COGNITIVE DYNAMIC SYSTEM FOR SMART GRID

COGNITIVE DYNAMIC SYSTEM FOR CONTROL AND CYBER SECURITY IN SMART GRID

BY

MOHAMMAD IRSHAAD OOZEER, B.Eng., M.Eng.

A THESIS

SUBMITTED TO THE DEPARTMENT OF COMPUTATIONAL SCIENCE AND ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Ph.D (2020) McMaster University

(Computational Science and Engineering) Hamilton, Ontario, Canada

TITLE:	COGNITIVE DYNAMIC SYSTEM FOR CONTROL AND CYBER SECURITY IN SMART GRID
AUTHOR:	MOHAMMAD IRSHAAD OOZEER B.Eng. (Electrical & Biomedical Engineering), McMaster University, Hamilton, Canada M.Eng. (Electrical & Computer Engineering), McMaster University, Hamilton, Canada
SUPERVISOR:	Prof. Simon Haykin
CO-SUPERVISOR:	Dr. Thomas Hurd
NUMBER OF PAGES:	xxiii, 266

Lay Abstract

The smart grid is forecasted to be the future of the grid by integrating the traditional grid with information and communication technology. However, the use of this technology has not only brought its benefits but also the vulnerability to cyber-attacks. False data injection attacks is a new category of attacks targeting the smart grid that can cause serious damage by manipulating the state estimation process and starting a chain of incorrect control decisions. The cognitive dynamic system is a powerful research tool inspired by the brain that can be used to study real time cyber physical systems. The key goal of this thesis is to apply cognitive dynamic systems to the smart grid to improve the state estimation process, detect cyber-attacks and mitigate their effects. Simulation results show that the proposed methods have robust performance in both state estimation and cyber-attack mitigation under various challenging scenarios.

Abstract

The smart grid is forecasted to be the future of the grid by integrating the traditional grid with information and communication technology. However, the use of this technology has not only brought its benefits but also the vulnerability to cyber-attacks. False data injection (FDI) attacks are a new category of attacks targeting the smart grid that manipulates the state estimation process to trigger a chain of incorrect control decisions leading to severe impacts.

This research proposes the use of cognitive dynamic systems (CDS) to address the cyber-security issue and improve state estimation. CDS is a powerful research tool inspired by certain features of the brain that can be used to study complex systems. As two of its special features, Cognitive Control (CC) is concerned with control in the absence of uncertainty, Cognitive Risk Control (CRC) uses the concept of predictive adaptation to bring risk under control in the presence of unexpected uncertainty.

The primary research objective of this thesis is to apply the CDS for the SG with emphasis on state estimation and cyber-security. The main objective of CC is to improve the state estimation process while CRC is concerned with mitigating cyber-attacks. Simulation results show that the proposed methods have robust performance for both state estimation and cyber-attack mitigation under various challenging scenarios.

This thesis contributes to the body of knowledge by achieving the following objectives: proposes the first theoretical work that integrates the CDS with the DC model of the SG for control and cyber-attack detection; demonstrates the first experimental work that brings a new concept of CRC for cyber-attack mitigation for the DC state estimator; introduces a new CDS architecture adapted for the AC model of the SG for state estimation and cyber-attack mitigation which builds upon all the research efforts made previously.

Preface

A. Included Manuscripts

This thesis contains three manuscripts, as listed below:

1. M. I. Oozeer, and S. Haykin, "Cognitive dynamic system for control and cyber-attack detection in smart grid," IEEE Access. 2019 Jun 12;7:78320-35.

2. M. I. Oozeer, and S. Haykin, "Cognitive risk control for mitigating cyber-attack in smart grid," IEEE Access. 2019 Sep 2;7:125806-26.

3. M. I. Oozeer, and S. Haykin, "Cognitive Dynamic System for AC State Estimation and Cyber-Attack Mitigation in Smart Grid," IEEE Access. under review, 2020.

B. Details on Included Manuscripts

(1) Manuscript 1 Presented in Chapter 2

The work was completed between December 2018 and March 2019. The manuscript was submitted in March 2019, revised in April 2019, and accepted in May 2019. My

contributions are:

- Development of cognitive dynamic system architecture emphasizing on cognitive control for DC model of smart grid
- Algorithm design, simulation implementation, and performance analysis.
- Manuscript authorship and journal submission corresponding author.

(2) Manuscript 2 Presented in Chapter 3

The work was completed between May 2019 and July 2019. The manuscript was submitted in July 2019 and accepted in August 2019. My contributions are:

- Extending the previous research in the earlier publication by developing cognitive risk control for DC model of smart grid to mitigate cyber-attacks
- Algorithm design, simulation implementation, and performance analysis.
- Manuscript authorship and journal submission corresponding author.

(3) Manuscript 3 Presented in Chapter 4

The work was completed between Septempber 2019 and January 2020. The manuscript was submitted in January 2020 and revised in June 2020. My contributions are:

- Development of cognitive dynamic system architecture emphasizing on cognitive control and cognitive risk control for AC model of smart grid
- Algorithm design, simulation implementation, and performance analysis.
- Manuscript authorship and journal submission corresponding author.

Copyright Permission

I have secured permission to include copyright material in this Ph.D. thesis from the copyright holders. The permission includes a grant of an irrevocable, non-exclusive license to McMaster University and to Library and Archives Canada to reproduce the material as part of the thesis.

To my parents, Iqbal and Shenaz, to whom I owe all the accomplishments in my life.

To my sister, Shabnaz, for all the support during my journey.

Acknowledgements

First and foremost, I would like to present my sincere thanks and gratitude to my supervisor Professor Simon Haykin for providing me the opportunity to pursue my Ph.D. degree under his guidance. I am very grateful to him for always believing in me, providing encouragement and support during my research in his lab and helping me in other ways whenever I needed it. He has always been there for me. During these three years, he has been a great source of inspiration and dedication, which has encouraged me to move forward and always to try my best in everything. It has been an extreme privilege to work by his side and learn so much from him.

I would also like to thank the three members of my supervisory committee - Dr. Thomas Hurd, Dr. Wenbo He and Dr. Thia Kirubarajan - for providing me with their valuable feedback and helpful support during my Ph.D journey.

I would like to express my deepest gratitude to Dr. Bartosz Protas, Dr. Nicholas Kevlahan and Dr. Thomas Hurd once again for providing me with invaluable assistance in meeting the requirements of the Ph.D program. I would also like to thank Ms. Diana Holmes and Ms. Julie Fogarty from the Department of Mathematics and Statistics and the School of Computational Science and Engineering. From the Department of Electrical and Computer Engineering, I would like to present my sincere thanks to Ms. Tracey Coop, Ms. Cheryl Gies and Ms. Kerri Hastings. I would also

like present a special thanks to Ms. Ginny Riddell from the Engineering Support Services (The Hub) for her assistance when I first started this journey.

Many thanks to my friends in the Cognitive Systems Laboratory, Dr. Eduardo Santos and Dr. Shuo Feng, for their advice, suggestions and discussions during my research in the lab.

Finally, I would like to deeply thank my family for their love, faith and unyielding support without which completion of this Ph.D. study would not have been possible. I will never be able to repay in my entire life for all you have done for me.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviations

AI	Artificial intelligence
BDD	Bad Data Detector
BDI	Bad Data Injection
CA	Contingency Analysis
CC	Cognitive Control
CDS	Cognitive Dynamic System
CP	Cognitive Predictor
CPS	Cyber-Physical Systems
CRC	Cognitive Risk Control
EHV	Extra High Voltage
EMS	Energy Management System

FDI	False Data Injection
HV	High Voltage
IDS	Intrusion-Detection System
IoT	Internet of Things
PAC	Perception-Action Cycle
PLC	Programmable Logic Controller
PMU	Phasor Measurement Units
RL	Reinforcement Learning
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition system
SE	State Estimator
SG	Smart Grid
TSC	Task-Switch Control
UCB	Upper Confidence Bound
WDT	Watch Dog Timer
WLS	Weighted Least Squares

Chapter 1

Introduction

1.1 Cyber-Physical System

The current fourth industrial revolution has been marked by the new generation of systems involving Internet of Things (IoT) and advances in Information and Communication Technology (ICT) [1]. Those have in turn contributed to the rise of modern engineering systems known as Cyber-Physical Systems (CPS). As those systems have been increasingly deployed in critical infrastructures that impact the daily lives of people such as the electrical power grids and health care, more and more concerns have been raised regarding the cyber security aspects of those systems and technologies [2]. In fact, in [3], the authors argue that these advances have made those technologies more vulnerable to cyber-attacks. A cyber-attack can be defined as any kind of offensive operation, executed by one or more computers to target other ICT infrastructures such as information systems and network infrastructures [4]. Although CPS are expected to be shielded from those kind of attacks, this is not a reality for

real-world systems. Consequently, there is a major need of new cyber security technologies for the mitigation and elimination of the negative repercussions caused by those attacks.

A CPS arises from a variety of combinations from physical hardware elements, such as sensors and actuators, and ICT components, such as IoT. However, each of those building blocks can be the source of CPS attacks. Thus, the complexity of those components and their arrangements have contributed towards the complexity of security and privacy issues surrounding CPS. The intricate interactions between the modules involved have made the assessment of vulnerabilities and identification of attacks more difficult. As a result, it is increasingly important to grasp the blueprints of those attacks and develop in-depth understanding on how those attacks are perpetuated in order to develop effective defense mechanisms. In the context of this thesis, we will be examining the smart grid (SG), which is a constantly developing revolutionary concept concerning the traditional power grid, and its greatest known threat known as False Data Injection (FDI) attacks, which is a new category of attacks targeting state estimation in different categories of CPS.

1.2 Smart Grid

The power grid is a complex network that connects multiple electric power generators from suppliers to consumers via power transmission and distribution units. Moreover, the SG is projected to be the next generation of the power grid that will be providing additional advantages and utilities [2]. Compared to the traditional power grid, the SG is expected to be more reliable and efficient since it will be making use of

the latest technological progress in the fields of sensing, monitoring, control strategies, computing and ICT [5]. By integrating additional intelligence received from the analysis of the measurements obtained at different phases in the grid, this will improve the resiliency and reliability of the SG [6]. Consequently, while allowing smart energy generation and improved emission control nationally, it will also allow local consumers to have enhanced energy savings since they will be able to better regulate the amount of energy consumed. The SG can be broken down into two major infrastructures which relate to power application and supporting infrastructure namely [7]. The power application component is concerned with the main function of the SG such as electricity generation and transmission. On the other hand, the supporting infrastructure is concerned with the intelligence of the SG. Thus, it is involved in the control and monitoring of the power application infrastructure using other components such as software and hardware. Hence, human intervention is greatly minimized with the help of the Supervisory Control and Data Acquisition systems (SCADA), which is the supporting communication infrastructure that is mainly concerned with the collection of measurements and transmission of necessary control actions to the actuators involved in the SG.

SCADA systems collect measurements from Remote terminal units (RTUs), such as field devices and sensors, and transmit them to the control center where processing and analysis are performed on those measurements for errors and irregularities. Above all, they are used by the Energy Management Systems (EMS) to calculate the states of the grid through a process known as state estimation. Since the estimated states play an essential role in the operational and economic function of the SG, it is very important that those estimates give a good depiction of the actual state of the

grid. Moreover, the states have to be estimated as they cannot be measured through any direct means. State estimators can be categorized as DC state estimators and AC state estimators. The DC state estimator uses a linear system model using measurements consisting of real power flows and injections while the AC state estimator uses a nonlinear system model and the measurements comprise of real and reactive power flows and injections. The states in the DC model consist of bus angles only. On the other hand, the states in the AC model comprise of bus magnitudes and angles.

The SG can be classified into different layers such as the physical layer, control layer and network layer [8]. While the physical and control layers model the physical environment of the system, the data communication layer and network layer interfaces the cyber and physical environments together. Lastly, the supervisory layer along with the management layer makes up the higher level application where interactions between human and machine occur. Several devices are employed for the proper functioning of the control layer. The important components are [9]:

(i) SCADA: This device is critical in order to monitor and perform control for the control layer. The SCADA infrastructure is present in the control center of all major U.S. utilities so that data can be processed and commands coordinated for power generation management and delivery within the extra high voltage (EHV) and high voltage (HV) section of their respective electric power system [10].

(ii) Observers/Sensors: Those refer to any devices used that perform measurements and collects data from the physical layer so that the physical state of system can be estimated. In this category, RTUs and substations are mainly the components which act as those observers. Data from sensors are collected by RTUs

and sent to the control center. On the other hand, the substation can be viewed as a central point of electrical connection connecting electrical equipment, such as transformers and generators, and control equipment together. Sensors are allowed to collect redundant readings to make sure that correct state measurements are obtained.

(iii) Intrusion-Detection Systems (IDSs): Those systems are responsible for the protection of both the physical and communication layers by applying either signature-based or anomaly-based intrusion detection algorithms. Whenever an anomaly or intrusion has occurred, the IDS will raise an alarm to inform the supervisor or work together with built-in-intrusion-prevention systems to apply remedial actions.

(iv) Phasor Measurement Units (PMUs): Compared to the previously mentioned sensors, PMUs are fairly new in origin and depend on GPS signals so that field measurements, such as phase-angles and voltages, can be obtained at a high frequency. However, traditional sensors do not rely on GPS to perform the same measurements. Furthermore, PMUs have higher sampling rate than those sensors. Lastly, PMUs can measure phase-angles directly while the traditional sensors cannot.

(v) Actuators and other intelligent control modules: In this category, Programmable Logic Controllers (PLCs) are the most popular components used for that purpose. Those devices can execute commands originating from the control center and apply changes to the physical infrastructures.

The application of SCADA systems is accompanied by the Energy Management

System (EMS) which help support other applications such as power network monitoring and control [11]. Thus, as a software package, the EMS's main goal is to work together with operators to improve the reliability and cost-effectiveness of supply of power. The state estimator (SE) of the power network is a real-time application that supplies the EMS with accurate estimated states using multiple measurements and the network model. Since the state estimate forms the fundamental foundation for other software applications and control, the SE has gained much importance in the power grid. It is also applied in contingency analysis (CA), where the estimated states are used to identify the possible dire consequences in the event that the electrical equipment could experience failures. In case of those theoretical contingencies, CA will determine whether those will void the current steady-state operating limits. The SE is also complimented with the bad data detector to identify and remove measurement errors that could impact the state estimation process negatively.

1.2.1 Weighted Least Squares State Estimation

While there are many techniques which have been applied for state estimation in the power grid such as Maximum Likelihood Criterion and Dynamic State estimation [12], Weighted Least Squares (WLS) State Estimation remains as one of the most popular state estimation algorithms applied to this day. It is based on the concept of using the quadratic function of the difference between the real measurements and an estimate of the measurements if they were accurate and errors were absent. Schweppe was the first to introduce WLS state estimation in 1970 [13]. As the WLS technique is covered to a greater extent in chapters 2, 3 and 4, at this point in this thesis, WLS will be covered very briefly so as to provide some background into False Data

Injection (FDI), which will be covered at a later point.

In the DC model, the measured real power flow from bus k to m can be simplified using the first order Taylor expansion around $\theta = 0$ with the following formula [14]:

$$P_{km} = \frac{\theta_k - \theta_m}{x_{km}} + e \tag{1.2.1}$$

where x_{km} corresponds to the reactance (in per unit values) of the branch k-m, θ_k is the phase angle(in radians) at bus k and e is the measurement error. In the same model, the power injection at a specific bus i is equal to the summation of all the power flows along incident branches to that bus:

$$P_i = \sum_{j \in N_j} P_{ij} + e \tag{1.2.2}$$

The DC measurement model for this estimation problem involves an overdetermined system of linear equations whereby the measurements and states are related by:

$$\mathbf{z} = \mathbf{Hx} + \mathbf{e} \tag{1.2.3}$$

where

- $\mathbf{x}$ is the n vector of the true states (unknown)
- $\mathbf{z}$ is the m vector of measurements (known)
- $\mathbf{H}$ is the m x n Jacobian matrix
- $\mathbf{Hx}$ is the m vector of linear function linking measurements to states
- $\mathbf{e}$ is the m vector of random errors

- m is the number of measurements
- n is the number of variables

In (1.2.3), $\mathbf{H}$ is a matrix which consists of power flow equations, which are described as vectors in its entries. These can be perceived as the theoretical calculations that relate the states to the measurement vector $\mathbf{z}$. In the AC model, the entries of $\mathbf{H}$ consist of a set of nonlinear functions of the state variables. However, in the DC model, the functions are linear. The state estimation problem can be expressed as a weighted- least-squares problem where the main goal is to calculate the best estimated states that matches the measurements. It was found that the best state estimates can be obtained by minimizing the following objective function [13]:

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})'\mathbf{W}(\mathbf{z} - \mathbf{H}\mathbf{x})' \tag{1.2.4}$$

In the above equation, the matrix $\mathbf{W}$ is a diagonal matrix which comprises of the measurement weights. $\mathbf{W}$ is usually based on the reciprocals of the measurement error σ for the different sensors.
$J(\mathbf{x})$ is then differentiated to obtain the first order optimal conditions:

$$\mathbf{G}\hat{\mathbf{x}} = \mathbf{H}'\mathbf{W}\mathbf{z} \tag{1.2.5}$$

where the estimate of the state $\hat{\mathbf{x}}$ is obtained by:

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} \tag{1.2.6}$$

In the above equations, $\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H}$ is the state estimation gain.

1.2.2 Numerical Example of WLS State Estimation

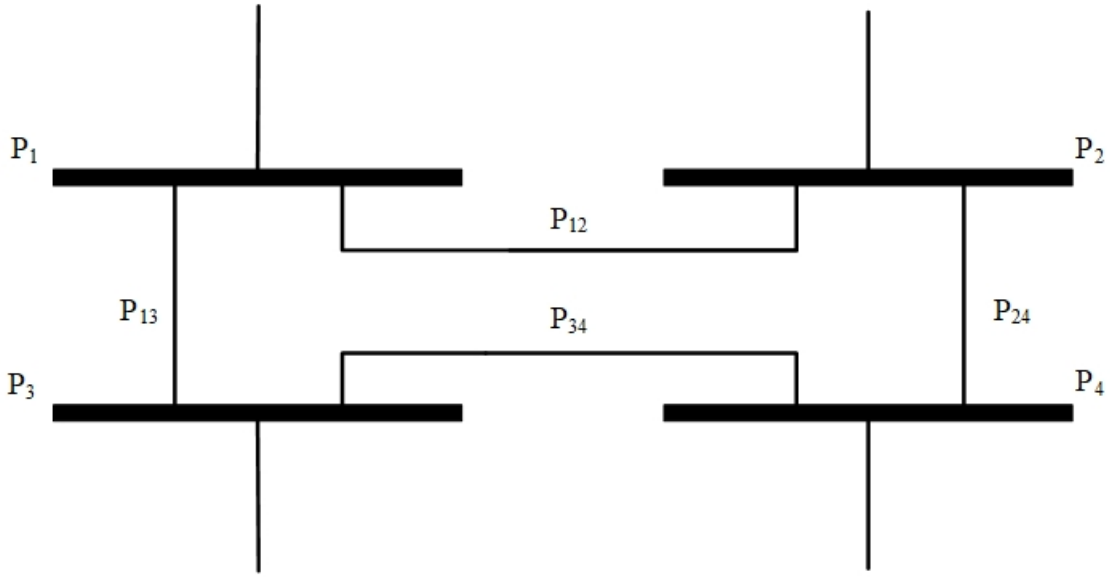

Figure 1.1: Line diagram for 4-bus, 2-generator transmission network case from [15].

From Bus	To Bus	X(p.u)
1	2	0.0504
1	3	0.0372
2	4	0.0372
3	4	0.0636

Table 1.1: Branch Impedance Values

For this example, we will be looking at the case4gs from Matpower [16] which deals with a 4 bus, 2 generator case from Grainger and Stevenson [15]. The data concerning the branch impedances are given in Table 1.1. In this example, we will be applying WLS to determine the states of the system, which in this case, are the phase angles of the buses. The measurement values are then given as follows:

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{pmatrix} = \begin{pmatrix} P_2 \\ P_3 \\ P_4 \\ P_{12} \\ P_{13} \\ P_{24} \\ P_{34} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{pmatrix} = \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ -1.31 \\ -1.06 \end{pmatrix}$$

The measurement equations are obtained, using eq. (1.2.1) and (1.2.2) as shown:

$$P_2 = -P_{12} + P_{24} = -\left(\frac{\theta_1-\theta_2}{x_{12}}\right) + \frac{\theta_2-\theta_4}{x_{24}} = 46.72\theta_2 - 26.88\theta_4$$
$$P_3 = -P_{13} + P_{34} = -\left(\frac{\theta_1-\theta_3}{x_{13}}\right) + \frac{\theta_3-\theta_4}{x_{34}} = 42.6\theta_3 - 15.72\theta_4$$
$$P_4 = -P_{24} + -P_{34} = -\left(\frac{\theta_2-\theta_4}{x_{24}}\right) + -\left(\frac{\theta_3-\theta_4}{x_{34}}\right) = -26.88\theta_2 - 15.72\theta_3 + 42.6\theta_4$$
$$P_{12} = \frac{\theta_1-\theta_2}{x_{12}} = -19.84\theta_2$$
$$P_{13} = \frac{\theta_1-\theta_3}{x_{13}} = -26.88\theta_3$$
$$P_{24} = \frac{\theta_2-\theta_4}{x_{24}} = 26.88\theta_2 - 26.88\theta_4$$
$$P_{34} = \frac{\theta_3-\theta_4}{x_{34}} = 15.72\theta_3 - 15.72\theta_4$$

Since we are using bus 1 as the reference bus, the phase angle associated with that bus is set to zero. The **H** matrix in eq. (1.2.3) is constructed. The weight matrix

from the data in Matpower is rounded and shown below.

$$\mathbf{H} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} 67 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 67 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 67 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 50 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 50 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 50 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 50 \end{bmatrix}$$

The gain matrix is obtained using $\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H}$:

$$\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H} = 10^5 \times \begin{pmatrix} 2.505 & 0.283 & -1.97 \\ 0.2831 & 1.8663 & -1.0209 \\ -1.97 & -1.02 & 2.35 \end{pmatrix}$$

The state estimate $\hat{\mathbf{x}}$ is calculated using eq. (1.2.6):

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} = \begin{pmatrix} -0.0187 \\ -0.0352 \\ 0.0312 \end{pmatrix}, \text{ thus } \hat{\mathbf{x}} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{\theta}_4 \end{pmatrix} = \begin{pmatrix} -0.0187 \\ -0.0352 \\ 0.0312 \end{pmatrix}$$

In hindsight, the real state values in Matpower were:

$$\mathbf{x} = \begin{pmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} = \begin{pmatrix} -0.01848 \\ -0.03547 \\ 0.0311 \end{pmatrix}$$

As we can see, WLS gives us a good estimate of the real state values.

1.2.3 Bad Data Detection

Bad data in state estimation refers to any kind of natural errors that can affect the state estimation negatively. The presence of bad data can therefore lead to bad consequences as a result of incorrect state estimation. The statistical properties of those errors allow us to use statistical techniques to identify them. Traditional methods for detecting bad data are based on the residue test. Referring to the previous section, the estimated measurements $\hat{\mathbf{z}}$ can be obtained using:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} \tag{1.2.7}$$

The residual $\mathbf{r}$ is the difference between $\hat{\mathbf{z}}$ and the actual measurements $\mathbf{z}$:

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} \tag{1.2.8}$$

In the DC model, the presence of bad data is indicated if the sum of the squared errors $\hat{f}$ is greater than a certain threshold τ, which is determined using the Chi-squared test. Normally, there is no bad data if:

$$\hat{f} < \chi^2_{(k,\alpha)} \tag{1.2.9}$$

where k refers to the appropriate degrees of freedom and α is a specified probability. If bad data is present, then the measurement that has the largest standardized error is removed from the state estimation procedure. State estimation and bad data

detection is re-iterated until bad data is absent.

1.2.4 Numerical Example of Bad Data Detection

Referring back to the previous example, using eq. (1.2.7) and (1.2.8), the estimated measurements and the estimated errors are:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} = \begin{pmatrix} -1.71 \\ -1.99 \\ 2.38 \\ 0.370 \\ 0.946 \\ -1.34 \\ -1.04 \end{pmatrix}, \quad \mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = \begin{pmatrix} -0.0099 \\ 0.0098 \\ 0.0068 \\ -0.0033 \\ 0.0057 \\ 0.0297 \\ -0.0165 \end{pmatrix}$$

In this case of the sum of squares $\hat{f}$ is found to be 0.0761. Performing the Chi-squares test in eq. (1.2.9) with 4 degrees of freedom and confidence level of 99%:

$$\hat{f} < \chi^2_{(k,\alpha)} \rightarrow 0.0761 < 13.28$$

Consequently, the test shows that bad data is absent. Now will consider the case where bad data is present. In Fig. 1.2, the leads of the power meter for P_{24} has been

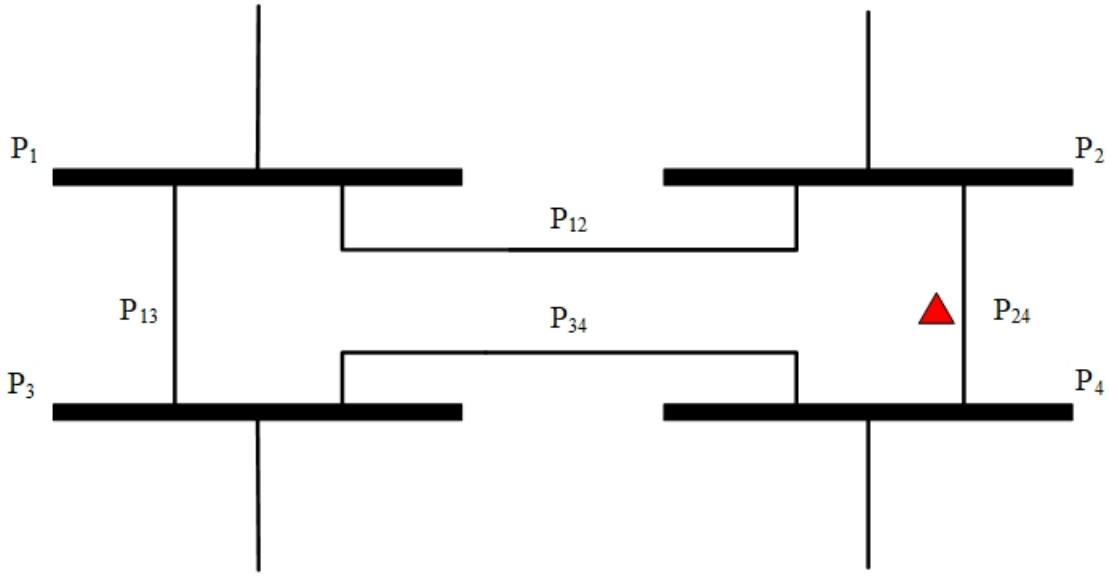

Figure 1.2: Line diagram for 4-bus, 2-generator transmission network case from [15] to test for Bad Data.

reversed such that its reading is now positive as shown:

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{pmatrix} = \begin{pmatrix} P_2 \\ P_3 \\ P_4 \\ P_{12} \\ P_{13} \\ P_{24} \\ P_{34} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{pmatrix} = \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ 1.31 \\ -1.06 \end{pmatrix}$$

The state estimate $\hat{\mathbf{x}}$ calculated using eq. (1.2.6) is:

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} = \begin{pmatrix} -0.0194 \\ -0.0462 \\ 0.0108 \end{pmatrix}$$

The estimated measurements and the estimated errors are:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} = \begin{pmatrix} -1.12 \\ -2.14 \\ 1.71 \\ 0.385 \\ 1.24 \\ -0.811 \\ -0.0897 \end{pmatrix}, \ \hat{\mathbf{e}} = \mathbf{z} - \hat{\mathbf{z}} = \begin{pmatrix} -0.5239 \\ 0.1597 \\ 0.6822 \\ -0.0178 \\ -0.291 \\ 2.12 \\ -0.0163 \end{pmatrix}$$

Now the sum of squares $\hat{f}$ is found to be 282, which exceeds the threshold of 13.28 of the Chi-squares test. Thus, the test suggests the presence of bad data. The standardized estimates of the errors calculated using the diagonal elements of $\mathbf{W}$ and

results are:

$$\begin{pmatrix} -7.57 \\ 2.29 \\ 9.15 \\ -0.152 \\ -2.55 \\ 16.8 \\ -1.26 \end{pmatrix}$$

In the above vector, the magnitude of the largest standardized error is linked to measurement z_6. Thus z_6 has been identified as a potential bad measurement. In this case, this measurement is omitted from state estimation calculations and the state estimate is calculated using the remaining measurements until the Chi-squares test is satisfied. After the removal of z_6, $\mathbf{H}$ and the state estimate are now:

$$\mathbf{H} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 0 & 15.72 & -15.72 \end{bmatrix}, \ \hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} = \begin{pmatrix} -0.0187 \\ -0.0350 \\ 0.0315 \end{pmatrix}$$

The sum of squares $\hat{f}$ is found to be 0.0207, which is below the threshold of 11.3 of the Chi-squares test with 3 degrees of freedom and confidence level of 99%. The test indicates absence of erroneous measurement and thus the estimated states are

deemed acceptable.

1.2.5 False Data Injection Attacks

The new technologies deployed in the transformation of the traditional power grid into the SG has made the new infrastructure increasingly dependent on cyber resources making them more prone to cyber-attacks. Information security has developed solutions for normal networking systems that can be used to detect and mitigate attacks against control systems. However, in [17], the authors argue that cyber security researchers have failed to recognize the importance of the attacks on the relationship between state estimation and control algorithms and the actual physical system. By hacking into the communication systems, hackers can affect state estimation algorithms in order to lead the operator or control center to apply incorrect control decisions thereby starting a cascading effect of bad decisions.

False Data Injections attacks (also known as Bad Data Injection (BDI) attacks) [8] are currently the greatest threat to the SG. Compared to bad data, which was discussed earlier, FDI attacks involve the addition of specially crafted bad data or offsets to the field measurements such that the state estimation process will output incorrect state estimates. Moreover, FDI attacks can bypass the traditional bad data detection schemes mentioned previously. Consequently, FDI attacks have been the most researched type of attack involving the SG. In the first paper detailing this kind of attack [8], the authors explain that when an attack vector $\mathbf{a}$ is injected to the original measurement vector $\mathbf{z}$, then we can write the new corrupted measurement vector $\mathbf{z}'$

as follows:

$$\mathbf{z}' = \mathbf{z} + \mathbf{a} \tag{1.2.10}$$

This will subsequently cause the state estimator to calculate an incorrect system estimate $\mathbf{x}'$ instead of the original $\mathbf{x}$. The difference between the corrupted states and new states can be denoted as $\mathbf{c}$ and written as:

$$\mathbf{x}' = \mathbf{x} + \mathbf{c} \tag{1.2.11}$$

The authors show that as long as the attack vector $\mathbf{a}$ satisfies the condition $\mathbf{a} = \mathbf{Hc}$, the attack will remain undetected as the residual remains unchanged. This is proven as follows [19]:

$$\begin{aligned} ||\mathbf{z}' - \mathbf{H}\mathbf{x}'|| &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}(\mathbf{x} + \mathbf{c})|| \\ &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}\mathbf{x} - \mathbf{H}\mathbf{c}|| \\ &= ||\mathbf{z} - \mathbf{H}\mathbf{x}||(\text{since,}\mathbf{a} = \mathbf{H}\mathbf{c}) \\ \mathbf{r}_{normal} &= \mathbf{r}_{attack} \end{aligned} \tag{1.2.12}$$

As shown by the proof, the attack will be undetected as current bad data detection techniques rely on the statistical properties of the residuals. If hackers have knowledge of the jacobian matrix, they can perform attacks by hacking field devices such as sensors or physically tampering with them. Since the estimated states are corrupted, the operator will be misled into applying incorrect control decisions thereby starting a domino effect of bad applied actions that could lead to disastrous consequences.

Fig. 1.3 taken from [20] gives a good depiction of the situation. FDI attacks will be covered in greater extent in chapters 2, 3 and 4. In the next section, we will demonstrate a numerical example involving FDI attack.

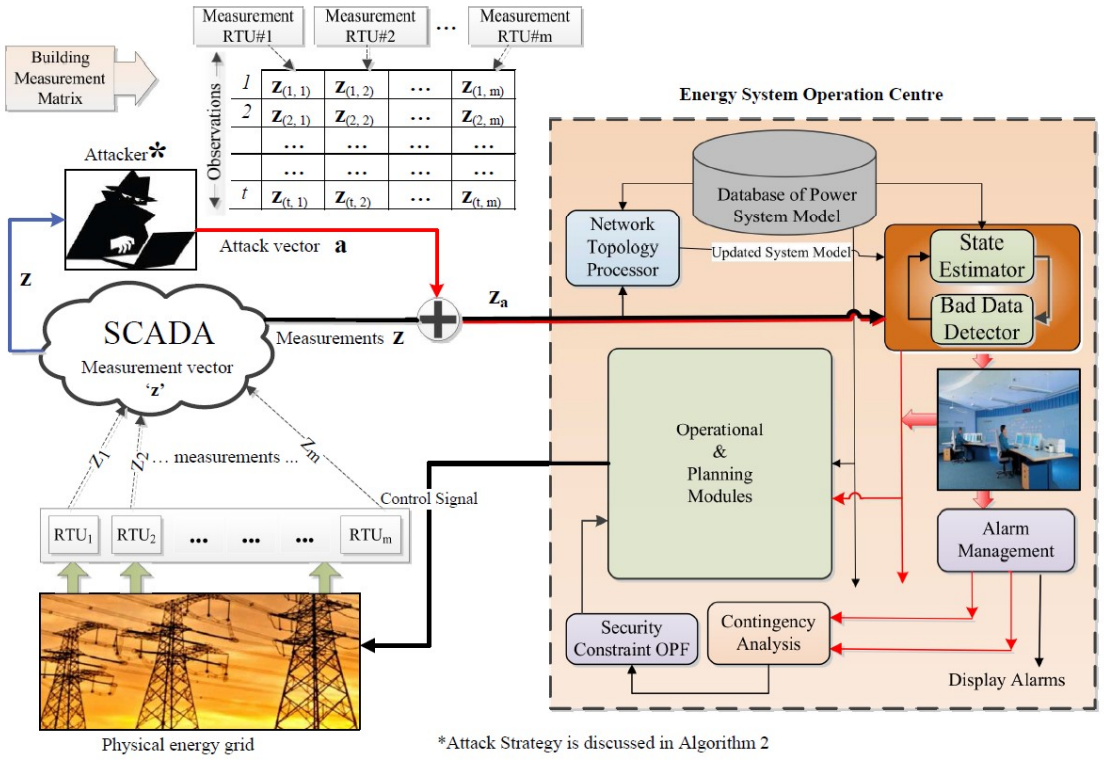

Figure 1.3: Overall architecture of the measurement signal based data-driven FDI attacks (taken from [20])

1.2.6 Numerical Example of FDI Attack

Referring back to example 1, consider the case whereby the attacker would want to change the estimated value of θ_2 by 0.5 radians. Previously, we solved the following calculation:

$$\mathbf{z} = \mathbf{Hx} + \mathbf{e}$$

where,

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{pmatrix} = \begin{pmatrix} P_2 \\ P_3 \\ P_4 \\ P_{12} \\ P_{13} \\ P_{24} \\ P_{34} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{pmatrix} = \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ -1.31 \\ -1.06 \end{pmatrix}, \ \mathbf{H} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix}, \ \mathbf{x} = \begin{pmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix}$$

In order to construct the attack vector, the problem can be rewritten as follows:

$$\rightarrow \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ -1.31 \\ -1.06 \end{pmatrix} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix} \begin{pmatrix} \theta_2 + 0.5 \\ \theta_3 \\ \theta_4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ -1.31 \\ -1.06 \end{pmatrix} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} + \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix} \begin{pmatrix} 0.5 \\ 0 \\ 0 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} -1.72 \\ -1.98 \\ 2.39 \\ 0.367 \\ 0.952 \\ -1.31 \\ -1.06 \end{pmatrix} - \begin{pmatrix} 23.36 \\ 0 \\ -13.44 \\ -9.92 \\ 0 \\ 13.44 \\ 0 \end{pmatrix} = \begin{bmatrix} 7.04 & 0 & -26.88 \\ 0 & -11.16 & -15.72 \\ 26.88 & 15.72 & -42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix} \begin{pmatrix} \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix}$$

$$\rightarrow \quad \mathbf{z} \quad + \quad \mathbf{a} \quad = \quad \mathbf{H} \quad \mathbf{x}$$

We will denote the attacked measurement resulting from the addition of the attack vector $\mathbf{a}$, to the initial measurement vector $\mathbf{z}$, as $\mathbf{z}_{attack}$. We can also see that the attack vector is relatively sparse. In this example, the measurements linked to z_1,z_3, z_4, and z_6 will be compromised. In a practical scenario, the attacker will attack the

meters providing those readings specifically. The attacked measurements are now:

$$\mathbf{z}_{attack} = \begin{pmatrix} -25.08 \\ -1.98 \\ 15.83 \\ 10.29 \\ 0.952 \\ -14.75 \\ -1.06 \end{pmatrix}$$

We will now repeat the same state estimation procedure that was carried out in the earlier example. The gain matrix is obtained using $\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H}$:

$$\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H} = 10^5 \times \begin{pmatrix} 2.505 & 0.283 & -1.97 \\ 0.2831 & 1.8663 & -1.0209 \\ -1.97 & -1.02 & 2.35 \end{pmatrix}$$

The new state estimate $\hat{\mathbf{x}}$ is calculated using eq. (1.2.6):

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z}_{attack} = \begin{pmatrix} -0.5187 \\ -0.0352 \\ 0.0312 \end{pmatrix}, \text{ hence } \hat{\mathbf{x}} = \begin{pmatrix} \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{\theta}_4 \end{pmatrix} = \begin{pmatrix} -0.5187 \\ -0.0352 \\ 0.0312 \end{pmatrix}$$

The estimated measurements and the estimated errors are:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} = \begin{pmatrix} -25.07 \\ -1.990 \\ 15.82 \\ 10.29 \\ 0.9463 \\ -14.78 \\ -1.044 \end{pmatrix}, \quad \hat{\mathbf{e}} = \mathbf{z} - \hat{\mathbf{z}} = \begin{pmatrix} -0.0099 \\ 0.0098 \\ 0.0068 \\ -0.0033 \\ 0.0057 \\ 0.0297 \\ -0.0165 \end{pmatrix}$$

The estimated errors (residuals) are the same as the ones obtained in second example whereby the attack was absent. Consequently, we can see that both cases yield the same residuals, thereby sustaining the points raised by Liu et al. In this case, the sum of squares $\hat{f}$ is found to be 0.0761, which is the same as in first example. Performing the Chi-squares test in eq. (1.2.9) with 4 degrees of freedom and confidence level of 99%:

$$\hat{f} < \chi^2_{(k-\alpha)} \rightarrow 0.0761 < 13.28$$

Consequently, the test shows that bad data is absent. As a result, the attack has successfully bypassed detection and changed the state estimate of one of the phase angles by 0.5 radians.

