power grid. *Energy*, 263:125865, 2023. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2022.125865. URL https://www.sciencedirect.com/science/article/pii/S0360544222027517.

[19] Xuan Liu and Zuyi Li. False data attacks against ac state estimation with incomplete network information. *IEEE Transactions on Smart Grid*, 8(5):2239–2248, 2017. doi: 10.1109/TSG.2016.2521178.

[20] Shiyu Jin. False data injection attack against smart power grid based on incomplete network information. *Electric Power Systems Research*, 230:110294, 2024. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2024.110294. URL https://www.sciencedirect.com/science/article/pii/S0378779624001822.

[21] Gabriela Hug and Joseph Andrew Giampapa. Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid*, 3(3):1362–1370, 2012. doi: 10.1109/TSG.2012.2195338.

[22] Junbo Zhao, Gexiang Zhang, Zhao Yang Dong, and Kit Po Wong. Forecasting-aided imperfect false data injection attacks against power system nonlinear state estimation. *IEEE Transactions on Smart Grid*, 7(1):6–8, 2016. doi: 10.1109/TSG.2015.2490603.

[23] Junbo Zhao, Lamine Mili, and Meng Wang. A generalized false data injection attacks against power system nonlinear state estimator and countermeasures. *IEEE Transactions on Power Systems*, 33(5):4868–4877, 2018. doi: 10.1109/TPWRS.2018.2794468.

[24] Mohsen Khalaf, Amr Youssef, and Ehab El-Saadany. Joint detection and mitigation of false data injection attacks in agc systems. *IEEE Transactions on Smart Grid*, 10(5):4985–4995, 2019. doi: 10.1109/TSG.2018.2872120.

[25] Jeong-Won Kang, Il-Young Joo, and Dae-Hyun Choi. False data injection attacks on contingency analysis: Attack strategies and impact assessment. *IEEE Access*, 6:8841–8851, 2018. doi: 10.1109/ACCESS.2018.2801861.

[26] Ali Tajer. False data injection attacks in electricity markets by limited adversaries: Stochastic robustness. *IEEE Transactions on Smart Grid*, 10(1):128–138, 2019. doi: 10.1109/TSG.2017.2733346.

[27] Ioannis Zografopoulos, Nikos D. Hatziargyriou, and Charalambos Konstantinou. Distributed energy resources cybersecurity outlook: Vulnerabilities, attacks, impacts, and mitigations. *IEEE Systems Journal*, 17(4):6695–6709, 2023. doi: 10.1109/JSYST.2023.3305757.

[28] Liang Che, Xuan Liu, Tao Ding, and Zuyi Li. Revealing impacts of cyber attacks on power grids vulnerability to cascading failures. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(6):1058–1062, 2019. doi: 10.1109/TCSII.2018.2869941.

[29] Neetesh Saxena, Leilei Xiong, Victor Chukwuka, and Santiago Grijalva. Impact evaluation of malicious control commands in cyber-physical smart grids. *IEEE Transactions on Sustainable Computing*, 6(2):208–220, 2021. doi: 10.1109/TSUSC.2018.2879670.

[30] Jichao Bi, Fengji Luo, Gaoqi Liang, Xiaofan Yang, Shibo He, and Zhao Yang Dong. Impact assessment and defense for smart grids with fdia against ami. *IEEE Transactions on Network Science and Engineering*, 10(2):578–591, 2023. doi: 10.1109/TNSE.2022.3197682.

[31] Yakup Koç, Martijn Warnier, Piet Van Mieghem, Robert E. Kooij, and Frances M.T. Brazier. The impact of the topology on cascading failures in a power grid model. *Physica A: Statistical Mechanics and its Applications*, 402:169–179, 2014. ISSN 0378-4371. doi: https://doi.org/10.1016/j.physa.2014.01.056.

[32] White House. Economic benefits of increasing electric grid resilience to weather outages. *U.S. Dept.Energy, Washington, DC, USA, Tech. Rep.,*, 2013.

[33] Banghua Xie, Changfan Li, Zili Wu, and Weiming Chen. Topological modeling research on the functional vulnerability of power grid under extreme weather. *Energies*, 14(16), 2021. ISSN 1996-1073. doi: 10.3390/en14165183.

[34] Nir Kshetri and Jeffrey Voas. Hacking power grids: A current problem. *Computer*, 50(12):91–95, 2017. doi: 10.1109/MC.2017.4451203.

[35] Gaoqi Liang, Steven R. Weller, Junhua Zhao, Fengji Luo, and Zhao Yang Dong. The 2015 ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, 32(4):3317–3318, 2017. doi: 10.1109/TPWRS.2016.2631891.

[36] Hang Zhang, Bo Liu, and Hongyu Wu. Smart grid cyber-physical attack and defense: A review. *IEEE Access*, 9:29641–29659, 2021. doi: 10.1109/ACCESS.2021.3058628.

[37] B.M. Ruhul Amin, Seyedfoad Taghizadeh, Md. Shihanur Rahman, Md. Jahangir Hossain, Vijay Varadharajan, and Zhiyong Chen. Cyber attacks in smart grid – dynamic impacts, analyses and recommendations. *IET Cyber-Physical Systems: Theory & Applications*, 5(4):321–329, 2020. doi: https://doi.org/10.1049/iet-cps.2019.0103.

[38] Min Zhou, Chensheng Liu, Amir Abiri Jahromi, Deepa Kundur, Jing Wu, and Chengnian Long. Revealing vulnerability of n-1 secure power systems to coordinated cyber-physical attacks. *IEEE Transactions on Power Systems*, 38(2):1044–1057, 2023. doi: 10.1109/TPWRS.2022.3169482.

[39] Kang Yan, Xuan Liu, Yidan Lu, and Fanglu Qin. A cyber-physical power system risk assessment model against cyberattacks. *IEEE Systems Journal*, 17(2):2018–2028, 2023. doi: 10.1109/JSYST.2022.3215591.

[40] Anurag Srivastava, Thomas Morris, Timothy Ernster, Ceeman Vellaithurai, Shengyi Pan, and Uttam Adhikari. Modeling cyber-physical vulnerability of the smart grid with incomplete information. *IEEE Transactions on Smart Grid*, 4(1):235–244, 2013. doi: 10.1109/TSG.2012.2232318.

[41] Timothy A. Ernster and Anurag K. Srivastava. Power system vulnerability analysis - towards validation of centrality measures. In *PES T & D 2012*, pages 1–6, 2012. doi: 10.1109/TDC.2012.6281483.

[42] Edward J. Oughton, Daniel Ralph, Raghav Pant, Eireann Leverett, Jennifer Copic, Scott Thacker, Rabia Dada, Simon Ruffle, Michelle Tuveson, and Jim W Hall. Stochastic counterfactual risk analysis for the vulnerability assessment of cyber-physical attacks on electricity distribution infrastructure networks. *Risk Analysis*, 39(9):2012–2031, 2019. doi: https://doi.org/10.1111/risa.13291.

[43] Yingmeng Xiang, Lingfeng Wang, and Yichi Zhang. Adequacy evaluation of electric power grids considering substation cyber vulnerabilities. *International Journal of Electrical Power & Energy Systems*, 96:368–379, 2018. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2017.10.004.

[44] Jia Guo, Yuqi Han, Chuangxin Guo, Fengdan Lou, and Yanbo Wang. Modeling and vulnerability analysis of cyber-physical power systems considering network topology and power flow properties. *Energies*, 10(1), 2017. ISSN 1996-1073. doi: 10.3390/en10010087.

[45] Wei Wang, Gregorio Cova, and Enrico Zio. A clustering-based framework for searching vulnerabilities in the operation dynamics of cyber-physical energy systems. *Reliability Engineering & System Safety*, 222:108400, 2022. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2022.108400.

[46] Sergio Gómez. *Centrality in Networks: Finding the Most Important Nodes*, pages 401–433. Springer International Publishing, 2019. ISBN 978-3-030-06222-4. doi: 10.1007/978-3-030-06222-4_8.

[47] Anuj Singh, Rishi Ranjan Singh, and S. R. S. Iyengar. Node-weighted centrality: a new way of centrality hybridization. *Computational Social Networks*, 7(1):6, Nov 2020. ISSN 2197-4314. doi: 10.1186/s40649-020-00081-w.

[48] Akrati Saxena and Sudarshan Iyengar. Centrality measures in complex networks: A survey. *CoRR*, abs/2011.07190, 2020. URL https://arxiv.org/abs/2011.07190.

[49] A. B. M. Nasiruzzaman, H. R. Pota, and M. A. Mahmud. Application of centrality measures of complex network framework in power grid. In *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, pages 4660–4665, 2011. doi: 10.1109/IECON.2011.6120079.

[50] Premananda Panigrahi and Somnath Maity. Topological analysis of power grid to identify vulnerable transmission lines and nodes, 2013.

[51] Xiaoguang Wei, Shibin Gao, Tao Huang, Ettore Bompard, Renjian Pi, and Tao Wang. Complex network-based cascading faults graph for the analysis of transmission network vulnerability. *IEEE Transactions on Industrial Informatics*, 15(3):1265–1276, 2019. doi: 10.1109/TII.2018.2840429.

[52] V. Rosato, S. Bologna, and F. Tiriticco. Topological properties of high-voltage electrical transmission networks. *Electric Power Systems Research*, 77(2):99–105, 2007. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2005.05.013.

[53] Gabriel J. Correa and José M. Yusta. Grid vulnerability analysis based on scale-free graphs versus power flow models. *Electric Power Systems Research*, 101:71–79, 2013. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2013.04.003.

[54] X. Liu and G. Joos. Transmission grid vulnerability assessment by eigen-sensitivity and cut-set screening. In *IEEE PES T&D 2010*, pages 1–8, 2010. doi: 10.1109/TDC.2010.5484204.

[55] Anurag K. SRIVASTAVA, Timothy A. ERNSTER, Ren LIU, and Vignesh G. KRISHNAN. Graph-theoretic algorithms for cyber-physical vulnerability analysis of power grid with incomplete information. *Journal of Modern Power Systems and Clean Energy*, 6(5):887–899, Sep 2018. ISSN 2196-5420. doi: 10.1007/s40565-018-0448-7.

[56] Ettore Bompard, Enrico Pons, and Di Wu. Extended topological metrics for the analysis of power grid vulnerability. *IEEE Systems Journal*, 6(3):481–487, 2012. doi: 10.1109/JSYST.2012.2190688.