1.3 Cyber-Security for Smart Grid

As illustrated with the previous numerical examples, the traditional state estimator and bad data detector are useful for normally occurring anomalies. However, they are ineffective when cyber-attacks targeting the state estimator in the form of FDI attacks take place. Consequently, there is a need for a new and improved state estimator that is able to perform state estimation in the presence of normally occurring uncertainties, such as sensor noise, and also artificial uncertainties, such as FDI attacks. In this thesis, it will be shown how the Cognitive Dynamic Systems (CDS) [21] can be integrated with both the DC and AC state estimator in order to make them more powerful. The new architecture provides improvement in the realm of state estimation for both cases by using the entropic state of the CDS as basis. Moreover, the entropic state also serves as attack detector for attacks such as the FDI attacks. Once the FDI attack is detected, it will be shown how a new and improved version of Cognitive Risk Control (CRC) [22] can be used to mitigate those cyber-attacks. In the following sections, the CDS will be elaborated briefly and later expanded on in chapters 2,3 and 4.

1.3.1 Cognitive Dynamic System

Although the CDS was first introduced in 2006 [21], its fundamental concepts were already applied before its conception in the fields of engineering, communication and tracking through the Cognitive Radar [23] and Cognitive Radio [24]. Inspired by a new way of thinking, the CDS brings together the worlds of engineering and cognitive neuroscience whereby they both learn from each other. Research on the CDS has known much progress over the past years leading to Cognitive Control (CC) [25] and

Cognitive Risk Control (CRC) [26] as its two main functions. At its core, the CDS is based on Fuster's paradigm of cognition [27] which is founded on the following the five principles: perception-action cycle (PAC), memory, attention, intelligence and language. These principles are implemented in a "divide and conquer" scheme to establish the CDS.

1.3.2 Perception-Action Cycle (PAC)

From a neuroscience point of view, the PAC is the cybernetic loop that allows the organism to adapt to its environment for goal-oriented behavior. Moreover, the concept of PAC is also common to Artificial Intelligence (AI) [28] where the agent follows a continuous cyclic loop involving observables, actions and rewards so that the agent becomes optimal over time. However, unlike the PAC of AI, the PAC of the CDS brings involves a perceptor, which is responsible for perception, and the executive, which is tasked with performing action on the environment. The PAC of CDS starts with the processing of incoming observables (or measurements) originating from the environment by the perceptor. The feedback information from the perceptor in regards to the environment is then sent to the executive, which is tasked in performing the optimal action. The action performed then sets the stage for the next PAC which goes through the same cycle again [29]. This processing continues until further information gain from the environment becomes negligible assuming the environment is stationary [30].

1.3.3 Memory

Memory in the CDS builds on the PAC and consists of two parts; perceptual memory and executive memory respectively. Memory plays an important role in the sense that it is tasked with storing past learnt experiences and updating them continually on a cycle-by-cycle basis as the environment changes. Perceptual memory is concerned with the storage of experiences acquired from the senses (episodic, semantic, etc) from the perceptor while the executive memory stores actions performed on the environment by the executive. Reciprocally coupled to each other, both memories assist the CDS in predicting the consequences of actions performed by the executive on the environment. However, although executive memory was recently introduced for the CDS in [26] in the context of risk control, perceptual memory has not been present in the CDS architecture since its conception. In this thesis, it will be shown that we were able to integrate the perceptual memory, along with other supporting structures, with the CDS for the first time to mimic mechanisms in the field of cognitive predictive adaptation [31] as closely as possible. Consequently, the works presented in the later parts lays the foundation for a new way of approaching risk control from a cognitive neuroscience perspective.

1.3.4 Attention

Attention is an algorithmic principle that is distributed throughout the CDS in the form of perceptual attention and executive attention respectively. From a neuroscience point of view, attention is an essential mechanism for enabling adaptive behavior in the perceptor and executive of the CDS [30], similar to how those two principles exist as far the brain is concerned. By building on memory and the PAC,

attention allows the CDS to distribute and manage resources more efficiently so to deliver better performance.

1.3.5 Intelligence

Compared to the previous principles discussed, intelligence is manifested throughout the whole CDS and does not reside inside a particular physical structure. Being an algorithmic principle that builds on attention, memory and the PAC, intelligence is the most powerful principle involved [30] as it allows the CDS to perform optimal control by assisting the perceptor and executive in efficient decision-making. This is further reinforced with the multiple local and global feedback loops existing throughout the CDS to ensure that intelligent decisions are taken in the presence or absence of uncertainties in the environment.

1.4 Architecture of CDS

Being the entity that matches Fuster's paradigm the closest [32], the CDS is made up of four main components arranged in a specific order: environment, perceptor, executive and feedback channel. While the feedback channel brings together the perceptor and the executive, the environment closes the feedback channel to enclose the PAC of the CDS. In the following subsections, these constituents will be covered very briefly and later expanded on in the subsequent chapters.

1.4.1 Perceptor

The perceptor is responsible for performing the perception process on incoming noisy measurements in order to extract useful information to be processed by the whole CDS. In order to do so, the latter is equipped with the Bayesian generative model, Bayesian filter and entropic information processor.

(i) Bayesian Generative Model

With Bayesian dynamics being the technique of choice [33] [34], the generative model is based on Bayes rule to model the incoming observables originating from the environment. In this first processing state for the CDS, the posterior computed using Bayes rule becomes the prior during the next PAC and this goes on until the CDS is brought to rest.

(ii) Bayesian Filter

While the Bayesian Generative model is concerned with modelling the incoming measurements, the Bayesian filter is involved with filtering for estimating the state of the physical system at play conditional on the generative posterior. From a universal perspective, the Bayesian filter [35] is chosen as the optimal solution for the filter required by the perceptor. However, when the system can be modelled by a linear state and Gaussian distribution, the Bayesian filter can be simplified to the well-known Kalman filter [36]. By reciprocally coupling the Bayesian filter to the Bayesian generative model, a local feedback loop is set up that allows perceptual attention between these two components to take place. Consequently, this allows the Bayesian filter to extract useful information from the generative model while suppressing irrelevant information.

1.4.2 Feedback-Channel

The feedback channel has a special role within the CDS as it links the perceptor and executive. By doing so, this closes the global loop involving those two components around the environment, thereby completing the PAC. The feedback channel comprises of entropic-information processor, which is tasked with calculating the entropic state of the perceptor and internal rewards during reinforcement learning in the executive. The entropic state of the perceptor can be perceived as the embodiment of the directed cyclic flow of information during each PAC from the perceptor to the executive. Moreover, the entropic state of the CDS is built on the incoming filtered posterior, which brings together information from the generative model, Bayesian filter and entropy, inspired from Shannon's information theory [37]. While the entropic state was originally only used for control, in this thesis, it will be shown that the latter is very versatile and can be exploited to detect and mitigate cyber-attacks.

1.4.3 Executive

Being the most important module of the CDS, the executive is concerned with performing control. In order to perform this function, it equipped with reinforcement learning and cognitive control, which can be further broken down into the action space, planner, working memory and policy. The tasks of each those components will be covered shortly as follows:

(i) Reinforcement Learning

As the feedback channel feeds the reinforcement learning (RL) component of the CDS with internal rewards, the latter converts those rewards into the value-to-go function during the planning stages of RL. Since the outcome of the

actions performed on the environment are probabilistic as it is influenced by uncertainties present in that environment from one PAC to the another, the objective function of RL is to optimize the entropic state of the CDS for the function for which it was designed.

(ii) Planner and Action Space

Being the recipient of the value-to-go function, the planner extracts a set of prospective actions from the action space to be evaluated during the planning/shunt cycles. Moreover, the selection of those actions are continually improved under the effect of attention as the shunt cycles progresses. Similar to the coupling in the perceptor between the Bayesian generative model and the Bayesian filter, which involves top-down attention, bottom-up attention exists in the executive through the reciprocal coupling between the planner and RL resulting into another feedback loop. This feedback loop in the executive allows the enhancement of relevant information extraction and inhibition of irrelevant information from one PAC to the next. Moreover, the shunt cycles further amplifies this information extraction ability by taking into consideration both top-down and bottom-up attention during the shunt cycles for optimal control. From a general perspective, the contents of the action space will depend on the application at hand and varies in both size and complexity.

(iii) Policy and Working Memory

In the context of CDS, policy is involved with decision-making for cognitive action. As the CDS learns in a Bayesian fashion from the impacts of past actions to pick the best actions, this process is facilitated by the working memory. The working memory is a short-term memory that contains the set of past actions

performed by the executive in the previous PAC and is used as a set point from which the system can be possibly improved. Thus, during every PAC, the contents of working memory is put against the set of prospective actions evaluated by the planner.

With all the contents of the executive discussed, we can now define the principal role of the executive. Rooted in the cortical functions and mechanisms of cognitive neuroscience [39], the executive is responsible for CC [40], which is considered as is over-arching function as it applies actions on the environment on a cycle-to-cycle basis. However, while CC is good for dealing for situations involving the absence of uncertainties, in the presence of those uncertainties, the CDS has to expand its structure in order to tackle those. To this end, CRC [26] has to be introduced to handle those cases. As CRC is a very versatile concept relying on the principle of cognitive predictive adaptation, it will be shown in chapters 3 and 4 how it can take different forms depending on the application of interest. Nevertheless, there is one last important component, known as Task-Switch Control that can needs to be covered since it mediates between these modes of control.

1.4.4 Task-Switch Control

Task-Switch control is charge in determining whether CC or CRC should be on. By using the feedback channel as an indicator, task-switch control actuates a network of switches to alternate between the two previously mentioned control types. As it does so, it also turns on some other supporting components, which will be elaborated on further into this thesis. At this point, it is important to highlight that in a normal conditions, CC will be the prevalent form of control. CRC takes over the

system, through task-switch control, whenever uncertainties such as cyber-attacks are present.

Lastly, all the structures mentioned in this section pertains to the standard base structure of the CDS for general applications. However, depending on the application of interest, that structure will vary and will be accompanied with other supporting modules to adapt to those applications. In this thesis, it will be shown how CRC can take different forms to address the same problem under different underlying conditions.

1.5 Research Motivation and Objectives

Since the FDI attacks pose a real threat to the current form of the SG, a practical solution needs to be implemented to mitigate those attacks. Current research on the topic have focused on the DC model of the SG as it is a simplified version of the AC model. To our knowledge, there has been very few works published as far as the problem of FDI attacks is concerned for the AC model. Moreover, while there are many solutions proposed for the DC model such as K-means clustering [41] and Singular Value Decomposition (SVD) and convex-optimization [42], most of them are very "static" in nature. In other words, these solutions rely on other conditions such as prior data in order to be efficient. In fact, many solutions proposed from the AI community have this weakness. Since the performance of those methods rely on the quality of the training data, they cannot adapt to changing conditions, especially for a system as complex as the SG. Besides from having issues like high rate of false positives, they are also very computationally expensive, both from a training and application perspective respectively. This problem becomes even worse when the network is scaled up or modified. All the solutions proposed act as detectors.

None of them has attempted to approach the problem of FDI attacks from a control point of view. In fact, although it is known that those attacks bypass the statistical based bad data detection techniques, there has not been no attempt to innovate state estimation in the SG to bring together FDI attack detection, bad data detection and the state estimation together under a new light. It is our firm belief that we need this innovation so that the SG can arm itself to face the new threats of the 20th century that were not present during the original conception of the power grid in the early 1900s. The final solution should be dynamic and adaptive to the properties of the signal over time without relying on prior training data sets or prior determined thresholds.

Since its first introduction in 2006, Cognitive Radar [23] has remained a very dynamic field of research to this day, whereby researchers have looked towards its application in other applications such as automotive [43]. It has also been combined with newer techniques such as deep learning [44][45] for different end results in the tracking field. Moreover, the concepts of Cognitive Radar have led towards the architecture of the CDS as we know it today. Inspired from the success of Cognitive Radar in the area of tracking, where state estimation is involved in one way or the other, the same principles were sought to improve state estimation in the SG and also bring it under control in the face of cyber-attacks. As the CDS emerged from those principles as well as cognitive neuroscience, the CDS is a powerful research tool that can be used to study real time systems in the presence or absence of adversarial environments. In order to address the problem of FDI attacks in the SG, the primary research objective of this thesis is to bring together the CDS and the state estimator together under a new paradigm with emphasis on improving the state estimation process and detecting the

FDI attack. Secondary objectives that will assist in achieving the primary objectives involve finding answers to the following questions:

(i) How to integrate the CDS with state estimation and bad data detection process in the DC model of the SG? The DC model is chosen first as it is simpler than the AC model.

(ii) How can the new construct be used to detect FDI attacks?

(iii) Once the cyber-attack has been detected, how can the CDS and CRC be used to mitigate the attack?

(iv) After a successful implementation for the DC model, how can the architecture be modified to cater for the AC model of the SG? The AC is far more complex and computationally intensive than the DC model.

(v) Once it is possible to bring together the AC model and CDS, how can CRC be re-invented to mitigate attacks at a reasonable computational cost?

1.6 Research Outline

Fig. 1.4 shows the research outline of the thesis. The main body of this thesis consists of three technical works that follow a logical workflow whereby it starts with the simpler DC model and the problems surrounding it in the context of the material presented earlier. The subsequent presented works then builds upon the previous one with increasing complexity until we end up with chapter 4 where the harder AC model is the focus.

1.7 Thesis Organization

This thesis consists of the following chapters:

Chapter 1 provides a brief introduction of the SG and the background of state estimation. It further shows how FDI exploits the current state estimation mathematics to remain hidden. Both concepts are illustrated through numerical examples. CDS is then introduced as the research tool to address those problems. Lastly, the research motivation and associated objectives are discussed followed by the research outline.

Chapter 2 introduces the first published work where the CDS was merged with the DC state estimator of the SG to provide an overall improvement in state estimation. As this is also the first time that a generative model was applied in the CDS, the overall mathematics regarding the calculation of the entropic state was modified so as to bring information from the environment, perceptor and executive together. This is first application of CC for the SG whereby the CDS acts as the supervisor of the network. The new entropic state for the SG has many qualities, among which it is used as an FDI detector and also part of the objective function of the executive. To the best of the author' knowledge, the scholarly work presented herein is the first experimental work of CC being applied to the DC model of the SG.

Chapter 3 expands on the previous work presented in chapter 2 by introducing CRC to mitigate cyber-attacks in the SG. A more powerful architecture together with some supporting structures is used to provide stronger CRC compared to the earlier cited work in literature. Once the FDI attack has been detected using the entropic state, the latter actuates task-switch control to switch the CDS from performing CC to CRC. During CRC, the system is driven through different mechanics rooted in past experiences and predictive adaptation. Those past experiences are used to identify

the attacked states and mitigate undesired effects. Finally, the entropic state is used once again together with those past experiences when the attack is over to switch back to CC. To the best of the author's knowledge, the scholarly work presented herein is the first experimental work of CRC being applied to the DC model of the SG.

Chapter 4 uses concepts from both chapter 2 and chapter 3 in order to introduce both CC and CRC for the AC model of the SG. Compared to the DC model, the mathematics regarding state estimation in the AC model is more complex and computer intensive. A new flexible algorithm for CC is derived and its adaptation to other recursive state estimation algorithms in the SG is highlighted. Moreover, a new CRC infrastructure is used to mitigate FDI attacks using past experiences and predictive adaptation. Similar to chapter 3, past experiences are used to identify which states are being attacked and predictive adaptation is used to mitigate the unwanted outcomes of those attacks. To the best of the author's knowledge, the scholarly work presented herein is the first experimental work of CC and CRC being applied to the AC model of the SG.

Chapter 5 concludes the thesis by summarizing the works presented, its main contributions, limitations and prospective areas for future work.

Wang Y, Amin MM, Fu J, Moussa HB. A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids. IEEE Access. 2017 Nov 2;5:26022-33.

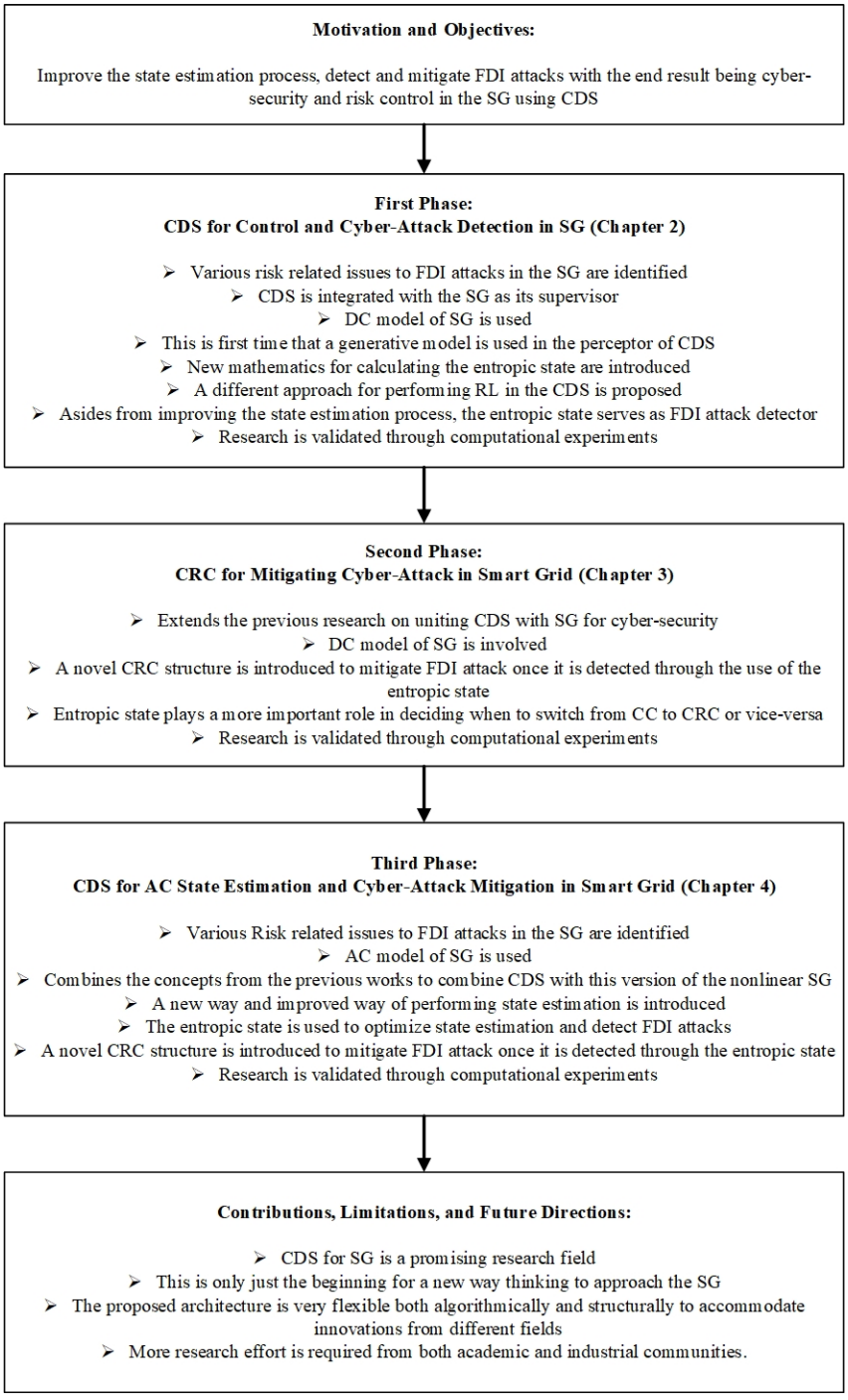

Figure 1.4: Research Outline of Thesis

Bibliography

[1] Wang Y, Amin MM, Fu J, Moussa HB. "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids." IEEE Access. 2017 Nov 2;5:26022-33.

[2] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security—A survey.", vol. PP, no. 99, pp. 1-1, 2017.

[3] Geluvaraj, B., P. M. Satwik, and TA Ashok Kumar. "The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace." International Conference on Computer Networks and Communication Technologies. Springer, Singapore, 2019.

[4] Nguyen TT, Reddi VJ, "Deep Reinforcement Learning for Cyber Security." arXiv preprint arXiv:1906.05799 (2019).

[5] Hao, Jinping, et al., "Sparse malicious false data injection attacks and defense mechanisms in smart grids.", IEEE Transactions on Industrial Informatics 11.5 (2015): 1-12.

[6] KP, V. P., and Bapat, J., "Bad data detection in smart grid for AC model.", 2014 Annual IEEE India Conference (INDICON). IEEE, 2014.

[7] Sridhar, S., Hahn, A. and Govindarasu, M.. "Cyber–physical system security for the electric power grid.", Proceedings of the IEEE 100.1 (2012): 210-224.

[8] Stoustrup J, Annaswamy A, Chakrabortty A, Qu Z. Smart Grid Control. Springer Nature Switzerland AG; 2019.

[9] Giannini M. Improving Cyber-Security of Power System State Estimators [Internet] [Dissertation]. 2014. (EES Examensarbete / Master Thesis). Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-142843

[10] Wu, F.F., Moslehi, K. and Bose, A., "Power system control centers: Past, present, and future.", Proceedings of the IEEE 93.11 (2005): 1890-1908.

[11] Teixeira, André, et al., "A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator.", IFAC Proceedings Volumes 44.1 (2011): 11271-11277.

[12] Monticelli, Alcir., "Electric power system state estimation.", Proceedings of the IEEE 88.2 (2000): 262-282.

[13] F. C. Scheweppe and J.Wildes, "Power system static-state estimation, Part I: Exact model," IEEE Trans. Power App. Syst. , vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.

[14] X. Fang, S. Misra, G. Xue, D. Yang, "Smart grid - the new and improved power grid: A survey", IEEE Commun. Surveys Tutorials 2012.

[15] J J. Grainger and W D. Stevenson JR.,"Power System Analysis 1st Edition", McGraw-Hill Series in Electrical and Computer Engineering, 1994.

[16] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12-19, Feb. 2011.(Digital Object Identifier: 10.1109/TPWRS.2010.2051168)

[17] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in Workshop on future directions in cyber-physical systems security, 2009.

[18] Y. Liu, P. Ning, M. Reiter, "False data injection attacks against state estimation in electric power grids", ACM CCS pp. 21-32 2009.

[19] A. Anwar, A. N. Mahmood, M. Pickering, "Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements", J. Comput. Syst. Sci. vol. 83 no. 1 pp. 58-72 2016.

[20] A. Anwara, A. Naser, M. Pickering,"Modelling and Performance Evaluation of Stealthy False Data Injection Attacks on Smart Grid in the Presence of Corrupted Measurements", arXiv. May 2016

[21] S. Haykin, "Cognitive dynamic systems [Point of view]," Proc. IEEE, vol. 94, no. 11, pp. 1910–1911, Nov. 2006.

[22] S. Haykin, J. M. Fuster, D. Findlay, and S. Feng, "Cognitive risk control for physical systems," IEEE Access, vol. 5, pp. 14 664–14 679, Jul. 2017.

[23] S. Haykin, "Cognitive radar: a way of the future.", IEEE signal processing magazine. 2006 Feb 13;23(1):30-40.

[24] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Sel. Areas Commun., vol. 23, no. 2, pp. 201–220, Feb. 2005.

[25] M. Fatemi and S. Haykin, "Cognitive control: Theory and application," IEEE Access, vol. 2, pp. 698–710, Jun. 2014.

[26] S. Haykin, J. M. Fuster, D. Findlay, and S. Feng, "Cognitive risk control for physical systems," IEEE Access, vol. 5, pp. 14 664–14 679, Jul. 2017.

[27] J. M. Fuster, "Cortex and Mind: Unifying Cognition", Oxford University Press, 2003.

[28] R. S. Sutton and A. G. Barto, "Reinforcement Learning", Cambridge, MA, USA: MIT Press, 1998.

[29] Haykin S, Amiri A, Fatemi M., "Cognitive control in cognitive dynamic systems: a new way of thinking inspired by the brain.", In 2014 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL) (pp. 1-7). IEEE.

[30] Haykin, S. and Fuster, J.M., "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other.", Proceedings of the IEEE 102.4 (2014): 608-628.

[31] J. M. Fuster, "The prefrontal cortex makes the brain a preadaptive system", Proc. IEEE, vol. 102, no. 4, pp. 417-426, Apr. 2014.

[32] J. M. Fuster, "Cortex and Mind: Unifying Cognition", Oxford University Press, 2003.

[33] C. Robert, The Bayesian Choice. New York, NY, USA: Springer-Verlag, 2001.

[34] D. J. MacKay, Information Theory, Inference and Learning Algorithms. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[35] Y. C. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," IEEE Trans. Autom. Control, vol. 9, no. 4, pp. 333-339, Oct. 1964.

[36] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems,". Journal of Basic Engineering, 82: 34–45, 1960

[37] C. E. Shannon, "A mathematical theory of communication", Bell Syst. Tech. J., vol. 27, no. 3, pp. 379-423, Jul./Oct. 1948.

[38] Richard S Sutton and Andrew G Barto. Reinforcement Learning: An Introduction. 2nd edition. MIT Press Cambridge, 2018 (cit. on pp. 13, 50, 52, 57, 79, 120).

[39] Earl K Miller and Jonathan D Cohen. "An integrative theory of prefrontal cortex function". In: Annual Review of Neuroscience 24.1 (2001), pp. 167–202 (cit. on p. 14).

[40] M. Fatemi and S. Haykin, "Cognitive control: Theory and application," IEEE Access, vol. 2, pp. 698-710, 2014.

[41] R. Xu, R. Wang, Z. Guan, et al. "Achieving Efficient Detection Against False Data Injection Attacks in Smart Grid," IEEE Access, vol. 5, pp.13787-13798, 2017.

[42] P. Gao, M. Wang, J. Chow, et al., "Identification of Successive "Unobservable" Cyber Data Attacks in Power Systems," IEEE Transactions on Signal Processing, 2016, 64 (21): 5557-5570.

[43] Liu P, Liu Y, Huang T, Lu Y, Wang X, "Cognitive Radar Using Reinforcement Learning in Automotive Applications," arXiv preprint arXiv:1904.10739. 2019 Apr 24.

[44] Elbir AM, Mishra KV, Eldar YC., "Cognitive radar antenna selection via deep learning," IET Radar, Sonar and Navigation. 2019 Jan 22;13(6):871-80.

[45] Mendis GJ, Wei J, Madanayake A., "Deep learning cognitive radar for micro UAS detection and classification," In 2017 Cognitive Communications for Aerospace Applications Workshop (CCAA) 2017 Jun 27 (pp. 1-5). IEEE.

Chapter 2

Cognitive Dynamic System for Control and Cyber-Attack Detection in Smart Grid

2.1 Preceding Introduction

The SG is a constantly evolving concept concerning the integration of the traditional power grid with ICT. Although ICT brings many benefits for the SG, it has also open the doors for cyber-attacks to occur. The FDI attack is currently the most dangerous attack targeting the grid as it can bypass the traditional bad data detection techniques. In order to make the SG safe, in this chapter, an architecture incorporating the CDS and the DC state estimator of the SG is proposed. The new construct, inspired by a new way of thinking, has much potential for the SG such as improving the state estimation process and detecting FDI attack.

To the best of the author's knowledge, the scholarly work presented herein is the first

experimental work of CDS being applied to the DC model of the SG for control and cyber-attack detection.

The publication included in this chapter is:

M. I. Oozeer, and S. Haykin, "Cognitive Dynamic System for Control and Cyber-Attack Detection in Smart Grid," IEEE Access. 2019 Jun 12;7:78320-35.

The co-author's contributions to the above work include:

1. Technical supervision and financial support of the study presented in this work.

2. Manuscript revising and editing.

Abstract

This paper introduces a new way of thinking that characterizes itself by uniting two entities, namely, State estimation in the Smart Grid (SG) on one hand, and Cognitive Dynamic System (CDS) on the other. False Data Injection (FDI) attacks are a family of new attacks that have been considered to be the most dangerous cyber-attack as it leads to cascaded bad decision making throughout the SG network, which can lead to severe repercussions. The conventional State Estimation and Bad Data Detection techniques, which have been applied to reduce observation errors and detect bad data in energy system state estimators, cannot detect FDI attacks. Here, we bring into play an objective-seeking system to act as the supervisor of the SG network. To this end, we propose to introduce a new metric for the SG: the entropic state. The entropic state has two purposes: 1) it provides an indication of the grid's health on a cycle-to-cycle basis and 2) it can be used to detect FDI attacks. Consequently, improving the entropic state is the goal of the supervisor. To achieve that objective, the supervisor dynamically optimizes the state estimation process by reconfiguring the weights of the sensors in the network. With optimality in mind, the CDS is the superior choice for the supervisory system. In this structure, the CDS interacts with the SG network, which is considered as the environment. Computer simulations are carried out on a 4-bus and the IEEE 14-bus systems to highlight the performance of the proposed approach in detecting both bad data and FDI attacks in the SG respectively.

2.2 Introduction

2.2.1 Cognitive Dynamic System

The Cognitive Dynamic System (CDS) is an organized physical model and research tool that simulates certain features of the brain. CDS was first introduced to the engineering world in [1] and then expanded on in [2]. Since its first applications in cognitive radio [3] and cognitive radar [4], CDS has evolved tremendously over the course of time to give rise to Cognitive Control (CC) [5] and Cognitive Risk Control (CRC) [6] as two of its special functions. While the CRC involves the principle of predictive adaptation, which is new to engineering literature [7], the main focus of this paper will be targeted towards the integration of CC, considered as the over-arching function of the CDS, with the Smart Grid (SG). From a neuroscience point of view, the CDS is based on Fuster's paradigm of cognition involving the following five principles: perception-action cycle (PAC), memory, attention, intelligence, and language [5]. In its purest form, the CDS is made of two main components: the perceptor, on one side, and the executive on the other with the feedback channel bringing them together. From an engineering perspective, CC is well structured to handle a slow progressing cyber-physical system such as the SG. Furthermore the architecture proposed in this paper is the first of its kind whereby a generative model has been incorporated in the perceptor and control performed by viewing the environment indirectly through that same perceptor. A new way to calculate the entropic state, with the SG as the main application, is also introduced. We will show how this entropic state will be fundamental to implement a control-sensing mechanism in the SG to identify and account for bad measurements while also laying the foundation for the detection of

False Data Injection (FDI) attacks in the SG.

2.2.2 Smart Grid

The current fourth industrial revolution has been marked by the emergence of the Internet of Things (IoTs) and Cyber-physical systems (CPSs) as a portrayal of the new upcoming generation of engineering systems [8]. These have in turn brought significant impacts to almost all aspects of our daily life, such as in electrical power grids, transportation systems, health-care etc. Being deployed in critical infrastructures, CPSs are expected to be safe from vulnerabilities and attacks [9]. Consequently, we can see the growing importance of cybersecurity for these systems. In this paper, we will focus on one such system which is the SG and its greatest current threat known as the FDI attacks (also known as Bad Data Injection (BDI) attacks).

The SG, compared to the traditional power grid, is forecasted to be more powerful in terms of reliability, efficiency and intelligence due to the fact that it will be making use of all the recent breakthroughs in sensing, monitoring, and control strategies [10][11]. In [12], the authors highlight that the SG can be broken down into two main parts namely the power application and the supporting infrastructure. The power application is the part dealing with the fundamental functions of the smart grid, which is electricity generation, transmission and distribution. On the other hand, the supporting infrastructure is the part equipped with intelligence, dealing mainly with the control and monitoring aspect of the fundamental operations of the SG using software, hardware and communication networks.

In the SG, Supervisory Control and Data Acquisition systems (SCADA) monitors and processes mainly the important control actions. SCADA systems collect meter

measurements from remote terminal units (RTUs) which consist of different field devices or sensors. The gathered measurements are then transmitted to a control center to be processed and analyzed for errors and inconsistencies. This is done mainly through a process known as state estimation [13][14]. In power systems, state estimation is used to estimate the system states using the available measurements at any point in time [15]. In the AC model, state variables are usually the voltage magnitudes and angles at the different buses in the system. The measurements used for state estimation are the real and reactive power flows, power injections and voltage magnitudes and angles from those buses. The current through transmission lines are also taken. Schweppe [13] was the first to introduce the concept of power system state estimation and used the Weighted Least Squares (WLS) method. As the number of measurements is greater than the number of states to be estimated, bad measurements can be discarded whilst still being able to obtain an estimation of the states of the system. Bad measurements are erroneous measurement readings that can hinder the state estimation process. The procedure of identifying those bad measurements is known as bad data identification and is also carried out during the state estimation process. The most commonly used bad data identification (detection) techniques are the Chi-Squared tests and Largest Normalized Residual Test [14][16]. These tests rely on the residuals between the estimated state variables and the measurement residuals. State estimators can be classified into DC and AC state estimators. In the DC model, a linear system model is used while in the AC model, a nonlinear model is employed. In the DC model, the measurements comprise of the real power flows and injections and states consist of bus angles [15][17][18]. The introduction of bad data, which evades the previously mentioned tests, can result in modifying the systems states.

Bad data are maliciously crafted offsets to measurements, which are injected to the transmitted sensor readings prior to state estimation to affect the estimated states in a certain way. Consequently, incorrect control decisions may be applied.

2.2.3 Contribution and Organization

The main contributions of this paper can be summarized as follows:

i. The architectural architecture of the CDS tailored for the SG is presented. To stay as true as possible to the brain, we bring into play a generative model in the perception part of the CDS. This is the first time where we demonstrate the potential of incorporating a generative model in the perceptor for an application such as the SG.

ii. A new way to calculate the entropic state is illustrated. In order to account for the latter and the generative model, a novel algorithm for optimal state estimation, based on CC is presented. The cognitive controller, which resides in the executive, is responsible for picking the right actions that will maximize the available information to the perceptor from one PAC to the next. We show through simulations that it only takes a few cycles for the system to learn which measurements to prioritize and which ones to neglect. The entropic state lays the foundation for a new way of control for the grid for bad data correction and FDI attack detection.

The rest of the paper is organized as follows: Section 2.3 reviews the basic concepts of state estimation, bad data detection and FDI attacks in the power grid. Section 2.4 expands on the structure of the CDS for the SG. We show how the SG can be

viewed as the environment with which the CDS interacts. Section 2.5 discusses the simulation results for two cases namely: bad data detection in a 4-bus network, and FDI attack detection in the IEEE 14-bus network. Finally, Section 2.6 concludes this paper by highlighting the key results and presenting new avenues for research.

2.3 Preliminaries

2.3.1 Weighted Least Squares State Estimation

The real-time operation of the Energy Management System (EMS) is dependent on the measurement data received from the SCADA. In order for the smart-grid to carry out its various tasks, it requires knowledge of the power system states for making decisions in real-time. Since the measurement signals are often corrupted by noise or erroneous, a reliable process is required to filter out the data. In the EMS, the state estimator and bad data detector are responsible for such tasks. Static state estimation refers to the process of obtaining the voltage phasors at all the system buses at discrete time intervals. A set of redundant measurements is taken in order to calculate the optimal state while filtering out the previously mentioned errors. This definition of system state implies only steady state bus voltage phasors most of the time. The states of a power system refer to the bus voltage angle θ and bus voltage magnitudes V. When using the DC model, measurements consist of real power flows and injections and the states are restricted to bus angles only. For this model, the bus magnitudes are assumed to be known beforehand and taken to be close to unity. Moreover, the phase angle at the reference bus is set to zero radians. Consequently, we estimate the n bus voltage angles $[\theta_1, \theta_2, ..., \theta_n]^T$only. In general,

the DC power flow model is broadly employed by power engineers and smart grid cyber-security researchers [13][19; 20; 21] as a way to linearize and approximate the Alternative Current (AC) power flow model [8]. The DC approximation has been widely accepted as a substitute for the AC model for the following reasons [22]:

i. Faster convergence is guaranteed

ii. Reduced algorithmic complexities related to power flow analysis

iii. Highly accurate results are obtained when used for transmission system analysis

The DC estimation model assumes that the bus voltage magnitudes are already known and are close to or equal to 1.0 per unit. Shunt elements and branch resistances are neglected. The measured real power flow from bus k to m can be approximated by the first order Taylor expansion around $\theta = 0$ with the following formula [11]:

$$P_{km} = \frac{\theta_k - \theta_m}{x_{km}} + e \tag{2.3.1}$$

where x_{km} corresponds to the reactance(in per unit values) of the branch k-m, θ_k is the phase angle(in radians) at bus k and e is the measurement error. The power injection at a specified bus i can be obtained by adding up all the flows along incident branches to that bus:

$$P_i = \sum_{j \in N_j} P_{ij} + e \tag{2.3.2}$$

In the DC state estimation problem, an overdetermined system of linear equations, which is known as the measurement model, is solved using the Weighted Least-Squares (WLS) problem. The state variables are related to the measurement using

the following measurement model:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e} \tag{2.3.3}$$

where

- $\mathbf{x}$ is the n vector of the true states (unknown)
- $\mathbf{z}$ is the m vector of measurements (known)
- $\mathbf{H}$ is the m x n Jacobian matrix
- $\mathbf{Hx}$ is the m vector of linear function linking measurements to states
- $\mathbf{e}$ is the m vector of random errors
- m is the number of measurements
- n is the number of variables

In (2.3.3), $\mathbf{H}$ is a matrix that describes the topology of the power system. It consists of power flow equations, which are described as vectors in its entries. These can be perceived as the theoretical calculations that relate the states to the measurement vector $\mathbf{z}$. In the AC model, the entries of $\mathbf{H}$ consists of a set of non linear functions of the state variables. However, in the DC model, the functions are linear.

In order to solve the weighted least-squares problem for the overdetermined system presented in (2.3.3), we need to find the n-vector $\mathbf{x}$ that minimizes the index $J(\mathbf{x})$, which is described as follows:

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})'\mathbf{W}(\mathbf{z} - \mathbf{H}\mathbf{x})' \tag{2.3.4}$$

In the above equation, the matrix $\mathbf{W}$ is a diagonal matrix which comprises of the measurement weights. These weights may represent qualities such as accuracy of the meter, reliability or an engineering judgement to express how important each individual measurement is to each other. Most of the time, $\mathbf{W}$, is founded on the reciprocals of the measurement error variance σ:

$$\mathbf{W} = \mathbf{R_z^{-1}} = \begin{bmatrix} \sigma_1^{-2} & \dots & \dots & \dots \\ \dots & \sigma_2^{-2} & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \sigma_m^{-2} \end{bmatrix} \tag{2.3.5}$$

where $\mathbf{R_z}$ is the covariance matrix of the measurement. We can then differentiate the performance index $J(\mathbf{X})$ to obtain the first order optimal conditions:

$$\mathbf{G\hat{x}} = \mathbf{H'Wz} \tag{2.3.6}$$

where the estimate of the state $\hat{\mathbf{x}}$ is obtained by:

$$\hat{\mathbf{x}} = \mathbf{G^{-1}H'Wz} \tag{2.3.7}$$

In the above equations, $\mathbf{G} = \mathbf{H'WH}$ is the state estimation gain.

2.3.2 Bad Data Detection

When the measurement values contain errors, these should be detected and identified so that they can be removed from the state estimation calculations. The statistical properties of these errors simplify their detection and identification. The estimated

measurements are obtained from the estimated measurements in (2.3.7) using the following equation:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} \tag{2.3.8}$$

The individual estimated measurement error is then obtained using:

$$\hat{\mathrm{e}}_j = (z_j - \hat{z}_j) \tag{2.3.9}$$

These errors follow a Gaussian distribution with zero mean [15]. Many methods for bad data detection exist. The Chi-Squares test and normalized residual are the most common ones. When Chi-squares test is performed, we assume that the states variables are mutually independent from each other and that the errors follow the normal distribution. It is shown in [15] that $\sum_{j=1}^{m} \frac{\left(z_j - \hat{z}_j\right)^2}{\sigma_j^2}$ follows a $\chi^2_{(m-n)}$ distribution, where m-n is the degree of freedom. The number of degrees of freedom (m-n) is defined as the difference between number of measurements, m, and the number of independent variables, n. The test procedure is performed as follows:

i. Calculate $\hat{\mathbf{x}}$ using (2.3.7)

ii. Using (2.3.8), calculate the corresponding estimated errors through (2.3.9)

iii. Use $\hat{f} = \sum_{j=1}^{N_m} \frac{\hat{e}_j^2}{\sigma_j^2}$ to evaluate the sum of squares.

iv. Using the appropriate number of degrees of freedom $k = (N_m - N_s)$ and a specified probability α, find whether or not the value of is less that the critical value corresponding to α. This means that we check if the following inequality has been satisfied:

$$\hat{f} < \chi^2_{(k,\alpha)} \tag{2.3.10}$$

v. If the above criterion is met, then the state estimates are considered accurate. Otherwise, we can suspect the presence of bad data in the measurement. When that happens, we remove the measurement related to the largest standardized error. The steps are then re-iterated until the conditions are met.

When erroneous measurements are present, the sum of squares $\hat{f}$ of the estimated errors will be large. In practical power system applications, the number of degrees of freedom is large. Consequently, this allows for the removal of a set of measurements, which relate to the largest standardized residuals. Nevertheless, this does not always mean that the largest standardized errors are always linked to bad measurements. Hence the technique allows for the identification of gross errors. The confidence level α of the Chi-squares test shares similar characteristics to a false alarm probability. The higher α is, the higher its sensitivity to errors. If α is lower, the lower its detection ability to bad data.

2.3.3 Bad Data Injection Attacks

BDI or FDI refers to those category of cyber-attacks where the attacker injects bad measurements such that they are not detected by the methods mentioned previously. In [23], the authors call it the stealthy attacks. Various BDI attacks have been identified and their impacts investigated in [9; 10; 14; 16; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32]. In [22], the adversary models for BDI attacks are broken into two main categories namely:

i. A model whereby the system parameters and topology (system Jacobian) is known to the attackers.

ii. A model where the system configuration is not known to the attackers.

The first category has been shown to have more damaging consequences. In [16], Liu et al. have shown that an attacker, with the knowledge of the system $\mathbf{H}_{m\times n}$, can maliciously inject an attack vector $\mathbf{a}_{m\times 1}$ along with the measurement vector $\mathbf{z}_{m\times 1}$ which cannot be detected and identified by the traditional state estimation and bad data detection techniques mentioned previously. When the attack vector is injected, the new corrupted measurement signal $\mathbf{z'}_{m\times 1}$ is now written as:

$$\mathbf{z'}_{m\times 1} = \mathbf{z}_{m\times 1} + \mathbf{a}_{m\times 1} \tag{2.3.11}$$

As a result, state estimation will produce the wrong system state $\mathbf{x'}_{m\times 1}$ instead of the original states $\mathbf{x}_{m\times 1}$. The difference in the states of the system will be given as $\mathbf{c}$, where

$$\mathbf{x'} = \mathbf{x} + \mathbf{c} \tag{2.3.12}$$

Liu et al. explain the theory behind this hidden attack and demonstrates his results through experiments. It is shown that if the attack vector satisfies the condition $\mathbf{a} = \mathbf{Hc}$, then the residual of the estimation process becomes[22]:

$$\begin{aligned} ||\mathbf{z'} - \mathbf{Hx'}|| &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}(\mathbf{x} + \mathbf{c})|| \\ &= ||\mathbf{z} + \mathbf{a} - \mathbf{Hx} - \mathbf{Hc}|| \\ &= ||\mathbf{z} - \mathbf{Hx}||(\text{since},\mathbf{a} = \mathbf{Hc}) \\ \mathbf{r}_{normal} &= \mathbf{r}_{attack} \end{aligned} \tag{2.3.13}$$

As shown in the mathematical proof, the residuals related to the attack vector and the residuals without attack are considered the same. Consequently bad data detection

will fail to distinguish the good measurement vector and the one contaminated with the attack due to the fact that it relies on statistical methods to calculate the residuals. As a result, the attack is undetected and will affect the determination of system states, which in turn can cascade into more damaging consequences. In (2.3.13), it inferred that this kind of attack targets state estimation directly in the model of the SG. Thus **a** can applied either physically by tampering with the targeted meters or wirelessly by injecting the vector when readings are sent to the SCADA. Consequently, any critical component of the SG model involved in state estimation, such as the substation state estimator (SSE) at the substations are not safe from such attacks.

2.4 Architectural Structure of CDS for Smart Grid

From a neuroscience point of view, the CDS is the closest system that matches Fuster's paradigm [7] when it comes to cognition. In the overall sense, the CDS consists of four basic components: the perceptor on one side and the executive on the other side; the executive is linked to the perceptor via a feedback channel, and the environment closes a global feedback loop whereby the entire CDS is embraced within it. In the context of this paper, the DC state estimator, being the recipient of the measurements in the network, is considered as the environment for which the CDS acts as the supervisor. Furthermore, through CC, this new system empowers the state estimator by equipping it with the cognition ability. More specifically, the CDS learns during every PAC which measurements to prioritize for optimal state estimation and which ones to disregard. The complex diagram depicting the unison of CDS and the DC state estimator is shown in Fig. 2.1. In the next subsections, the main constituents of the diagram will be elaborated.

2.4.1 Perception-Action Cycle

Assuming that the environment is free of uncertainty, the PAC is responsible for information gain for every cycle. Consequently, the global feedback loop of the PAC successively improves the information extraction ability of the perceptor during the each successive cycles. As a result, a continuous cyclic directed flow of information from the perceptor to the executive is set up. In a goal-focused scenario, a hypothesis, originating from the memory, guides the current PAC for each action performed on the environment by the executive. This hypothesis is then modified at every cycle depending on the information extracted from the perceptor.

2.4.2 Perceptor

Both in the human brain, and in the CDS, a perception process is performed on sensory measurements. The role of perception is to extract the available information out of noisy measurements which in response the human performs actions to continually enhance this information in subsequent cycles. These actions are called the cognitive actions. While the perceptor is able to see the environment directly and extract the relevant information about the environment from the observables, the controller senses the environment indirectly via the same perceptor. Unlike the CDS model proposed in [6], the perceptor for the SG consists of two components namely the generative model and the Bayesian filter, which are reciprocally coupled to each other.

2.4.2.1 Generative Model

Conceptually, the perceptor of the CDS originates with the *Bayesian generative model*[6], which classifies the observables from the environment. However, due to the dynamic nature of the SG, the Bayesian generative model is not a suitable choice. As the SG is highly complex in structure, it is crucial to detect any anomalies and deal with them as soon as possible before they can spread over the network and consequently lead to further cascaded problems throughout the system. Inspired by quickest detection theory, the generative model that will be put into action in the perceptor is based on cumulative sum (CUSUM) and is given by:

$$\mathbf{B}_k = \sum_{i=k-L}^{k} \mathbf{x}_i \tag{2.4.1}$$

where k refers to the current cycle number, L is the window over which the past states is being accumulated, $\mathbf{B}_k$ is the vector retaining the cumulative sum for each cycle and $\mathbf{x}_i$ is the vector of the states output from the DC state estimator for the cycle i. While CUSUM–based schemes have mostly been applied for detection only, whereby a suitable threshold has to be chosen as baseline for presence of attacks [33][34], it will be shown later on in this paper how the accumulator in (2.4.1) can be used for control as well when employed in the CDS structure. This generative model also has some desirable properties such as being able to smooth out the noise under the slow dynamics of the system.

2.4.2.2 Bayesian Filter

The second component of the perceptor, which is coupled to the generative model, is the Bayesian filter. Since the system is linear in nature and has additive white Gaussian noise, the well-known *Kalman filter* [35] is opted for the modelling the incoming inputs. The Kalman filter involves the state-space model which consists of a pair of equations known as the Process equation and the Measurement equation. Under the assumption that the power system is quasi-static in nature, it is expected that the state variable $\mathbf{x}$ at the time k+1 will be close to its values at its previous cycle k and might be subject only to very small deviations [36; 37; 38]. Mathematically this can be simplified as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \omega_k \tag{2.4.2}$$

where ω_k is independent Gaussian noise vector with zero mean. Consequently, the measurement equation of the Kalman filter for the perceptor is:

$$\mathbf{Y}_k = \mathbf{L}_k\mathbf{B}_k + \omega_k \tag{2.4.3}$$

where $\mathbf{Y}_k$ is the measurement vector of the Kalman filter at cycle k. The covariance matrix of ω_k is:

$$\mathbf{R} = \text{diag}[\sigma_\omega^2], \sigma_\omega^2 = \text{var}[\omega_i] \tag{2.4.4}$$

Under the quasi-static postulation, the process equation employed will be a random walk model which is given as:

$$\mathbf{B}_{k+1} = \mathbf{F}_k\mathbf{B}_k + \mathbf{v}_k \tag{2.4.5}$$

where $\mathbf{v}_k$ is the process noise vector which is assumed to be statistically independent and zero mean as well. The covariance matrix of $\mathbf{v}_k$ is:

$$\mathbf{Q} = \text{diag}[\sigma_{\text{v}}^2], \sigma_{\text{v}}^2 = \text{var}[\text{v}_i] \tag{2.4.6}$$

With respect to (2.4.2), in (2.4.3) and (2.4.5) the system matrix $\mathbf{L}_k$ and the predictive transition matrix $\mathbf{F}_k$ are assumed to be identity respectively. With the reference to the two previous equations, the computational steps involving the Kalman filter starts with some initial estimates of the states, $\hat{\mathbf{B}}_{k|k}$, and predicted error covariance, $\mathbf{P}_{k|k}$, which are used for the time update steps as follows:
The predicted estimated states for the next cycle, $\hat{\mathbf{B}}_{k+1|k}$, is calculated using

$$\hat{\mathbf{B}}_{k+1|k} = \mathbf{F}_{k+1,k}\hat{\mathbf{B}}_{k|k} + \mathbf{v}_k \tag{2.4.7}$$

and the predicted error covariance, $\mathbf{P}_{k+1|k}$, is found using

$$\mathbf{P}_{k+1|k} = \mathbf{F}_{k+1,k}\mathbf{P}_{k|k}\mathbf{F}_{k+1,k}^T + \mathbf{Q} \tag{2.4.8}$$

When the next cycle starts, those are then used for the measurement update stages: The Kalman gain, $\mathbf{K}_k$, is expressed as

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{L}_k^T(\mathbf{L}_k\mathbf{P}_{k|k-1}\mathbf{L}_k^T + \mathbf{R})^{-1} \tag{2.4.9}$$

The filtered estimate, $\hat{\mathbf{B}}_{k|k}$, is formulated as

$$\hat{\mathbf{B}}_{k|k} = \hat{\mathbf{B}}_{k|k-1} + \mathbf{K}_k(\mathbf{Y}_k - \mathbf{L}_k\hat{\mathbf{B}}_{k|k-1}) \tag{2.4.10}$$

The process covariance matrix is then updated using

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{L}_k \mathbf{P}_{k|k-1} \tag{2.4.11}$$

Thus after each time and measurement updated steps, this procedure is repeated with the preceding *a posteriori* estimates used to predict new *a priori* estimates.

2.4.3 Feedback Channel

The *feedback channel* [6] plays a distinctive role within the CDS as it links the perceptor to the executive, thus concluding the PAC. For the CDS to act as the supervisor for the SG, the feedback channel is equipped with entropic-information processor, which is responsible for the calculation of the so-called *entropic state* and internal rewards during reinforcement learning in the executive. The internal rewards calculation will be elaborated in section 2.4.4 (Executive) where it is more relevant to the role of the executive during planning.

2.4.3.1 Entropic-Information Processor

The directed cyclic flow of information from the perceptor to the executive is known as the *entropic state of the perceptor*. The latter is built on the principles of the perceptual posterior, which can be as viewed as the incoming filtered posterior incorporating the essence of the generative model and the Kalman filter, and entropy, which is derived from *Shannon's information theory* [39]. According to Shannon's

information theory, the entropic state at time k can be formulated as:

$$h_{k|k} = \int_{\mathbb{R}} p(\mathbf{B}_k|\mathbf{Y}_k) \log \frac{1}{p(\mathbf{B}_k|\mathbf{Y}_k)} d\mathbf{B}_k \tag{2.4.12}$$

where $p(\mathbf{B}_k|\mathbf{Y}_k)$ is the perceptual posterior of the Kalman filter and $\mathbb{R}$ denotes the entire space where the state $\mathbf{B}_k$ lies. Under the assumption that the noise terms in (2.4.3) and (2.4.5) are Gaussian, the posterior $p(\mathbf{B}_k|\mathbf{Y}_k)$ can be simplified to its mean and covariance matrix at each cycle. Hence the entropic state can be reduced to:

$$h_{k|k} = \frac{1}{2}\log(\det\{(2\pi e)\mathbf{P}_{k|k}\}) \tag{2.4.13}$$

where det{.} is the determinant operator. In [5], this is further simplified to:

$$h_{k|k} = \det\{\mathbf{P}_{k|k}\} \tag{2.4.14}$$

The application of $det\{.\}$ in (2.4.13) and (2.4.14) is originally intended to capture the whole information of the matrix into a single number. However, in the case of optimal control and attack detection in mind, (2.4.13) and (2.4.14) were not suitable to quantize the grid's performance as it was not sensitive enough to react to the actual changes in the environment. Therefore, the following equation is instead used to calculate the entropic state in the entropic information processor:

$$h_{k|k} = \frac{\det\{\mathbf{P}_{k|k-1} - (\text{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{Y}_k\}^2)\}}{\det\{\mathbf{P}_{k|k-1}\}} \tag{2.4.15}$$

where diag{.} refers to the diagonal operator. In a general sense, (2.4.15) aims to condense the information predicted from the previous cycle $k-1$ and information

from the current cycle k into a single number. (2.4.15) compares the previously predicted filtering-error covariance $\mathbf{P}_{k|k-1}$ with the actual error between the state estimate $\hat{\mathbf{B}}_{k|k-1}$ and the current measurement at cycle k, $\mathbf{Y}_k$. The denominator (2.4.15) serves as a normalizing operator where the entropic state, $h_{k|k}$, is confined to take values between 0 and 1, whereby 1 indicates full control and values below 1 indicate presence of disturbance or uncertainty. Moreover it is to be added that, unlike Shannon's information theory, (2.4.13), (2.4.14) and (2.4.15) will never assume the value of zero due to the fact that imperfections will always be present in the perceptor in one form or the other. As (2.4.15) involves the det{.} operator, the following condition has to be imposed on $h_{k|k}$:

$$h_{k|k} = \begin{cases} -h_{k|k}, & \text{if } d_k > 1. \\ h_{k|k}, & \text{otherwise.} \end{cases} \tag{2.4.16}$$

where d_k is the number of negative elements along the diagonal of $h_{k|k}$ during cycle k. Due to the nature of the environment being in the intermediate presence and absence of uncertainties at different times, the entropic state will share completely different properties explained as follows:

i. When the environment is in the absence of uncertainty, $h_{k|k}$, will always be positive because of the probabilistic representation of the uncertainties.

ii. Under the presence of uncertainties or cyber-attack, $h_{k|k}$, can become negative. To that end a suitable threshold γ can be chosen for which, when $h_{k|k}$ will be less than γ, this would imply presence of attack.

The *trace* operation can also be instead of using the determinant. In that case, the equation is then

$$h_{k|k} = \frac{\mathrm{Tr}\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{Y}_k\}^2)\}}{\mathrm{Tr}\{\mathbf{P}_{k|k-1}\}} \tag{2.4.17}$$

where Tr represents trace. When (2.4.17) is used, (2.4.16) does not need to be implemented since it involves the sum of the diagonal elements.

2.4.4 Executive

Conceptually the executive is the most important aspect of the CDS as it is solely responsible for control. To that end, it consists of reinforcement learning and cognitive control, which can be further subdivided into the action space, planner, working memory and policy.

2.4.4.1 Reinforcement Learning: Bayes-UCB

While the output of the feedback channel is the entropic state during the PAC, it also produces another output known as the internal rewards during the planning stages involved during Reinforcement Learning (RL) [40]. Before elaborating further, it is to be highlighted that the RL in the CDS relies on the current entropic state. This in turn is used to optimize an objective function for optimal control in the network. In order to explain how every other component of the CDS comes together in the executive, the Bayes-UCB [41] RL algorithm will be briefly covered.

Bayes-UCB represents the current state of the art from a class of multi-armed bandit algorithms called UCB algorithms [42], which are founded on the principle of optimism

in the face of uncertainty. In this Bayesian approach to the multi-armed bandit model, the estimate of the reward distribution for each action is updated using the usual Bayesian method. The action to be performed is then selected based on the action which has the highest reward. Thus, the Bayes-UCB algorithm is an index policy that uses the prior distribution to pick a dynamic quantile of the posterior estimates for the index for each action. Therefore, at each time t, Bayes-UCB selects the action A_t that satisfies the following condition:

$$A_t = \underset{a}{\operatorname{argmax}}\, q_a(t) = Q(1 - \frac{1}{t(\log(t))^c}, \lambda_a^{t-1}) \tag{2.4.18}$$

where $Q(\alpha, \pi)$ refers to the quantile of order α of the distribution π. By assuming that the rewards follow a Bernoulli distribution, and when the prior distribution of each action is Beta(1,1), we can rewrite (2.4.18) as [43]:

$$\begin{aligned} A_t &= \underset{a}{\operatorname{argmax}}\, q_a(t) = Q(1 - \frac{1}{t(\log(t))^c}; \\ &\quad \text{Beta}(S_a(t) + 1, N_a(t) - S_a(t) + 1) \end{aligned} \tag{2.4.19}$$

To keep (2.4.19) consistent with the notation used so far, it will be re-written as:

$$\begin{aligned} A_k &= \underset{a}{\operatorname{argmax}}\, q_a(k) = Q(1 - \frac{1}{k(\log(k))^c}; \\ &\quad \text{Beta}(S_a(k) + 1, N_a(k) - S_a(k) + 1) \end{aligned} \tag{2.4.20}$$

where k is the PAC cycle number, S_a is the cumulative reward for action a, N_a is the number of times action a has been chosen and c is real parameter. From a computational viewpoint, to be consistent with the CDS, the quantile values of the

different picked actions that are updated during each PAC share the same consistency as the value function which is described in [6] for the CDS. While the CDS is an entity that aims to mimic the brain as close as possible, it is to be pointed out that Bayes-UCB shares many traits to decision making in humans [44]. Realizing that our architecture comprises of three different models namely, the system configuration **H** of the power grid, the generative model of the perceptor and the process model in the Kalman filter, the next question that we ask ourselves is how do we bring all three altogether for optimal state estimation with RL? This will be explained in the next section relating to Cognitive Control.

2.4.4.2 Cognitive Control

CC is a special part of the CDS as it builds on every other components, described so far, for a goal-oriented action on the system, which is the SG in this instance. In order to do so, CC consists of two important modules namely the *planner* and the *policy*. The function of the planner is to extract a set of prospective actions from the action-space A to be evaluated during the planning cycles (i.e., shunt cycles [6] in CDS terminology). This is turn allows the policy to grasp the better actions under the influence of attention from one PAC to the next. Similar to how it is done in the human brain, those shunt cycles involve both the perceptor and the executive to account for all prospective actions within each PAC. In the context of the SG, the action space consists of discrete weight values that can be attributed to the different meters. Consequently, under the influence of attention, the CDS will learn which meters are crucial for optimal state estimation and which are detrimental. Hence, referring back to section 2.3, the important meters will be allocated higher weight

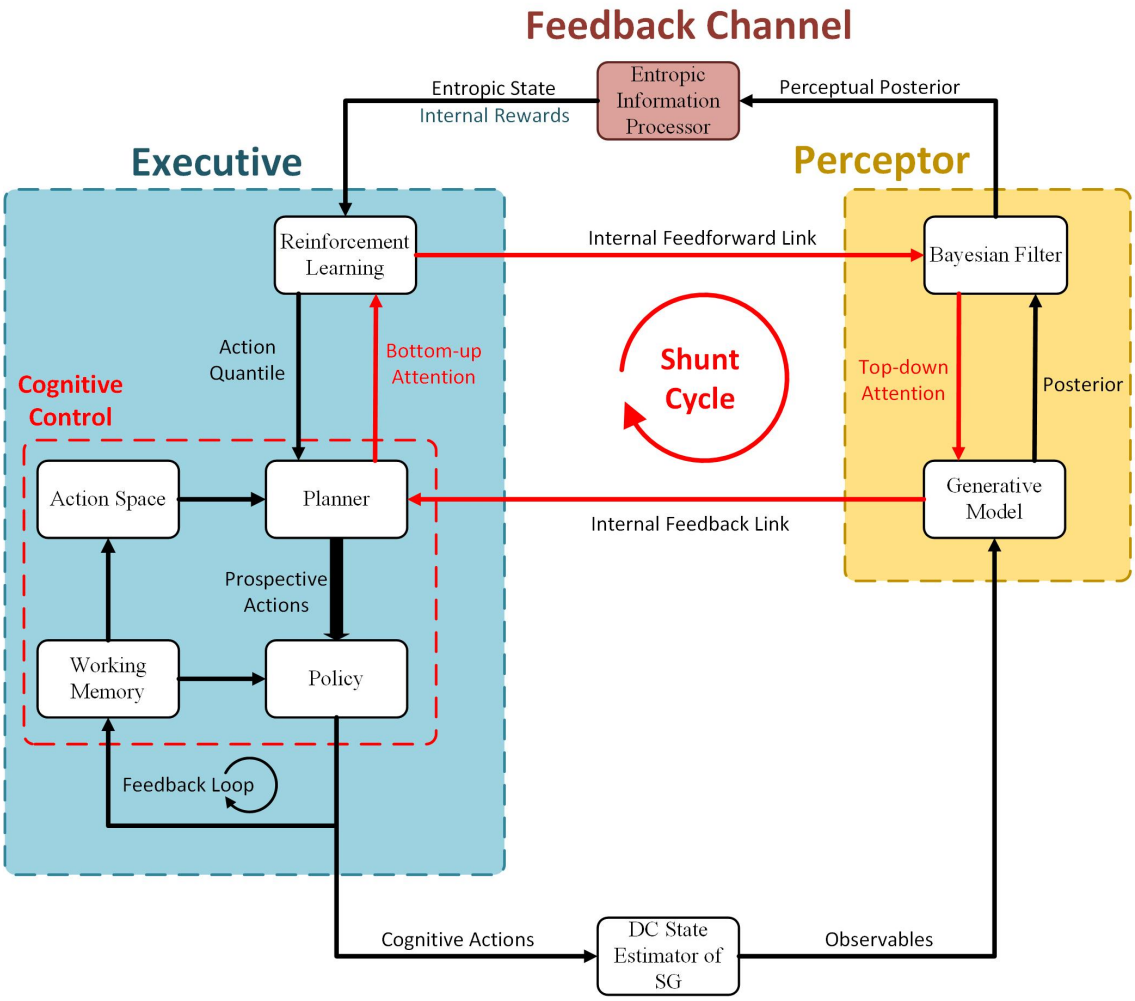

Figure 2.1: Architectural Structure of CDS for the SG.

values and vice-versa for the faulty ones.

Planning for the cognitive control of the system involves all the three models described previously. The planning starts with a prospective action $a_k^{i,j}$ which represents weight value a^i for meter j during cycle k. This action is then applied virtually to the weight matrix $\mathbf{W}$ in (2.3.5). Using this modified weight matrix, the gain is then re-calculated:

$$\mathbf{G}_k^p = \mathbf{H'}_k \mathbf{W}_k^{i,j} \mathbf{H}_k \tag{2.4.21}$$

where $\mathbf{G}_k^p$ denotes the planned gain and $\mathbf{W}_k^{i,j}$ is modified weight matrix where meter j's weight value has been replaced by a^i. The new predicted state estimate is then:

$$\hat{\mathbf{x}}_k^p = (\mathbf{G}_k^p)^{-1}\mathbf{H'}_k\mathbf{W}_k^{i,j}\mathbf{z}_k \tag{2.4.22}$$

where $\hat{\mathbf{x}}_k^p$ refers to the planned state estimate if the weight matrix had those values. A new estimate of the planned cumulative sum involving $\hat{\mathbf{x}}_k^p$ is then calculated:

$$\mathbf{B}_k^p = \sum_{i=k-L}^{k-1} \mathbf{x}_i + \hat{\mathbf{x}}_k^p \tag{2.4.23}$$

where $\mathbf{B}_k^p$ is planned cumulative sum involving $\hat{\mathbf{x}}_k^p$ instead of $\hat{\mathbf{x}}_k$. On the basis of this new cumulative sum, a planned entropic state $h_{k|k}^p$ is calculated using:

$$h_{k|k}^p = \frac{\det\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{B}_k^p\}^2)\}}{\det\{\mathbf{P}_{k|k-1}\}} \tag{2.4.24}$$

As mentioned previously, this is followed by:

$$h_{k|k}^p = \begin{cases} -h_{k|k}^p, & \text{if } d_k^p > 1. \\ h_{k|k}^p, & \text{otherwise.} \end{cases} \tag{2.4.25}$$

where d_k^p is the number of negative elements along the diagonal of $h_{k|k}$ during cycle k. Alternatively the *trace* operation from (2.4.17) can also be used instead. In that case *trace* should replace *det* throughout the structure. From (2.4.24) it is inferred that any uncertainties in the environment, whether it is stochastic or deterministic, will cause the output the generative model to deviate from the hidden state estimated by

the Kalman filter. As a result, $h^p_{k|k}$ will be closer to the optimal value of 1 when the output state of the DC state estimator is closer to the divergence calculated by the Kalman filter.

2.4.4.3 Internal Rewards

With equations that define the different steps involved during the shunt cycles, which are reciprocally coupled to the PAC, the stage is now set to describe the calculation of internal rewards for RL. The hypothesized internal reward, $r^{i,j}_k$, associated with each prospective action $a^{i,j}_k$, for cycle k can be written as:

$$r^{i,j}_k = h^p_{k|k} - h_{k|k} \tag{2.4.26}$$

Consequently, the RL algorithm attempts to minimize the disturbance or uncertainty present in the system from cycle to cycle. This can also be viewed from the Kalman filter viewpoint whereby the system is aiming to restrict the amount of uncertainty during the state estimation process within the range of the uncertainty computed by the filter. Hence, it can be seen that the CDS, as defined in this specific architecture, learns from the past and present actions to pick the best actions for the future. This is facilitated through the role of the working memory. The working memory holds temporarily the actions that holds the highest quantile as defined in (2.4.19) for the different meters and applies it to the system. During the shunt cycles, as the prospective actions are evaluated and their respective quantiles are updated, if one those actions achieves a higher quantile value than the one for its respective meter in the working memory, then this particular action replaces the previously best action stored in the working memory. From a Bandit perspective, this can be considered as

a Contextual Bandit problem whereby each cycle presents new situations to be faced and each action performed on the system brings the system configuration to a new set point that the RL algorithm will have to adapt to and so on.

2.4.4.4 Complete Algorithm

With the detailed description of the different components of the CDS for the SG, the stage is now set for an overall definition of the algorithm for the cognitive controller. Algorithm 1 demonstrates the steps involved during each PAC of the system. For convenience, Table 2.1 shows all the notations mentioned so far with respect to the algorithm. Due to the frequentist approach of the Bayes UCB algorithm and bounded rewards of the Bernoulli reward distribution, some modifications need to be made to keep the algorithm consistent with CC. To that end, in order to update the quantile of the prospective actions, another conditional variable must be used instead of the real cumulative reward. The variable defined as *BayesReward* is thus used to fullfill this role as shown in lines 34 to 39 where different conditions are specified. In lines 34 and 35, the threshold, f, is used to prevent the the cumulative rewards from becoming overly negatively saturated. Consequently, consider an action which was considered inappropriate most of the time because of the accumulated unbounded negative rewards over time. If a situation arises whereby that said action is now appropriate, it will take a longer time for the Bayes UCB to build up the quantile to default to that action. By introducing f to bound the cumulative negative rewards, it will take a shorter time for its quantile to build up and achieve the highest quantile to be picked for that particular situation. This f grants a come-back opportunity to those actions. As mentioned previously, since the rewards are bounded in [0,1], the

purpose of line 35 is to update the quantile with a zero-valued *BayesReward* instead of the current negative cumulative reward which is not compabitible with the algorithm. For the same reason, this is implementated again on line 37 for that different condition where the cumulative reward is negative but not yet saturated. Lastly, in the advent that the cumulative reward is positive, then this value is used to update the quantile as shown in line 39. The use of the entropic state formula in (2.4.24) to calculate the internal rewards maintains the consistency of the Bernoulli bounded rewards. Referring to line 39, in the case that the rewards are positive, it is guaranteed they will lie in [0,1]. However, if the internal rewards are negative, lines 34 to 39 are meant to deal with the consequences of the accumulated negative internal rewards during planning. In order to provide some initial stability to the cognitive control algorithm when it starts out, the quantiles relating to the default configuration of the weights are initially biased with a value of α. Moreover, the concept of cognitive confidence cycles, n_{cc}, is also introduced. During the n_{cc} cycles, the cognitive controller will learn from the prospective actions and gain a first impression of what the best actions are by calculating their respective quantile values. However those actions are not applied during those cycles. When the n_{cc} cycles are elapsed, the algorithm will start applying those actions to the system using the initial information gained during these n_{cc} cycles. Lastly, the tailoring of this algorithm for the SG allows us to bypass one of the major limitations of RL algorithms as multiple actions relating to the different to meters can applied during each PAC.