[57] Tianlei Zang, Shibin Gao, Tao Huang, Xiaoguang Wei, and Tao Wang. Complex network-based transmission network vulnerability assessment using adjacent graphs. *IEEE Systems Journal*, 14(1):572–581, 2020. doi: 10.1109/JSYST.2019.2934317.

[58] Hale Cetinay, Karel Devriendt, and Piet Mieghem. Nodal vulnerability to targeted attacks in power grids. *Applied Network Science*, 3, 08 2018. doi: 10.1007/s41109-018-0089-9.

[59] Qingyu Yang, Dou An, Rui Min, Wei Yu, Xinyu Yang, and Wei Zhao. On optimal pmu placement-based defense against data integrity attacks in smart grid. *IEEE Transactions on Information Forensics and Security*, 12(7):1735–1750, 2017. doi: 10.1109/TIFS.2017.2686367.

[60] Sanjay Kumar, Barjeev Tyagi, Vishal Kumar, and Sunita Chohan. Optimization of phasor measurement units placement under contingency using reliability of network components. *IEEE Transactions on Instrumentation and Measurement*, 69(12):9893–9906, 2020. doi: 10.1109/TIM.2020.3004680.

[61] Jinping Hao, Robert J. Piechocki, Dritan Kaleshi, Woon Hau Chin, and Zhong Fan. Sparse malicious false data injection attacks and defense mechanisms in smart grids. *IEEE Transactions on Industrial Informatics*, 11(5):1–12, 2015. doi: 10.1109/TII.2015.2475695.

[62] Shamsun Nahar Edib, Yuzhang Lin, Vinod M. Vokkarane, Feng Qiu, Rui Yao, and Dongbo Zhao. Optimal pmu restoration for power system observability recovery

after massive attacks. *IEEE Transactions on Smart Grid*, 12(2):1565–1576, 2021. doi: 10.1109/TSG.2020.3028761.

[63] Mohammad Hossein Rezaeian and Saeid Esmaeili. Power system monitoring ensuring direct observation of critical buses and transmission lines using a bi-level approach. In *2016 Smart Grids Conference (SGC)*, pages 1–6, 2016. doi: 10.1109/SGC.2016.7882948.

[64] Mohammad Hossein Rezaeian Koochi and Mohammad Hasan Hemmatpour. A general pmu placement approach considering both topology and system aspects of contingencies. *International Journal of Electrical Power & Energy Systems*, 118:105774, 2020. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2019.105774.

[65] Chao Pei, Yang Xiao, Wei Liang, and Xiaojia Han. Pmu placement protection against coordinated false data injection attacks in smart grid. *IEEE Transactions on Industry Applications*, 56(4):4381–4393, 2020. doi: 10.1109/TIA.2020.2979793.

[66] Ahmed S. Musleh, Haris M. Khalid, S. M. Muyeen, and Ahmed Al-Durra. A prediction algorithm to enhance grid resilience toward cyber attacks in wamcs applications. *IEEE Systems Journal*, 13(1):710–719, 2019. doi: 10.1109/JSYST.2017.2741483.

[67] Aditya Ashok, Manimaran Govindarasu, and Jianhui Wang. Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid. *Proceedings of the IEEE*, 105(7):1389–1407, 2017. doi: 10.1109/JPROC.2017.2686394.

[68] Zakaria El Mrabet, Naima Kaabouch, Hassan El Ghazi, and Hamid El Ghazi. Cyber-security in smart grid: Survey and challenges. *Computers & Electrical Engineering*, 67:469–482, 2018. ISSN 0045-7906. doi: https://doi.org/10.1016/j.compeleceng.2018.01.015.

[69] Lei Su, Dan Ye, and Xin-Gang Zhao. Distributed secure state estimation for cyber-physical systems against replay attacks via multisensor method. *IEEE Systems Journal*, 16(4):5720–5728, 2022. doi: 10.1109/JSYST.2021.3123617.

[70] Jing Wang, Dongji Wang, Huaicheng Yan, and Hao Shen. Composite anti-disturbance $\mathcal{H}_\infty$ control for hidden markov jump systems with multi-sensor against replay attacks. *IEEE Transactions on Automatic Control*, pages 1–7, 2023. doi: 10.1109/TAC.2023.3326861.

[71] Bo Chen, Daniel W. C. Ho, Guoqiang Hu, and Li Yu. Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks. *IEEE Transactions on Cybernetics*, 48(6):1862–1876, 2018. doi: 10.1109/TCYB.2017.2716115.

[72] Dan Li, Nagi Gebraeel, and Kamran Paynabar. Detection and differentiation of replay attack and equipment faults in scada systems. *IEEE Transactions on Automation Science and Engineering*, 18(4):1626–1639, 2021. doi: 10.1109/TASE.2020.3013760.

[73] Minghui Zhu and Sonia Martínez. On the performance analysis of resilient networked control systems under replay attacks. *IEEE Transactions on Automatic Control*, 59(3):804–808, 2014. doi: 10.1109/TAC.2013.2279896.

[74] Ying Sun, Yamei Ju, Derui Ding, and Hongjian Liu. Distributed $h\infty$ filtering of replay attacks over sensor networks. *ISA Transactions*, 141:113–120, 2023. ISSN 0019-0578. doi: https://doi.org/10.1016/j.isatra.2023.04.018.

[75] Minal Chougule and Shreevardhan A Soman. Real-time data-assisted replay attack detection in wide-area protection system. *IET Generation, Transmission & Distribution*, 14(19):4021–4032, 2020. doi: https://doi.org/10.1049/iet-gtd.2020.0215.

[76] Kaustav Chatterjee and S. A. Khaparde. Data-driven online detection of replay attacks on wide-area measurement systems. In *2018 20th National Power Systems Conference (NPSC)*, pages 1–6, 2018. doi: 10.1109/NPSC.2018.8771807.

[77] Giuseppe Franzè, Francesco Tedesco, and Walter Lucia. Resilient control for cyber-physical systems subject to replay attacks. *IEEE Control Systems Letters*, 3(4):984–989, 2019. doi: 10.1109/LCSYS.2019.2920507.

[78] Paritosh Ramanan, Dan Li, and Nagi Gebraeel. Blockchain-based decentralized replay attack detection for large-scale power systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(8):4727–4739, 2022. doi: 10.1109/TSMC.2021.3104087.

[79] Thien-Toan Tran, Oh-Soon Shin, and Jong-Ho Lee. Detection of replay attacks in smart grid systems. In *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*, pages 298–302, 2013. doi: 10.1109/ComManTel.2013.6482409.

[80] Arunava Naha, André Teixeira, Anders Ahlén, and Subhrakanti Dey. Sequential detection of replay attacks. *IEEE Transactions on Automatic Control*, 68(3):1941–1948, 2023. doi: 10.1109/TAC.2022.3174004.

[81] Andreas Hoehn and Ping Zhang. Detection of replay attacks in cyber-physical systems. In *2016 American Control Conference (ACC)*, pages 290–295, 2016. doi: 10.1109/ACC.2016.7524930.

[82] Helem Sabina Sánchez, Damiano Rotondo, Teresa Escobet, Vicenç Puig, Jordi Saludes, and Joseba Quevedo. Detection of replay attacks in cyber-physical systems

using a frequency-based signature. *Journal of the Franklin Institute*, 356(5): 2798–2824, 2019. ISSN 0016-0032. doi: https://doi.org/10.1016/j.jfranklin.2019.01.005.

[83] Bharadwaj Satchidanandan and P. R. Kumar. Dynamic watermarking: Active defense of networked cyber–physical systems. *Proceedings of the IEEE*, 105(2): 219–240, 2017. doi: 10.1109/JPROC.2016.2575064.

[84] Yilin Mo, Sean Weerakkody, and Bruno Sinopoli. Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems Magazine*, 35(1):93–109, 2015. doi: 10.1109/MCS.2014.2364724.

[85] Riccardo M.G. Ferrari and André M.H. Teixeira. Detection and isolation of routing attacks through sensor watermarking. In *2017 American Control Conference (ACC)*, pages 5436–5442, 2017. doi: 10.23919/ACC.2017.7963800.

[86] Chongrong Fang, Yifei Qi, Peng Cheng, and Wei Xing Zheng. Optimal periodic watermarking schedule for replay attack detection in cyber–physical systems. *Automatica*, 112:108698, 2020. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2019.108698.

[87] Omid Palizban and Kimmo Kauhaniemi. Hierarchical control structure in microgrids with distributed generation: Island and grid-connected mode. *Renewable and Sustainable Energy Reviews*, 44:797–813, 2015. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2015.01.008.

[88] E.S.N. Raju P and Trapti Jain. Chapter 2 - distributed energy resources and control. In Rajeev Kumar Chauhan and Kalpana Chauhan, editors, *Distributed Energy Resources in Microgrids*, pages 33–56. Academic Press, 2019. ISBN 978-0-12-817774-7. doi: https://doi.org/10.1016/B978-0-12-817774-7.00002-8.

[89] Feixiong Chen, Minyou Chen, Qiang Li, Kaikai Meng, Josep M. Guerrero, and Derek Abbott. Multiagent-based reactive power sharing and control model for islanded microgrids. *IEEE Transactions on Sustainable Energy*, 7(3):1232–1244, 2016. doi: 10.1109/TSTE.2016.2539213.

[90] Subham Sahoo, Jimmy Chih-Hsien Peng, Sukumar Mishra, and Tomislav Dragičević. Distributed screening of hijacking attacks in dc microgrids. *IEEE Transactions on Power Electronics*, 35(7):7574–7582, 2020. doi: 10.1109/TPEL.2019.2957071.

[91] Subham Sahoo, Tomislav Dragičević, and Frede Blaabjerg. Cyber security in control of grid-tied power electronic converters—challenges and vulnerabilities. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(5):5326–5340, 2021. doi: 10.1109/JESTPE.2019.2953480.

[92] Subham Sahoo, Sukumar Mishra, Jimmy Chih-Hsien Peng, and Tomislav Dragičević. A stealth cyber-attack detection strategy for dc microgrids. *IEEE Transactions on Power Electronics*, 34(8):8162–8174, 2019. doi: 10.1109/TPEL.2018.2879886.