There are currently many popular detection methods that have been developped that relies on special meaurements. Among those we will contrast our method with [48], which is a detection scheme using K-means clustering on CSSVC (Control Signal

from the controller to the Static Var Compensator) and NVSI (Node Voltage Stability Index) measurements, and [49], that involves the use of SVD (Singular Value Decomposition) and convex-optimization in order exploit the low-rank property of the measurement matrix comprising of PMU measurements. Compared to those methods, the technique, proposed in this paper, has higher accuracy and is less prone to false positives since the entropic state will decrease as the attack is continuously propagated throughout the generative model. However, the tradeoff is a slight increase in dectection time. This will be shown in the second part of computational experiments section. The method is also less complex and less computationally intensive as the perceptor operates on an indirect model of the states themselves compared to [48] and [49], which rely on some special measurements rather than the traditional readings from RTUs on the field. By performing the detection through the workflow described in Fig. 2.1, our method is more resistant to the curse of dimensionality which becomes a problem as the size of measurement matrix for K-means clustering or exploiting the low-rank property is scaled up for bigger networks. Lastly, our method also has an adaptive property through the Kalman filter which makes the entropic state dynamic as the system evolves. Thus, this provides higher accuracy at a small cost of increased detection time such as 5 to 15 PAC depending on the intensity of the attack. In the next section of this paper, we will show that the control and FDI attack detection aspects can work independent of each other in two separate experiments. Compared to [48] and [49], the main parameters of interest for the attack detection criteria are the parameters in the $\mathbf{Q}$ matrix of the Kalman filter that dictates its sensitivity to fluctuations or disturbances when applied in the way described in this paper. Consequently it is less complex compared to the two methods referenced. Finally, by

iteratively optimizing the weights for state estimation, this adds some non-linearity to the state estimation process and makes it more sensitive towards non-probabilistic disturbances, such as the FDI attack in the case of this paper. Hence, the method is very robust for cyber-attack detection.

2.5 Computational Experiments

In this section, two different experiments were carried out to demonstrate the capability of the architecture which was just described. The first experiment pertains to CC as a Bad Data Detector (BDD) and corrector. In the second experiment, it will be shown how the entropic state can be a metric for cyber-attack detection. In both experiments, the data used to simulate both network configurations comes from the case files in *MATPOWER* [46] which is an Electric Power System Simulation and Optimization Tools for MATLAB and Octave. In the first experiment, a 4-bus network will be considered. Since this is a small network with a small number of states, this experiment provides a greater insight of how the generative models, states, entropic states, weight values and mean squared errors are evolving through Fig. 2.3, 2.4 and 2.5. While the first experiment is dedicated to the control aspect of the architecture, the second experiment is focused on FDI attack detection in a bigger grid. Since IEEE bus networks have been used as benchmarks for simulations in the other papers referenced in this paper, the IEEE 14-bus network was chosen for the second experiment.

Table 2.1: Summary of Notations (Part 1 of 2)

Notation	Definition
M	Total number of PAC cycles
N	Total number of shunt/planning cycles
n_{cc}	Number of cognitive confidence cycles
$\mathbf{z}_k$	Vector of measurements taken at cycle k
$\mathbf{H}_k$	System configuration matrix at cycle k
$\mathbf{x}_k$	Vector of calculated states by DC state estimator at cycle k
$\mathbf{W}_k$	Weight matrix at cycle k
$\mathbf{B}_k$	Vector retaining the cumulative sum of the states at cycle k
L	Window over which the past states is being accumulated
$\hat{\mathbf{B}}_{k\|k}$	Filtered estimate of the cumulative sum from the generative model at cycle k
$\mathbf{P}_{k\|k}$	Process error covariance matrix at cycle k
$\hat{\mathbf{B}}_{k+1\|k}$	Predicted estimate of the cumulative sum for cycle k+1
$\hat{\mathbf{P}}_{k+1\|k}$	Predicted error covariance matrix for cycle k+1
$\hat{\mathbf{B}}_{k\|k-1}$	Predicted estimate of the cumulative sum for current cycle k which was calculated during the previous cycle $k-1$
$\mathbf{P}_{k\|k-1}$	Predicted error covariance matrix for current cycle k which was calculated during the previous cycle $k-1$
$h_{k\|k}$	Entropic state at cycle k
d_k	Number of negative elements along the diagonal of $h_{k\|k}$
γ	Threshold for attack detection
A	Set of all possible actions stored in action space
A_1	Set of selected prospective actions for planning

Table 2.2: Summary of Notations (Part 2 of 2)

Notation	Definition
$\mathbf{W}_k^{i,j}$	Modified weight matrix where meter j's weight value has been replaced by a^i during planning
$\mathbf{G}_k^p$	Hypothesized gain during planning
$\hat{\mathbf{x}}_k^p$	Hypothesized state estimate during plannings
$\mathbf{B}_k^p$	Hypothesized cumulative sum involving $\hat{\mathbf{x}}_k^p$ during planning
$h_{k\|k}^p$	Hypothesized entropic state during planning
d_k^p	Number of negative elements along the diagonal of $h_{k\|k}^p$
$r_k^{i,j}$	Internal reward associated with each prospective action $a_k^{i,j}$
a_m	Action stored in working memory
$S_a^{i,j}$	Cumulative reward for action $a^{i,j}$
$N_a^{i,j}$	Number of times action $a^{i,j}$ has been chosen
$Q_a^{i,j}$	Quantile of action $a^{i,j}$
c	real parameter for Bayes-UCB
u	Number of prospective actions to select from action space
f	Negative rewards saturation threshold
α	Quantile initial bias

Algorithm 1: Complete Algorithm for implementation of Cognitive Control in the defined Structure

Initialization:
memstate := short-term memory for storing L past outputs from the DC state estimator
A_1 := set of selected unique cognitive actions for planning
a_m := load working memory with default configuration of weights for all meters
BayesReward := short-term memory for storing cumulative entropic reward
Set action quantiles of the default configuration to α
$\mathbf{B}_0$, $\hat{\mathbf{B}}_{1|0}$, $\mathbf{P}_{1|0}$, $\mathbf{W}_0$, n_{cc}, f, γ
$a_m \leftarrow a_0$
$k \leftarrow 1$

Begin **while** $k \leq M$ **:**
    DC State Estimation:
    Caculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
    $memstate \leftarrow \mathbf{z}_k$

    Generative Model:
    Calculate $\mathbf{B}_k$ using the past L outputs from *memstate*

    Bayesian Filter/Kalman Filter:
    Update Steps:
    Calculate $\mathbf{K}_k$, Calculate $\hat{\mathbf{B}}_{k|k}$, $\mathbf{P}_{k|k}$
    Prediction Steps:
    Calculate $\hat{\mathbf{B}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$

    Feedback Channel:
    Calcualte $h_{k|k}$

    Executive:
    Planning:
    for *all cognitive actions* $a_k^{i,j} \in A_1$
        Calculate corresponding $\mathbf{B}_k^p$
        Calculate corresponding $h_{k|k}^p$

        Internal reward:
        Calculate $r_k^{i,j}$

        (Continued on next page)

(*Continuing as from line 29 onwards*)
Learning:
Update $S_a^{i,j}$ for $a_k^{i,j}$
Update $N_a^{i,j}$ for $a_k^{i,j}$
if $S_a^{i,j} < f$ **:**
    $S_a^{i,j} = f$
    $BayesReward = 0$
else if $S_a^{i,j} < 0$ **:**
    $BayesReward = 0$
else:
    $BayesReward = S_a^{i,j}$
Update $Q_a^{i,j}$ for $a_k^{i,j}$ using $BayesReward$
if $Q_a^{i,j} > Q_{a_m}^{i,j}$ **:**
    $a_m^{i,j} \leftarrow a_k^{i,j}$
if $k > n_{cc}$ **:**
    Policy:
    Apply a_m to the DC state estimator
$k = k + 1$

2.5.1 Cognitive Control for BDD

In this first experiment a 4 bus, 2-generator transmission network case from[15] will be simulated as shown in Fig. 2.2. The state values relating to the three buses was obtained by solving the power flow equations using *MATPOWER* [46] for the case data. Using those state values, the seven meter signals were then obtained through the use of (2.3.3). Using the mean of those signals, their noisy counterparts were generated with a signal-to-noise ratio (SNR) of 20 dB to create $\mathbf{z}$. For the 4 bus, 2-generator transmission network in Fig. 2.2, the network configuration matrix is:

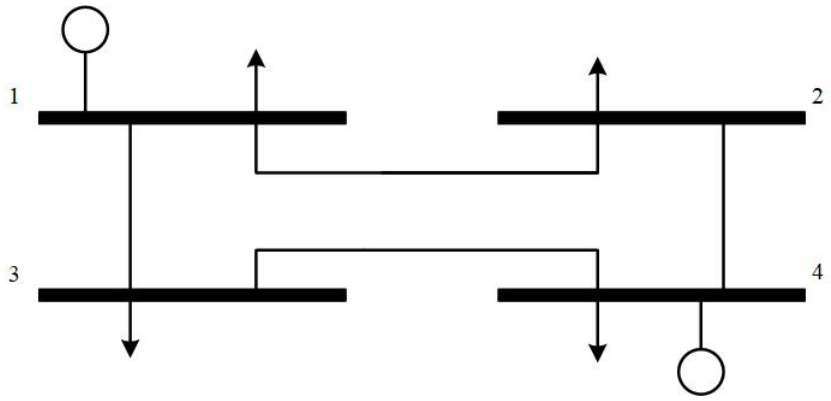

Figure 2.2: Line diagram for 4-bus, 2-generator transmission network case from[15].

$$\mathbf{H} = \begin{bmatrix} 46.72 & 0 & -26.88 \\ 0 & 42.6 & -15.72 \\ -26.88 & -15.72 & 42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix}$$

For the generative model in the perceptor, L was assigned a value of 20. The initial estimates for the Kalman filter is set to 0 for all the three incoming state states from the DC state estimator. The diagonal elements of the diagonal matrix $\mathbf{Q}$ were set to 6.25e-04 and the diagonal elements of the $\mathbf{R}$ to 0.01. The goal of this experiment is to show the adaptability of this architecture to changing conditions while still carrying out optimal state estimation. To that end, the system is slowly driven to unobservability by changing the SNR of the following meters to 10 dB at the mentioned times: $\mathbf{t = 400s}$ for meter 3, $\mathbf{t = 700s}$ for meter 1, $\mathbf{t = 1000s}$ for meter 7, $\mathbf{t = 1300s}$ for meter 2 and $\mathbf{t = 1600s}$ for meter 2. This setup can seen as meter

malfunction or a random attack, whereby the attacker has access to limited meters to accomplish his task. The action space consists of a total of 28 actions whereby each meter's weight can set to any of the following values: 1 50 100 150. f was assigned a value of -0.05 and $\alpha = 0.2$. Lastly 15 planning/shunt cycles are evaluated during every PAC. In the experiment CC is started at $\mathbf{t = 200s}$ with 30 cognitive confidence cycles. The results are shown in Fig. 2.3, 2.4 and 2.5. CC is not started immediately at $\mathbf{t = 0s}$ as the Kalman filter needs to to settle on the track first for the algorithm to be effective.

As it can be seen in the figures, CC allows the network to be dynamic while choosing the best set of meters simultaneously and assigning the best weights for optimal state estimation. Thus, a direct consequence of using the cognitive controller is that it gains the distinctive capability to learn from the current and past cycles to pick the best possible set of actions for the future. Hence in Fig. 2.4, it is shown that when the first meter malfunction occurs at $\mathbf{t = 400s}$, it takes only a couple of PAC cycles to learn from the new situation and decrease the weight attributed to meter 3. Adding to that, the weight values for all the meters are not the same; the cognitive controller adapts to the probabilistic characteristic of the noisy signals. Moreover the algorithm also allows the application of more than one action at each PAC as shown by the simultaneous change of the weight values of the different meters as the disturbance is inflicted to the system. In Fig. 2.3, we can evaluate the impacts of the actions taken. At $\mathbf{t = 400s}$, the increase in the amount on noise has the greatest effect on state 3 and state 1 to a lesser extent. However, through the learning process of the cognitive controller, the actions applied to the system restores the states signal to how it was previously before the malfunction occured. Furthermore, because of the dynamic

nature of this new architecture, when the meter 1 malfunction occurs at $\mathbf{t = 700s}$, we can see that the same course of action is applied to meter 1 similar to meter 3. Although we can see its weight is lifted again after that, CC is quick to bring it under control. The reason for the earlier increase of the weight of meter 1 can be attributed to the probabilistic origin of the noise coupled with the frequentist approach of the Bayes UCB. Addtionally, prior to this malfunction, the weights of the other meters have also been altered as representation of the cognitive ability of the controller to trust certain meters more than the others. As the other consequent malfunctions are triggered, the appropriate actions are applied to keep the system under control, even during the last case involving unobservability. As a matter of fact, CC learns to pick the best set of meters for state estimation on the go. The state of the grid can also be evaluated using $h_{k|k}$. $h_{k|k}$ decreases gradually throughout this experiment as more and more disturbances are applied to the grid. Nevertheless, the cognitive controller reacts quickly to this situation by taking the necessary steps to bring the entropic state as close as possible to 1. Fig. 2.5 shows the mean-squared error (MSE) of the estimated state compared to the real state without noise. It can be seen that the cognitive controller achieves a lower MSE than the conventional state detector without the bad data detector on the overall. Lastly, the Chi-squares test was not implemented in this experiment as the latter is rooted in the statistical properties of the signals while the approach of the architecture presented is based on the principle of cognition rooted in the brain.

Since fluctuations in the voltage angles are very common disturbances in power systems, the experiment gives some insight on how the algorithm differentiates among these. When a disturbance is inflicted on the states, the latter is propagated to the

generative model, which consequently affects the determination of the entropic state. As the entropic state is an embodiment of how the grid is performing, it was shown in the experiment that those fluctuations would cause a decrease in the entropic state. However, the goal of the algorithm is to always optimize the entropic state towards a value of 1. By optimizing $h_{k|k}$, CC is trying to minimize the amount of fluctuations in the system and maintain the evolution of the states in a controlled manner. Referring back to Fig. 2.3, it can be seen that when the meter malfunctions occurred, this in turn caused larger deviations in some of the states. At the same time, it can be seen that there was a dip in $h_{k|k}$. However, CC is able to quickly raise $h_{k|k}$ and bring state estimation under control. The last part of Fig. 2.5 shows that the algorithm is able to keep the amount of fluctuations under control even in a case where the system would be considered unobservable.

2.5.2 Cyber-Attack Detection

In this section, the dual property of the entropic state for FDI cyber-attack detection will be demonstrated. Previously, it was shown how the latter is an objective function for the normal running of CC under the absence of uncertainty whereby it is always positive. However, when the presense of uncertainties are no longer probabilistic, such as when an attack takes place, the entropic state will also enable early detection of such attacks. Although many specialized attacks such as replay attack or Distributed Denial of Service (DDoS) attack exist, four broad categories of FDI attacks will be considered as follows:

i. Case 1: Here we assume that the intruder has perfect knowledge of the network configuration **H** and full access to meters to commit the perfect FDI attack as

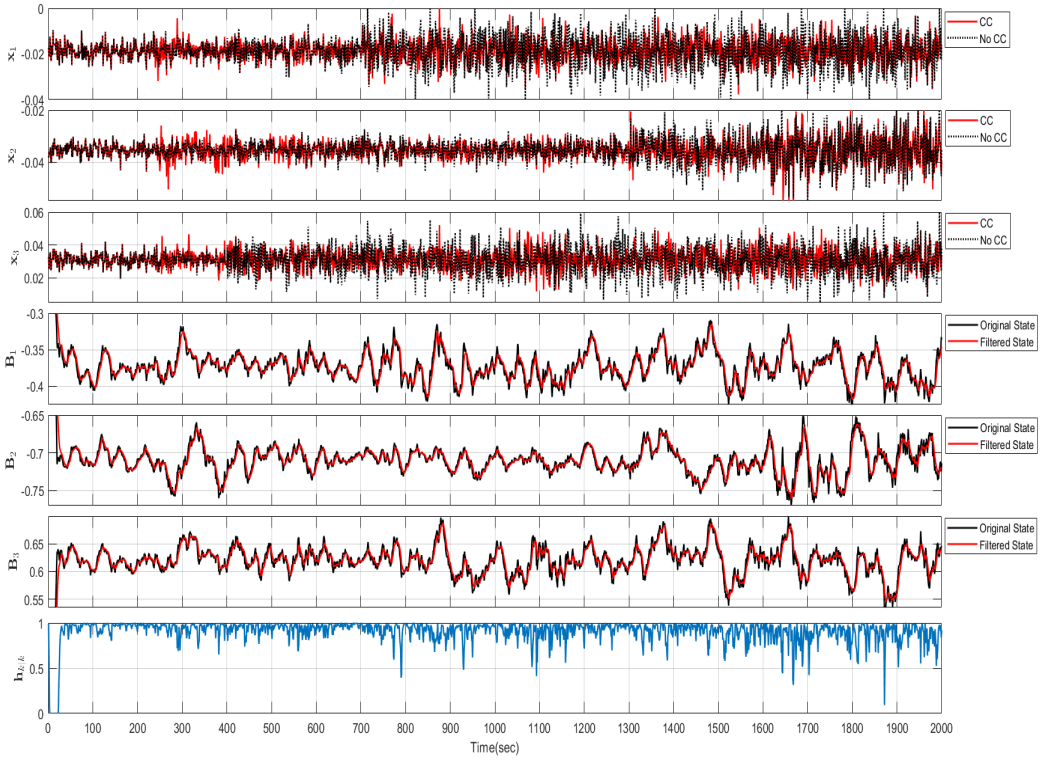

Figure 2.3: Graphs of States, Generative models and Entropic State

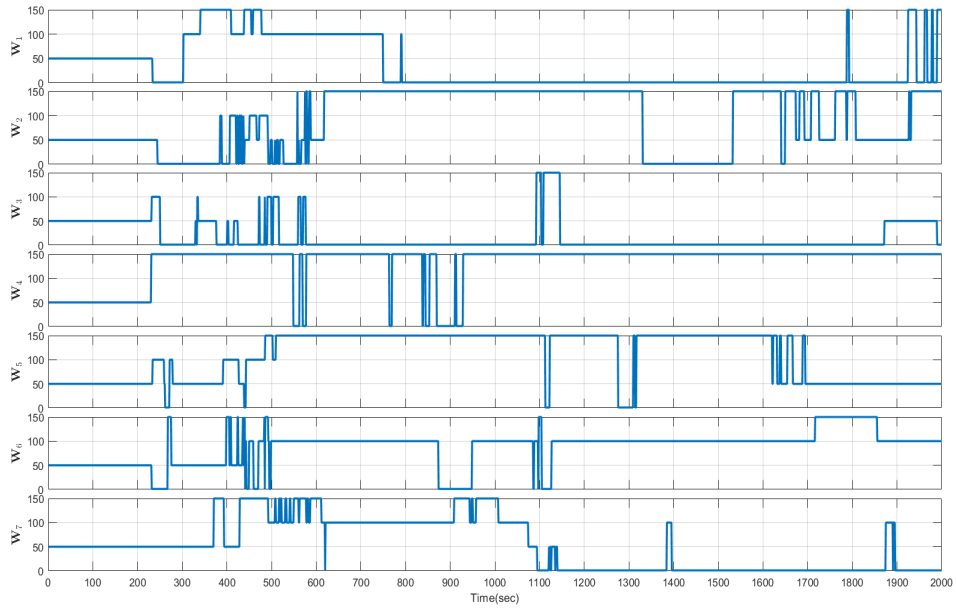

Figure 2.4: Graphs of weight values of the meters with time

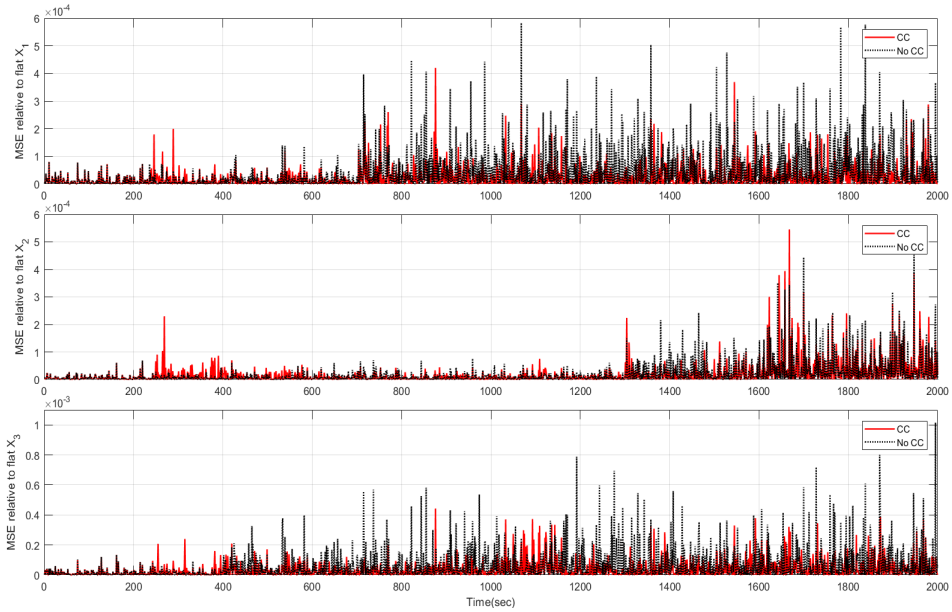

Figure 2.5: Graphs of the MSE of the states compared to flat mean without noise

mentioned in section 2.3. The remaining cases consider more realistic scenarios whereby the hacker faces some constrictions.

ii. Case 2: In this scenario, the intruder still has full knowledge of **H** but has limited access to meters in the grid. To simulate this attack, some of the rows of the attack vector **a** are zeroed to represent the inability to access those sensors.

iii. Case 3: Here the circumstances of case 2 are flipped around; the intruder has access to all the meters but incomplete knowledge of **H**. To carry out the attack, some of the rows of **H** are zeroed as an indication of the lack of information.

iv. Case 4: Finally, a rogue attack combining case 2 and 3 is considered. The attacker has both imperfect knowledge of **H** and constrained access to the sensors in the grid. In order to simulate this attack, some rows of **a** and **H** are zeroed out to represent the attacker's limitations.

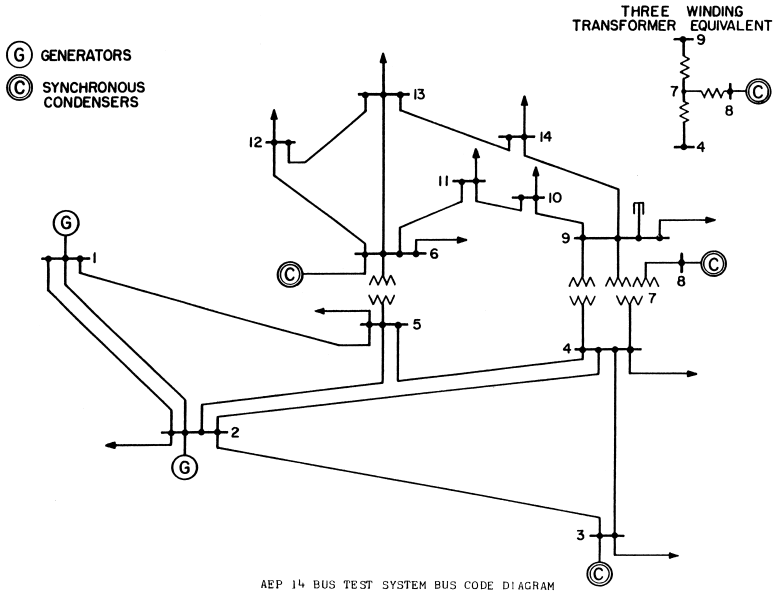

Figure 2.6: Line diagram for the IEEE 14 Bus network.

The mentioned attacks in those different situations will be simulated on the IEEE 14 bus network as shown in Fig. 2.6. In all of the mentioned cases, the hacker's goal is to deflect the value of one the states by a value of 0.15 radians. Since attack data is not publicly available, the parameters in the *MATPOWER* package will be used to simulate the IEEE 14 bus network.

In all four attack cases, the attack is started at $\mathbf{t = 500s}$. The same conditions were used as in the precedent 4 bus case except for the following: the diagonal elements of the diagonal matrix $\mathbf{Q}$ was set to 4.9e-03, $\gamma = 0.25$ and the *trace* operation will be applied instead of *det* as it leads to higher values of h_k which is better for the sake of illustrability. Additionally, the property of h_k will demonstrated as a stand-alone utility in the absence of CC. While CC is originally defined for tackling control when the uncertainties are probabilistic and h_k is positive, the CDS has to expand its structure its to include CRC to be able to bring risk under the control in the

presence of the cyber-attacks. The discussion relating to the implementation of the CRC to this architecture will be left for another day. The results pertaining to the simulation of the attacks presented earlier are shown in Fig.2.7, 2.8, 2.9 and 2.10. In all four cases, by assigning a suitable γ, the attack was detected. Furthermore, it can also be seen that as the hacker has less and less information on the current grid, it becomes easier to detect the deflection as the entropic state becomes more negative. The results also display the efficiency of the generative model, whereby the attack propagates thoughout the cumulative sum upto a certain point before the Kalman filter gets back on the current track. This propagation causes h_k to become increasingly negative which consequently lends the property of detection.

All the computational experiments were carried out on a system running Windows 10 with an Intel i7-8750H processor. The computational running time of the first experiment was **1.8s** and the second experiment took around **0.13s** for each of the cases mentioned. From the experiments, it was found that the use of *trace* for the calculations was faster than the one involving the determinant of the matrix. If the CDS architecture proposed in this paper is applied in a medium or large-scale power system, the computational complexity will be lesser compared to the other current detection methods, such as the ones mentioned earlier. Moreover, the application of the CDS for an application such as the SG is revolutionary as it is a dual system catering to both the control and attack detection aspects of the SG. The main parameter of interest that needs to be scaled up for a more complex grid will be the number of shunt cycles since more meters will have to be evaluated. Nevertheless, it recommended to keep the action space small so as to make planned rewards, during planning, distinguishable from each other. Another important hyper-parameter in

the system, especially for FDI attack detection, are the values in the $\mathbf{Q}$ matrix. In the first experiment, the diagonal elements of matrix $\mathbf{Q}$ was set as 6.25e-4 while in the second experiment, the values were set to 4.9e-3. The choice of using determinant or the *trace* operation for the calculations involved in the CDS is very important when choosing $\mathbf{Q}$. Furthermore, unlike many tracking applications such as the simulation carried out in [5], which was supported by a mathematical formulation [50], this is not the case in our system. Thus, the contents of $\mathbf{Q}$ has be defined by the designer depending on the required sensitivity of the system towards disturbances. In order to find proper values for $\mathbf{Q}$, prior simulations can be carried out using past historical data. Usually, it is recommended to start with very small values, like the ones used in the simulations carried out in this paper, and then tuning until the desired performance is obtained. Lastly, as the SG is scaled up, that hyper-parameter will have to be increased to reflect the circumstances of a bigger power system.

In the second experiment involving the four attack cases, the detection time was around 5-10 PAC. However, depending on the intensity of the attack, this detection time can be as small as 5-10 PAC and can go up to 10-15 PAC for slowly evolving attacks such as a ramp attack. Thus, the PAC will depend on the sampling time of the DC state estimator; if the states are being calculated every 5s, then it can take 25s to 50s to detect the attack. While this may slightly long, it is a small cost to pay for higher detection accuracy. As mentioned previously, this time can be cut down either by tuning the Q values or by increasing the frequency at which the state is sampled. The threshold for attack detection can also be increased.

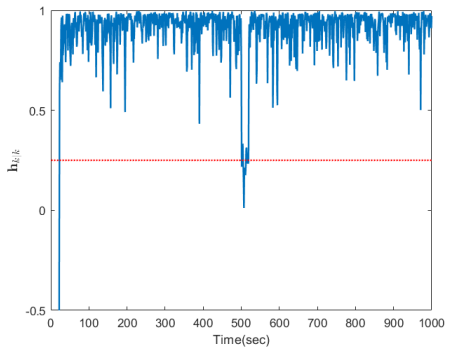

Figure 2.7: Case 1.

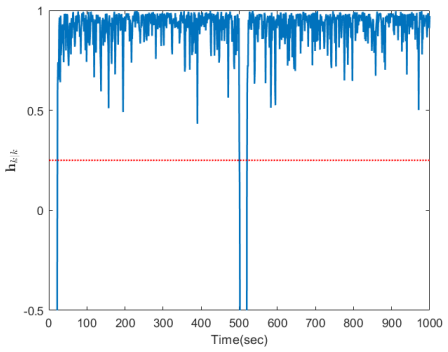

Figure 2.8: Case 2.

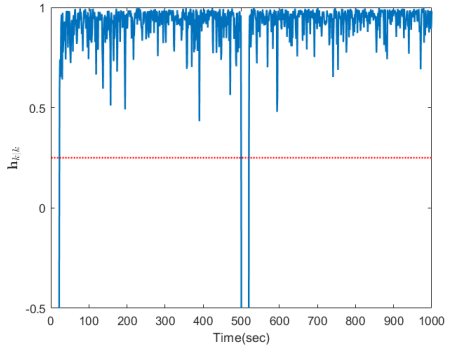

Figure 2.9: Case 3.

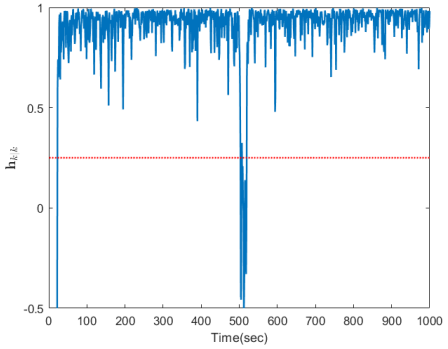

Figure 2.10: Case 4.

2.6 Conclusion

This paper is innovative on the following three accounts:

i. This is the first time where we were able the bring the CDS, which is rooted in the brain, and the SG together for a new architecture that is able to handle the new problems that will be facing the grid in the coming years as everything becomes more and more interconnected.

ii. The cognitive control algorithm presented is novel in the sense that multiple

actions can be applied to the system during each PAC while maintaining stability. Moreover, learning in the RL algorithm is carried out immediately in a Bayesian fashion utilizing information from the current cycle and the successive past cycles. The CDS which was described in the earlier papers referenced was targeted towards simpler systems whereby learning was assessed only once and was limited to one performed action during every PAC. The computational experiments and results in this paper show that this new algorithm, specific towards the SG, can overcome the previous limitations.

iii. The architecture serves the dual purpose of a new kind control, which is based on cognition, and cyber-attack detection.

Just like **Q** represents the process uncertainty in the Kalman filter equations, the RL algorithm in this architecture also inherits this property. More specifically, the RL algorithm regulates the amount of the uncertainty in the system dynamically from cycle to cycle. In a real SG network involving thousands of meters, the inverse operation during state estimation and planning can be costly. To that end, a function approximator such as a neural network [47] can be used to accelerate the computation. On the other hand, the use of a RL algorithm such as the Bayes UCB presented this paper, which is based on a frequentist approach, will increase the amount of time for the cognitive controller to learn the best configuration as the network is scaled up and malfunctions occur. As a solution, the RL algorithm can be tweaked to make it more sensitive to changes to decrease its response time. Additonally, although this was touched lightly in the paper, the number of planning (shunt) cycles is very important for the learning process. However, the number of those cycles should neither be too small nor too big to prevent confusion in the cognitive controller as

there might be more than one optimal configuration of a huge grid. Lastly, the focus of this paper was to give an insight of this architecture from a control point of view. Nevertheless, once a cyber-attack is detected, the CDS has to expand its structure to include CRC to handle control in such situations. This discussion will be left for another day. Since the DC estimation model was considered in this paper, future research in this topic can be oriented towards the application of this architecture in the AC estimation model, which involves the reactive components. Since the AC state estimation procedure is a recursive procedure that is different from the DC model, a new process that embodies the AC state estimator and the perceptor will have to be formulated to make the process more computer efficient. Additionally, as mentioned in the introduction, the architecture of the CDS in Fig. 2.1, tailored for the SG, has also been previously applied in other areas such as cognitive radar and cognitive radio. Nevertheless, the architecture proposed can also be extended to other applications subjected to disturbances and/or uncertainties, such as Vehicular Radar Systems where optimal state estimation and tracking are very important. However, for these different applications, the mathematical formulations of the perceptor and the executive will have to be tailored accordingly depending on the end goal of the system.

Bibliography

[1] S. Haykin, "Cognitive dynamic systems [Point of view]," Proc. IEEE, vol. 94, no. 11, pp. 1910–1911, Nov. 2006.

[2] S. Haykin, "Cognitive Dynamic Systems: Radar, Control, and Radio", Proc. IEEE, Point of View Article, vol. 100, no. 7, pp. 2095-2103, July 2012.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Sel. Areas Commun., vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] S. Haykin, "Cognitive radar: a way of the future." IEEE signal processing magazine 23.1 (2006): 30-40.

[5] M. Fatemi and S. Haykin, "Cognitive control: Theory and application," IEEE Access, vol. 2, pp. 698–710, Jun. 2014.

[6] S. Haykin, J. M. Fuster, D. Findlay, and S. Feng, "Cognitive risk control for physical systems," IEEE Access, vol. 5, pp. 14 664–14 679, Jul. 2017.

[7] J. M. Fuster, "Cortex and Mind: Unifying Cognition", Oxford University Press, 2003.

[8] Y. Wang, M. Amin, J. Fu, H. Moussa, "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids", IEEE Access 2017.

[9] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-Physical Systems Security – A Survey", IEEE Internet of Things Journal, vol. PP, no. 99, pp. 1-1, 2017.

[10] J. Hao, R.J Piechocki, D. Kaleshi, et al: 'Sparse malicious false data injection attacks and defense mechanisms in smart grids', IEEE Trans. Ind. Inf.s, 2015, 11, (5), pp. 1–12 (doi: 10.1109/TII.2015.2475695).

[11] X. Fang, S. Misra, G. Xue, D. Yang, "Smart grid - the new and improved power grid: A survey", IEEE Commun. Surveys Tutorials 2012.

[12] S. Sridhar, A. Hahn, M. Govindarasu, "Cyber-physical system security for the electric power grid", Proc. IEEE vol. 99 no. 1 pp. 1-15 Jan. 2012.

[13] F. C. Scheweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model," IEEE Trans. Power App. Syst. , vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.

[14] K. P. V. Priya, J. Bapat, "Bad Data Detection in Smart Grid for AC model", IEEE Indicon 2014.

[15] J J. Grainger and W D. Stevenson JR.,"Power System Analysis 1st Edition", McGraw-Hill Series in Electrical and Computer Engineering, 1994.

[16] Y. Liu, P. Ning, M. Reiter, "False data injection attacks against state estimation in electric power grids", ACM CCS pp. 21-32 2009.

[17] A. Abur and A. Gómez-Expósito,"Power System State Estimation Theory and Implementation", 2004.

[18] A. Monticelli, "State Estimation in Electric Power System A Generalized Approach", Springer Science+Business Media New York, 1999.

[19] Z. Yu, W. Chin, "Blind false data injection attack using PCA approximation method in smart grid", IEEE Trans. Smart Grid vol. 6 no. 3 pp. 1219-1226 May 2015.

[20] J. Kim, L. Tong, and R. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," IEEE Transactions on Signal Processing, vol. 63, no. 5, pp. 1102–1114, March 2015.

[21] L. Liu, M. Esmalifalak, Q. Ding, V. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," IEEE Transactions on Smart Grid, vol. 5, no. 2, pp. 612–621, March 2014.

[22] A. Anwar, A. N. Mahmood, M. Pickering, "Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements", J. Comput. Syst. Sci. vol. 83 no. 1 pp. 58-72 2016.

[23] J. Jiang, Y. Qian, "Defense mechanisms against data injection attacks in smart grid networks", IEEE Commun. Mag. vol. 55 no. 10 pp. 76-82 Oct. 2017.

[24] P. McDaniel, S. McLaughlin, "Security and privacy challenges in the smart grid", IEEE Security Privacy vol. 7 no. 3 pp. 75-77 May/Jun. 2009.

[25] R. Deng, G. Xiao, R. Lu, H. Liang, A. V. Vasilakos, "False data injection on

state estimation in power systems—Attacks impacts and defense: A survey", IEEE Trans. Ind. Informat. vol. 13 no. 2 pp. 411-423 Apr. 2017.

[26] D. Wang, X. Guan, T. Liu, Y. Gu, Y. Sun, Y. Liu, "A survey on bad data injection attack in smart grid", Proc. IEEE PES Asia-Pac. Power Energy Eng. Conf. pp. 1-6 2013.

[27] K. Manandhar, X. J. Cao, F. Hu, Y. Liu, "Combating false data injection attacks in smart grid using kalman filter", Proceedings of International Conference on Computing Networking and Communications Communications and Information Security Symposium pp. 16-20 2014.

[28] K. Manandhar, X. Cao, F. Hu, Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter", IEEE Trans. Control Netw. Syst. vol. 1 no. 4 pp. 370-379 Dec. 2014.

[29] P.Y. Chen, S. Yang, J. A. McCann, J. Lin, X. Yang, "Detection of false data injection attacks in smart-grid systems", IEEE Commun. Mag. vol. 53 no. 2 pp. 206-213 Feb. 2015.

[30] Y. Liu, L. Yan, J. Ren, D. Su, "Research on efficient detection methods for false data injection in smart grid", International Conference on Wireless Communication and Sensor Network (WCSN) pp. 188-192 December 2014.

[31] D. B. Rawat, C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems", IEEE Signal Process. Lett. vol. 22 no. 10 pp. 1652-1656 Oct. 2015.

[32] Y. Gu, T. Liu, D. Wang, X. Guan, Z. Xu, "Bad data detection method for smart grids based on distributed state estimation", Proc. IEEE Int. Conf. Commun. pp. 4483-4487 2013.

[33] Y. Huang et al., "Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis", IEEE Syst. J. vol. 10 no. 2 pp. 532-543 Jun. 2016.

[34] S. Li, Y. Yilmaz, X. Wang, "Quickest detection of false data injection attack in wide-area smart grids", IEEE Trans. Smart Grid vol. 6 no. 6 pp. 2715-2735 Nov. 2015.

[35] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems,". Journal of Basic Engineering, 82: 34–45, 1960

[36] A. S. Debs, R. E. Larson "A dynamic estimator for tracking the state of a power system", IEEE Trans. PAS vol. PAS-89 pp. 1670-1673 September/October 1970.

[37] E. A. Blood, M. D. Ilic, J. Ilic, B. H. Krogh, "A Kalman filter approach to quasi-static state estimation in electric power systems", 38th North American Power Symposium pp. 417-422 2006 2006.

[38] A Saikia, RK Mehta, "Power system static state estimation using Kalman filter algorithm", EDP Sciences. 2016; 7: 1-7.

[39] C. E. Shannon, "A mathematical theory of communication", Bell Syst. Tech. J., vol. 27, no. 3, pp. 379-423, Jul./Oct. 1948.

[40] R. S. Sutton and A. G. Barto, "Reinforcement Learning", Cambridge, MA, USA: MIT Press, 1998.

[41] E. Kaufmann, O. Cappé and A. Garivier, "On Bayesian upper confidence bounds for bandit problems" Proc. Int. Conf. Artif. Intell. Stat. pp. 592-600 2012.

[42] G. Burtini, J. Loeppky, and R. Lawrence,"A survey of online experiment design with the stochastic multi-armed bandit", CoRR, abs/1510.00757, 2015.

[43] E. Kaufmann, "Analysis of bayesian and frequentist strategies for sequential resource allocation", Machine Learning [cs.LG]. Télécom ParisTech, 2014. English. ¡NNT : 2014ENST0056¿. ¡tel-01413183¿

[44] P. Reverdy, V. Srivastava ,N. E. Leonard, "Modeling human decision-making in generalized Gaussian multi-armed bandits", Proc. IEEE vol. 102 no. 4 pp. 544-571 Apr. 2014.

[45] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," IEEE Trans. Autom. Control, vol. 54, no. 6, pp. 1254-1269, Jun. 2009.

[46] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12-19, Feb. 2011.(Digital Object Identifier: 10.1109/TPWRS.2010.2051168)

[47] S. Haykin, "Neural Networks and Learning Machines", 3rd ed. Prentice-Hall, 2009.

[48] R. Xu, R. Wang, Z. Guan, et al. "Achieving Efficient Detection Against False Data Injection Attacks in Smart Grid", IEEE Access, vol. 5, pp.13787-13798, 2017.

[49] P. Gao, M. Wang, J. Chow, et al. Identification of Successive "Unobservable" Cyber Data Attacks in Power Systems. IEEE Transactions on Signal Processing, 2016, 64 (21): 5557-5570.

[50] D. J. Kershaw and R. J. Evans, "Optimal waveform selection for tracking systems," IEEE Trans. Inf. Theory, vol. 40, no. 5, pp. 1536-1550, Sep. 1994.

Chapter 3

Cognitive Risk Control for Mitigating Cyber-Attack in Smart Grid

3.1 Preceding Introduction

FDI attacks is the greatest known threat as it can bypass the normal data detection methods that have been used for more than a decade. By modifying the estimated states, this can start a chain of incorrectly performed control decisions that can lead to serious effects such as massive blackouts. In the previous chapter, CDS was introduced for cognitive control in the SG to improve state estimation in the DC model. It was also shown how the entropic state can be a good FDI attack detector. This chapter extends the CDS architecture tailored for the DC state estimator by introducing CRC to mitigate the undesired effects of cyber-attacks once it is detected by the entropic state. Being a brain-inspired model, the CRC aspect of this CDS

is more powerful than the one from which it was adapted from by integrating more concepts concerning cognitive predictive adaptation.

To the best of the author's knowledge, the scholarly work presented herein is the first experimental work of CDS being applied to the the DC model of the SG for cyber-attack mitigation.

The publication included in this chapter is:

M. I. Oozeer, and S. Haykin, "Cognitive Risk Control for Mitigating Cyber-Attack in Smart Grid," IEEE Access. 2019 Sep 2;7:125806-26.

The co-author's contributions to the above work include:

i. Technical supervision and financial support of the study presented in this work.

ii. Manuscript revising and editing.

Abstract

In this paper, we extend our previous research on uniting the Cognitive Dynamic Systems (CDS) and the Smart Grid (SG) by introducing Cognitive Risk Control (CRC). The CDS is a structured physical model and research tool inspired by certain features of the brain. The CRC is an advanced feature of the CDS that embodies the concept of predictive adaptation allowing it to bring risk under control in situations involving unexpected or abnormal uncertainty such as a cyber-attack. The False Data Injection (FDI) attack is a special class of cyber-attack targeting the SG that is able to bypass the traditional bad data detection techniques. Here we will demonstrate how the entropic state, which is the objective function of the CDS, is able to detect and bring FDI attacks under control under the action of CRC. Through Task-Switch control, the CDS is able to switch on a new executive with different set of actions that affects the system configuration to bring the risk under control during an attack. With the CDS acting as the supervisor of the SG, simulations are carried out on a 4 bus-system and IEEE 14-bus system to demonstrate the capability of CRC when faced with FDI attacks. The results show that this system has great potential for future SG systems.

3.2 Introduction

The Cognitive Dynamic System (CDS) is an organized physical model and research tool that is rooted in certain aspects of the brain. With its first introduction to the engineering world in [1] and its expansion in [2], the first practical applications of this construct was involved in Cognitive Radio [3] and Cognitive Radar [4]. Over the past few years, CDS has known tremendous progress, giving rise to Cognitive Control (CC) [5] and Cognitive Risk Control (CRC) [6] as two of its special functions. In [7], it was the first time that the CDS and the Smart Grid (SG) were united to form a new architecture that showed great potential in handling the new problems that will be facing the grid in the future. From a neuroscience perspective, the CDS is based on Fuster's paradigm of cognition involving the following five principles: perception-action cycle (PAC), memory, attention, intelligence, and language [5]. In its simplest form, the CDS consists of two main constituents: the perceptor, on one side, and the executive on the other with the feedback channel bringing them together.

In [7], it was shown how the integration of CC, considered as the over-arching function of CDS, with the SG, was well structured to handle those kind of cyber-physical systems which are slow progressing. In this paper, we extend this previous research to bring into play CRC, which is based on the principle of predictive adaptation and is new to engineering literature [8]. Although both CC and CRC are inspired by the prefrontal cortex, they are both geared towards handling different situations involving the environment. To be more specific, CC conforms well to a system operating under normal uncertainty while CRC is another subsystem that is dedicated to dealing to situations involving abnormal uncertainty such as cyber-attack. In [6], this subsystem consisted of the executive memory and classifier. However, in this paper, we propose

a more powerful architecture, that is even closer to neuroscience and risk control theory, to bring out CRC as a special function of the CDS. The proposed architecture will be evaluated on the SG in the presence of False Data Injection attacks. It will be shown how this novel construct has great potential to bring the risk associated with cyber-attacks in the SG under control and lays the foundation for a new generation of SG systems.

3.2.1 Smart Grid

The arrival of the Industry 4.0 era has been marked by the next generation of engineering systems involving Internet of Things (IoTs) and Cyber-Physical Systems (CPSs) [9]. Over the past few years, the cyber-security aspects of those systems have grown in importance as they have been increasingly deployed in critical infrastructures, affecting the daily life of people through systems such as the electrical power grids, transportation systems, health-care etc [10]. In the context of this paper, the greatest threat targeting the SG is known as the False Data Injection (FDI) attacks (also known as Bad Data Injection (BDI) attacks).

By collecting meter measurements from Remote Terminal Units (RTUs), consisting of different field devices or sensors, the Supervisory Control and Data Acquisition (SCADA) systems are able to monitor and process the important control actions implicating the SG. Those measurements are then transmitted to a control center to be processed and analyzed for errors and inconsistencies through a process known as state estimation [11][12]. In the SG, the state variables calculated by state estimation usually comprises of the voltage magnitudes and angles of the different buses in the system [13]. In the AC model, the voltage magnitudes and angles of the different buses

in the network are usually considered as the state variables. On the other hand, in the DC model, the state variables are limited to the bus angles only. The power system state estimation is carried out using the Weighted Least Squares (WLS) method [11] and the measurements involved in the process are typically the real and reactive power flows, power injections, voltage magnitudes and angles and current. During state estimation, bad data identification is carried out to discard bad measurements to enhance the accuracy of the estimated state. This is usually possible due to the fact that the number of measurements in a power system usually outnumbers the amount of states to be estimated. Bad measurements are typically erroneous measurement readings that are detrimental for state estimation. These are detected by a process known as bad data identification. Chi-Squared tests and Largest Normalized Residual Test are the most common bad data identification methods currently applied [12][14]. Those statistical techniques depend on the residuals between the estimated state variables and the measurement residuals in order to detect the bad data. Moreover, state estimators can be subdivided into DC and AC state estimators. The DC estimator relies on a linear system model while the AC estimator uses a nonlinear model. In the DC model, the measurements consist of the real power flows and injections and states consist of bus angles [13][15][16]. However, the states of the power system can be modified by the introduction of bad data, which is able to bypass the previously mentioned tests. Bad data are maliciously crafted offsets to measurements that are injected to transmitted sensor readings to influence the state estimation of the states in a certain way. As a result, this can result in bad control decisions being applied.

The tight integration of the cyber and physical infrastructures for the formation of the SG has opened the doors of power systems to cyber-attacks. Attacks, like

information tampering that have been a pest for the internet, are now threatening the security and stability of smart grids. In [17], the authors consider the BDI attack in the SG as the most dangerous cyber-attack as it can lead to further complications such as energy theft or device breakdown on the power generation side. Moreover, they characterize a cyber-physical attack as one consisting of cyber side and a physical side. From the cyber point of view, bad data is injected into the information system through the use of information techniques for intrusion. On the other hand, on the physical side, the bad data is meticulously fabricated, by the attackers, to bypass the error detection techniques previously mentioned. Recent studies have shown that the state estimator and bad data detection are vulnerable to this class of cyber-attack [14][18][19]. In [14], it was shown that an attacker, who is equipped with knowledge of the network configuration, could carry out the BDI attack on a DC state estimator without being detected. Due to this resulting in the estimation of wrong states, the system operator will be misled into performing improper operation decisions thereby inducing a domino effect of cascaded incorrect control decisions with disastrous consequences [20]. This is reminiscent to the 2010 cyber-attack in Iran's Bushehr whereby the Stuxnet worm was providing false system state to the SCADA to hinder system protection strategies.

Current literature on research on FDI attacks in the grid are focused on the detection aspect only of the problem [12; 18; 19; 21; 22; 23; 24; 25; 26; 27; 28]. Few researchers have looked at the other side of the problem; bringing the attack under control [21; 29; 30]. It is our firm belief that more effort must be focused on this other aspect since cyber-attacks are getting more and more sophisticated. Additionally, all the detection techniques referenced earlier all share a major weakness; they all depend

on some sort of pre-defined threshold in order to work. Consequently, if the attacker has knowledge on the detection method and the related threshold, all these detection algorithms become useless. In [7], we developed a new system for control and attack detection for the SG through a dynamic threshold that evolves during every PAC. In this paper, we will expand that structure with a new mechanism that is able to bring the attack under control once it is detected.

3.2.2 Contribution and Organization

The main contributions of this paper can be summarized as follows:

i. The architectural architecture of CRC, tailored for the SG, is presented. As the CDS is a construct that is rooted in the brain, the neuroscience behind the principle of predictive adaptation will be expanded on and implemented in the structure. With risk control also being the topic of this paper, risk control theory and neuroscience will be united to give rise to a new CRC subsystem which is far more powerful than the one described in [6].

ii. Using the entropic state as basis for control and cyber-attack detection, a novel algorithm of CRC for the SG is introduced. While the cognitive controller residing in the frontal executive is responsible for CC under the presence of normal uncertainty, a second executive dedicated for risk control is introduced to handle cases involving abnormal uncertainty such as a cyber-attack. It is shown that both systems can work together by having different action spaces that are able to handle the sections of the grid without attack and those under attack. We show through simulations that this new architecture is capable of bringing the risk associated to the attack under control for different situations.

Moreover, this system also has the ability to detect when a threat is no longer a risk and to switch off the CRC. Consequently, CRC lays the foundation for a new approach for risk control in power grid systems. Lastly the importance of past experiences, rooted in the brain, will be expanded and highlighted throughout the last main sections of this paper.

The rest of the paper is organized as follows: Section 3.3 covers briefly the basic concepts associated with state estimation, bad data detection and FDI attacks in the power grid. Section 3.4 expands on the structure of the CDS for the SG. Since material in this section is based on [7], this will be covered concisely. Readers are advised to through [7] for more background information. Section 3.5 covers the neuroscience concepts of predictive adaption and risk control. An algorithm based on both theories is then derived and illustrated as an expansion of the CC algorithm mentioned in [7]. Section 3.6 discusses the simulation results of this system in the presence of FDI attacks in a 4-bus network and the IEEE 14-bus network. Finally, Section 3.7 concludes this paper by highlighting the key results and presenting new avenues of research for this new structure.

3.3 Preliminaries

3.3.1 Weighted Least Squares State Estimation

The measurement data sent from the SCADA is essential for the real-time operation of the Energy Management System (EMS). However, since those signals are often corrupted by noise, the state estimator and the bad data detector are responsible for filtering out the data for optimal state estimation. As power systems involve an

overdetermined system consisting of redundant measurements, this process discards the measurements that will be detrimental for estimating the optimal state. The states of a power system refer to the bus voltage angle θ and bus voltage magnitudes V. In the context of this paper, the DC model will be applied for our methodologies. In the DC model, measurement data comprise of real power flows and injections and the states are limited to bus angles only. It is assumed that prior information relating to the bus magnitudes is available and taken to be close to unity. Additionally, a reference bus is chosen and is set to zero radians. Therefore, state estimation in the DC model simply involves estimating the n bus voltage angles $[\theta_1, \theta_2, ..., \theta_n]^T$. The DC power flow model has been commonly employed by power engineers and smart grid cyber-security researchers [11; 31; 32; 33] as a linearization and approximation to the AC model. This substitution to the AC model has been widely acknowledged for reasons such as guaranteed faster convergence [34].

In the DC estimation model, shunt elements and branch resistances are neglected. Moreover, it assumes that the bus voltage magnitudes are already known and are close to or equal to 1.0 per unit. Consequently, by approximating the first order Taylor expansion around $\theta = 0$, the measured real power flow from bus k to m is calculated as follows [35]:

$$P_{km} = \frac{\theta_k - \theta_m}{x_{km}} + e \tag{3.3.1}$$

where x_{km} corresponds to the reactance (in per unit values) of the branch k-m, θ_k is the phase angle(in radians) at bus k and e is the measurement error. The power injection at a specified bus i can be obtained by summing up all the flows along

incident branches to that bus:

$$P_i = \sum_{j \in N_j} P_{ij} + e \tag{3.3.2}$$

The measurement model, consisting of an overdetermined system of linear equations, is solved using the Weighted Least-Squares (WLS) problem using the following formula:

$$\mathbf{z} = \mathbf{Hx} + \mathbf{e} \tag{3.3.3}$$

where

- $\mathbf{x}$ is the n vector of the true states (unknown)
- $\mathbf{z}$ is the m vector of measurements (known)
- $\mathbf{H}$ is the m x n Jacobian matrix
- $\mathbf{Hx}$ is the m vector of linear function linking measurements to states
- $\mathbf{e}$ is the m vector of random errors
- m is the number of measurements
- n is the number of variables

$\mathbf{H}$ in is a matrix describing the power system. It consists of theoretical calculations that link the states to the measurement vector $\mathbf{z}$. These are known as power flow equations and these are described as vectors inside $\mathbf{H}$. In the DC model, those entries are made up of a set of linear functions of the state variables while in the AC model, those functions are nonlinear. In order to solve the WLS problem in , we need to find

the n-vector $\mathbf{x}$ that minimizes the index $J(\mathbf{X})$, which is described as follows:

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})'\mathbf{W}(\mathbf{z} - \mathbf{H}\mathbf{x})' \tag{3.3.4}$$

In the above equation, matrix $\mathbf{W}$ is a diagonal matrix consisting of the measurement weights. These weights are usually based on the reciprocals of the measurement error variance σ:

$$\mathbf{W} = \mathbf{R}_\mathbf{z}^{-\mathbf{1}} = \begin{bmatrix} \sigma_1^{-2} & \dots & \dots & \dots \\ \dots & \sigma_2^{-2} & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \sigma_m^{-2} \end{bmatrix} \tag{3.3.5}$$

where $\mathbf{R}_\mathbf{z}$ is the covariance matrix of the measurement. The performance index $J(\mathbf{X})$ can then be differentiated to obtain the first order optimal conditions:

$$\mathbf{G}\hat{\mathbf{x}} = \mathbf{H}'\mathbf{W}\mathbf{z} \tag{3.3.6}$$

where the estimate of the state $\hat{\mathbf{x}}$ is calculated by:

$$\hat{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{H}'\mathbf{W}\mathbf{z} \tag{3.3.7}$$

In the above equations, $\mathbf{G} = \mathbf{H}'\mathbf{W}\mathbf{H}$ is the state estimation gain.

3.3.2 Bad Data Detection

Erroneous measurements must be detected and identified to be removed from the state estimation process. However, the statistical properties of these errors facilitate

their detection and identification. The estimated measurements are obtained from the estimated measurements in using the following equation:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} \tag{3.3.8}$$

The individual estimated measurement error is then obtained using:

$$\hat{\mathrm{e}}_j = (z_j - \hat{z}_j) \tag{3.3.9}$$

These errors follow a Gaussian distribution with zero mean [15]. The Chi-squares test and normalized residual have been most commonly applied techniques for data detection [33]. Chi-squares test assumes that the state variables are mutually independent from each other and the errors follow a normal distribution. The test consist of a number of recursive steps involving the number of degrees of freedom of the system, sum of squares $\hat{f}$ and a critical value corresponding to α to satisfy the following inequality:

$$\hat{f} < \chi^2_{(k,\alpha)} \tag{3.3.10}$$

where k is the appropriate number of degrees of freedom and α is a specified probability. According to , $\hat{f}$ will be large when there is a large number of bad measurements. However, since α is large in practical power applications, this technique can remove those measurements that contribute the most to the largest standardized residuals.

3.3.3 False Data Injection Attacks

FDI is category of cyber-attacks targeting the SG, which are able to bypass the bad data detection approaches mentioned previously. Numerous forms of these attacks and their consequences have been investigated in [10; 12; 14; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 36]. In the context of this paper, the FDI attacks simulated will be based from [20], whereby the adversary model for FDI assumes that the attackers have full knowledge of the system parameters and topology (system Jacobian). FDI attacks involving full knowledge of the system have been proven to have more disastrous consequences. Furthermore, in [14], the authors demonstrate how an attacker, armed with perfect knowledge of the system matrix $\mathbf{H}_{m\times n}$, can maliciously inject an attack vector $\mathbf{a}_{m\times 1}$ to the measurement vector $\mathbf{z}_{m\times 1}$ that is capable of bypassing the bad data detection techniques currently employed. Hence, when $\mathbf{a}_{m\times 1}$ is inserted, the new corrupted measurement vector $\mathbf{z'}_{m\times 1}$ can be summarised as:

$$\mathbf{z'}_{m\times 1} = \mathbf{z}_{m\times 1} + \mathbf{a}_{m\times 1} \tag{3.3.11}$$

Consequently, the state estimator will calculate a corrupted system state $\mathbf{x'}_{m\times 1}$ instead of the original state $\mathbf{x}_{m\times 1}$. The difference between these two states can be denoted as $\mathbf{c}$ as follows:

$$\mathbf{x'} = \mathbf{x} + \mathbf{c} \tag{3.3.12}$$

Moreover, in the same paper, Liu et al. show the mathematical proof behind the attacks and validate his results through the experiments. Indeed, the authors prove that as long as $\mathbf{a}$ satisfies the condition $\mathbf{a} = \mathbf{Hc}$, the attack will go undetected by the

bad data detector. The residual of the estimation process is demontrated as [20]:

$$\begin{aligned} ||\mathbf{z}' - \mathbf{H}\mathbf{x}'|| &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}(\mathbf{x} + \mathbf{c})|| \\ &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}\mathbf{x} - \mathbf{H}\mathbf{c}|| \\ &= ||\mathbf{z} - \mathbf{H}\mathbf{x}||(\text{since}, \mathbf{a} = \mathbf{H}\mathbf{c}) \\ \mathbf{r}_{normal} &= \mathbf{r}_{attack} \end{aligned} \tag{3.3.13}$$

As a result, as shown by the mathematical proof, the residuals due to the attack and those due to normal conditions are considered the same. For this reason, since the bad data detection techniques rely on statistical methods for the calculation of these residuals, they are unable to detect that the measurement vector has been maliciously falsified. Therefore, this results in the calculation of wrong system states that can in turn start a domino effect with disastrous consequences. Since this type of attack is aimed towards state estimation in the SG model, this implies that $\mathbf{a}$ can be inserted either physically by tampering with some selected meters or wirelessly by injecting the vector when the readings are transmitted to the SCADA. As a result, important components of the SG, such as the substation state estimator (SSE) at the substations, involved in state estimation are at the mercy of this type of attacks.

3.4 Architectural Structure of CDS for Smart Grid

From a neuroscience perspective, the CDS is the closest system that matches Fuster's paradigm [8] when it comes to cognition. The new advanced architecture proposed in this paper is largely based on [37], where the adaptation of the brain to various conditions is discussed. To this end, in order to be able tackle the issue of bringing

risk under control, the CDS consists of five major components. Firstly it consists of two executives; namely the frontal executive and the posterior executive. Similar to how the frontal cortex deals with information regarding the motor organ and the posterior cortex deals with information concerning the sensory organs [37], it will be shown in their respective sections how the two mentioned executives attempt to replicate these mechanisms, like the brain does. Two other components of the CDS are the perceptor and the feedback channel which link those two executives. The last module is the environment, which closes a global feedback channel whereby the entire CDS is contained within it. In this architecture, the DC state estimator of the SG is considered as the environment for which the CDS acts as the supervisor. Referring to [7], the frontal executive is responsible for CC in the system. Hence, during every PAC, it learns which measurements to prioritize for optimal state estimation. On the other hand, the posterior executive deals with risk control when an attack takes place. The diagram depicting the architecture of CRC for the SG is shown in Fig. 3.1. In the next subsections, it will be elaborated how the embodiment of each of the major components mentioned comes together for goal-oriented action on the SG.

3.4.1 Perception-Action Cycle

When the environment is governed under normal uncertainty, the global feedback loop of the PAC gradually improves the information extraction ability of the perceptor during each consecutive cycle. Consequently, a continuous cyclic directed flow of information from the perceptor to the executive is set up to guide the current PAC for the actions to be performed on the environment by the executive. Thus, depending on the information extracted from the perceptor at every cycle, the executive will

evaluate its current ability in order to better achieve its set goal that it was designed for.

3.4.2 Perceptor

The role of the perceptor is to extract the available information out of incoming noisy measurements, which in response the frontal executive performs actions on the environment to enhance the information gain in subsequent cycles. These actions, performed under CC, are called cognitive actions. The perceptor is made up of the generative model and the Bayesian filter, which are reciprocally coupled to each other. Thus, although the perceptor senses the environment directly, the cognitive controller, in the frontal executive, senses the environment indirectly through the information extracted from the perceptor.

3.4.2.1 Generative Model

Originally, as defined in [6], the first constituent of the perceptor is the *Bayesian generative model*, which acts as a classifier for the observables originating from the environment. However, in [7], it was elaborated that due to the dynamic nature of the SG, the Bayesian generative model would be impractical. With security in mind and inspiration from quickest detection theory, the generative model used is based on cumulative sum (CUSUM) and is written as follows:

$$\mathbf{B}_k = \sum_{i=k-L}^{k} \mathbf{x}_i \tag{3.4.1}$$

where k refers to the current cycle number, L is the window over which the past states is being accumulated, $\mathbf{B}_k$ is the vector retaining the cumulative sum for each cycle and $\mathbf{x}_i$ is the vector of the states output from the DC state estimator for the cycle i. In [38; 39], CUSUM-based detection methods have been investigated and showed great potential at detecting FDI attacks. Nevertheless, these techniques rely on a pre-defined threshold in order to be effective. If the cyber-attacker has knowledge of the threshold, then the latter can design an attack in order to stay undetected. In the CDS, as will be shown later, it is possible to circumvent this issue by making the threshold dynamic using the *entropic state*. The entropic state is the primary unit of control and attack detection in the CDS structure, as investigated in [7]. Lastly, the use of a CUSUM based generative model has desirable properties such as the smoothing out of noise operating under the slow dynamics of the SG.

3.4.2.2 Bayesian Filter

Reciprocally coupled to the generative model, the second component of the perceptor is the Bayesian filter. Since the DC state estimator is linear in nature and the noise can be assumed as white additive gaussian noise, the *Kalman filter* [40] will be utilized as the Bayesian filter for modelling the incoming inputs. Since in this paper, we are assuming that the power system is quasi-static in nature [41; 42; 43], we can postulate that the state variable $\mathbf{x}$ at time k+1 will only deviate by a small margin from its previous values at its previous cycle k. This can be written as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \omega_k \tag{3.4.2}$$

where ω_k is independent Gaussian noise vector with zero mean. Using , the measurement equation used for the Kalman filter is formulated as:

$$\mathbf{Y}_k = \mathbf{L}_k\mathbf{B}_k + \omega_k \tag{3.4.3}$$

and the covariance matrix of ω_k is

$$\mathbf{R} = \text{diag}[\sigma_\omega^2], \sigma_\omega^2 = \text{var}[\omega_i] \tag{3.4.4}$$

Assuming that the system is slow moving or quasi-static, the process equation of the Kalman filter can be postulated as a random walk model as follows:

$$\mathbf{B}_{k+1} = \mathbf{F}_k\mathbf{B}_k + \mathbf{v}_k \tag{3.4.5}$$

where $\mathbf{v}_k$ is the process noise covariance vector which is assumed to be statistically independent and zero mean. The covariance matrix of $\mathbf{v}_k$ is:

$$\mathbf{Q} = \text{diag}[\sigma_\text{v}^2], \sigma_\text{v}^2 = \text{var}[\text{v}_i] \tag{3.4.6}$$

In both and the system matrix $\mathbf{L}_k$ and the predictive transition matrix $\mathbf{F}_k$ are the identity matrices. Using the measurement and the process equations mentioned previously, the computational steps of the Kalman filter use predefined initial estimates of the state $\hat{\mathbf{B}}_{k|k}$, and predicted error covariance, $\mathbf{P}_{k|k}$, to calculate the predicted state estimate $\hat{\mathbf{x}}_{k+1|k}$ and predicted error covariance, $\mathbf{P}_{k+1|k}$ for the next cycle during the

time update steps using the following equations:

$$\hat{\mathbf{B}}_{k+1|k} = \mathbf{F}_{k+1,k}\hat{\mathbf{B}}_{k|k} + \mathbf{v}_k \tag{3.4.7}$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_{k+1,k}\mathbf{P}_{k|k}\mathbf{F}_{k+1,k}^T + \mathbf{Q} \tag{3.4.8}$$

During the next cycle, those two estimates are used for the measurement stages to calculate the Kalman gain, $\mathbf{K}_k$, filtered estimate, $\hat{\mathbf{x}}_{k|k}$, and to update the process covariance matrix $\mathbf{P}_{k|k}$ as follows:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{L}_k^T(\mathbf{L}_k\mathbf{P}_{k|k-1}\mathbf{L}_k^T + \mathbf{R})^{-1} \tag{3.4.9}$$

$$\hat{\mathbf{B}}_{k|k} = \hat{\mathbf{B}}_{k|k-1} + \mathbf{K}_k(\mathbf{Y}_k - \mathbf{L}_k\hat{\mathbf{B}}_{k|k-1}) \tag{3.4.10}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{L}_k\mathbf{P}_{k|k-1} \tag{3.4.11}$$

Consequently, this computational iteration uses the preceding *a posteriori* estimates to predict new *a priori* estimates.

3.4.3 Feedback Channel

The feedback channel serves two special purposes in this architecture proposed. At this point in the paper, its first purpose relating to the proper functioning of CC will be elaborated. In the later sections of this paper, its second purpose relating to CRC will be expanded on. Equipped with the Entropic information processor, the

feedback channel links the frontal and posterior executives together. However, only one executive can be active at the same time. In order to select which executive will be active, the entropic information processor is linked to Task-Switch Control, which is the second component of the feedback channel. Task-Switch control is responsible for determining the executive that will be operational using the input from the entropic information processor. The feedback channel is responsible for the calculation of the entropic state and internal rewards during reinforcement learning in the frontal executive. This calculation will be elaborated in section 3.4.4 (Frontal Executive), where it is more relevant to the latter's role during planning.

3.4.3.1 Entropic-Information Processor

The directed cyclic flow of information from the perceptor to the executive is known as the *entropic state of the perceptor*. The entropic state is based on the principles of the perceptual posterior, which can be regarded as the filtered posterior embodying the conglomeration of the generative model, Kalman filter and entropy, that is derived from *Shannon's information theory* [44]. The entropic state at time k, in this architecture [7], is written as:

$$h_{k|k} = \frac{\mathrm{Tr}\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{Y}_k\}^2\}}{\mathrm{Tr}\{\mathbf{P}_{k|k-1}\}} \tag{3.4.12}$$

where Tr represents the trace operator and diag{.} refers to the diagonal operator. In a generic sense, condenses information between the filtering-error covariance $\mathbf{P}_{k|k-1}$ and the real error between the state estimate $\hat{\mathbf{x}}_{k|k-1}$ and the current measurement at cycle k. The denominator in the equation acts as a normalizer that confines $h_{k|k}$ to attain a maximum value of 1 when the environment is under normal uncertainty. Any

value below this maximum indicates the degree of disturbance or uncertainty affecting the SG. Consequently, since the SG will be facing varying levels of uncertainty during its operation, we will now define two important properties of the entropic state that will be crucial for the proper functioning of CC and CRC:

i. When the environment is operating under normal environmental uncertainty, $h_{k|k}$ will always be positive because of the probabilistic representation of the uncertainties.

ii. When uncertainties are present, $h_{k|k}$ will fluctuate around values which are less than 1. Thus, to distinguish between normal uncertainties, due to the probabilistic nature of the environment, and abnormal uncertainties, such as cyber-attack, a suitable threshold γ can be chosen such that if $h_{k|k}$ is below γ, then this would indicate presence of attack and to switch on CRC.

3.4.3.2 Task-Switch Control

Task-Switch Control (TSC) plays a distinctive role in the CDS in the activation of two pair of switches namely the F switches for the frontal executive and P switches for the posterior executive, whereby they deal with CC and CRC respectively. As mentioned in the previous section, we do not have control over the amount uncertainty affecting the state estimation. However, we can quantify this amount of uncertainty through the entropic state. Consequently, using the entropic state as an attack detector and risk raiser, we can define the following:

$$\zeta_k = \begin{cases} 0, & \text{if } h_{k|k} > \gamma. \\ 1, & \text{otherwise.} \end{cases} \tag{3.4.13}$$

where ζ_k is the result of detection. Thus when ζ_k is 0, the F switches are ON and the frontal executive, responsible for CC, will be operating. When ζ_k is 1, the P switches are ON and the posterior executive, responsible for CRC, will take over. So far, (3.4.13) explains how the system will transition from CC to CRC. However, in order to shift from CRC to CC, another mechanism is required since the entropic state cannot be relied upon when under the presence of abnormal uncertainty or cyber-attack. To that end, TSC is also equipped with a watchdog timer and an internal memory, which we will define as TSC memory. This will be elaborated in the section on CRC where we will highlight the principles of predictive adaptation and risk control theory. In the next section, the role of the frontal executive will be described. Under normal uncertainty, the latter will be the dominant active executive for control in the SG.

3.4.4 Frontal Executive

Theoretically, the Frontal Executive is one of the most important parts of the CDS as it is the entity responsible for the control of the SG under normal uncertainty. In order to be able to perform its duties, it is equipped with Reinforcement Learning (RL) and Cognitive Control, both of which can be further split into the action space, planner, working memory and policy.

3.4.4.1 Reinforcement Learning: Bayes-UCB

As mentioned previously, the feedback channel is also involved in the calculation of internal rewards during the planning stages of the RL algorithm [45] in both the frontal and posterior executives. At this point in the paper, the emphasis will be

on the involvement of RL for CC in the frontal executive. The RL for CC is based on the entropic state, which in turn is used to optimize an objective function for optimal control in the network. Before elaborating further, a brief description of Bayes-UCB [46] RL algorithm will be given in the next paragraph. Bayes-UCB is the RL algorithm that is applied throughout the CDS.