[93] Satabdy Jena, Narayana Prasad Padhy, and Josep M. Guerrero. Cyber-resilient cooperative control of dc microgrid clusters. *IEEE Systems Journal*, 16(2): 1996–2007, 2022. doi: 10.1109/JSYST.2021.3059445.

[94] Omar Ali Beg, Taylor T. Johnson, and Ali Davoudi. Detection of false-data injection attacks in cyber-physical dc microgrids. *IEEE Transactions on Industrial Informatics*, 13(5):2693–2703, 2017. doi: 10.1109/TII.2017.2656905.

[95] Subham Sahoo, Jimmy Chih-Hsien Peng, Annavaram Devakumar, Sukumar Mishra, and Tomislav Dragičević. On detection of false data in cooperative dc microgrids—a discordant element approach. *IEEE Transactions on Industrial Electronics*, 67(8): 6562–6571, 2020. doi: 10.1109/TIE.2019.2938497.

[96] Omar Ali Beg, Luan Viet Nguyen, Taylor T. Johnson, and Ali Davoudi. Cyber-physical anomaly detection in microgrids using time-frequency logic formalism. *IEEE Access*, 9:20012–20021, 2021. doi: 10.1109/ACCESS.2021.3055229.

[97] Omar Ali Beg, Luan V. Nguyen, Taylor T. Johnson, and Ali Davoudi. Signal temporal logic-based attack detection in dc microgrids. *IEEE Transactions on Smart Grid*, 10(4):3585–3595, 2019. doi: 10.1109/TSG.2018.2832544.

[98] Mohammad Sadegh Ghafoori and Jafar Soltani. Designing a robust cyber-attack detection and identification algorithm for dc microgrids based on kalman filter with unknown input observer. *IET Generation, Transmission & Distribution*, 16(16): 3230–3244, 2022. doi: https://doi.org/10.1049/gtd2.12517.

[99] Ahmed H. El-Ebiary, Mahmoud A. Attia, Fathy H. Awad, Mostafa I. Marei, and Mohamed Mokhtar. Kalman filters based distributed cyber-attack mitigation layers for dc microgrids. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 71(3):1358–1370, 2024. doi: 10.1109/TCSI.2023.3348928.

[100] Mohammad Reza Habibi, Subham Sahoo, Sebastián Rivera, Tomislav Dragičević, and Frede Blaabjerg. Decentralized coordinated cyberattack detection and mitigation strategy in dc microgrids based on artificial neural networks. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(4):4629–4638, 2021. doi: 10.1109/JESTPE.2021.3050851.

[101] Mohammad Reza Habibi, Hamid Reza Baghaee, Frede Blaabjerg, and Tomislav Dragičević. Secure control of dc microgrids for instant detection and mitigation of cyber-attacks based on artificial intelligence. *IEEE Systems Journal*, 16(2): 2580–2591, 2022. doi: 10.1109/JSYST.2021.3119355.

[102] Mohammad Reza Habibi, Hamid Reza Baghaee, Tomislav Dragičević, and Frede Blaabjerg. Detection of false data injection cyber-attacks in dc microgrids based on recurrent neural networks. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 9(5):5294–5310, 2021. doi: 10.1109/JESTPE.2020.2968243.

[103] Ali Jafarian Abianeh, Yihao Wan, Farzad Ferdowsi, Nenad Mijatovic, and Tomislav Dragičević. Vulnerability identification and remediation of fdi attacks in islanded dc microgrids using multiagent reinforcement learning. *IEEE Transactions on Power Electronics*, 37(6):6359–6370, 2022. doi: 10.1109/TPEL.2021.3132028.

[104] Yu Wang and Bikash C. Pal. Destabilizing attack and robust defense for inverter-based microgrids by adversarial deep reinforcement learning. *IEEE Transactions on Smart Grid*, 14(6):4839–4850, 2023. doi: 10.1109/TSG.2023.3263243.

[105] Yihao Wan and Tomislav Dragičević. Data-driven cyber-attack detection of intelligent attacks in islanded dc microgrids. *IEEE Transactions on Industrial Electronics*, 70(4):4293–4299, 2023. doi: 10.1109/TIE.2022.3176301.

[106] Hao Cui, Xiaorui Dong, Hongyan Deng, Moslem Dehghani, Khalid Alsubhi, and Hani Moaiteq Abdullah Aljahdali. Cyber attack detection process in sensor of dc micro-grids under electric vehicle based on hilbert–huang transform and deep learning. *IEEE Sensors Journal*, 21(14):15885–15894, 2021. doi: 10.1109/JSEN.2020.3027778.

[107] Moslem Dehghani, Taher Niknam, Mohammad Ghiasi, Navid Bayati, and Mehdi Savaghebi. Cyber-attack detection in dc microgrids based on deep machine learning and wavelet singular values approach. *Electronics*, 10(16), 2021. ISSN 2079-9292. doi: 10.3390/electronics10161914. URL https://www.mdpi.com/2079-9292/10/16/1914.

[108] Suman Rath, Diptak Pal, Parth Sarthi Sharma, and Bijaya Ketan Panigrahi. A cyber-secure distributed control architecture for autonomous ac microgrid. *IEEE Systems Journal*, 15(3):3324–3335, 2021. doi: 10.1109/JSYST.2020.3020968.

[109] Mohammed Masum Siraj Khan, Jairo A. Giraldo, and Masood Parvania. Attack detection in power distribution systems using a cyber-physical real-time reference model. *IEEE Transactions on Smart Grid*, 13(2):1490–1499, 2022. doi: 10.1109/TSG.2021.3128034.

[110] Ioannis Zografopoulos and Charalambos Konstantinou. Detection of malicious attacks in autonomous cyber-physical inverter-based microgrids. *IEEE Transactions on Industrial Informatics*, pages 1–1, 2021. doi: 10.1109/TII.2021.3132131.

[111] Ge Cao, Wei Gu, Guannan Lou, Wanxing Sheng, and Keyan Liu. Distributed synchronous detection for false data injection attack in cyber-physical microgrids.

International Journal of Electrical Power & Energy Systems, 137:107788, 2022. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2021.107788.

[112] Kirti Gupta, Subham Sahoo, Rabindra Mohanty, Bijaya Ketan Panigrahi, and Frede Blaabjerg. Distinguishing between cyber attacks and faults in power electronic systems—a noninvasive approach. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 11(2):1578–1588, 2023. doi: 10.1109/JESTPE.2022.3221867.

[113] Hamdi M. Albunashee, Chris Farnell, Andrew Suchanek, Kelby Haulmark, Roy A. McCann, Jia Di, and Alan Mantooth. A test bed for detecting false data injection attacks in systems with distributed energy resources. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 10(1):1303–1315, 2022. doi: 10.1109/JESTPE.2019.2948216.

[114] Anuoluwapo O. Aluko, Rudiren Pillay Carpanen, David G. Dorrell, and Evans E. Ojo. Real-time cyber attack detection scheme for standalone microgrids. *IEEE Internet of Things Journal*, 9(21):21481–21492, 2022. doi: 10.1109/JIOT.2022.3180939.

[115] Mohammad Reza Khalghani, Jignesh Solanki, Sarika Khushalani Solanki, Mohammad Hassan Khooban, and Arman Sargolzaei. Resilient frequency control design for microgrids under false data injection. *IEEE Transactions on Industrial Electronics*, 68(3):2151–2162, 2021. doi: 10.1109/TIE.2020.2975494.

[116] Aquib Mustafa, Binod Poudel, Ali Bidram, and Hamidreza Modares. Detection and mitigation of data manipulation attacks in ac microgrids. *IEEE Transactions on Smart Grid*, 11(3):2588–2603, 2020. doi: 10.1109/TSG.2019.2958014.

[117] Kebina Manandhar, Xiaojun Cao, Fei Hu, and Yao Liu. Detection of faults and attacks including false data injection attack in smart grid using kalman filter. *IEEE Transactions on Control of Network Systems*, 1(4):370–379, 2014. doi: 10.1109/TCNS.2014.2357531.

[118] Ruirui Liu, Hao Yu, and Fei Hao. Stochastic stealthy false data injection attacks against cyber-physical systems. *IEEE Systems Journal*, 16(4):6009–6020, 2022. doi: 10.1109/JSYST.2022.3171786.

[119] Shreyas Sundaram and Christoforos N. Hadjicostis. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control*, 56(7):1495–1508, 2011. doi: 10.1109/TAC.2010.2088690.

[120] Yulin Chen, Donglian Qi, Hangning Dong, Chaoyong Li, Zhenming Li, and Jianliang Zhang. A fdi attack-resilient distributed secondary control strategy for islanded microgrids. *IEEE Transactions on Smart Grid*, 12(3):1929–1938, 2021. doi: 10.1109/TSG.2020.3047949.

[121] Shan Zuo, Omar Ali Beg, Frank L. Lewis, and Ali Davoudi. Resilient networked ac microgrids under unbounded cyber attacks. *IEEE Transactions on Smart Grid*, 11 (5):3785–3794, 2020. doi: 10.1109/TSG.2020.2984266.

[122] Mahmood Jamali, Mahdieh S. Sadabadi, Masoud Davari, Subham Sahoo, and Frede Blaabjerg. Resilient cooperative secondary control of islanded ac microgrids utilizing inverter-based resources against state-dependent false data injection attacks. *IEEE Transactions on Industrial Electronics*, pages 1–12, 2023. doi: 10.1109/TIE.2023.3281698.

[123] Quan Zhou, Mohammad Shahidehpour, Ahmed Alabdulwahab, Abdullah Abusorrah, Liang Che, and Xuan Liu. Cross-layer distributed control strategy for cyber resilient microgrids. *IEEE Transactions on Smart Grid*, 12(5):3705–3717, 2021. doi: 10.1109/TSG.2021.3069331.

[124] Azwirman Gusrialdi, Zhihua Qu, and Marwan A. Simaan. Competitive interaction design of cooperative systems against attacks. *IEEE Transactions on Automatic Control*, 63(9):3159–3166, 2018. doi: 10.1109/TAC.2018.2793164.

[125] Zhiqiang Zuo, Xiong Cao, Yijing Wang, and Wentao Zhang. Resilient consensus of multiagent systems against denial-of-service attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(4):2664–2675, 2022. doi: 10.1109/TSMC.2021.3051730.

[126] Ali Bidram, Binod Poudel, Lakshmisree Damodaran, Rafael Fierro, and Josep M. Guerrero. Resilient and cybersecure distributed control of inverter-based islanded microgrids. *IEEE Transactions on Industrial Informatics*, 16(6):3881–3894, 2020. doi: 10.1109/TII.2019.2941748.