Bayes-UCB originates from a category of multi-armed bandit algorithms called UCB algorithms, which are based on the principle of optimism in the face of uncertainty [47]. In this approach to the multi-armed bandit problem, the algorithm uses a Bayesian approach for estimating the reward distribution of the different available actions. The chosen action is then selected based on the action that brings the highest reward. Hence, Bayes-UCB is an index policy that makes use of the prior distribution to select a dynamic quantile of the posterior estimates for the index of each action. Consequently, during each discrete time t, the algorithm will choose an action A_t that satisfies the following condition:

$$A_t = \underset{a}{\operatorname{argmax}}\, q_a(t) = Q(1 - \frac{1}{t(\log(t))^c}, \lambda_a^{t-1}) \tag{3.4.14}$$

where $Q(\alpha, \pi)$ refers to the quantile of order α of the distribution π. Furthermore, it is shown in [48] that if we assume that the rewards follow a Bernoulli distribution, and when the prior distribution of each action is Beta(1,1), can be simplified. In order to be consistent with the mentioned notations so far, can therefore be written as:

$$\begin{aligned} A_k &= \underset{a}{\operatorname{argmax}}\, q_a(k) = Q(1 - \frac{1}{k(\log(k))^c}; \\ &\quad \text{Beta}(S_a(k) + 1, N_a(k) - S_a(k) + 1) \end{aligned} \tag{3.4.15}$$

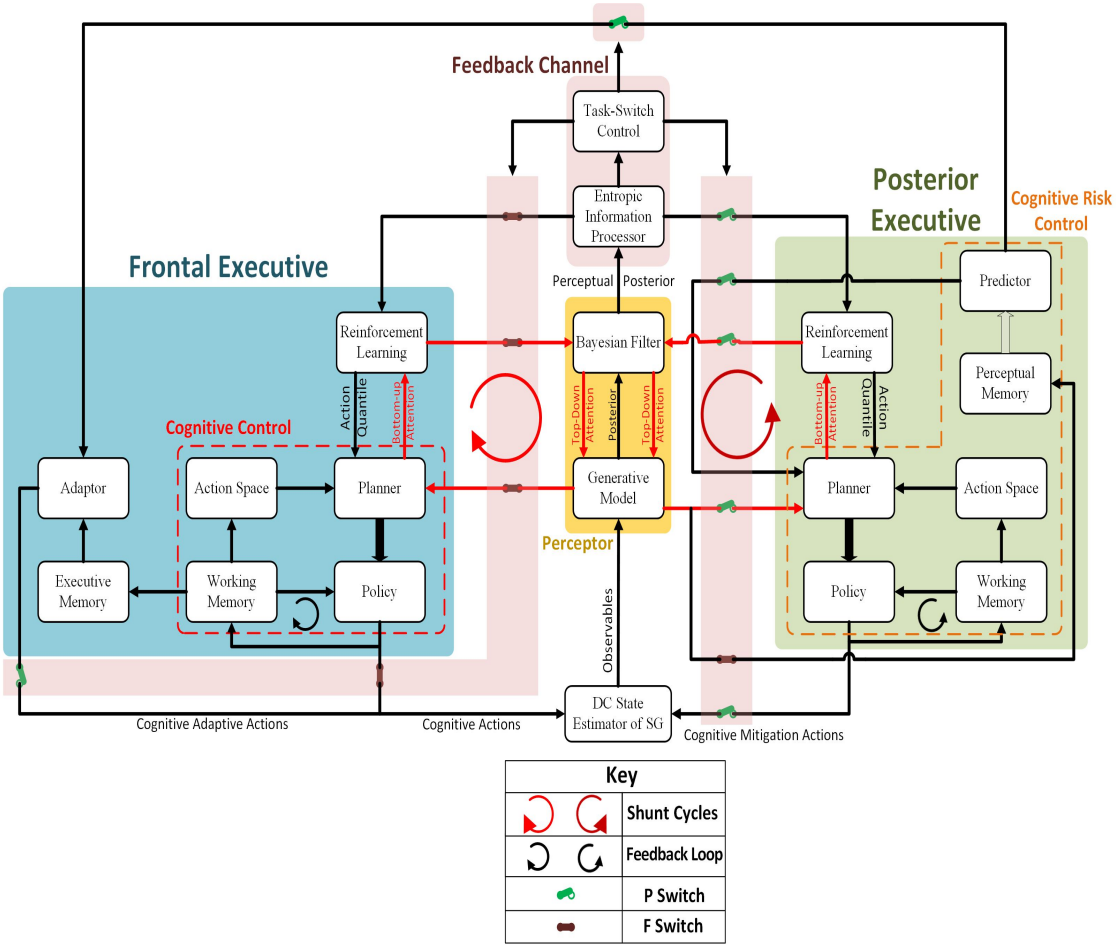

Figure 3.1: Architectural Structure of CDS, embodying CC and CRC, for SG.

where k is the PAC cycle number, S_a is the cumulative reward for action a, N_a is the number of times action a has been chosen and c is real parameter. Since the CDS is a construct that aims to be as close as possible to the brain, it is to be added that Bayes-UCB shares many similarities to the Bayesian approach of decision making in human brains [49]. With this small introduction to Bayes-UCB, it will be shown in the next section, relating to Cognitive Control, how the algorithm brings together the system configuration **H** of the power grid, the generative model of the perceptor and the process model in the Kalman filter together for optimal state estimation.

3.4.4.2 Cognitive Control

Building on the components described so far, CC plays a key role in the CDS for goal-oriented action on the SG. The two most important aspects of CC is the *planner* and the *policy*. Throughout the PAC, the planner is involved in the extraction of a set of prospective actions from the action-space A and their evaluation during the shunt cycles (i.e., planning cycles). Under the influence of attention from one PAC to the next, the policy is then able to learn the best actions to be applied on the system. In the case of the SG, the action space consists of discrete weight values that can be assigned to the meters. As a result, through the influence of attention, the CDS will select the optimal weight values of the meters for state estimation. During the planning steps, the latter starts with a prospective action $a_k^{i,j}$ which represents weight value a^i for meter j during cycle k. This hypothesized weight value for a specific meter is then used to calculate a predicted gain as follows:

$$\mathbf{G}_k^p = \mathbf{H'}_k \mathbf{W}_k^{i,j} \mathbf{H}_k \tag{3.4.16}$$

where $\mathbf{G}_k^p$ denotes the planned gain and $W_k^{i,j}$ is the modified weight matrix where meter j's weight value has been replaced by a^i. This is turn allows us to calculate a new predicted state estimate:

$$\hat{\mathbf{x}}_k^p = (\mathbf{G}_k^p)^{-1} \mathbf{H'}_k \mathbf{W}_k^{i,j} \mathbf{z}_k \tag{3.4.17}$$

where $\hat{\mathbf{x}}_k^p$ refers to the planned state estimate using the modified weight matrix with the hypothesized weight value. The projected cumulative sum involving $\hat{\mathbf{x}}_k^p$ can then

be calculated:

$$\mathbf{B}_k^p = \sum_{i=k-L}^{k-1} \mathbf{x}_i + \hat{\mathbf{x}}_k^p \tag{3.4.18}$$

where $\mathbf{B}_k^p$ is the projected cumulative sum involving $\hat{\mathbf{x}}_k^p$ instead of $\hat{\mathbf{x}}_k$. Consequently, a planned entropic state $h_{k|k}^p$ is found using:

$$h_{k|k}^p = \frac{\mathrm{Tr}\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{B}_k^p\}^2\}}{\mathrm{Tr}\{\mathbf{P}_{k|k-1}\}} \tag{3.4.19}$$

The influence of normal uncertainties, stochastic or deterministic, will cause the output of the generative model of the perceptor to diverge from the estimated hidden state of the Kalman filter. Thus, the objective of is to find a weight configuration that reduces this deviation. Consequently, any hypothesized weight that predicts $h_{k|k}^p$ closer to the optimal value of 1 satisfies this condition, whereby the estimated state of the DC state estimator reduces the propagated variation in the generative model.

3.4.4.3 Internal Rewards

Building on the previous equations involved during the planning stages, we can now define the predicted internal reward, $r_k^{i,j}$, associated with each prospective action $a_k^{i,j}$, for cycle k as:

$$r_k^{i,j} = h_{k|k}^p - h_{k|k} \tag{3.4.20}$$

Referring to , we can define that the ultimate goal of RL, in this context, is to minimize the amount of uncertainty or disturbance in the system every PAC by always searching for an optimal weight configuration that will outperform the current entropic state. As a result of all the steps mentioned so far, the CDS learns from the past and current actions to pick the best set of actions for the future. To facilitate

this task, the working memory holds temporarily the actions that have achieved the highest quantile from Bayes-UCB in after the shunt cycles have elapsed and applies them to the system before starting the next PAC. Once the next cycle starts, a new set of prospective actions are evaluated. Should any of those actions achieve a higher quantile than the quantile of its respective meter in working memory, the greater yielding action replaces that previously best action. This approach towards control in the SG can also be perceived as a Contextual Bandit problem, whereby the actions performed shift the system to a new set point that the RL algorithm will have to adapt and this continues throughout the operation of the CDS.

3.5 Cognitive Risk Control

When the SG is under the influence of abnormal uncertainty, such as an FDI attack, the latter can be detected through the use of an appropriate threshold γ on the entropic state. In such situations, CC must expand its functionality to be able to deal with those unexpected adverse events [6], collectively known as risks. As mentioned in the section pertaining to Task-Switch Control, the posterior executive is introduced to handle those hostile conditions when they occur. Under CRC, as will be shown later, we now have a mechanism, rooted in predictive adaptation and past experiences, that is able to tackle those uncertainties and bring risk under control. Before proceeding further, we will begin by a providing a brief overview of the relationship between risk control and the neuroscience of predictive adaptation before elaborating how those two key principles are brought together in the posterior executive.

3.5.0.1 Risk Control

Risk refers to the possibility that an undesired outcome disrupts our system [50]. Thus, risk control (or management) relates to the ensemble of actions taken to identify and control those unwanted outcomes proactively. Consequently, we can take one step further and identify three critical aspects of risk as follows [50]:

i. Uncertainty: Managing risk implies always involving oneself with uncertainties. It is unclear if a situation involving risk will happen until it occurs and after it has been brought under control. Consequently, this intrinsic uncertainty cannot be avoided but can be narrowed to a tolerable level through other means such as simplifying the probability of occurrence of the risk. Hence, risk control does not guarantee perfect solutions for all situations.

ii. Loss: Risk always comprise the probability of some form of loss. Risk is managed to minimize this loss, though this likelihood could be small. However when faced with risk, it is also possible that a positive outcome could result out of it. Nevertheless, risk control focusses on events that can have negative consequences on the system.

iii. Time Component: Every situation involving risk is associated with a time when it no longer exists. This means that either a loss was incurred or the risk was brought under control to a tolerable level. Therefore, it is crucial to identify when this time arrives so that resources, which were allocated to resolving the risk, can be re-assigned to other tasks. The determination of this termination time is very important as in some cases, it is well-defined, and in other cases, it is ongoing. The "time component" might not be always stated as time but it

can also be expressed as a specific condition.

3.5.0.2 Risk Control Process and Mitigation

A typical risk management process can be broken down into five steps [50]. The first step involves the identification of possible risks to be faced throughout the life cycle of the system. In the second step, the identified risks are analyzed to uncover their driving factors, potential impacts and their likelihood. In the next step, those risks are short-listed and the important ones to resolve are chosen. In the fourth step, action plans to counter those risks are designed. In the last step, the progress of the applied action plans are monitored. Those action plans that have been able to resolve the risks are terminated and then step one is re-initiated in search of new risks. However, mitigation plans are possibly the most powerful category of action plans that can be applied as they target the root causes of risk directly. Mitigation is the backbone of effective risk management. Mitigation actions are crafted to resolve the effects of risk events and impact drivers directly. Mitigation actions are usually associated with trigger points. Similar to the time component, trigger points signify the conditions for which to trigger the actions for the mitigation plans. Hence, it is important to have a means to detect if the risk event has occurred.

3.5.0.3 Predictive Adaptation

The principle of *predictive adaptation* as addressed in [51] is the next important topic to be discussed. In more advanced organisms, especially humans, there is an ability that manifests itself to anticipate changes within the organism itself or its current environment and to adapt to them before they are predicted to happen. Consequently,

it can be said that this predictive ability permits the organism to preadapt to these expected changes. This preadaptation capacity is based in its ability to reorganize past and deep-rooted individual experience to shape new adaptive structures of goal-directed behavior according to their temporal relationship. It is also further proposed that this is the predominant function of the prefrontal cortex, whereby it is proactive and rooted in prediction and preparation. Moreover, in both the CDS and primates, the internal PAC becomes the neural framework for prediction and cognitive control. According to the cognit model [52], cognits are hierarchically structured to embody a hierarchy of perceptual memory/experiences in posterior cortex and another of executive memory/experiences in frontal cortex. These hierarchies are also connected to each other at all levels in the cortical PAC. After their formation in the two moieties of the left cortical hemisphere, the perceptual memory acquires experiences through information from the senses-in posterior (PTO) cortex while the executive memory acquires experiences through action-in from frontal cortex. We will refer these stored experiences as past experiences in the next section. From a neuroscience perspective, those stored past experiences and emotional inputs will drive PAC to new adaptive behavior. Moreover, in [51], the author clarifies that in the case of brain dynamics, there is no prospective future without a consolidated past. Before and during pursuing a new goal, there will be an intervention of the prefrontal cortex for the selection of past knowledge to guide and correct the course of action to that goal. Thus, the use of past information is the basis of prediction and error correction in the cognitive functions of the prefrontal cortex. The internal cerebral cycle running from the prefrontal to posterior cortices is essential for that predictive and preadaptive ability of cognition. Furthermore, in [53], the author states that this dynamics requires two

critical components. The first component relates that interactions between the organism and the environment involve Bayesian probability. The second element suggests that prediction is an important part of those interactions. So far in our CDS, the first postulation has been satisfied through the use of the Bayes-UCB, which is also supported by [49] on the same topic. While the second component can be seen as the embodiment of the shunt cycles between the perceptor and the executive. In cases of abnormal uncertainty, such as the FDI attack, we need a new way for error correction based on past experiences. The frontal executive is useless as it is perturbed and not equipped to face for those situations. In the next section, dealing with the posterior executive, we will show how we can modify the CDS to bring risk control using the material covered in the previous and current sections.

3.5.0.4 Posterior Executive

Conceptually, the posterior executive is the second most important component of the CDS. In the presence of an FDI attack, the entropic state will decrease below γ, which causes the F set of switches to be turned off. When this occurs the frontal executive is turned OFF and thus cognitive control is no longer operating. At the same moment, the P set of switches for the posterior executive is now ON. This indicates that CRC is now in control of the SG. In order for the posterior executive to bring risk under control, it is equipped with supporting structures and CRC module. The supporting structures can be expanded as follows:

i. Executive memory and Perceptual memory: From a neuroscience perspective, those two memories differ from the working memory by a temporal parameter. More specifically, while the working memory is concerned with the storing

of actions to be performed for the current PAC, the executive and perceptual memory store past experiences over a longer span of time. The executive memory stores the motor actions that were present in the working memory over the past L_{PE} perception action cycles. Similarly the perceptual memory stores the past L_{PE} sensory information from the generative model of the perceptor. In many ways, this is reminiscent to how these two memories have been described in their respective roles in the brain. In the case of the CDS, each index (row or column) of the executive memory recalls the actions taken while the same index in the perceptual memory relates to the context in which these same actions were taken. Lastly, it should be emphasized that the accommodation of the experiences in both memories takes place only when cognitive control is operating. When CRC takes over, the F switches are OFF and experiences are no longer stored because the environment is no longer viable.

ii. Predictor: Similar to how prediction is important for adaptation and error correction in the brain, the predictor is concerned with the prediction aspect of the predictive adaptation concept. To that end it's role is two fold:

(a) When the P switches are ON, the predictor extracts past experiences from the perceptual library to identify the affected states by the FDI attack based on their statistical properties before the attack took place. Although advanced signal properties [54] can be used, in the context of this paper, a re-constructed mean and absolute maximum standard deviation of the stored values output from the generative model in perceptual library is used to calculate the mean predicted states and safety bounds for the states from the DC state estimator. Thus an estimate of the mean of the predicted

states, $\tilde{\mathbf{x}}_{PE}$, from the DC state estimator is calculated indirectly from the accumulator values from the generative model stored in the perceptual library as follows,

$$\tilde{\mathbf{x}}_{PE} = \frac{\sum_{j=1}^{L_{PE}} \mathbf{B}_j}{L \times L_{PE}} \tag{3.5.1}$$

where $\mathbf{B}_j$ is the stored accumulator values in the perceptual memory from and L_{PE} is the length of time over which the executive and perceptual memories extend. Using $\tilde{\mathbf{x}}_{PE}$, the safety limits, $\mathbf{x}_{\text{lim}}$, of $\mathbf{x}_k$ is then defined as the maximum deviation from $\tilde{\mathbf{x}}_{PE}$ in the perceptual memory. Although in a practical scenario $\mathbf{x}_k$ will be varying in time, in a situation dealing with risk whereby the states are perturbed and the real state values are unknown, $\mathbf{x}_{\text{lim}}$ define the range within which $\mathbf{x}_k$ has the highest probability to lie. Furthermore $\tilde{\mathbf{x}}_{PE}$ and $\mathbf{x}_{\text{lim}}$ can be brought together to identify the attacked states as follows:

$$\varepsilon_k^x = [(\tau^x \times \mathbf{x}_{\text{lim}}) - |\mathbf{x}_k - \tilde{\mathbf{x}}_{PE}|] < 0 \tag{3.5.2}$$

where ε_k^x is a logical vector that identifies the targeted states and τ^x is a vector that defines the volatility or deviation tendency of the different states. In the case of the SG, where prior history of the states is available, a more suitable τ^x can be estimated from the maximum known possible deviation of the states using that history. This will then lower the probability of false alarms during an attack as according to (3.5.2), the particular state will not be considered as attacked as long it falls within a certain range defined by past experiences and history of the states.

(b) Once the states that are currently under attack have been identified, the predictor sends the corrected values of those states, according to past experiences, to the planner. This predicted state will then be integrated to the objective function of the RL in CRC as will be explained later. In this architecture, the CRC resolves the different risks raised by the targeted states through the shunt cycles. The higher the number of shunt cycles, the more effective the CRC becomes at handling the amount risks being faced.

iii. Adaptor: The adaptor is involved with the adaptation part of predictive adaptation. Although not shown in Fig. 3.1, the adaptor is also coupled to the perceptual memory. As the name suggests it, the adaptor is concerned with the adaptation process of the CDS to the current perturbed situation by picking the best action according to past experiences. During CRC, the adaptor receives $\tilde{\mathbf{B}}_k$ from the predictor. $\tilde{\mathbf{B}}_k$ is a vector whereby the identified affected states have been replaced with their closest match in the perceptual memory. From there on, the adaptor is now tasked to find the closest experience in the perceptual memory that matches $\tilde{\mathbf{B}}_k$. This can be done through metrics such as sum of squared errors or Euclidean distance. Once the nearest match is found, its indices is used to find the corresponding counterpart in the executive memory. As mentioned previously, each row or column of this executive memory contains stored past actions applied when the system was under normal uncertainty and the same row or column in perceptual memory refers to their sensory context. In some ways, it is a matching of inputs to outputs similar to how a Neural network [55] is applied. The closest matched set of actions is then applied to

the system. Nevertheless, those actions, which are the closest weight combination in executive memory for the current situation, only aim at keeping state estimation for unaffected states and matched states under control. It does not directly address the issue that FDI attacks goal's are to drive the states towards some predetermined values by the attackers. This section covered how to bring CC under risk control using past experiences according to the principles of the brain. In the next section, pertaining to CRC, we will show how we can address these attacks using mitigations plans while still building on the material presented on risk control, predictive adaptation and the supporting structures introduced up to now.

3.5.0.5 Cognitive Risk Control

CRC is the dual of CC in the CDS. Although they consist of the same components, they are driven very differently. CRC extends from past experiences and predictor for risk sensitive goal oriented action on the SG. Realizing that the ultimate goal of the FDI cyber-attack is the deviation of specific predetermined states which subsequently triggers a domino effect of bad control decisions, the action space is drastically different from its counter part in CC. Here the latter consists of carefully selected tuning parameters to be applied on the system configuration **H** to counter the mal-intent deviations. Referring to Fig. 3.1, the planner in the posterior executive receives input from both the predictor and the generative model. Consequently, the signal originating from the perceptor represents the current situation while the predictor's input is the desired situation according to past experiences. Since the predictor identified the states under the highest risk, the planner's mitigation plan consists of tuning the

parameters in $\mathbf{H}$ without disrupting the estimation of the other states. Although the true root causes of FDI attacks is due to attacks on unprotected sensors or false data injected, in our approach, we are treating $\mathbf{H}$ as a realistic root cause because of its important role for state estimation based on input meter readings. Moreover, during an FDI attack, the hackers will tend to target specific states and drive them to particular values. As a result, the parameters of the system is core for the effectiveness of such attacks. Since the essence of those attacks is the manipulation of the states, we can simplify the objective function of the RL algorithm as the minimization of the absolute Euclidean distance or errors between the current affected states and the desired predictor's states until they fall within the safety limits mentioned earlier. Since only one mode can be operational at a time (either CC or CRC), the notations used in the planning and rewards calculation will be similar to the ones used in the section pertaining to CC. However, their actual definitions during CRC will vary slightly. Thus, the planning steps proceed as follows:
Apply action $a_k^{i,j}$ to the j^{th} column of $\mathbf{H}_k$:

$$\tilde{\mathbf{H}}_k = a_k^{i,j} \cdot \mathbf{H}_k \tag{3.5.3}$$

where j refers to the index of the one of the affected states provided by the predictor, which has been compromised. For example, if the second state is attacked, the $j = 2$ and the second column of $\mathbf{H}$ will be tuned. $\tilde{\mathbf{H}}_k$ is now the new configuration matrix where the j^{th} column has been altered. $\tilde{\mathbf{H}}_k$ is then used to calculate a planned state estimate $\hat{\mathbf{x}}_k^p$ for the current PAC:

$$\hat{\mathbf{x}}_k^p = (\tilde{\mathbf{G}}_k^p)^{-1} \tilde{\mathbf{H}}_k \mathbf{W}_k \mathbf{z}_k \tag{3.5.4}$$

where the estimated gain $\tilde{\mathbf{G}}_p^k$ is calulated from

$$\tilde{\mathbf{G}}_p^k = \tilde{\mathbf{H}}_k' \mathbf{W}_k \tilde{\mathbf{H}}_k \tag{3.5.5}$$

Since each individual shunt cycles during CRC is dedicated to solving one of the risks at a time, the reward associated with a particular action a_k^i is calculated as

$$r_k^i = \frac{(|x_{PE}^i - x_k^i|) - \left(|\hat{x}_{PE}^i - \hat{x}_k^{p,i}|\right)}{|x_{PE}^i - x_k^i|} \tag{3.5.6}$$

where x_k^i is the current affected state i at instant k, x_{PE}^i is its predicted value according to past experience and $\hat{x}_k^{p,i}$ is the planned value of that state if action a_k^i is applied. The purpose of the denominator in is to scale the value of the reward between 0 and 1. Thus, this maintains the consistency of the Bayes-UCB RL algorithm which is applied. In fact, the RL will prioritize actions that will bring the current state under attack to the one closest to past memories. Similar to CC, the working memory holds temporarily the actions having the highest quantile according to Bayes-UCB. However, once the affected state is brought within a range $\mathbf{x}_{\text{lim}}$ of the state's past experience, it is considered that the risk is under control and hence no action will be applied to that particular column of $\tilde{\mathbf{H}}_k$ since it will no longer be considered as a threat during the consequent PAC. Nevertheless, $\tilde{\mathbf{H}}_k$ will remain the current configuration as long as the CDS is operating under CRC. Referring back to the time component aspect of risk control which was mentioned earlier, there will be a situation when all the risks have been neutralized and no longer exist. In the case of the SG, it means that there will be a time when all the attacks will stop. Consequently, a mechanism is needed to identify that condition and restore the CDS back to CC. That mechanism

should also be able to restore the system configuration back to its original unaltered $\mathbf{H}_k$. The methodology describing this process is explained in the next section.

3.5.0.6 Task-Switch Control Restoration

In order to restore the CDS back to its primary purpose, which is CC, the TSC is equipped with a memory and a watchdog timer (WDT). Asides from $\tilde{\mathbf{x}}_{PE}$ and $\mathbf{x}_{\text{lim}}$, the predictor also calculates $\tilde{\mathbf{B}}_{PE}$ and $\mathbf{B}_{\text{lim}}$, which are the predicted output from generative model and safety limits for these values respectively. However unlike (3.5.1), $\tilde{\mathbf{B}}_{PE}$ can be calculated directly by using the average of the values stored for the different stored $\mathbf{B}_j$ values in the perceptual memory. $\mathbf{B}_{\text{lim}}$ is then calculated in a similar fashion as $\mathbf{x}_{\text{lim}}$ as was mentioned earlier. When the risk(s) is first detected, the TSC memory will store the current system configuration $\mathbf{H}_k$ in its memory as reference set point. During every PAC, the predictor will use this stored configuration in TSC memory, which we will denote as $\mathbf{H}_{TSC}$, to estimate the states. When these accumulator values are now in conformity of $\tilde{\mathbf{B}}_{PE}$ and $\mathbf{B}_{\text{lim}}$, a WDT will trigger for the next T_{TSC} perception action cycles. This process can be performed in a similar process to eqref3eq51:

$$\varepsilon_k^b = \Big[\big(\tau^b \times \mathbf{B}_{\text{lim}}\big) - |\mathbf{B}_k - \tilde{\mathbf{B}}_{PE}|\Big] < 0 \tag{3.5.7}$$

where ε_k^b is a logical vector that checks if the previously identified attacked states ε_k^x are within a certain safety range according to past experiences, and τ^b is a vector that defines the deviation tendency of the accumulator values as τ^x was for the DC estimator. Once the timer starts, the CDS will still operate under CRC but no mitigation actions will be performed. Moreover the stored configuration $\mathbf{H}_{TSC}$ will

replace $\tilde{\mathbf{H}}_k$ temporarily during the state estimation process. This allows the CDS to be proactive in case an attack occurs during this transition state. Once the WDT reaches T_{TSC} and if $h_{k|k}$ is greater than γ, F switches are then switched ON and the P switches will be OFF. Consequently, the abnormal uncertainty due to the attack is considered no longer existent and thus the CDS will be restored back to CC. When this occurs, the system configuration in the DC state estimator will change back to the original configuration $\mathbf{H}_{TSC}$ prior to the attack. In a practical scenario, if the configuration had changed during the attack, then the latter can be used.

3.5.0.7 Complete Algorithm

After a detailed description of the posterior executive for CRC, we can now extend the CDS algorithm in [7] to incorporate the latter. Table 3.2, taken from [7], shows all the notations associated with CC in the Frontal executive while Table 3.1 contains those associated with CRC in the posterior executive as well as the new functions of TSC. Since the focus of this paper is the CRC component of the CDS for FDI attack in the SG, the CC aspect of the algorithm will be covered briefly. For a more elaborate account of CC algorithm for control and cyber-attack detection in the SG, the reader is advised to [7]. Algorithm 2 describes the whole process of the CDS during every PAC. Since the Bayes UCB algorithm uses a frequentist approach and relies on bounded rewards of the Bernoulli reward distribution, the variable *BayesReward*, defined in lines 49 to 54, is used to maintain the consistency of the CC algorithm. As the cumulative rewards can become negative during the planning process, the threshold, f, is used in conjunction with *BayesReward* to alleviate the effects of negative saturation of rewards. Thus, in the event that an action, which was

previously considered inadequate for a long time, has now become the best action for the current situation, the use of f to bound the cumulative negative rewards, allows the RL algorithm to build up the quantile for that action faster to eventually be picked by achieving the highest quantile. Consequently, f gives an opportunity for unselected actions to redeem themselves whenever the right situations arise. Since Bayes UCB relies on bounded rewards in [0,1], the application of *BayesReward* on lines 50 and 52, is meant to deal with those situations. Hence, referring to the context presented on line 50, whereby the cumulative rewards has been negatively saturated, *BayesReward* is assigned a reward of 0 and used to update the respective quantile. For the same reason, on line 52, where negative saturation has not yet been attained, *BayesReward* will serve a similar purpose. Finally, on line 54, if the rewards have been positive, then *BayesReward* will be assigned those values and used for the update of the quantile values. In order to stabilize the algorithm at startup, the quantiles relating to the default configuration of the weights are initially biased with a value of α. Furthermore, cognitive confidence cycles, n_{cc}, on line 58, are the number of PAC cycles, during which the cognitive controller of the Frontal executive will gain an initial impression of the quantile values of the different prospective actions. However, those actions are not applied during those cycles. Once the n_{cc} cycles have elapsed, then RL algorithm will start applying the best actions learnt during those cycles. As a result, this helps to alleviate the random chaotic behavior of RL algorithms at startup. In the case of the SG, this kind of behavior can be detrimental, unless it is constrained in some way. Lastly, it was shown in [7] that this CC algorithm, tailored for the SG, has the ability to bypass one of key limitations of RL algorithms by being able to apply multiple actions across the different meters in a controlled approach.

In [7], the ability of the entropic state to detect FDI attacks was also demonstrated. The mechanism of the TSC, which is part of the feedback channel, to switch between the Frontal executive and posterior executive is demonstrated on lines 24 to 36. The reason for having the two conditions on line 28 is to provide a smooth transition from CRC to CC. This will be shown in the computational experiments in the next section. On line 36, resetting the Predictor, whenever a transition from CC to CRC has occurred, allows it to adapt to new situations using the evolving executive and perceptual memories. Moreover, the Predictor, on line 65, is run only once during this shift, thereby saving computational resources and making the algorithm very efficient. Similar to line 36, the tables relating to Bayes UCB are reset whenever a new transition has occurred. This is because the FDI attacks could be targeting other states at a different instance and thus new different mitigation actions have to favored compared to actions which have been used to resolve threats in the past. Consequently, CRC is very robust and has good adaptive properties similar to how the brain adapts to its ever-changing environment. Line 73 of the algorithm allows us to cut down the size of the action space effectively in half by using the output of the Predictor as reference. Line 74 enables the CDS to quickly switch the system configuration back to its original configuration. However, while WDT has not reached T_{TSC} PACs, CRC will still be the dominant mode, ready to counter-attack if a new threat emerges. Lastly, since ε_k^x is calculated during each PAC, once an affected state is brought within the safety limits calculated by the Predictor, the latter will no longer be considered a threat in the subsequent cycles and no actions will be performed on it unless the attack changes form.

Table 3.1: Summary of Notations for CRC

Notation	Definition
$\tilde{\mathbf{x}}_{PE}$	Predicted states by predictor
L_{PE}	Time span over which the executive and perceptual memories extend.
$\mathbf{B}_j$	j^{th} accumulator value stored in perceptual memory
$\mathbf{x}_{\text{lim}}$	Safety limits of $\tilde{\mathbf{x}}_{PE}$
ε_k^x	Logical vector where the current states, under attack, have been indentified
τ^x	Vector that defines the deviation tendency of the states
$\tilde{\mathbf{B}}_k$	Vector where the index of the attacked states in ε_k^x have been used to replace the relative $\mathbf{B}$ values in $\mathbf{B}_k$ with their closest match in perceptual memory
$\tilde{\mathbf{H}}_k$	Modified system configuration, $\mathbf{H}_k$, applied in DC state estimation during CRC
$a_k^{i,j}$	Prospective action involving the virtual application of a^i to the j^{th} of $\mathbf{H}_k$ or $\tilde{\mathbf{H}}_k$
$\hat{\mathbf{x}}_k^p$	Hypothesized state estimate during planning
$\tilde{\mathbf{G}}_p^k$	Hypothesized gain during planning
x_k^i	i^{th} affected state at instant k
x_{PE}^i	Predicted value of the i^{th} affected state according to past experiences
$\hat{x}_k^{p,i}$	Hypothesized value of state x_k^i if action a_k^i is applied
$\tilde{\mathbf{B}}_{PE}$	Predicted output of generative model
$\mathbf{B}_{\text{lim}}$	Safety limits of $\tilde{\mathbf{B}}_{PE}$
ε_k^b	Logical vector that checks if the accumalor values in $\mathbf{B}_k$, relevant to the identified states in ε_k^x, are within a certain safety range
τ^b	Vector that defines the deviation tendency of the accumalor values, $\mathbf{B}_k$
$\mathbf{H}_{TSC}$	System configuration, $\mathbf{H}$, of the DC state estimator before switching to CRC mode
T_{TSC}	Maximum recorded elapsed time for the WDT

Table 3.2: Summary of Notations for CC from [7] (Part 1 of 2)

Notation	Definition
M	Total number of PAC cycles
N	Total number of shunt/planning cycles
n_{cc}	Number of cognitive confidence cycles
$\mathbf{z}_k$	Vector of measurements taken at cycle k
$\mathbf{H}_k$	System configuration matrix at cycle k
$\mathbf{x}_k$	Vector of calculated states by DC state estimator at cycle k
$\mathbf{W}_k$	Weight matrix at cycle k
$\mathbf{B}_k$	Vector retaining the cumulative sum of the states at cycle k
L	Window over which the past states is being accumulated
$\hat{\mathbf{B}}_{k\|k}$	Filtered estimate of the cumulative sum from the generative model at cycle k
$\mathbf{P}_{k\|k}$	Process error covariance matrix at cycle k
$\hat{\mathbf{B}}_{k+1\|k}$	Predicted estimate of the cumulative sum for cycle k+1
$\hat{\mathbf{P}}_{k+1\|k}$	Predicted error covariance matrix for cycle k+1
$\hat{\mathbf{B}}_{k\|k-1}$	Predicted estimate of the cumulative sum for current cycle k which was calculated during the previous cycle k−1
$\mathbf{P}_{k\|k-1}$	Predicted error covariance matrix for current cycle k which was calculated during the previous cycle k−1
$h_{k\|k}$	Entropic state at cycle k
d_k	Number of negative elements along the diagonal of $h_{k\|k}$
γ	Threshold for attack detection
A	Set of all possible actions stored in action space
A_1	Set of selected prospective actions for planning
$a_k^{i,j}$	Prospective action involving the virtual application of a^i to meter j during planning
$\mathbf{W}_k^{i,j}$	Modified weight matrix where meter j's weight value has been replaced by a^i during planning
$\mathbf{G}_k^p$	Hypothesized gain during planning

Algorithm 2: Complete Algorithm for implementation of CC and CRC in the defined structure

Initialization:
memstate := short-term memory for storing L past outputs from the DC state estimator
memexec := Executive memory for storing L_{PE} past contents of CC working memory
mempercept := Perceptual memory for storing the respective L_{PE} past outputs of generative model for the different states
P and F Switches := P switches for CC and F switches for CRC
A_1 := set of selected unique cognitive actions for planning
a_m := load working memory with default configuration of weights for all meters
BayesReward := short-term memory for storing cummulative entropic reward
Set action quantiles of the default configuration to α
$\mathbf{B}_0$, $\hat{\mathbf{B}}_{1|0}$, $\mathbf{P}_{1|0}$, $\mathbf{W}_0$, n_{cc}, f, γ, τ^x, τ^b, T_{TSC}
$a_m \leftarrow a_0$
$k \leftarrow 1$

Begin **while** $k \leq M$ **:**
    DC State Estimation:
    Caculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
    $memstate \leftarrow \mathbf{z}_k$
    Generative Model:
    Calculate $\mathbf{B}_k$ using the past L outputs from *memstate*
    Bayesian Filter/Kalman Filter:
    Update Steps:
    Calculate $\mathbf{K}_k$, Calculate $\hat{\mathbf{B}}_{k|k}$, $\mathbf{P}_{k|k}$
    Prediction Steps:
    Calculate $\hat{\mathbf{B}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$
    Feedback Channel:
    Calcualte $h_{k|k}$
    if *P Switches are ON* **:**
        Caluculate ε_k^b, raise $flag$ if it conforms with past experiences
        Continued on next page

(*Continuing as from line 27 onwards*)
Caluculate ε_k^b, raise $flag$ if it conforms with past experiences
if *flag is raised and* $h_{k|k} > \gamma$ **:**
    Increament WDT
    if *WDT is equal to* T_{TSC} **:**
        Switch OFF $PSwitches$, Turn ON $FSwitches$
        Reset WDT, Set $\mathbf{H}_k$ to $\mathbf{H}_{TSC}$
if $h_{k|k} < \gamma$ *and F Switches are ON* **:**
    Switch OFF $FSwitches$, Switch ON $PSwitches$
    Store current $\mathbf{H}_k$ as $\mathbf{H}_{TSC}$
    Reset Predictor
Frontal Executive:
if *F Switches are ON* **:**
    Planning:
    all cognitive actions $a_k^{i,j} \in A_1$
        Calculate corresponding $\mathbf{B}_k^p$
        Calculate corresponding $h_{k|k}^p$

        Internal reward:
        Calculate $r_k^{i,j}$

        Learning:
        Update $S_a^{i,j}$ for $a_k^{i,j}$
        Update $N_a^{i,j}$ for $a_k^{i,j}$
        if $S_a^{i,j} < f$ **:**
            $S_a^{i,j} = f$
            $BayesReward = 0$
        else if $S_a^{i,j} < 0$ **:**
            $BayesReward = 0$
        else:
            $BayesReward = S_a^{i,j}$
        Update $Q_a^{i,j}$ for $a_k^{i,j}$ using $BayesReward$
        Continued on next page

(*Continuing as from line 55 onwards*)
Update $Q_a^{i,j}$ for $a_k^{i,j}$ using *BayesReward*
if $Q_a^{i,j} > Q_{a_m}^{i,j}$ **:**
  $a_m^{i,j} \leftarrow a_k^{i,j}$
if $k > n_{cc}$ **:**
  Policy:
  Apply a_m to the DC state estimator
  mempercept$\leftarrow \mathbf{B}_k$
  memexec$\leftarrow a_m$

Posterior Executive:
if *P Switches are ON* **:**
  if *CC transitioned to CRC* **:**
    Predictor:
    Calculate $\tilde{\mathbf{x}}_{PE}$, $\mathbf{x}_{\text{lim}}$, $\tilde{\mathbf{B}}_{PE}$, $\mathbf{B}_{\text{lim}}$ using *mempercept*
    Initialise tables for RL and working memory
  Calculate ε_k^x , $\tilde{\mathbf{B}}_k$
  Adaptor:
  Find best set of actions in *memexec* that matches $\tilde{\mathbf{B}}_k$ and apply them
  Using ε_k^x , check their polarity with their respective counterparts in $\tilde{\mathbf{x}}_{PE}$
  For those states which had their polarity flipped, multiply their respective column in $\mathbf{H}_k$ by -1
if *WDT > 0* **:**
  Set $\mathbf{H}_k$ to $\mathbf{H}_{TSC}$
else:
  Planning:
  all cognitive actions $a_k^{i,j} \in A_1$
    Internal reward:
    Calculate $r_k^{i,j}$
    $r_k^{i,j}$ is set to 0 if it is negative
    Learning: (Continued on next page)

(*Continuing as from line 82 onwards*)
Learning:
Update $S_a^{i,j}$ for $a_k^{i,j}$
Update $N_a^{i,j}$ for $a_k^{i,j}$
Update $Q_a^{i,j}$ for $a_k^{i,j}$
if $Q_a^{i,j} > Q_{a_m}^{i,j}$ **:**
    $a_m^{i,j} \leftarrow a_k^{i,j}$
Policy:
Apply a_m to $\mathbf{H}_k$ or $\tilde{\mathbf{H}}_k$
$k = k + 1$

Table 3.3: Summary of Notations for CC from [7] (Part 2 of 2)

Notation	Definition
$\hat{\mathbf{x}}_k^p$	Hypothesized state estimate during planning
$\mathbf{B}_k^p$	Hypothesized cumulative sum involving $\hat{\mathbf{x}}_k^p$ during planning
$h_{k\|k}^p$	Hypothesized entropic state during planning
d_k^p	Number of negative elements along the diagonal of $h_{k\|k}^p$
$r_k^{i,j}$	Internal reward associated with each prospective action $a_k^{i,j}$
a_m	Action stored in working memory
$S_a^{i,j}$	Cumulative reward for action $a^{i,j}$
$N_a^{i,j}$	Number of times action $a^{i,j}$ has been chosen
$Q_a^{i,j}$	Quantile of action $a^{i,j}$
c	real parameter for Bayes-UCB
u	Number of prospective actions to select from action space
f	Negative rewards saturation threshold
α	Quantile initial bias

3.6 Computational Experiments

In this section, CC and CRC will be applied to two different bus networks in the presence of FDI attacks. In the first experiment, involving a 4-bus network, it will be demontrated how CC and CRC work in conjecture to bring the attack under control. Since this is a small network consisting of a small number of states, it is shown in greater detail how the generative models, states, entropic state, weight values and switches evolve through Fig. 3.3, 3.4 and 3.5 both in presence and absence of attack. The second experiment is intended to show the scalability and robustness of the algorithm when it is applied to a bigger network whereby a larger number of states are under attack. As IEEE bus networks have been used as benchmarks for simulations in the other papers referenced in this paper and pertaining to this topic, the IEEE 14-bus network was chosen for the second experiment. Due to the larger number of states in this network compared to the first experiment, the results of this experiment will mostly focus on the states being attacked and how they are brought under control with CRC. For both experiments, *MATPOWER* [56], which is an Electric Power System Simulation and Optimization Tools for MATLAB and Octave, was used to extract the data and network configurations required for the simulations.