[127] Mahdieh S. Sadabadi, Subham Sahoo, and Frede Blaabjerg. A fully resilient cyber-secure synchronization strategy for ac microgrids. *IEEE Transactions on Power Electronics*, 36(12):13372–13378, 2021. doi: 10.1109/TPEL.2021.3091587.

[128] Xinyu Wang, Xiaoyuan Luo, Yuyan Zhang, and Xinping Guan. Detection and isolation of false data injection attacks in smart grids via nonlinear interval observer. *IEEE Internet of Things Journal*, 6(4):6498–6512, 2019. doi: 10.1109/JIOT.2019.2916670.

[129] Andres Intriago, Francesco Liberati, Nikos D. Hatziargyriou, and Charalambos Konstantinou. Residual-based detection of attacks in cyber-physical inverter-based microgrids. *IEEE Transactions on Power Systems*, 39(2):4020–4038, 2024. doi: 10.1109/TPWRS.2023.3286019.

[130] Rentao Lu, Jie Wang, and Ziqiang Wang. Distributed observer-based finite-time control of ac microgrid under attack. *IEEE Transactions on Smart Grid*, 12(1): 157–168, 2021. doi: 10.1109/TSG.2020.3017793.

[131] Shankar Abhinav, Hamidreza Modares, Frank L. Lewis, Frank Ferrese, and Ali Davoudi. Synchrony in networked microgrids under attacks. *IEEE Transactions on Smart Grid*, 9(6):6731–6741, 2018. doi: 10.1109/TSG.2017.2721382.

[132] Mengxuan Shi, Xia Chen, Mohammad Shahidehpour, Quan Zhou, and Jinyu Wen. Observer-based resilient integrated distributed control against cyberattacks on sensors and actuators in islanded ac microgrids. *IEEE Transactions on Smart Grid*, 12(3):1953–1963, 2021. doi: 10.1109/TSG.2021.3050203.

[133] Amir Afshari, Mehdi Karrari, Hamid Reza Baghaee, and G. B. Gharehpetian. Resilient synchronization of voltage/frequency in ac microgrids under deception attacks. *IEEE Systems Journal*, 15(2):2125–2136, 2021. doi: 10.1109/JSYST.2020.2992309.

[134] Ali Jafarian Abianeh, Mohammad Mehdi Mardani, Farzad Ferdowsi, Raju Gottumukkala, and Tomislav Dragičević. Cyber-resilient sliding-mode consensus secondary control scheme for islanded ac microgrids. *IEEE Transactions on Power Electronics*, 37(5):6074–6089, 2022. doi: 10.1109/TPEL.2021.3125985.

[135] Yajie Jiang, Yun Yang, Siew-Chong Tan, and Shu Yuen Hui. Distributed sliding mode observer-based secondary control for dc microgrids under cyber-attacks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 11(1):144–154, 2021. doi: 10.1109/JETCAS.2020.3046781.

[136] Yun Liu, Huanhai Xin, Zhihua Qu, and Deqiang Gan. An attack-resilient cooperative control strategy of multiple distributed generators in distribution networks. *IEEE Transactions on Smart Grid*, 7(6):2923–2932, 2016. doi: 10.1109/TSG.2016.2542111.

[137] Quan Zhou, Mohammad Shahidehpour, Ahmed Alabdulwahab, and Abdullah Abusorrah. A cyber-attack resilient distributed control strategy in islanded microgrids. *IEEE Transactions on Smart Grid*, 11(5):3690–3701, 2020. doi: 10.1109/TSG.2020.2979160.

[138] Alexander Julian Gallo, Mustafa Sahin Turan, Francesca Boem, Thomas Parisini, and Giancarlo Ferrari-Trecate. A distributed cyber-attack detection scheme with application to dc microgrids. *IEEE Transactions on Automatic Control*, 65(9): 3800–3815, 2020. doi: 10.1109/TAC.2020.2982577.

[139] Ying Wan, Guanghui Wen, Xinghuo Yu, and Jürgen Kurths. Distributed event-based resilient secondary control for ac microgrids: A trust-reputation approach. *IEEE Transactions on Smart Grid*, 15(2):2116–2128, 2024. doi: 10.1109/TSG.2023.3302902.

[140] Mingyu Huang, Li Ding, Zhi-Wei Liu, Yuan Ge, and Zhi-Hong Guan. Resilient secondary control of islanded ac microgrid under corrupted measurements. *Electric Power Systems Research*, 221:109428, 2023. ISSN 0378-7796. doi: https://doi.

org/10.1016/j.epsr.2023.109428. URL https://www.sciencedirect.com/science/article/pii/S0378779623003176.

[141] Subham Sahoo, Yongheng Yang, and Frede Blaabjerg. Resilient synchronization strategy for ac microgrids under cyber attacks. *IEEE Transactions on Power Electronics*, 36(1):73–77, 2021. doi: 10.1109/TPEL.2020.3005208.

[142] Yu Wang, Suman Mondal, Chao Deng, Kuntal Satpathi, Yan Xu, and Souvik Dasgupta. Cyber-resilient cooperative control of bidirectional interlinking converters in networked ac/dc microgrids. *IEEE Transactions on Industrial Electronics*, 68(10): 9707–9718, 2021. doi: 10.1109/TIE.2020.3020033.

[143] Jianzhe Liu, Xiaonan Lu, and Jianhui Wang. Resilience analysis of dc microgrids under denial of service threats. *IEEE Transactions on Power Systems*, 34(4): 3199–3208, 2019. doi: 10.1109/TPWRS.2019.2897499.

[144] Songlin Hu, Fuyi Yang, Sergey Gorbachev, Dong Yue, Victor Kuzin, and Chao Deng. Resilient control design for networked dc microgrids under time-constrained dos attacks. *ISA Transactions*, 127:197–205, 2022. ISSN 0019-0578. doi: https://doi.org/10.1016/j.isatra.2022.02.022.

[145] Binod P. Poudel, Aquib Mustafa, Ali Bidram, and Hamidreza Modares. Detection and mitigation of cyber-threats in the dc microgrid distributed control system. *International Journal of Electrical Power and Energy Systems*, 120:105968, 2020. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2020.105968.

[146] Sheik M. Mohiuddin, Junjian Qi, Sasha Fung, Yu Huang, and Yufei Tang. Deep learning based multi-label attack detection for distributed control of ac microgrids. In *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 233–238, 2021. doi: 10.1109/SmartGridComm51999.2021.9631998.

[147] Ahmed S. Musleh, Guo Chen, Zhao Yang Dong, Chen Wang, and Shiping Chen. Spatio-temporal data-driven detection of false data injection attacks in power distribution systems. *International Journal of Electrical Power & Energy Systems*, 145:108612, 2023. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2022.108612.

[148] Lina Sheng, Wei Gu, and Ge Cao. Distributed detection mechanism and resilient consensus strategy for secure voltage control of ac microgrids. *CSEE Journal of Power and Energy Systems*, 9(3):1066–1077, 2023. doi: 10.17775/CSEEJPES.2020.07140.

[149] José Antonio Cebollero, David Cañete, Susana Martín-Arroyo, Miguel García-Gracia, and Helder Leite. A survey of islanding detection methods

for microgrids and assessment of non-detection zones in comparison with grid codes. *Energies*, 15(2), 2022. ISSN 1996-1073. doi: 10.3390/en15020460.

[150] Wilsun Xu, Guibin Zhang, Chun Li, Wencong Wang, Guangzhu Wang, and Jacek Kliber. A power line signaling based technique for anti-islanding protection of distributed generators—part i: Scheme and analysis. *IEEE Transactions on Power Delivery*, 22(3):1758–1766, 2007. doi: 10.1109/TPWRD.2007.899618.

[151] Yuwei SHANG, Shenxing SHI, and Xinzhou DONG. Islanding detection based on asymmetric tripping of feeder circuit breaker in ungrounded power distribution system. *Journal of Modern Power Systems and Clean Energy*, 3(4):526–532, Dec 2015. doi: 10.1007/s40565-015-0162-7.

[152] Brian Dob and Craig Palmer. Communications assisted islanding detection: Contrasting direct transfer trip and phase comparison methods. In *2018 71st Annual Conference for Protective Relay Engineers (CPRE)*, pages 1–6, 2018. doi: 10.1109/CPRE.2018.8349783.

[153] Ali Rostami, Amin Jalilian, Sasan Zabihi, Javad Olamaei, and Edris Pouresmaeil. Islanding detection of distributed generation based on parallel inductive impedance switching. *IEEE Systems Journal*, 14(1):813–823, 2020. doi: 10.1109/JSYST.2019.2923289.

[154] C.N. Papadimitriou, V.A. Kleftakis, and N.D. Hatziargyriou. A novel islanding detection method for microgrids based on variable impedance insertion. *Electric Power Systems Research*, 121:58–66, 2015. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2014.12.004.

[155] P. K. Ganivada and P. Jena. Active slip frequency based islanding detection technique for grid-tied inverters. *IEEE Transactions on Industrial Informatics*, 16 (7):4615–4626, 2020. doi: 10.1109/TII.2019.2949009.

[156] Reza Bakhshi-Jafarabadi, Javad Sadeh, and Marjan Popov. Maximum power point tracking injection method for islanding detection of grid-connected photovoltaic systems in microgrid. *IEEE Transactions on Power Delivery*, 36(1):168–179, 2021. doi: 10.1109/TPWRD.2020.2976739.

[157] Wen Cai, Bangyin Liu, Shanxu Duan, and Changyue Zou. An islanding detection method based on dual-frequency harmonic current injection under grid impedance unbalanced condition. *IEEE Transactions on Industrial Informatics*, 9 (2):1178–1187, 2013. doi: 10.1109/TII.2012.2209669.

[158] David Reigosa, Fernando Briz, Cristian Blanco, Pablo García, and Juan Manuel Guerrero. Active islanding detection for multiple parallel- connected inverter-based distributed generators using high frequency signal injection. In *2012 IEEE Energy*

Conversion Congress and Exposition (ECCE), pages 2719–2726, 2012. doi: 10.1109/ECCE.2012.6342534.

[159] B. Wen, D. Boroyevich, R. Burgos, Z. Shen, and P. Mattavelli. Impedance-based analysis of active frequency drift islanding detection for grid-tied inverter system. *IEEE Transactions on Industry Applications*, 52(1):332–341, Jan 2016. ISSN 0093-9994. doi: 10.1109/TIA.2015.2480847.

[160] H. Vahedi, M. Karrari, and G. B. Gharehpetian. Accurate sfs parameter design criterion for inverter-based distributed generation. *IEEE Transactions on Power Delivery*, 31(3):1050–1059, June 2016. ISSN 0885-8977. doi: 10.1109/TPWRD.2015.2391193.

[161] H. Muda and P. Jena. Phase angle-based PC technique for islanding detection of distributed generations. *IET Renewable Power Generation*, 12(6):735–746, 2018. ISSN 1752-1416. doi: 10.1049/iet-rpg.2017.0089.

[162] H. Pourbabak and A. Kazemi. Islanding detection method based on a new approach to voltage phase angle of constant power inverters. *IET Generation, Transmission Distribution*, 10(5):1190–1198, 2016. ISSN 1751-8687. doi: 10.1049/iet-gtd.2015.0776.

[163] Dionisis Voglitsis, Nick Peter Papanikolaou, and Anastasios C. Kyritsis. Active cross-correlation anti-islanding scheme for pv module-integrated converters in the prospect of high penetration levels and weak grid conditions. *IEEE Transactions on Power Electronics*, 34(3):2258–2274, 2019. doi: 10.1109/TPEL.2018.2836663.

[164] Yogesh M. Makwana, Bhavesh R. Bhalja, and Ramakrishna Gokaraju. Auto-correlation-based islanding detection technique verified through hardware-in-loop testing. *IET Generation, Transmission & Distribution*, 13 (17):3792–3802, 2019. doi: https://doi.org/10.1049/iet-gtd.2018.6370.

[165] Houshang Karimi, Amirnaser Yazdani, and Reza Iravani. Negative-sequence current injection for fast islanding detection of a distributed resource unit. *IEEE Transactions on Power Electronics*, 23(1):298–307, 2008. doi: 10.1109/TPEL.2007.911774.

[166] Hamid Reza Baghaee, Dragan Mlakić, Srete Nikolovski, and Tomislav Dragicević. Support vector machine-based islanding and grid fault detection in active distribution networks. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 8(3):2385–2403, 2020. doi: 10.1109/JESTPE.2019.2916621.

[167] Biljana Matic-Cuka and Mladen Kezunovic. Islanding detection for inverter-based distributed generation using support vector machine method. *IEEE Transactions on Smart Grid*, 5(6):2676–2686, 2014. doi: 10.1109/TSG.2014.2338736.

[168] Q. Cui, K. El-Arroudi, and G. Joós. Islanding detection of hybrid distributed generation under reduced non-detection zone. *IEEE Transactions on Smart Grid*, 9 (5):5027–5037, Sept 2018. ISSN 1949-3053. doi: 10.1109/TSG.2017.2679101.

[169] H. T. Do, X. Zhang, N. V. Nguyen, S. S. Li, and T. T. Chu. Passive-islanding detection method using the wavelet packet transform in grid-connected photovoltaic systems. *IEEE Transactions on Power Electronics*, 31(10):6955–6967, Oct 2016. ISSN 0885-8993. doi: 10.1109/TPEL.2015.2506464.

[170] V.L. Merlin, R.C. Santos, A.P. Grilo, J.C.M. Vieira, D.V. Coury, and M. Oleskovicz. A new artificial neural network based method for islanding detection of distributed generators. *International Journal of Electrical Power & Energy Systems*, 75:139–151, 2016. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2015.08.016.

[171] Dragan Mlakić, Hamid Reza Baghaee, and Srete Nikolovski. A novel anfis-based islanding detection for inverter-interfaced microgrids. *IEEE Transactions on Smart Grid*, 10(4):4411–4424, 2019. doi: 10.1109/TSG.2018.2859360.

[172] Xiaolong Chen and Yongli Li. An islanding detection algorithm for inverter-based distributed generation based on reactive power control. *IEEE Transactions on Power Electronics*, 29(9):4672–4683, 2014. doi: 10.1109/TPEL.2013.2284236.

[173] Xing Xie, Chun Huang, and Danni Li. A new passive islanding detection approach considering the dynamic behavior of load in microgrid. *International Journal of Electrical Power & Energy Systems*, 117:105619, 2020. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2019.105619.

[174] M. R. Alam, M. T. Ara Begum, and K. M. Muttaqi. Assessing the performance of rocof relay for anti-islanding protection of distributed generation under subcritical region of power imbalance. In *2018 IEEE Industry Applications Society Annual Meeting (IAS)*, pages 1–8, 2018. doi: 10.1109/IAS.2018.8544467.

[175] H. Samet, F. Hashemi, and T. Ghanbari. Islanding detection method for inverter-based distributed generation with negligible non-detection zone using energy of rate of change of voltage. *IET Generation, Transmission Distribution*, 9(15): 2337–2350, 2015. ISSN 1751-8687. doi: 10.1049/iet-gtd.2015.0638.

[176] M. A. A. Farhan and K. S. Swarup. Islanding detection using mathematical morphology for distributed generation. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6, 2017. doi: 10.1109/ISGTEurope.2017.8260173.

[177] Yogesh M. Makwana and Bhavesh R. Bhalja. Experimental performance of an islanding detection scheme based on modal components. *IEEE Transactions on Smart Grid*, 10(1):1025–1035, 2019. doi: 10.1109/TSG.2017.2757599.

[178] Ruchita Nale, Monalisa Biswal, and Nand Kishor. A transient component based approach for islanding detection in distributed generation. *IEEE Transactions on Sustainable Energy*, 10(3):1129–1138, 2019. doi: 10.1109/TSTE.2018.2861883.

[179] A. Samui and S. R. Samantaray. Assessment of rocpad relay for islanding detection in distributed generation. *IEEE Transactions on Smart Grid*, 2(2):391–398, June 2011. ISSN 1949-3053. doi: 10.1109/TSG.2011.2125804.

[180] Samuel Admasie, Syed Basit Ali Bukhari, Raza Haider, Teke Gush, and Chul-Hwan Kim. A passive islanding detection scheme using variational mode decomposition-based mode singular entropy for integrated microgrids. *Electric Power Systems Research*, 177:105983, 2019. ISSN 0378-7796.

[181] G. P. Kumar and P. Jena. Pearson's correlation coefficient for islanding detection using micro-pmu measurements. *IEEE Systems Journal*, pages 1–12, 2020. doi: 10.1109/JSYST.2020.3021922.

[182] S. Dhar and P. K. Dash. Harmonic profile injection-based hybrid active islanding detection technique for pv-vsc-based microgrid system. *IEEE Transactions on Sustainable Energy*, 7(4):1473–1481, 2016. doi: 10.1109/TSTE.2016.2515158.

[183] M. Khodaparastan, H. Vahedi, F. Khazaeli, and H. Oraee. A novel hybrid islanding detection method for inverter-based dgs using sfs and rocof. *IEEE Transactions on Power Delivery*, 32(5):2162–2170, Oct 2017. ISSN 0885-8977. doi: 10.1109/TPWRD.2015.2406577.

[184] D. Mlakić, H. R. Baghaee, and S. Nikolovski. Gibbs phenomenon-based hybrid islanding detection strategy for vsc-based microgrids using frequency shift, thd_U , and rms_U. *IEEE Transactions on Smart Grid*, 10(5):5479–5491, 2019. doi: 10.1109/TSG.2018.2883595.

[185] X. Chen, Y. Li, and P. Crossley. A novel hybrid islanding detection method for grid-connected microgrids with multiple inverter-based distributed generators based on adaptive reactive power disturbance and passive criteria. *IEEE Transactions on Power Electronics*, 34(9):9342–9356, 2019. doi: 10.1109/TPEL.2018.2886930.

[186] Masoumeh Seyedi, Seyed Abbas Taher, Babak Ganji, and Josep Guerrero. A hybrid islanding detection method based on the rates of changes in voltage and active power for the multi-inverter systems. *IEEE Transactions on Smart Grid*, 12(4):2800–2811, 2021. doi: 10.1109/TSG.2021.3061567.

[187] S. R. Mohanty, N. Kishor, P. K. Ray, and J. P. S. Catalo. Comparative study of advanced signal processing techniques for islanding detection in a hybrid distributed generation system. *IEEE Transactions on Sustainable Energy*, 6(1):122–131, 2015. doi: 10.1109/TSTE.2014.2362797.

[188] A. Khamis, Y. Xu, Z. Y. Dong, and R. Zhang. Faster detection of microgrid islanding events using an adaptive ensemble classifier. *IEEE Transactions on Smart Grid*, 9 (3):1889–1899, May 2018. ISSN 1949-3053. doi: 10.1109/TSG.2016.2601656.

[189] Apoorva Shukla, Soham Dutta, Sourav Kumar Sahu, and Pradip Kumar Sadhu. A narrative perspective of island detection methods under the lens of cyber-attack in data-driven smart grid. *Journal of Electrical Systems and Information Technology*, 10(1):14, Mar 2023. ISSN 2314-7172. doi: 10.1186/s43067-023-00083-4.

[190] Yashasvi Bansal and Ranjana Sodhi. A novel μpmus assisted loss-of-mains detection technique for active distribution systems. *Electric Power Systems Research*, 202: 107578, 2022. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2021.107578.

[191] Soham Dutta, Maddikara Jaya Bharata Reddy, Dusmanta Kumar Mohanta, Makrand Sing Kushwah, and Pradip Kumar Sadhu. μpmu-based intelligent island detection – the first crucial step toward enhancing grid resilience with mg. *IET Smart Grid*, 3(2):162–173, 2020. doi: https://doi.org/10.1049/iet-stg.2019.0161.

[192] Apoorva Shukla, Soham Dutta, and Pradip Kumar Sadhu. An island detection approach by μ-pmu with reduced chances of cyber attack. *International Journal of Electrical Power & Energy Systems*, 126:106599, 2021. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2020.106599.

[193] Apoorva Shukla, Soham Dutta, Pradip Kumar Sadhu, and Bishwajit Dey. An island detection methodology with protection against cyber attack. *Microsystem Technologies*, Jan 2024. ISSN 1432-1858. doi: 10.1007/s00542-023-05596-6.