3.6.1 Experiment on 4-bus network

The first experiment comprises of a 4-bus, 2-generator transmission network case from [13] as shown in Fig. 3.2. Initially, the state values for the three buses was calculated by the solving the power flow equations of the network case data in *MATPOWER*. Once the state values were obtained, the signals relating to the seven meters were

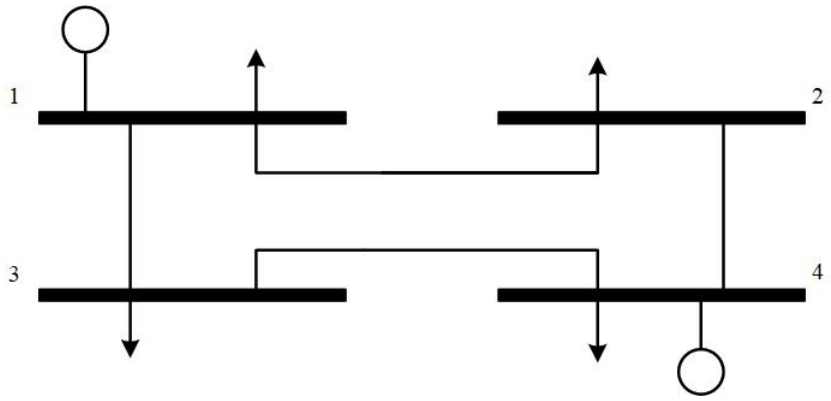

Figure 3.2: Line diagram for 4-bus, 2-generator transmission network case from[13].

then calculated using (3.3.1). From the mean of those signals, their noisy counterparts were then generated with a signal-to-noise ratio(SNR) of 20 dB to create $\mathbf{z}$. For the current network in Fig. 3.2, the network configuration matrix is:

$$\mathbf{H} = \begin{bmatrix} 46.72 & 0 & -26.88 \\ 0 & 42.6 & -15.72 \\ -26.88 & -15.72 & 42.6 \\ -19.84 & 0 & 0 \\ 0 & -26.88 & 0 \\ 26.88 & 0 & -26.88 \\ 0 & 15.72 & -15.72 \end{bmatrix}$$

The simulation is run for 2000s. L, which relates to the generative model in the perceptor was set to 20. The initial estimates for all the incoming accumulator values of the generative model for the Kalman filter is set to 0 and the elements of the diagonal matrix $\mathbf{Q}$ were assigned to 6.25e-04. The diagonal elements of the $\mathbf{R}$ matrix were set to 0.01. From the CC side of the CDS, the action space is made up of a total of 28 actions whereby each meter's weight can be assigned any of the following values:

1 50 100 150. Moreover f was assigned a value of -0.55, $\alpha = 0.5$ and the number of planning/shunt cycles to be evaluated during each PAC was set to 15. CC is started at $\mathbf{t = 300s}$ with 5 cognitive confidence cycles. In the simulation, CC is not started initially at $\mathbf{t = 0s}$ since some time must be allowed for the Kalman filter to settle on the track in order for the algorithm to work properly. The goal of this experiment is to show how CC and CRC work together in the CDS to bring the attack under control. To this end, CRC started at $\mathbf{t = 500s}$ with $\gamma = 0.4$. L_{PE} was set to 60 and τ^x for the different states to: 4 5 7. τ^b was assigned a uniform value of 10 for the relevant accumulator values. T_{TSC} was set to 40s. The action space for CRC consists of a total of 63 possible tuner values whereby the relevant column can tuned with values ranging from 0.9 to 1.2, evenly spaced by 0.015. Lastly, the number of CRC shunt cycles to be evaluated during each PAC when under attack was also set to 15. In this experiment, at $\mathbf{t = 1000s}$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are under attack whereby $\mathbf{x}_1$ is decreased by 0.5 radians and $\mathbf{x}_2$ is incremented by 0.4 radians. The total duration of the attack is 300s. For this simulation, we are considering the most dangerous type of FDI attack, that is the attacker has perfect prior knowledge of $\mathbf{H}$. The results of the experiment is shown in Fig. 3.3, 3.4 and 3.5.

As it can be seen from Fig. 3.5, although the CRC module is enabled at $\mathbf{t = 500s}$, the CDS is still operating under CC until the attack occurs at $\mathbf{t = 1000s}$. Prior to $\mathbf{t = 1000s}$, under the absence of the attack, we can see in Fig. 3.3 and 3.4 that CC chooses the best set weights for the meters for optimal state estimation. By using the cognitive controller, the system gains the special ability of learning from the past and current cycles to select the best set of weights for the future. Once the attacks starts at $\mathbf{t = 1000s}$, there is an immediate drop in entropic state to below the threshold for

detection, $\gamma = 0.4$, which was set initially. In fact, as mentioned and demonstrated in several attack cases in [7], the greater the attack, the greater will be the decrease in $h_{k|k}$. Consequently, the P switches for CRC are turned ON and F Switches for CC are turned off by TSC. During CRC, the CDS operates entirely on the past experiences acquired in the executive and perceptual memories over the past L_{PE} cycles before the attack took place. This also allows the CDS to pay attention and identify the states under attack. On the other hand, the states not under attack are controlled through a different version of CC, rooted in past experiences and the principle of predictive adaptation, as mentioned previously in this paper. Consequently, the CDS is able to correctly recognize that $\mathbf{x}_1$ and $\mathbf{x}_2$ are indeed the states under attack. Under the actions of CRC, the risk associated with the deviation of those states are quickly brought under control. In a realistic scenario, when this kind of attacks occurs, the real values of those states under attack are not available. Hence, the best that we can do is to bring those states to a tolerable threshold before the attack occurred. Thus as shown in Fig. 3.3, $\mathbf{x}_1$ and $\mathbf{x}_2$ are brought to within the thresholds mentioned earlier in about 20 cycles through the actions of CRC. The same can also be said for $\mathbf{B}_1$ and $\mathbf{B}_2$. Furthermore, referring to Fig. 3.4, the weights selection for the different meters is still being done optimally during the attack through the use of past experiences. When the attack is over at $\mathbf{t = 1300s}$, we can see from Fig. 3.5 that the CDS still operates under CRC for 39 additional cycles. During those cycles, the system will wait until certain conditions such as the running out of the watchdog timer and the matching of the past experiences and current experiences before switching back to CC. At $\mathbf{t = 1300s}$, there is a slight inflection in $h_{k|k}$ as the state estimator was operating under a modified $\mathbf{H}$ during CRC to bring the attack under control. Moreover, there

is a small spike in the values of $\mathbf{x}_1$ and $\mathbf{x}_2$ at that instant. Nevertheless, since CRC is intended to bring the system to a condition close to how to it was operating prior to the attack, this also allows the system to recover quickly once the attack is over. When the attack initially occurred, TSC stored the current $\mathbf{H}$ before CRC took over. Hence as shown in Fig. 3.3, TSC is able to switch back to that previous $\mathbf{H}$ within a few cycles after the attack is over and the conditions mentioned previously are met. Lastly, this experiment also demonstrates the impact of the entropic state for control and attack detection in the architecture described. Similar to the principle of cognition rooted in the brain, the architecture proposed shows that by mimicking such properties, it is possible to bring the problem of risk associated with cyber-attacks under the control.

As fluctuations in the voltage angles are common disturbances in power systems, this experiment provides some greater detail on how the CDS differentiates those normal disturbances which are probabilistic in nature and the disturbances associated with an attack which are deterministic in nature. Any disturbance, that affects the states, propagates to the generative model as a consequence of their relationship in this architecture. As a result, this then influences the computed value of the entropic state. Since the entropic state is an embodiment of the grid's performance, the results of the simulation in Fig. 3.3 shows how the entropic state is expected to behave under normal conditions and how it decreases when an attack takes place. However, since the goal of the algorithm is the optimization of the entropic state to a value of 1, it tries to keep those disturbances to a minimum and maintain the evolution of the states in a controlled manner. As shown previously, the deviation of the states when the attack takes place propagates through the generative model which consequently

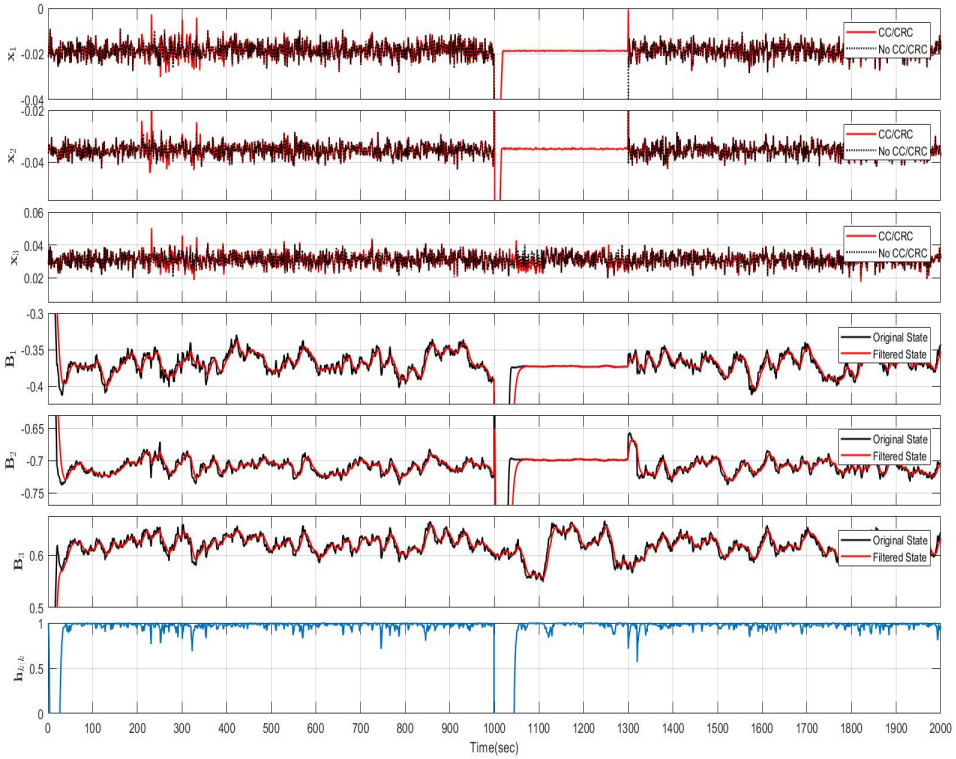

Figure 3.3: Graphs of States, Generative Models and Entropic State

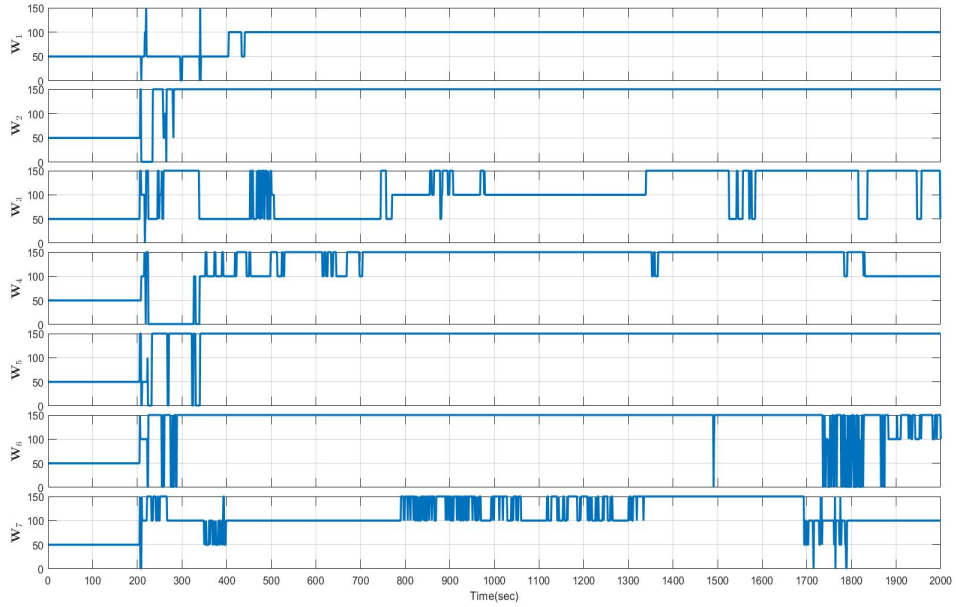

Figure 3.4: Graphs of weight values of the meters with time

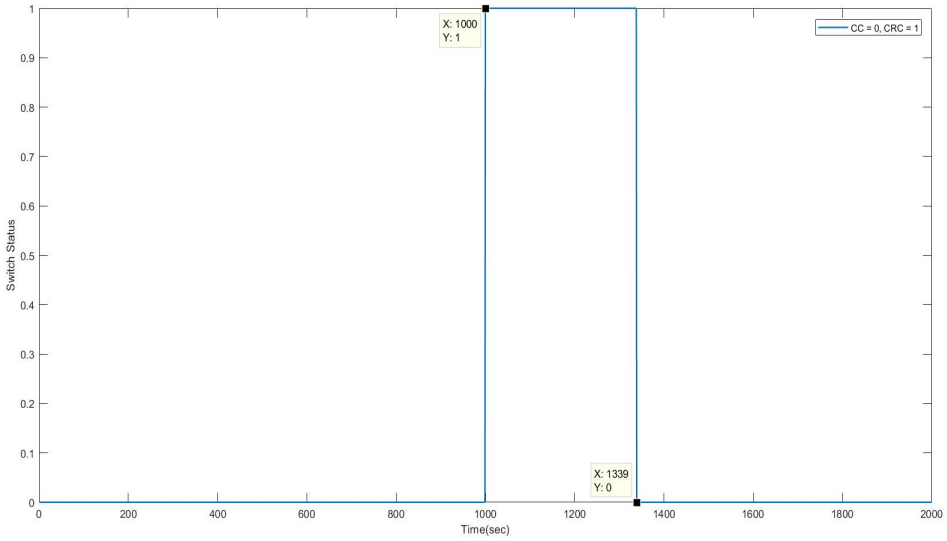

Figure 3.5: Graphs of status of Switches

causes a significant decrease in $h_{k|k}$. Nevertheless, during CRC, through the use of past experiences, the architecture is able to bring $h_{k|k}$ under control as shown in Fig. 3.3.

3.6.2 Experiment on IEEE 14 bus network

In this second experiment, the scalability of this algorithm towards a bigger network and larger attack is investigated. To this end, an attack will be targeted at six of the thirteen states of the network. Similar to the first simulation, the running time is 2000s and L have the same values. 30 cognitive confidence cycles were evaluated. The diagonal elements of the $\mathbf{Q}$ matrix were assigned to 0.0144 and those of the $\mathbf{R}$ matrix were set to 0.01. For the CC side of the CDS, the action space consists of the same possible weight values for the different meters as mentioned previously. Additionally, f was assigned the same value of -0.55, $\alpha = 0.35$ and the number of planning/shunt

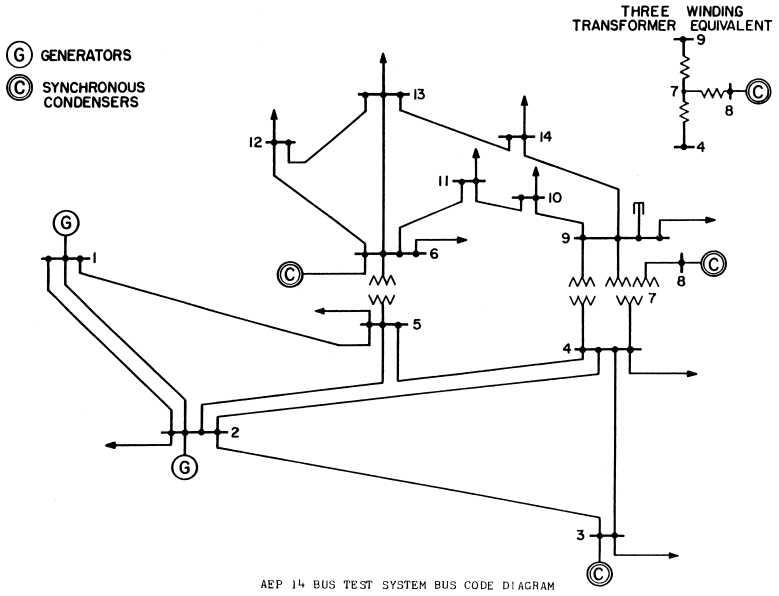

Figure 3.6: Line diagram for the IEEE 14 Bus network.

cycles to be evaluated during each PAC was set to 30. Referring to the CRC side of the CDS for the experiment, γ was set to 0.4 and L_{PE} to 60. τ^x was a uniform vector of 5 for the different states while τ^b was similar to the first experiment. Moreover T_{TSC} and the action space were also the same. Since the network is bigger, the number of CRC shunt/planning cycles was increased to 20. Asides, the other unmentioned parameters were the same as the first experiment. The attack occurs at $\mathbf{t = 1000s}$ and affects the states in the following way: decrease $\mathbf{x}_1$ by 0.5 radians, decrease $\mathbf{x}_2$ by 0.6 radians, increase $\mathbf{x}_3$ by 0.3 radians, decrease $\mathbf{x}_6$ by 0.8 radians, decrease $\mathbf{x}_8$ by 0.6 radians and increase $\mathbf{x}_9$ by 0.5 radians. Thus, the FDI attack in this case targets almost half of the total number of states. The attack time and duration follow the same pattern as in experiment one. The results of the simulation for the IEEE 14 bus network is shown in Fig. 3.7 and 3.8. As the network is bigger, the results will pertain only to the affected states and the switch status. Referring to Fig. 3.7, it can

be seen that the attack is detected immediately as it occurs at $\mathbf{t = 1000s}$ with the drop in the entropic state and that the system is effectively able to bring the risk, associated with the deviations of attacked states, under control through the use of past experiences. Moreover when the attack is stopped at $\mathbf{t = 1300s}$, it takes the CDS around 61 cycles to switch back to CC mode as shown in Fig. 3.8. Thus, the algorithm is very robust and can be effectively scaled to tackle FDI attacks in bigger networks.

The two computational experiments demonstrated in this section were carried out on a system running Windows 10 with an Intel i7-8750H processor. The computational running time of the first experiment was $\mathbf{1.8s}$ and $\mathbf{3.2s}$ for the second experiment. Referring to [7], where it was discussed how the use of the entropic state of the CDS as detection method for cyber-attack has lower computational complexity compared to other detection methods if applied in a medium or large-scale power system, it was shown in this paper and simulations how indeed it is possible to upgrade the architecture in [7] to be able to handle attacks by using the brain as inspiration. Furthermore, the application of CRC in the CDS for the SG is novel since it is a revolutionary system catering to the triple purpose of control, attack detection and attack mitigation in the grid. In order to scale up the architecture described in this paper for bigger networks, the number of CC shunt cycles will have to be increased as a larger number of meters will be involved and required to be evaluated. Furthermore, it is beneficial to still keep the action space small to allow the planned rewards, during planning, to be relatively differentiable from each other. Moreover the diagonal elements of $\mathbf{Q}$ also have to be scaled up when applied to larger grids. In the first experiment, the diagonal elements of that matrix was 6.25e-04 while in the second

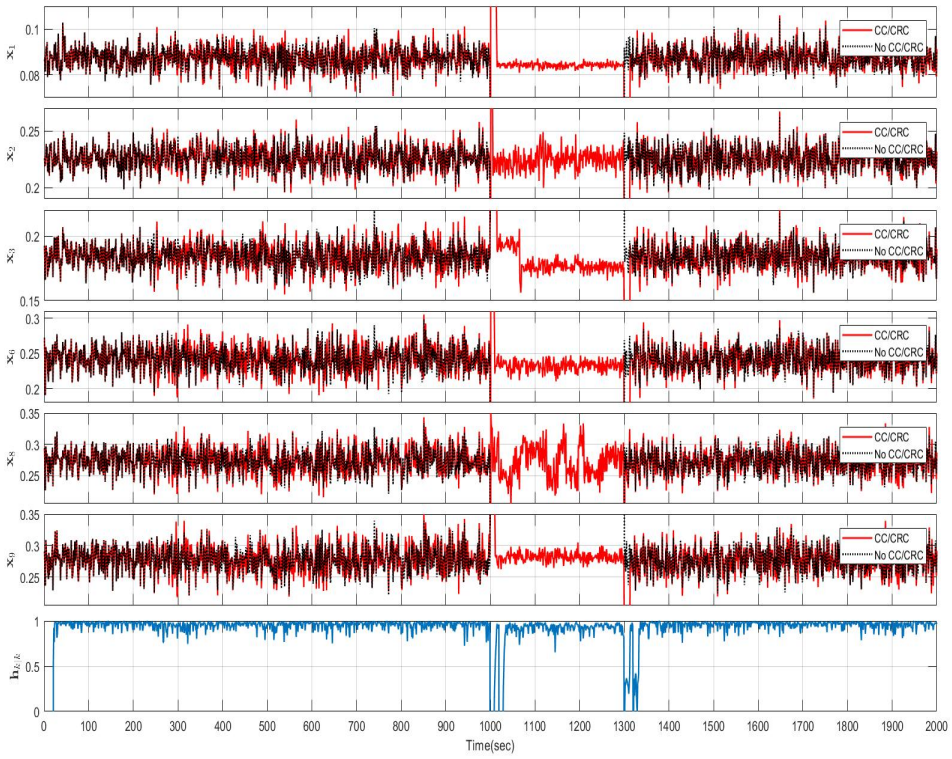

Figure 3.7: Graphs of States, Generative models and Entropic State

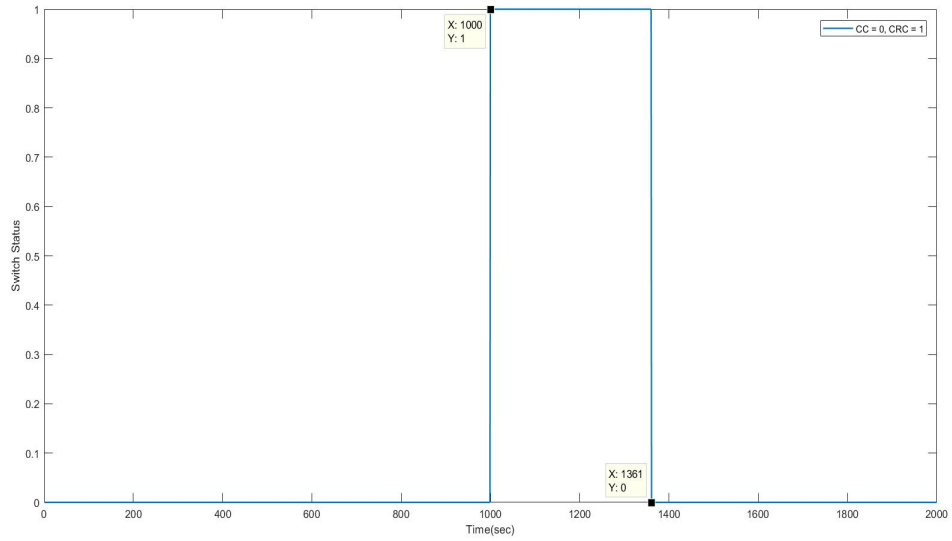

Figure 3.8: Graphs of status of Switches

experiment it was 0.0144. However unlike many applications where the elements of the **Q** matrix are supported by a mathematical formulation [56], the constituents of **Q** in this system have to be defined by the designer as **Q** serves the dual purpose of control for CC and sensitivity of attack detection. Consequently, prior simulations using past historical data can be used to determine the right value of **Q** for the application. [7] goes into more detail in their determination. For bigger networks, the number of CRC shunt cycles can also be scaled up according to the number of states and response time that the designer will find more convenient contingent to past available historical data or simulations.

In both simulations, the CDS was quick to detect the attack in a matter of a few cycles because of the nature of the attack. Nevertheless, the detection time can increase slightly when other types of FDI attack such as the slowly evolving ramp attack is applied. Consequently, in this case, the sampling time of the DC state estimator will have a large impact on the overall performance. For example, if the states are estimated every 4s, then it can take up to 40s to detect such kinds of attack depending on their intensity. However, in a realistic scenario, many states will have to be targeted at the same time to drive the grid to a bad situation. The detection property of the architecture excels in such situations; the greater the amount of the states under attack, the greater will be the deflection in the entropic state. As a matter of fact, the system will be very robust for practical applications. On the other hand, the detection accuracy can also be improved at the cost of higher sampling frequency. The elements of **Q** can also be tuned or the threshold for attack detection increased. It was also shown how the CDS embedded with the state estimator to give rise to the new construct explained in this paper is able to mitigate the attack as it is

able to modify the current **H** used in state estimation in such a way that it nullifies the hacker's attack which was designed using his prior knowledge of **H**. From the CRC side, it is possible to mitigate the attack faster by choosing more appropriate tuner values. Lastly, the determination of τ^x and τ^b using past historical will be convenient if the system is applied in a practical scenario. In the first and the second experiment, τ^x was different as the deviation tendency is bigger in a smaller network consisting of a smaller number of meters, compared to a larger network where the larger number of meters provides a smoother deviation tendency.

3.7 Conclusion

This paper is novel for the following reasons:

i. This is the first time that we have been able to incorporate the CRC with the previous the structure referenced in our earlier paper to give rise to a new construct that is able to bring the problem of attack in the SG under control. We believe that the architecture proposed has great potential in handling the risks associated with attacks that the grid will face in the future as the networks become more interconnected.

ii. The CRC algorithm presented in this paper, tailored for the SG, is revolutionary as it is possible to unite the important topics of risk control theory, neuroscience and control theory to bring about a new way of tackling risk the way the brain would approach it. Moreover, the algorithm is more elaborate, extensive and powerful that the one mentioned in the CRC paper referenced in the introduction.

iii. The construct described is able to serve the triple purpose of a new kind of control and attack mitigation, which are both based on cognition, and cyber-attack detection.

In this paper, it was shown how the previous structure in [7] was united with CRC to bring the issue of attack in the SG under control. With the CDS, being a construct rooted in the brain, it was demonstrated how it was possible to bring together neuroscience and risk control theory to give rise to the architecture discussed in the paper. The computational experiments were able to validate the effectiveness, robustness and scalability of the algorithm in bigger networks. A discussion regarding the choice of parameters was also provided. Since the DC estimation model was the main model in this paper, future orientation of research in this topic can be geared towards the application of this construct in AC state estimation, where reactive components are involved. Compared to the DC model, the AC estimation model is recursive in nature. Consequently, a new procedure to unite the AC estimation model and the perceptor will have to be designed to make the process computationally efficient. Although CRC has also been applied in other fields such as Vehicular Radar Systems [57], where optimal state estimation and tracking are crucial, CRC generally has to be tailored for the desired application. To that end, the mathematics involving the perceptor and both executives have to be modified according to the goals of the different intended applications.

Another potential of the CDS for this application, which was not explored in this paper, is the identification of the attacked sensors. During CRC, the estimated measurements, rooted in past experience, can calculated using an a modified version

of as follows:

$$\hat{\mathbf{z}} = \mathbf{H_{TSC}}\hat{\mathbf{x}} \tag{3.7.1}$$

and the absolute estimated errors, $\hat{\mathbf{e}}$, associated with the measurements can be calculated using

$$\hat{\mathbf{e}} = |\mathbf{z} - \hat{\mathbf{z}}| \tag{3.7.2}$$

A suitable threshold can then be defined to identify the attacked meters. Thus, if the absolute error associated with a particular meter is above that threshold, then it is highly likely that the sensor is under attack. Consequently, through this identification process, the operator can initiate corrective measures accordingly.

Bibliography

[1] S. Haykin, "Cognitive dynamic systems [Point of view]," Proc. IEEE, vol. 94, no. 11, pp. 1910–1911, Nov. 2006.

[2] S. Haykin, "Cognitive Dynamic Systems: Radar, Control, and Radio", Proc. IEEE, Point of View Article, vol. 100, no. 7, pp. 2095-2103, July 2012.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Sel. Areas Commun., vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] S. Haykin, "Cognitive radar: a way of the future." IEEE signal processing magazine 23.1 (2006): 30-40.

[5] M. Fatemi and S. Haykin, "Cognitive control: Theory and application," IEEE Access, vol. 2, pp. 698–710, Jun. 2014.

[6] S. Haykin, J. M. Fuster, D. Findlay, and S. Feng, "Cognitive risk control for physical systems," IEEE Access, vol. 5, pp. 14 664–14 679, Jul. 2017.

[7] M. I. Oozeer and S. Haykin, "Cognitive Dynamic System for Control and Cyber-Attack Detection in Smart Grid," in IEEE Access, vol. 7, pp. 78320-78335, 2019. doi: 10.1109/ACCESS.2019.2922410

[8] J. M. Fuster, Cortex and Mind: Unifying Cognition. Oxford University Press, 2003.

[9] Y. Wang M. Amin J. Fu H. Moussa "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids" IEEE Access 2017.

[10] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-Physical Systems Security – A Survey", IEEE Internet of Things Journal, vol. PP, no. 99, pp. 1-1, 2017.

[11] F. C. Scheweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model," IEEE Trans. Power App. Syst. , vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.

[12] K. P. V. Priya, J. Bapat, "Bad Data Detection in Smart Grid for AC model", IEEE Indicon 2014.

[13] J. J. Grainger and W. D. Stevenson, Power System Analysis. 1st ed., New York, NY, USA: McGraw-Hill 1994.

[14] Y. Liu P. Ning M. Reiter "False data injection attacks against state estimation in electric power grids" ACM CCS pp. 21-32 2009.

[15] A. Abur and A. Gómez-Expósito, Power System State Estimation Theory and Implementation. Boca Raton, FL, USA: CRC Press, 2004.

[16] A. Monticelli, "State Estimation in Electric Power System A Generalized Approach", Springer Science+Business Media New York, 1999.

[17] Y. Gu, T. Liu, D. Wang, X. Guan, Z. Xu, "Bad data detection method for smart grids based on distributed state estimation", Proc. IEEE Int. Conf. Commun. pp. 4483-4487 2013.

[18] R. Deng, G. Xiao, R. Lu, H. Liang, A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks impacts and defense: A survey", IEEE Trans. Ind. Informat. vol. 13 no. 2 pp. 411-423 Apr. 2017.

[19] D. Wang, X. Guan, T. Liu, Y. Gu, Y. Sun, Y. Liu, "A survey on bad data injection attack in smart grid", Proc. IEEE PES Asia-Pac. Power Energy Eng. Conf. pp. 1-6 2013.

[20] A. Anwar, A. N. Mahmood, M. Pickering, "Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements", J. Comput. Syst. Sci. vol. 83 no. 1 pp. 58-72 2016.

[21] J. Hao, R.J Piechocki, D. Kaleshi, et al: 'Sparse malicious false data injection attacks and defense mechanisms in smart grids', IEEE Trans. Ind. Inf.s, 2015, 11, (5), pp. 1–12 (doi: 10.1109/TII.2015.2475695).

[22] J. Jiang, Y. Qian, "Defense mechanisms against data injection attacks in smart grid networks", IEEE Commun. Mag. vol. 55 no. 10 pp. 76-82 Oct. 2017.

[23] K. Manandhar, X. J. Cao, F. Hu, Y. Liu, "Combating false data injection attacks in smart grid using kalman filter", Proceedings of International Conference on Computing Networking and Communications Communications and Information Security Symposium pp. 16-20 2014.

[24] K. Manandhar, X. Cao, F. Hu, Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter", IEEE Trans. Control Netw. Syst. vol. 1 no. 4 pp. 370-379 Dec. 2014.

[25] P.Y. Chen, S. Yang, J. A. McCann, J. Lin, X. Yang, "Detection of false data injection attacks in smart-grid systems", IEEE Commun. Mag. vol. 53 no. 2 pp. 206-213 Feb. 2015.

[26] Y. Liu, L. Yan, J. Ren, D. Su, "Research on efficient detection methods for false data injection in smart grid", International Conference on Wireless Communication and Sensor Network (WCSN) pp. 188-192 December 2014.

[27] D. B. Rawat, C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems", IEEE Signal Process. Lett. vol. 22 no. 10 pp. 1652-1656 Oct. 2015.

[28] Y. Gu, T. Liu, D. Wang, X. Guan, Z. Xu, "Bad data detection method for smart grids based on distributed state estimation", Proc. IEEE Int. Conf. Commun. pp. 4483-4487 2013.

[29] A. A. Cárdenas, S. Amin, S. S. Sastry, "Secure control: Towards survivable cyber-physical systems" Proc. 28th Int. Conf. Distrib. Comput. Syst. Workshops pp. 495-500 2008-Jun.

[30] D. I. Urbina et al., "Limiting the impact of stealthy attacks on industrial control systems", Proc. ACM SIGSAC Conf. Comput. Commun. Secur. pp. 1092-1105 Oct. 2016.

[31] Z. Yu, W. Chin, "Blind false data injection attack using PCA approximation method in smart grid", IEEE Trans. Smart Grid vol. 6 no. 3 pp. 1219-1226 May 2015.

[32] J. Kim, L. Tong, and R. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," IEEE Transactions on Signal Processing, vol. 63, no. 5, pp. 1102–1114, March 2015.

[33] L. Liu, M. Esmalifalak, Q. Ding, V. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization", IEEE Transactions on Smart Grid, vol. 5, no. 2, pp. 612–621, March 2014.

[34] A. Anwar, A. N. Mahmood, M. Pickering, "Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements", J. Comput. Syst. Sci. vol. 83 no. 1 pp. 58-72 2016.