[194] Ruchita Nale, Monalisa Biswal, and Nand Kishor. A passive communication based islanding detection technique for ac microgrid. *International Journal of Electrical Power & Energy Systems*, 137:107657, 2022. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2021.107657.

[195] Annarita Giani, Russell Bent, and Feng Pan. Phasor measurement unit selection for unobservable electric power data integrity attack detection. *International Journal of Critical Infrastructure Protection*, 7(3):155–164, 2014. ISSN 1874-5482. doi: https://doi.org/10.1016/j.ijcip.2014.06.001.

[196] Shaik Mullapathi Farooq, S. M. Suhail Hussain, Siddavaram Kiran, and Taha Selim Ustun. Certificate based authentication mechanism for pmu communication networks based on iec 61850-90-5. *Electronics*, 7(12), 2018. ISSN 2079-9292. doi: 10.3390/electronics7120370.

[197] Hui Lin, Chen Chen, Jianhui Wang, Junjian Qi, Dong Jin, Zbigniew T. Kalbarczyk, and Ravishankar K. Iyer. Self-healing attack-resilient pmu network for power system operation. *IEEE Transactions on Smart Grid*, 9(3):1551–1565, 2018. doi: 10.1109/TSG.2016.2593021.

[198] Bei Gou. Optimal placement of pmus by integer linear programming. *IEEE Transactions on Power Systems*, 23(3):1525–1526, 2008. doi: 10.1109/TPWRS.2008.926723.

[199] Zhenhua Wang, Haibo He, Zhiqiang Wan, and Yan Sun. Detection of false data injection attacks in ac state estimation using phasor measurements. *IEEE Transactions on Smart Grid*, pages 1–1, 2020. doi: 10.1109/TSG.2020.2972781.

[200] Shiwen Sun, Xiaoxiao Liu, Li Wang, and Chengyi Xia. New link attack strategies of complex networks based on k-core decomposition. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(12):3157–3161, 2020. doi: 10.1109/TCSII.2020.2973668.

[201] Ranjana Sodhi, S. C. Srivastava, and S. N. Singh. Optimal pmu placement to ensure system observability under contingencies. In *2009 IEEE Power Energy Society General Meeting*, pages 1–6, 2009. doi: 10.1109/PES.2009.5275618.

[202] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven R. Weller, and Zhao Yang Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 2017. doi: 10.1109/TSG.2015.2495133.

[203] Diyun Huang. *Dynamic PTDF Implementation in the Market Model*. PhD thesis, 2011.

[204] Junbo Zhao, Gexiang Zhang, and Rabih A. Jabr. Robust detection of cyber attacks on state estimators using phasor measurements. *IEEE Transactions on Power Systems*, 32(3):2468–2470, 2017. doi: 10.1109/TPWRS.2016.2603447.

[205] A. Abur and A.G. Expósito. *Power System State Estimation: Theory and Implementation*. Power Engineering (Willis). CRC Press, 2004. ISBN 9780203913673.

[206] Ranjana Sodhi, S. C. Srivastava, and S. N. Singh. Phasor-assisted hybrid state estimator. *Electric Power Components and Systems*, 38(5):533–544, 2010. doi: 10.1080/15325000903376925.

[207] Sourav De and Ranjana Sodhi. A pmu assisted cyber attack resilient framework against power systems structural vulnerabilities. *Electric Power Systems Research*, 206:107805, 2022. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2022.107805.

[208] Ali Bidram, Frank L. Lewis, and Ali Davoudi. Distributed control systems for small-scale power networks: Using multiagent cooperative control theory. *IEEE Control Systems Magazine*, 34(6):56–77, 2014. doi: 10.1109/MCS.2014.2350571.

[209] Ali Bidram, Ali Davoudi, and Frank L. Lewis. A multiobjective distributed control framework for islanded ac microgrids. *IEEE Transactions on Industrial Informatics*, 10(3):1785–1798, 2014. doi: 10.1109/TII.2014.2326917.

[210] Ali Bidram, Ali Davoudi, Frank L. Lewis, and Zhihua Qu. Secondary control of microgrids based on distributed cooperative control of multi-agent systems. *IET Generation, Transmission & Distribution*, 7(8):822–831, 2013. doi: https://doi.org/10.1049/iet-gtd.2012.0576.

[211] Alex Smola, Arthur Gretton, Le Song, and Bernhard" Schölkopf. A hilbert space embedding for distributions. In *Algorithmic Learning Theory*, pages 13–31, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-75225-7.

[212] Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012.

[213] Yibin Li, Yan Song, Lei Jia, Shengyao Gao, Qiqiang Li, and Meikang Qiu. Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning. *IEEE Transactions on Industrial Informatics*, 17 (4):2833–2841, 2021. doi: 10.1109/TII.2020.3008010.

[214] Ndaga Mwakabuta and Arun Sekar. Comparative study of the ieee 34 node test feeder under practical simplifications. In *2007 39th North American Power Symposium*, pages 484–491, 2007. doi: 10.1109/NAPS.2007.4402354.

[215] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 785–794. Association for Computing Machinery, 2016. ISBN 9781450342322. doi: 10.1145/2939672.2939785.

[216] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: with Applications in R*. Springer New York, NY, 2013. ISBN 978-1-4614-7138-7.

[217] André Teixeira, Iman Shames, Henrik Sandberg, and Karl H. Johansson. Distributed fault detection and isolation resilient to network model uncertainties. *IEEE Transactions on Cybernetics*, 44(11):2024–2037, 2014. doi: 10.1109/TCYB.2014.2350335.

[218] J. Chen and R.J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. The International Series on Asian Studies in Computer and Information Science. Springer US, 2012. ISBN 9781461551492. URL https://books.google.co.in/books?id=_wvrBwAAQBAJ.

[219] Deyin Yao, Hongyi Li, and Yang Shi. Adaptive event-triggered sliding-mode control for consensus tracking of nonlinear multiagent systems with unknown perturbations. *IEEE Transactions on Cybernetics*, pages 1–13, 2022. doi: 10.1109/TCYB.2022.3172127.

[220] Jiahu Qin, Gaosheng Zhang, Wei Xing Zheng, and Yu Kang. Adaptive sliding mode consensus tracking for second-order nonlinear multiagent systems with actuator faults. *IEEE Transactions on Cybernetics*, 49(5):1605–1615, 2019. doi: 10.1109/TCYB.2018.2805167.

[221] Kaishun Xiahou, Yang Liu, and Q. H. Wu. Decentralized detection and mitigation of multiple false data injection attacks in multiarea power systems. *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, 3(1):101–112, 2022. doi: 10.1109/JESTIE.2021.3112919.

[222] Reynaldo Salcedo, Edward Corbett, Christopher Smith, and et al. Banshee distribution network benchmark and prototyping platform for hardware-in-the-loop integration of microgrid and device controllers. *The Journal of Engineering*, 2019 (8):5365–5373, 2019. doi: 10.1049/joe.2018.5174.

[223] Ali Rostami, Hamdi Abdi, Mansour Moradi, Javad Olamaei, and Ehsan Naderi. Islanding detection based on rocov and rocorp parameters in the presence of synchronous dg applying the capacitor connection strategy. *Electric Power Components and Systems*, 45(3):315–330, 2017. doi: 10.1080/15325008.2016.1250842.

[224] Mohsen Tajdinian, Hasan Khosravi, Haidar Samet, and Ziad M. Ali. Islanding detection scheme using potential energy function based criterion. *Electric Power Systems Research*, 209:108047, 2022. ISSN 0378-7796. doi: https://doi.org/10.1016/j.epsr.2022.108047.

[225] Masoumeh Seyedi, Seyed Abbas Taher, Babak Ganji, and Josep M. Guerrero. A hybrid islanding detection technique for inverter-based distributed generator units. *International Transactions on Electrical Energy Systems*, 29(11):e12113, 2019. doi: https://doi.org/10.1002/2050-7038.12113.

Chapter A

Test System Data

A.1 Modified IEEE 13-Bus Distribution Network

The IEEE 13-Bus system is a multi-phase radial distribution network, considered as a standard reference model in various power distribution studies, is basically characterized by short transmission lines, unbalanced structure and highly loaded feeders containing multi phase laterals with distributed and spot loads. The system has a nominal frequency and line-to-line voltage of 60 Hz and 4.16 kV respectively. The total active, reactive and apparent power of the system are 3.466 MW, 2.102 MVAr (Inductive), 0.7 MVAr (Capacitive)and 3.739 MVA respectively with power factor of 0.927 For the shake of simplicity in modeling and developing an attack resilient steady state control mechanism in Chapter - 5, this test system is modeled in RSCAD software of NovaCoR Real-Time Digital Simulator, following certain customization's in its structure to make it a balanced three phase distribution feeder. The major modifications and use of simplifying assumptions considered for modeling the modified version of the conventional IEEE 13-Bus radial network into a balanced standard test system are as follows:

- The utility is removed as the whole network is designed to be operated in islanded mode.
- Four Solar Photovoltaic (PV)-based grid forming DER units of equal one per unit active power and voltage rating are connected through a 1.0/0.5 MVA, 0.48/4.16 kV Yg-Yg transformer at 4 distinct locations i.e Node-650, Node-633, Node-671 and Node-692 respectively. The modified single diagram of the test feeder is shown in Fig. A.1.
- The IEEE 13-Node test feeder undergoes through the following simplified assumptions.
    - All laterals are transformed into the three phase section from whence they originate.
    - It is assumed that the self and mutual impedances of phases are equal to the averages of their respective self and mutual impedances, transposing the three-phase sections used to mitigate the unequal distribution of electromagnetic forces and impedance in overhead transmission lines. Thereafter, for each section, positive and zero sequence impedances are

computed as follows.

$$Z_{abc} = \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} \tag{A.1}$$

To transpose the system and compute the average self and mutual impedances, while taking into account the distributed nature of the line and the effect of ground, the modified Carson and Kron reduction equations are used as:

$$Z_s = \frac{1}{3}[Z_{aa} + Z_{bb} + Z_{cc}] \tag{A.2}$$

$$Z_m = \frac{1}{3}[Z_{ab} + Z_{bc} + Z_{ca}] \tag{A.3}$$

Therefore, the positive and zero sequence impedance are calculated as:

$$Z_{11} = \quad Z_s - Z_m \tag{A.4}$$

$$Z_{00} = \quad Z_s + 2Z_m \tag{A.5}$$

– To get the overall three phase balanced loads, the unbalanced phase loads in each of the three phase sections are added together.

- Tables A.1-A.4 provides component details for the IEEE 13-Bus distribution network, including transformers, spot and distributed loads, line segments and capacitor banks, and line configurations (impedance matrix as symmetric).

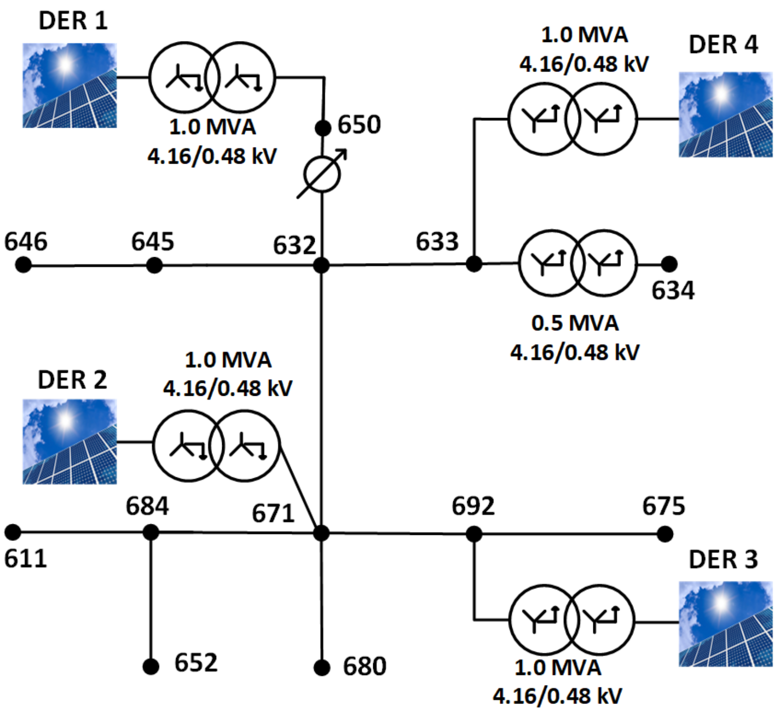


Figure A.1: Modified IEEE 13-Bus distribution feeder network with PV DERs

Table A.1: Overhead and underground line configuration data

Overhead Line Configurations				
Config.	Phasing	Phase	Neutral	Spacing
		ACSR	ACSR	ID
601	B A C N	556,500 26/7	4/0 6/1	500
602	C A B N	4/0 6/1	4/0 6/1	500
603	C B N	1/0	1/0	505
604	A C N	1/0	1/0	505
605	C N	1/0	1/0	510
Underground Line Configurations				
Config.	Phasing	Cable	Neutral	Space ID
606	A B C N	250,000 AA, CN	None	515
607	A N	1/0 AA, TS	1/0 Cu	520

Table A.2: Transformer details

	kVA	kV-high	kV-low	R - %	X - %
Substation	5,000	115 - D	4.16 Gr. Y	1	8
633-634	500	4.16 - Gr.W	0.48 - Gr.W	1.1	2

Table A.3: Both spot load and distributed load details

Spot Loads								
Node		Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
		Model	kW	kVAr	kW	kVAr	kW	kVAr
634		Y-PQ	160	110	120	90	120	90
645		Y-PQ	0	0	170	125	0	0
646		D-Z	0	0	230	132	0	0
652		Y-Z	128	86	0	0	0	0
671		D-PQ	385	220	385	220	385	220
675		Y-PQ	485	190	68	60	290	212
692		D-I	0	0	0	0	170	151
611		Y-I	0	0	0	0	170	80
		TOTAL	1158	606	973	627	1135	753
Distributed Loads								
Node A	Node B	Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
		Model	kW	kVAr	kW	kVAr	kW	kVAr
632	671	Y-PQ	17	10	66	38	117	68

Table A.4: Configuration Details of Line Parameters

Configuration 601			**Configuration 605**		
$Z = \mathbb{R} + j\mathbb{X}$ in ohms per mile			$Z = \mathbb{R} + j\mathbb{X}$ in ohms per mile		
0.3465 +j1.0179	0.1560 +j0.5017	0.1580 +j0.4236	0 +j0	0 +j0	0 +j0
	0.3375 +j 1.0478	0.1535 +j0.3849		0 +j0	0 +j0
		0.3414 +j1.0348			1.3292 +j1.3475
B in microsiemens per mile			B in microsiemens per mile		
6.2998	-1.9958	-1.2595	0	0	0
	5.9597	-0.7417		0	0
		5.6386			4.5193
Configuration 602			**Configuration 606**		
$Z = \mathbb{R} + j\mathbb{X}$ in ohms per mile			$Z = \mathbb{R} + j\mathbb{X}$ in ohms per mile		
0.7526 +j1.1814	0.1580 +j0.4236	0.1560 +j0.5017	0.7982+j 0.4463	0.3192 +j0.0328	0.2849 -j0.0143
	0.7475 +j1.1983	0.1535 +j0.3849		0.7891 +j0.4041	0.3192 +j0.0328
		0.7436 +j1.2112			0.7982 +j0.4463
B in microsiemens per mile			B in microsiemens per mile		
5.699	-1.0817	-1.6905	96.8897	0	0
	5.1795	-0.6588		96.8897	0
		5.4246			96.8897
Configuration 603			**Configuration 607**		
$Z = \mathbb{R} + j\mathbb{X}$ in ohms per mile			$Z = \mathbb{R} + j\mathbb{X}$in ohms per mile		
0 +j0	0 +j0	0 +j0	1.3425 +j0.5124	0 +j0	0 +j0
	1.3294 +j1.3471	0.2066+j 0.4591		0 +j0	0 +j0
		1.3238 +j1.3569			0 +j0
B in microsiemens per mile			B in microsiemens per mile		
0	0	0	88.9912	0	0
	4.7097	-0.8999		0	0
		4.6658			0
Configuration 604			**Configuration 604**		
$Z = \mathbb{R} + j\mathbb{X}$in ohms per mile			B in microsiemens per mile		
1.3238 +j1.3569	0 +j0	0.2066 +j0.4591	4.6658	-1.0817	-0.8999
	0 +j0	0 +j0		0	0
		1.3294 +j1.3471			4.7097

A.2 Banshee, A Real-Life Industrial Microgrid Network

The Banshee distribution network is a small scale, real-life, reconfigurable, industrial facility used as a pivotal benchmark model in the realm of microgrid research and development. Widely acknowledged for its real-world applicability and adaptability within the community microgrid, it is very popular as a quintessential standard and used by various academic scholars, industry frontrunners, practitioners and esteemed research laboratories for assessing microgrid performance. This model can be utilized for multifaceted analyses such as exploring various operational scenarios, evaluating control strategies, and assessing resilience and reliability measures. In this thesis, the Banshee system is used in Chapter 6 for evaluating the efficacy of the proposed islanding detection method. The system's layout and its different components depicted in its one line diagram is shown in Fig. 6.5. The major facilities and its key components are listed as follows:

1. Banshee MG is composed of three radial distribution feeders carrying load with minimum and maximum ranges between 5 MW to 14 MW and formed three areas with limited connectivity through normally open switches (NOS) to the utility grid.

2. Each mainstream feeders are rated with medium voltage of 13.8 kV at distribution level and service voltage levels of 4.16 kV, 480 V, and 208 V. There are total 18 aggregated dynamic loads with power factor of 0.9 lag, categorized as either critical, priority or interruptible. Additionally, there are two 200 hp induction motors that serves chiller compressor loads.

3. There are typical 4 different types of generation assets are available within the Banshee MG. A 4 MVA diesel generator (DieGen) in area-1, 2.5 MVA battery energy storage system (BESS) and 5 MW PV array designed via average modelling with time-varying irradiance profile and temperature is area-2, and 3.5 MVA natural gas fired combined heat and power plant (CHP) operating at 13.8 KV in area-3

4. Moreover, the system is further modified by integrating 3 more grid following average modeled VSC based DGs, i.e., DG1, DG2 and DG3, of equal 1.25 MW rating, are located at Bus-107, Bus-305 and Bus-209, respectively.

Tables A.5-A.13 provides the summary of the each component details of Banshee MG.

Table A.5: Short Circuit Levels Respective to Each Feeders

	3Ph(kA)	X/R	SLG (kA)	X/R
Feeder 1	14.58	4.6	10.57	0.9
Feeder 2	15.73	7.9	10.24	2.6
Feeder 3	15.73	4.6	10.57	0.9

Table A.6: Aggregated Load Details For Each Feeders of Banshee MG

Load ID	**Category**	**Feeder Number**	**kVA Demand**
C1	Critical	1	1200
C2	Critical	1	1500
C3	Critical	2	1000
C4	Critical	2	1000
C5	Critical	3	1000
C6	Critical	3	800
P1	Priority	1	1000
P2	Priority	2	1000
P3	Priority	2	1000
P4	Priority	3	600
P5	Priority	2	700
P6	Priority	3	1000
I1	Interruptible	1	300
I2	Interruptible	1	250
I3	Interruptible	2	300
I4	Interruptible	2	600
I5	Interruptible	2	400
I6	Interruptible	3	600

Table A.7: Parameter Details of Induction Machines Load

Name	**Description**	**Value**	**Unit**
vbsll	Rated Stator Voltage (L-L RMS)	0.48	kV
trato	Turns Ratio, Rotor over Stator	1	p.u.
pbase	Rated MVA	0.1597	MVA
hrtz	Rated Frequency	60	Hertz
ra	Stator Resistance	2.0110E-02	p.u.
xa	Stator Leakage Reactance	1.0448E-01	p.u.
xmd0	Unsaturated Magnetizing Reactance	9.0424E+00	p.u.
rfd	First Cage Rotor Resistance	4.5768E-02	p.u.
xfd	First Cage Rotor Leakage Reactance	1.0448E-01	p.u.
rkd	Second Cage Rotor Resistance	N/A	p.u.
xkd	Second Cage Rotor Leakage Reactance	N/A	p.u.
xkf	Rotor Mutual Leakage Reactance	N/A	p.u.

Table A.8: Parameter Details of Transformers

Name	**Rating [MVA]**	**Winding #1 Voltage [kV]**	**Winding #2 Voltage [kV]**	**Impedance [%]**	**X/R**	**Leakage Inductance [pu]**	**No load losses [pu]**
T101	0.5	13.8	0.48	5	4.9	0.04899	0.01000
T102	2.5	13.8	0.48	5.75	6.6	0.05685	0.00861
T103	3.75	13.8	4.16	4.75	11.4	0.04732	0.00415
T104	2	4.16	0.48	5.75	4.7	0.05624	0.01197
T105	2	4.16	0.48	5.75	4.7	0.05624	0.01197
T106	0.5	13.8	0.208	5	4.9	0.04899	0.01000
T107	2.5	13.8	0.48	5.75	6.6	0.05685	0.00861
T201	2.5	13.8	0.48	5.56	5.52	0.05471	0.00991
T202	0.5	13.8	0.208	5	4.9	0.04899	0.01000
T203	3.75	13.8	4.16	4.75	11.4	0.04732	0.00415
T204	1	4.16	0.48	5.75	4.21	0.05594	0.01329
T205	1.5	4.16	0.48	5.75	5.04	0.05640	0.01119
T206	2.5	13.8	0.48	5.75	6.6	0.05685	0.00861
T207	5	13.8	0.48	5	5.44	0.04918	0.00904
T208	2	13.8	0.48	5.75	4.7	0.05624	0.01197
T209	2	13.8	0.48	5.75	4.7	0.05624	0.01197
T210	1	13.8	0.48	5.75	4.21	0.05594	0.01329
T301	2	13.8	0.48	5.75	4.7	0.05624	0.01197
T302	2	13.8	0.48	5.75	4.7	0.05624	0.01197
T303	1	13.8	0.48	5.75	4.21	0.05594	0.01329
T304	1	13.8	0.48	5.75	4.21	0.05594	0.01329
T305	2.5	13.8	0.48	5.75	6.6	0.05685	0.00861

Table A.9: Parameter Details of Cables

Name	Length [ft]	AWG or kcmil	R [ohms/1000ft]	X [ohms/1000ft]
C101	1800	1-#500	0.0284	0.0351
C102	5500	1-#500	0.0284	0.0351
C103	1000	1-4/0	0.064	0.0389
C104	3000	1-#500	0.0284	0.0351
C105	3000	1-#500	0.0284	0.0351
C106	1500	1-#500	0.0284	0.0351
C107	2000	2-#500	0.0284	0.0351
C108	1000	1-#500	0.0294	0.0349
C109	2000	2-#500	0.0284	0.0351
C201	5500	1-4/0	0.064	0.0389
C202	2000	1-#500	0.0284	0.0351
C203	3000	1-#500	0.0284	0.0351
C204	1500	2-#500	0.0284	0.0351
C205	1500	2-#500	0.0284	0.0351
C206	1500	2-#500	0.0284	0.0351
C301	2500	1-#500	0.0284	0.0351
C302	2000	1-4/0	0.064	0.0389
C303	2000	1-#500	0.0284	0.0351
C304	1500	2-4/0	0.064	0.0389

Table A.10: Parameter Details of Natural Gas CHP and Diesel Generator Located at Bus 306 and Bus 103 Respectively

Parameter	Description	Units	Value
Mmva	Rated MVA of the Machine	MVA	3.5 [CHP] 4 [DieGen]
Vbsll	Rated RMS Line-to-Line Voltage	kV	13.8
HTZ	Base Angular Frequency	Hz	60
H	Inertia Constant	MWs/MVA	0.3468
D	Synchronous Mechanical Damping	pu/pu	0
XS1	Stator Leakage Reactance	pu	0.05
XMD0	D-axis Unsaturated Magnetization Reactance	pu	2.35
X230	D: Field-Damper Mutual Leakage Reactance	pu	0
X2D	D: Field Leakage Reactance	pu	0.511
X3D	D: Damper Leakage Reactance	pu	3.738
XMQ	Q-axis Magnetizing Reactance	pu	1.72
X2Q	1st Q-axis Damper Leakage Reactance	pu	0.2392
X3Q	2nd Q-axis Damper Leakage Reactance	pu	0.0942
RS1	Stator Resistance	pu	0.008979
R2D	Field Resistance	pu	0.00206
R3D	Direct-Axis Damper Resistance	pu	0.2826
R2Q	1st Q-axis Damper Resistance	pu	0.2392
R3Q	2nd Q-axis Damper Resistance	pu	0.0082

Table A.11: Technical Specifications of PV array Module Located at Bus 203

Sr. No	Description	Value
1	Insolation	1000 Watt/m^2
2	Temperature	25 Degree
3	Shading Effect	N/A
4	No. of Series Connected Cells in a Module	60
5	No. of Parallel Connected Cells in a Module	1
6	No. of Series Connected Modules	95
7	No. of Parallel Connected Modules	168
8	Open Circuit Voltage	45 V
9	Short Circuit Current	9.2 A
10	Voltage at Max Power @STD = 25 deg Centrigrade	37 V
11	Current at Max Power @STD = 25 deg Centrigrade	8.5 A
12	Open Circuit Series Resistance	0.349 Ohms
13	Short Circuit Shunt Resistance	111.55 Ohms
14	Rated Power Output	5 MW
15	Maximum Power Point Tracking	Enabled
16	AC Side Filter Resistance (Rf)	2.38E-3 Ohms
17	AC Side Filter Inductance (Lf)	200 UH
18	DC Voltage Set Point	1 kV
19	DC Bus Voltage Proportional Control (Gp)	0.899
20	DC Bus Voltage Integral Time Constant (TI)	0.0585 sec
21	Max DC Volatge Limit	6
22	Min DC Volatge Limit	-6 V
23	dq axis Proportional Gain For Current Controller (Gp)	0.2
24	dq axis Integral Time Constant For Current Controller (TI)	0.30675 sec

Table A.12: Technical Specifications of BESS Located at Bus 204

Sr. No	Description	Value
1	Battery Type	Lithium-ion
2	No. of Cells in Series in a Stack	460
3	N. of Stacks in Parallel	428
4	Capacity of a Single Cell (AH)	1
5	Initial State of Charge (SOC) of a Single Cell	85%
6	Nominal Voltage	0.48 kV
7	Power Rating	2.5 MVA
8	dq axis Proportional Gain For Current Controller (Gp)	0.2
9	dq axis Integral Time Constant For Current Controller (TI)	0.30675 sec
10	dq axis Proportional Gain For Voltage Controller (Gp)	1
11	dq axis Integral Time Constant For Voltage Controller (TI)	0.006 sec
12	d axis voltage reference	1.04614
13	q axis voltage reference	0

Table A.13: Parameter Details of Average Modeled DGs Located at Bus 107, Bus 305 and Bus 209

Sr. No		**Description**	**Value**
1	DC Bus Voltage Control	DC Voltage Set Point	1 kV
2		Proportional Control (Kp)	0.5
3		Integral Control (KI)	5
4		Max DC Volatge Limit	5 V
5		Min DC Volatge Limit	-5 V
6	Outer Loop Control	Reactive Power Reference	0
7		Reactive Power, Kp	1
8		Reactive Power, KI	5
9		Reactive Power, Upper Limit (UL)	5
10		Reactive Power, Upper Limit (LL)	-5
11	AC Bus Voltage Control	Reated Capacity	1.25 MW
12		Ref Voltage (Peak)	0.392 kV
13		Voltage, Kp	0.025
14		Voltage, KI	3
15	Inner Loop Control	dq axis current controller, Kp	0.025
16		dq axis current controller, KI	0.5
17		dq axis Current, UL	5
18		dq axis Current, LL	-5

Chapter B

Publications

Journal

1. S. De and R. Sodhi, ”A PMU Assisted Cyber Attack Resilient Framework Against Power Systems Structural Vulnerabilities”, in **Electric Power System Research (IF: 3.9),** vol. 206, ISSN 0378-7796, Jan. 2022.

2. S. De, M. V. Reddy and R. Sodhi, ”A Data-Driven Passive Islanding Detection Scheme,” in **IEEE Transactions on Industry Applications (IF: 4.4),**, vol. 60, no. 2, pp. 3698-3709, March-April 2024, doi: 10.1109/TIA.2023.3348425.

3. S. De and R. Sodhi, ”A Unified Cyber Attack Detection and Mitigation Framework for an AC Islanded Microgrids”, in **IEEE Transactions on Systems, Man, and Cybernetics: Systems (IF: 8.7)**, vol. 54, no. 9, pp. 5270-5282, Sept. 2024, doi: 10.1109/TSMC.2024.3403749.

4. S. De and R. Sodhi, “A Simple Replay-Attack-Resilient Power System State Estimation Scheme”, in **IEEE Transaction of Automation Science and Engineering (IF: 5.9),** 2024. *(2nd Revision Submitted)*

5. S. De, N. Kumar and R. Sodhi, “A Rule-based False Data Injection Attack Detection, Classification and Localization Scheme on AC Microgrid System Using XGBoost Machine Learning Classifier”, in **IEEE Transaction on Industrial Applications (IF: 4.4),** 2024. *(Under Review)*

6. S. De and R. Sodhi, ”Cyber Secured Passive Islanding Detection Based on Statistically Crafted Cyber Attack Detector”, in **Electric Power System Research (IF: 3.9),** 2024. *(Under Review)*.

Conference

1. S. De and R. Sodhi, ”A Simple Cyber Attack Detection Scheme for Smart Grid Cyber Security Enhancement,” **2020 21st National Power Systems Conference (NPSC),** Gandhinagar, India, 2020, pp. 1-6, doi: 10.1109/NPSC49263.2020.9331837

2. M. V. Reddy, S. De and R. Sodhi, ”A Data-Driven Passive Islanding Detection Scheme,” **2022 IEEE 10th Power India International**

Conference (PIICON), New Delhi, India, 2022, pp. 1-6, doi: 10.1109/PIICON56320.2022.10045174.

3. N. Kumar, S. De and R. Sodhi, "Comparative Assessment of Machine Learning (ML) Techniques for False Data Injection Attack (FDIA) Classification," **2023 IEEE 3rd International Conference on Smart Technologies for Power, Energy and Control (STPEC),** Bhubaneswar, India, 2023, pp. 1-6, doi: 10.1109/STPEC59253.2023.10430663.

4. S. De and R. Sodhi, "A Maximum Mean Discrepancy Estimator Enabled Cyber Attack Detection in an Autonomous AC Microgrid," **2024 23rd National Power Systems Conference (NPSC),** Indore, India, 2024. *(Accepted).*