[35] X. Fang, S. Misra, G. Xue, D. Yang, "Smart grid - the new and improved power grid: A survey", IEEE Commun. Surveys Tutorials 2012.

[36] P. McDaniel, S. McLaughlin, "Security and privacy challenges in the smart grid", IEEE Security Privacy vol. 7 no. 3 pp. 75-77 May/Jun. 2009.

[37] J. M. Fuster "The prefrontal cortex makes the brain a preadaptive system" Proc. IEEE vol. 102 no. 4 pp. 417-426 Apr. 2014.

[38] Y. Huang et al., "Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis", IEEE Syst. J. vol. 10 no. 2 pp. 532-543 Jun. 2016.

[39] S. Li, Y. Yilmaz, X. Wang, "Quickest detection of false data injection attack in wide-area smart grids", IEEE Trans. Smart Grid vol. 6 no. 6 pp. 2715-2735 Nov. 2015.

[40] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems,". Journal of Basic Engineering, 82: 34–45, 1960

[41] A. S. Debs, R. E. Larson "A dynamic estimator for tracking the state of a power system", IEEE Trans. PAS vol. PAS-89 pp. 1670-1673 September/October 1970.

[42] E. A. Blood, M. D. Ilic, J. Ilic, B. H. Krogh, "A Kalman filter approach to quasi-static state estimation in electric power systems", 38th North American Power Symposium pp. 417-422 2006 2006.

[43] A Saikia, RK Mehta, "Power system static state estimation using Kalman filter algorithm", EDP Sciences. 2016; 7: 1-7.

[44] C. E. Shannon, "A mathematical theory of communication", Bell Syst. Tech. J., vol. 27, no. 3, pp. 379-423, Jul./Oct. 1948.

[45] R. S. Sutton and A. G. Barto, "Reinforcement Learning", Cambridge, MA, USA: MIT Press, 1998.

[46] E. Kaufmann, O. Cappé and A. Garivier, "On Bayesian upper confidence bounds for bandit problems" Proc. Int. Conf. Artif. Intell. Stat. pp. 592-600 2012.

[47] G. Burtini, J. Loeppky, and R. Lawrence,"A survey of online experiment design with the stochastic multi-armed bandit", CoRR, abs/1510.00757, 2015.

[48] E. Kaufmann, "Analysis of Bayesian and frequentist strategies for sequential resource allocation," in Machine Learning. Télécom Paris-Tech, 2014. [Online]. Available: https://pastel.archives-ouvertes.fr/tel-01413183/document

[49] P. Reverdy, V. Srivastava ,N. E. Leonard, "Modeling human decision-making in generalized Gaussian multi-armed bandits", Proc. IEEE vol. 102 no. 4 pp. 544-571 Apr. 2014.

[50] P. G. Smith and G. M. Merritt , "Proactive Risk Management: Controlling Uncertainty in Product Development", Productivity Press, 2002

[51] J. M. Fuster, "The prefrontal cortex makes the brain a preadaptive system", Proc. IEEE, vol. 102, no. 4, pp. 417-426, Apr. 2014.

[52] J. M. Fuster, "Cortex and memory: Emergence of a new paradigm", J. Cogn. Neurosci., vol. 21, pp. 2047–2072, 2009.

[53] K. Friston, "Cognitive dynamics: From attractors to active inference", Proc. IEEE, vol. 102, no. 4, Apr. 2014, DOI: 10.1109/JPROC.2014.2306251.

[54] D.G. Manolakis, V.K. Ingle, S.M. Kogon, "Statistical and Adaptive Signal Processing: Spectral Estimation Signal Modeling Adaptive Filtering and Array Processing", MA Boston:McGraw-Hill 2000.

[55] S. Haykin, "Neural Networks and Learning Machines", 3rd ed. Prentice-Hall, 2009.

[56] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER:

Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education," Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12-19, Feb. 2011.(Digital Object Identifier: 10.1109/TPWRS.2010.2051168)

[57] S. Feng and S. Hakin, "Cognitive Risk Control for Transmit-Waveform Selection in Vehicular Radar Systems." IEEE Transactions on Vehicular Technology 67.10 (2018): 9542-9556.

Chapter 4

Cognitive Dynamic System for AC State Estimation and Cyber-Attack Mitigation in Smart Grid

4.1 Preceding Introduction

Little research has been conducted on the AC model of the SG and FDI attacks. While the AC model of the SG is a more realistic depiction of that entity, FDI attacks can still occur in that non-linear version of the SG, although it is harder to achieve compared to the DC model. Compared to the DC model, the AC model is both more complex and computationally expensive than its DC counterpart. In this chapter, the concepts from the previous two chapters are re-engineered to create a new architecture that embodies both the AC state estimator and the CDS for state estimation and cyber-security in the SG. The new construct brings forward CC and a new form of CRC to handle the limitations of the non-linear grid.

To the best of the author's knowledge, the scholarly work presented herein is the first experimental work of CDS being applied to the AC model of the SG for state estimation and cyber-attack mitigation.

The publication included in this chapter is:

M. I. Oozeer, and S. Haykin, "Cognitive Dynamic System for AC State Estimation and Cyber-Attack Mitigation in Smart Grid," IEEE Access. under review, 2020.

The co-author's contributions to the above work include:

i. Technical supervision and financial support of the study presented in this work.

ii. Manuscript revising and editing.

Abstract

This paper is an extension of our previous research of bringing together the Cognitive Dynamic System (CDS) and the Smart Grid (SG) by focusing on AC state estimation and Cyber-Attack mitigation. Under the nonlinear AC power flow model, state estimation is complex and computationally expensive as it relies on iterative procedures. On the other hand, the False Data Injection (FDI) attacks are a new category of cyber- attacks targeting the SG that can bypass the current bad data detection techniques in the SG. Due to the complexity of the nonlinear system involved, the amount of published works on AC based FDI attacks have been fewer compared to the DC based FDI attacks. In this paper, we will demonstrate how the entropic state, which is the objective function of the CDS, can be used as a metric to monitor the grid's health and detect FDI attacks as well. Furthermore, Cognitive Risk Control (CRC) will be introduced to bring the attack under control once it is detected. The CDS, acting as the supervisor of the system, improves the entropic state on cycle to cycle basis by dynamically optimizing the state estimation process through the reconfiguration of the weights of the sensors in the network. In order to showcase performance of this new structure, computer simulations are carried out on the IEEE 14-bus system for optimal state estimation and FDI attack mitigation.

4.2 Introduction

4.2.1 Cognitive Dynamic System

The Cognitive Dynamic System (CDS) is an organized physical model and research tool that is based on certain features of the brain. Following its first introduction in [1], it was later expanded in [2] leading to its first applications in cognitive radio [3] and cognitive radar [4]. Since then, CDS has progressed enormously to give rise to Cognitive Control (CC) [5] and Cognitive Risk Control (CRC)[6] as two of its particular functions. Using those principles, the CDS was first merged in [7] with the Smart Grid (SG) to form a new structure, based on the DC state estimation model, that shows tremendous potential for handling the possible problems that the SG will be facing in the near future. Furthermore, in [8], the construct presented in [7] was expanded to include a more complex CRC that is closer to the brain. In that paper, it was proven how this new approach can to be used to mitigate the problem of cyber-attack in the SG. From a neuroscience perspective, the CDS is founded on Fuster's paradigm of cognition comprising of the following five principles: perception-action cycle (PAC), memory, attention, intelligence and language [5]. In its simplest form, the CDS is built on two main components: the perceptor, on one side, and the executive on the other with the feedback channel uniting them together. In [7], it was shown that the integration of the over-arching function of CDS, CC, with the SG, is well adapted for slowly progressing cyber-physical systems. In this paper, we will re-engineer the construct presented in [7], where the DC-estimation model was involved, to be able to carry out AC state estimation optimally and also be able to detect cyber-attacks. In order to do so, the perceptor of the CDS will incorporate a generative

model that will allow it to sense and control the environment indirectly. Moreover, in order to bring forward the cognitive ability of the CDS and make it compatible with the current nonlinear state estimation in SG, the steps involved in the state estimation process will be re-engineered in a novel way. It will also be shown how the entropic state, which is the objective function of the CDS, will be instrumental in implementing a control-sensing mechanism that is capable of identifying and handling bad measurements. We will also show how this entropic state serves as the basis for detecting False Data Injection attacks (FDI) in SG. Moreover, once the cyber-attack is detected, CRC will be introduced to bring the attack under control. The CRC module, presented in this paper, has also been re-engineered from our previous work in [8] so as to be less computationally expensive and be able to cooperate with the heavy load of AC state estimation. Ultimately, this version of CRC is not only backwards compatible with the DC model, but can also serve as a general solution for applications where state estimation is critical.

4.2.2 Smart Grid

The next generation of engineering systems consisting of the Internet of Things (IoT) and Cyber-physical systems (CPSs) are currently paving the way towards the fourth industrial revolution [10]. As those systems are gradually occupying a more prominent role in our daily lives, through applications in critical infrastructures such as electrical power grids or transportation systems, the cyber-security aspects of those systems will also grow in importance [11]. In the context of this paper, emphasis will be laid upon the SG and its most dangerous threat known as False Data Injection (FDI) attacks. More specifically, compared to our previous research where the DC model for state

estimation was investigated [7], focus will be laid upon on the AC model, which is more a realistic representation of the smart grid, and introducing the CDS for a new way of control and FDI attack detection and mitigation.

Making use of all the new generation of sensing, monitoring and control strategies, the SG is forecasted to be a more powerful entity than the traditional power grid in many facets such as reliability and efficiency [12][13]. In the SG, the Supervisory Control and Data Acquisition systems (SCADA) is responsible for monitoring and processing the main control actions by collecting meter measurements from remote terminal units (RTUs) consisting of different field devices or sensors. Through a process known as state estimation, those measurements are then processed and analyzed for errors and inconsistencies after being transmitted to a control center [15][16]. The state variables that are calculated by this process usually consists of the voltage magnitudes and angles of the different buses in the system [17]. The measurements used for state estimation are the currents, real and reactive power flows, power injections and voltage magnitudes and angles. In the DC model, the state variables are the bus angles only while in the more complex AC model, the voltage magnitudes and angles of the different buses in the network are estimated. Weighted Least Squares (WLS), introduced by Schweppe [15], is the technique used for the power system state estimation using those measurements. In order to enhance the accuracy of the estimated states, another process, known as Bad Data Identification, is carried out to remove bad measurements. Bad measurements are erroneous measurement readings that will impact state estimation negatively. The most commonly applied bad data identification techniques are the Chi-Squared Tests and Largest Normalized Residual Test [16][18]. Those statistical tests rely on the residuals between the estimated states

and the measurement residuals to identify the bad data. In the case of an FDI attack, bad data, which can be bypass the previously mentioned tests, is introduced into the system such that the estimated states can be modified stealthily. Those bad data are maliciously crafted offsets to measurements that are injected to the sensor readings so as the state estimation process is influenced in a particular way. Consequently, with the incorrect calculated states, bad control decisions will be applied.

Although FDI attacks have been a popular topic of research over the past years [21], most of the works, e.g., in [10; 11; 12; 13; 14], investigated the FDI attacks on the DC model. Few works have been published on the AC model and those attacks [21; 25; 23]. Nevertheless, the DC model is just a simplified representation of the nonlinear AC state estimation model. There are major differences between the two models that could explain why the AC model has been unpopular. Firstly, in the nonlinear state estimation model, the estimated states are obtained after undergoing iterations while in the DC model, those states are obtained in closed-form. Moreover, the linear state estimation relies on active power flow analysis [17][19][20]. On the other hand, the AC model uses both active and reactive power flow analysis. Furthermore, the state variables in the DC model consists of the voltage angles only while the states in the AC model consists of both the voltage angles and magnitudes. Consequently, these differences raises the complexity and computational expense of nonlinear state estimation as a topic of research when it comes to FDI attacks [24]. In fact, DC based FDI attacks can be detected by AC-based data detection techniques [25]. Hence, since the AC model is commonly applied in power systems, finding a way to detect these attacks and mitigating them under that environment is going to be very important for the coming years.

4.2.3 Contribution and Organization

The main contributions of this paper can be summarized as follows:

i. The architectural architecture of the CDS, tailored for AC state estimation and FDI attack detection in the SG, is presented. Compared to our earlier work in [7], which was based on the DC model, we will show how that construct can re-engineered with the goal of nonlinear state estimation and computational efficiency in mind. Consequently, it will be shown how the CDS allows for optimal state estimation with relatively less computations, using the principles of cognition rooted in the brain.

ii. To expand on our previous research, the entropic state will be re-introduced for two purposes namely; 1) it serves as a metric of the grid's health on a cycle to cycle basis and 2) it is used in the detection of FDI attacks. The optimization of the entropic state is the goal of the cognitive controller residing in the executive of the CDS. The latter does this by selecting the most optimal actions that will maximize the available information from one PAC to the next. Thus, a novel algorithm, based on CC, is presented for nonlinear state estimation in the SG. Simulations are performed on the IEEE 14-bus network to show the efficiency of this new approach using the CDS. By learning which measurements to prioritize and which ones to neglect, the CDS showcases a new way of control for bad data correction and FDI attack detecton with the SG being the topic of application.

iii. Using the entropic state as the basis for control and cyber-attack detection, the CRC principles introduced in [8] for the SG will be used as basis to mitigate the attack. Since the AC model is more computationally expensive than its DC

counterpart, a simplified CRC subsystem will be added to the architecture in order to mitigate the attack once it is detected. It will be shown in simulations that this sub-system can remain as effective and efficient as its dual in[8] when it comes to bringing attack under control.

The rest of the paper is organized as follows: In Section 4.3, the basic concepts of state estimation and data detection for the DC and AC model will be presented and contrasted. The mathematics of FDI attacks for the linear and nonlinear will also be demonstrated. Section 4.4 expands on the structure of the CDS for the SG. Since this research is an extension of [7], the material presented in that paper will be re-engineered for this new application. In the context of the CDS, the SG is considered as the environment with which it interacts. Section 4.5 expands on FDI mitigation using the principles of predictive adaptation for the CRC module.The CC/CRC algorithm for nonlinear state estimation is also presented in that section. Section 4.6 gives a discussion on the application and simulation results of the algorithm on the IEEE 14-bus network. It will be shown how this new structure is able to handle the two problems of bad data detection and FDI attack detection simultaneously. After being detected, CRC will take over the CDS to bring the attack under control. Finally, Section 4.7 concludes this paper by highlighting the key results and presenting new avenues for research of this novel construct.

4.3 Preliminaries

4.3.1 Weighted Least Squares State Estimation

In order for the Energy Management System (EMS) to operate properly, it is important for the SCADA to provide the latter with the required measurement data so that correct control decisions can be applied in real-time. However, as those signals are often contaminated with noise, filtering is carried out by both the state estimator and the bad data detector to obtain the most accurate states. However, since power systems comprise of an overdetermined system whereby redundant measurements are taken, the filtering process allows the discarding of those erroneous measurements that will be detrimental for state estimation.

4.3.2 DC Model

The states of a power system refer to the bus voltages angles θ and bus voltage magnitudes V. In the case of the DC model, the states are restricted to the bus angles only and the measurements consist of the real power flows and injections. Additionally, it is assumed that prior knowledge relating to the bus magnitudes is available and those are taken to be close to unity. After choosing a reference bus and setting it to zero radians, the state estimation in the linear system is simplified to only estimating the n bus voltage angles $[\theta_1, \theta_2, ..., \theta_n]^T$. The DC power flow model has been a popular research tool for power engineers and smart grid cyber-security researchers as it serves as a linearization and approximation of the AC power flow model [15; 26; 27; 28]. In fact, this substitution to the AC model has been widely accepted for reasons such as guaranteed faster convergence and reduced algorithmic

complexities [29].

According to the DC estimation model, it is assumed that the bus voltage magnitudes are already known and are close to or equal to 1.0 per unit. Moreover, shunt elements and branch resistances are neglected. Thus, by approximating the first order Taylor expansion around $\theta = 0$, the measured real power flow from bus k to m can be approximated using the following formula [13]:

$$P_{km} = \frac{\theta_k - \theta_m}{x_{km}} + e \tag{4.3.1}$$

where x_{km} corresponds to the reactance(in per unit values) of the branch k-m, θ_k is the phase angle(in radians) at bus k and e is the measurement error. Therefore, the power injection at a specified bus i can be obtained through the summation of all the flows along incident branches to that bus:

$$P_i = \sum_{j \in N_j} P_{ij} + e \tag{4.3.2}$$

Weighted Least-Squares (WLS) minimization is used to solve the measurement model which is comprised of an overdetermined system of linear equations that relate the measurements to the states. The following formula is applied:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e} \tag{4.3.3}$$

where

- $\mathbf{x}$ is the n vector of the true states (voltate angles)
- $\mathbf{z}$ is the m vector of measurements (active power flows, active power injections,

voltage angles)

- **H** is the m x n Jacobian matrix (relates measurements to states)
- **Hx** is the m vector of linear function linking measurements to states
- **e** is the m vector of measurement errors
- m is the number of measurements
- n is the number of variables

H in (4.3.3), also known as the Jacobian matrix, is a matrix that defines the theoretical calculations that relates the states to the measurement vector **z** and therefore serves as a mathematical description of the power system. These equations are also referred to as the power flow equations and are described as vectors inside **H**. While in the DC model, those entries consists of a set of linear functions of the state variables, those functions are nonlinear as far as the AC model is concerned. In order to solve the weighted least-squares minimization problem, we need to find the n-vector **x** that minimizes the index $J(\mathbf{x})$, which is described as follows:

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H}\mathbf{x})'\mathbf{W}(\mathbf{z} - \mathbf{H}\mathbf{x})' \tag{4.3.4}$$

W in (4.3.4), is a diagonal matrix that contains the measurement weights. These are

based on the reciprocals of the measurement error variance σ:

$$\mathbf{W} = \mathbf{R_z^{-1}} = \begin{bmatrix} \sigma_1^{-2} & \dots & \dots & \dots \\ \dots & \sigma_2^{-2} & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \sigma_m^{-2} \end{bmatrix} \tag{4.3.5}$$

where $\mathbf{R_z}$ is the covariance matrix of the measurement. The performance index $J(\mathbf{X})$ is then differentiated to obtain the first order optimal conditions:

$$\mathbf{G\hat{x}} = \mathbf{H'Wz} \tag{4.3.6}$$

where the estimate of the state $\hat{\mathbf{x}}$ is calculated using:

$$\hat{\mathbf{x}} = \mathbf{G^{-1}H'Wz} \tag{4.3.7}$$

In the above equations, $\mathbf{G} = \mathbf{H'WH}$ is the state estimation gain. (4.3.7) represents the closed form solution to the least squares minimization problem.

4.3.3 AC Model

In the AC model, the nonlinear power flow equations are fundamental for state estimation since they indicate the link between the measurements and the estimated states. Compared to (4.3.1), the active and reactive power for the transmission line between buses k and m are given by

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos(\theta_{km}) - V_k V_m b_{km} \sin(\theta_{km}) \tag{4.3.8}$$

$$Q_{km} = -V_k^2 b_{km} + V_k V_m b_{km} \cos(\theta_{km}) - V_k V_m g_{km} \sin(\theta_{km}) \tag{4.3.9}$$

Additionally, for each bus k, it calculated using the following equations:

$$P_k = V_k \sum_{m \in S_k} V_m(-g_{km} \cos(\theta_{km}) - b_{km} \sin(\theta_{km})) + V_k^2 \sum_{m \in S_k} g_{km} \tag{4.3.10}$$

$$Q_k = V_k \sum_{m \in S_k} V_m(-g_{km} \sin(\theta_{km}) - b_{km} \cos(\theta_{km})) - V_k^2 \sum_{m \in S_k} b_{km} \tag{4.3.11}$$

where $S_k \subset S$ is the set of all buses that have lines connected to bus k and g_{km} and b_{km} are the conductance and susceptance of the line between buses k and m respectively. θ_{km} denotes and the phase angle difference between bus k and bus m. In the AC power flow estimation, the nonlinear relationship between the state variables and the measurement is described as follows:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \tag{4.3.12}$$

where

- $\mathbf{x}$ is the n vector of the true states (voltage magnitudes and angles)
- $\mathbf{z}$ is the m vector of measurements (active and reactive power flows, active and reactive power injections, voltage magnitudes and angles)
- $\mathbf{h}$ is the m x n Jacobian matrix (relates measurements to states)
- $\mathbf{h(x)}$ is the m vector of nonlinear function linking measurements to states

- $\mathbf{e}$ is the m vector of measurement errors
- m is the number of measurements
- n is the number of variables

Relative to the WLS method and (4.3.4), the determination of the state variables is done according to the following criteria:

$$\min J(\mathbf{x}) = (\mathbf{z} - \mathbf{h}(\mathbf{x}))'\mathbf{W}(\mathbf{z} - \mathbf{h}(\mathbf{x}))' \tag{4.3.13}$$

There are several iterative methods, such as Honest Gauss Newton method, Dishonest Gauss Newton method and Fast Decoupled State Estimator [20], which are used to solve (4.3.13). The first order optimality condition of (4.3.13) to be solved is then expressed as:

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} |_{\mathbf{x}=\hat{\mathbf{x}}} = -2\mathbf{F}_h^T(\hat{\mathbf{x}})\mathbf{W}(\mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}))' = 0 \tag{4.3.14}$$

where $\mathbf{F}_h$ is the Jacobian matrix derived from $\mathbf{h}(\mathbf{x})$ and the $\hat{\mathbf{x}}$ is the estimated state vector. In the case of the CDS, the state estimation process is modified slightly in order to remain compatible with the planning stages in the executive, which will be discussed later. Therefore, for the first t_s cycles, state estimation proceeds similar to the iterative procedures mentioned previously. As from t_s, the preceding calculated state of the AC state estimator, $\mathbf{x}_{k-1}$, is used as the initial guess for the current cycle with any of those iterative techniques. Moreover, the number of iterations is also limited to N_s iterations to save on computational resources.

4.3.4 Bad Data Detection

During the state estimation process, faulty measurements have to be detected and identified to be removed as they lead to erroneous calculated states. However, the statistical properties of these errors simplify their detection and identification. In order to determine those errors, the estimated measurements, $\hat{\mathbf{z}}$, are first calculated from (4.3.7) using the following equation for the DC case:

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\mathbf{x}} \tag{4.3.15}$$

In the AC model, this equation is written as follows:

$$\hat{\mathbf{z}} = \mathbf{h}(\hat{\mathbf{x}}) \tag{4.3.16}$$

The individual estimated measurement error is then obtained using:

$$\hat{\mathrm{e}}_j = (z_j - \hat{z}_j) \tag{4.3.17}$$

As these errors follow a zero mean Gaussian distribution [17], techinques such as the Chi-Squares test and normalized residual have been the most common ones applied for their detection [28]. When Chi-squares test is applied, it is assumed that the state variables are mutually independent from each other and the errors follow a normal distribution. The test involves a number of iterative steps that depend on the number of degrees of freedom of the system, sum of squares $\hat{f}$ and a critical value

corresponding to α satisfying the inequality:

$$\hat{f} < \chi^2_{(k,\alpha)} \tag{4.3.18}$$

where k is the appropriate number of degrees of freedom and α is a specified probability. Thus, $\hat{f}$ will be large when a large number of bad measurements are present. However, since k is large in power systems, this method allows for the removal of those measurements that are responsible for the largest standardized residuals.

4.3.5 False Data Injection Attacks

FDI attacks (also known as Bad Injection attacks) are a special category of attacks targeting the SG, whereby bad measurements are injected such that they are able to bypass the bad data detection methods discussed previously. While FDI attacks can also target other cyber-physical systems, various forms of these attacks and consequences have been investigated in [11; 12; 16; 18; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39]. In this paper, FDI attacks will be simulated using assumptions from [27], whereby it is assumed that the system parameters and topology (system Jacobian) is known to the attackers, and [21], where a mathematical formulation for simulating the FDI attack in the AC model is provided. Additionally, FDI attacks satisfying the first assumption regarding prior knowledge of the system have been proven to result in more disastrous consequences. Moreover, in [18], the authors demonstrate how an attacker, using that knowledge of the system matrix $\mathbf{H}_{m\times n}$, can inject an attack vector $\mathbf{a}_{m\times 1}$ to the measurement vector $\mathbf{z}_{m\times 1}$ that remains undetected from the detection techniques mentioned previously. Consequently, with the insertion of $\mathbf{a}_{m\times 1}$, the new

corrupted measurement signals $\mathbf{z'}_{m\times 1}$ takes the following form:

$$\mathbf{z'}_{m\times 1} = \mathbf{z}_{m\times 1} + \mathbf{a}_{m\times 1} \tag{4.3.19}$$

Hence, this will result in the calculation of an incorrect system state vector $\mathbf{x'}_{m\times 1}$ instead of the orginal state $\mathbf{x}_{m\times 1}$. The difference between those states is denoted as $\mathbf{c}$ and is calculated as follows:

$$\mathbf{x'} = \mathbf{x} + \mathbf{c} \tag{4.3.20}$$

Liu et al. provide the mathematical proof for the attacks in the DC model in [18] and demonstrate their results through simulations. Furthermore, it is shown that as long as the attack vector satisfies the condition $\mathbf{a} = \mathbf{Hc}$, then the attack will not be detected according to the following proof [29]:

$$\begin{aligned} ||\mathbf{z'} - \mathbf{Hx'}|| &= ||\mathbf{z} + \mathbf{a} - \mathbf{H}(\mathbf{x} + \mathbf{c})|| \\ &= ||\mathbf{z} + \mathbf{a} - \mathbf{Hx} - \mathbf{Hc}|| \\ &= ||\mathbf{z} - \mathbf{Hx}||(\text{since, } \mathbf{a} = \mathbf{Hc}) \\ \mathbf{r}_{attack} &= \mathbf{r}_{normal} \end{aligned} \tag{4.3.21}$$

As shown in (4.3.21), the residual of the estimation process related to the attack vector and the residual without attack are considered the same. Since the detection methods are based on statistical methods for the calculation of residuals, they are unable to detect the malicious measurement vector. For the AC model, it is shown

in [21] that the attack vector will remain undetected when it satisfies the condition:

$$\mathbf{a} = \mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\mathbf{x}) \tag{4.3.22}$$

It is then proven as follows:

$$\begin{aligned}
\mathbf{r}_{attack} &= \mathbf{z}' - \mathbf{h}(\mathbf{x}') \\
&= \mathbf{z} - \mathbf{h}(\mathbf{x}') + \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{x}) \\
&= \mathbf{z} + \mathbf{a} - \mathbf{h}(\mathbf{x}') + \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{x}) \\
&= \mathbf{r} + \mathbf{a} - \mathbf{h}(\mathbf{x}') + \mathbf{h}(\mathbf{x}) \\
\mathbf{r}_{attack} &= \mathbf{r}_{normal} \text{ (since, } \mathbf{a} = \mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\mathbf{x})\text{)}
\end{aligned} \tag{4.3.23}$$

Consequently, in the case of nonlinear state estimation, it is more complicated to implement the FDI attack. Compared to (4.3.21), where the attacker only required knowledge of the Jacobian matrix, in the AC model, the latter is now additionally required to have some prior knowledge of the current states of the system. While it is more complicated to meet those conditions, it is still shown in [21] that such an attack is possible and the consequences can be disastrous. In both the DC and AC model, the calculation of wrong state variables, caused by this attack, can start a domino effect of incorrect control decisions leading to dire consequences. As this type of attack targets state estimation in the SG predominantly, the vector $\mathbf{a}$ can be inserted physically by tampering with the meters or wirelessly by injecting the offsets when the readings are transmitted to the SCADA. Hence, the substation state estimator (SSE), which is also an important component of the SG, will also be the target of such attacks as it plays an essential role in state estimation at the substations.

4.4 Architectural Structure of CDS for Smart Grid

From a neuroscience perspective, the CDS is the entity that matches Fuster's paradigm [9] the closest as far as cognition is concerned. Basically, the CDS is made up of four components namely; environment, perceptor, executive and feedback channel. Moreover, they are arranged in a very particular way. The feedback channel links the perceptor and executive, which are situated on two opposite sides. The environment finally closes the global feedback channel whereby the entire CDS is contained within it. Since the focus of this paper is the nonlinear state estimation and FDI attack in the SG, the AC state estimator will be considered as the environment with which the CDS interacts since it is the recipient of the measurements in the network. By acting as the supervisor of the network, the CDS empowers the state estimator, through CC, with the cognitive ability to learn during every PAC which measurements to prioritize for optimal state estimation and which of them to discard. Fig. 4.1 shows the complex diagram whereby the CDS and the AC state estimator are brought together for meeting the goals mentioned previously. In the next subsections, it will be elaborated how the arrangement and the role of each constituent plays a major role for goal-oriented action on the SG.

4.4.1 Perception-Action Cycle

When the environment is free of uncertainty, the PAC is responsible for updating the CDS with new information from the environment for every cycle. Thus, with the continuous acquisition of new information from this global feedback loop, the information extraction ability of the perceptor is constantly being improved with each successive cycles. Consequently, this sets up an uninterrupted cyclic directed

flow of information from the perceptor to the executive to lead the PAC with the most optimal actions to be performed on the environment. As a result, this hypothesis for a goal-focused scenario is then modified with new information gained from the PAC to allow the executive to improve its current ability to achieve the primary goal that it was designed for.

4.4.2 Perceptor

Similar to the concept of Percept in the agent of AI [47], both in the brain and CDS, a perception process is performed on incoming measurements. The perceptor of the CDS extracts useful information from the noisy measurements, which subsequently the executive uses to optimize its actions and improve the information gain for the next cycles. Those actions, performed by the executive under CC, are called cognitive actions. However, unlike the role of the percept in AI, the perceptor perceives the environment directly and extracts relevant information from it which in turn the cognitive controller, residing in the executive, uses to sense the environment indirectly. In order to perform its function, the perceptor is made up of the generative model and the Bayesian filter, which are reciprocally coupled to each other.

4.4.2.1 Generative Model

As defined in [6], the first component of the perceptor for the CDS is conceptually the *Bayesian generative model*[6], which acts as classifier for the observables received from the environment. However, in [7], it was argued that due to the dynamic nature of the SG, the Bayesian generative model would not be suitable for this specific application. Due to the complexity of the SG and its adoption for almost all applications, it is

of utmost importance to detect anomalies or cyber-attacks as soon as possible before they can infect the network further, thereby starting a domino effect of cascaded problems throughout the entire network and end users. Therefore, inspired from quickest detection theory, the generative model proposed for the perceptor was based on cumulative sum (CUSUM) and is written as follows:

$$\mathbf{B}_k = \sum_{i=k-L}^{k} \mathbf{x}_i \tag{4.4.1}$$

where k refers to the current cycle number, L is the window over which the past states are being accumulated, $\mathbf{B}_k$ is the vector retaining the cumulative sum for each cycle and $\mathbf{x}_i$ is the vector of the states output from the DC state estimator for the cycle i. While CUSUM-based detection methods has been very effective in detecting FDI attacks in [40][41], they fall short when the attacker has prior knowledge of the threshold applied. Indeed, the latter can then craft an attack that remain undetected. However, the CDS allows us to bypass this problem through the use of the dynamic *entropic state*, as will be elaborated later. The entropic state is the foundation of control and attack detection in this CDS structure adapted for the SG. Lastly, the CUSUM based generative model also possesses some other desirable traits such as the smoothing out of noise operating under the slow dynamics of the SG.

4.4.2.2 Bayesian Filter

The second component of the perceptor is the Bayesian filter, which is coupled to the generative model. Although the equations describing the SG for state estimation are nonlinear in nature, we can linearize the state estimates using the *Kalman* filter and assuming that it is operating under additive white gaussian noise [42]. Since we are

assuming that the power system is quasi-static in nature in this paper [43; 44; 45], we can use the well-known Kalman filter as the Bayesian filter in the perceptor. The Kalman filter is based on the state-space model which operates on a pair of equations known as the Process equation and the Measurement equation respectively. Moreover, under quasi-static assumptions, we can assume that the state variables $\mathbf{x}$ at the time k+1 will only deviate by a small amount from its previous values at its previous cycle k. Consequently we can simplify this relationship to the following equation:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \omega_k \tag{4.4.2}$$

where ω_k is independent Gaussian noise vector with zero mean. Based on (4.4.2), we can propose the measurement equation as follows:

$$\mathbf{Y}_k = \mathbf{L}_k\mathbf{B}_k + \omega_k \tag{4.4.3}$$

and the covariance of matrix ω_k as:

$$\mathbf{R} = \text{diag}[\sigma_\omega^2], \sigma_\omega^2 = \text{var}[\omega_i] \tag{4.4.4}$$

As we are assuming that the system is operating under quasi-static conditions, a random walk model can be employed as the process equation as follows:

$$\mathbf{B}_{k+1} = \mathbf{F}_k\mathbf{B}_k + \mathbf{v}_k \tag{4.4.5}$$

where $\mathbf{v}_k$ is the process noise vector which is assumed to be statistically independent and zero mean. The covariance matrix of $\mathbf{v}_k$ is:

$$\mathbf{Q} = \text{diag}[\sigma_{\text{v}}^2], \sigma_{\text{v}}^2 = \text{var}[\text{v}_i] \tag{4.4.6}$$

Referring to (4.4.3) and (4.4.5), the system matrix $\mathbf{L}_k$ and the predictive transition matrix $\mathbf{F}_k$ are assumed to be identity respectively. In regards to the measurement and process equations mentioned previously, the computational steps of the Kalman filter starts with some predefined initial estimates of the states $\hat{\mathbf{B}}_{k|k}$, and predicted error covariance, $\mathbf{P}_{k|k}$, which are used for the time update steps as follows:
The predicted estimated states of the generative model and predicted error covariance, $\hat{\mathbf{B}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ respectively, are calculated using the following equations:

$$\hat{\mathbf{B}}_{k+1|k} = \mathbf{F}_{k+1,k}\hat{\mathbf{B}}_{k|k} + \mathbf{v}_k \tag{4.4.7}$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_{k+1,k}\mathbf{P}_{k|k}\mathbf{F}_{k+1,k}^T + \mathbf{Q} \tag{4.4.8}$$

When the next cycle starts, those two estimates are then used for the measurement update stages to calculate the Kalman gain, $\mathbf{K}_k$, filtered accumulated estimate , $\hat{\mathbf{B}}_{k|k}$, and to update the process covariance matrix, $\mathbf{P}_{k|k}$, according to the equations below:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{L}_k^T(\mathbf{L}_k\mathbf{P}_{k|k-1}\mathbf{L}_k^T + \mathbf{R})^{-1} \tag{4.4.9}$$

$$\hat{\mathbf{B}}_{k|k} = \hat{\mathbf{B}}_{k|k-1} + \mathbf{K}_k(\mathbf{Y}_k - \mathbf{L}_k\hat{\mathbf{B}}_{k|k-1}) \tag{4.4.10}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{L}_k \mathbf{P}_{k|k-1} \tag{4.4.11}$$

As a result, through the iteration of the time update and measurement update steps, the preceding *a posteriori* estimates are used to predict new *a priori* estimates.

4.4.3 Feedback Channel

The feedback channel has very distinctive roles in the CDS as it completes the PAC by bringing together the perceptor and the executive. At this point in this paper, its first purpose relating to control and cyber-attack detection in the SG will be elaborated. Its other relating purpose relating to bringing the attack under control will be expanded on at a later point in this manuscript. The entropic-information processor is also linked to the Task-Switch Control (TSC), which consists of a set of switches that triggers a sub-system that is responsible for bringing the cyber-attack under control once it is detected. In order for the CDS to supervise the SG, the feedback channel holds the entropic-information processor, which is tasked with calculating the *entropic state* and internal rewards during reinforcement learning in the executive. This will be elaborated in Subsection 4.4.4 (Executive) where it is more relevant to the role of the executive during planning.

4.4.3.1 Entropic-Information Processor

The directed cyclic flow of information from the perceptor to the executive is known as the *entropic state of the perceptor*. The entropic state is built on the principles of the perceptual posterior, which can be viewed as the incoming filtered posterior embodying the essence of the generative model, Kalman filter and entropy, which is

derived from *Shannon's information theory* [46]. According to Shannon's information theory, the entropic state at time k can be formulated as:

$$h_{k|k} = \int_{\mathbb{R}} p(\mathbf{B}_k|\mathbf{Y}_k) \log \frac{1}{p(\mathbf{B}_k|\mathbf{Y}_k)} d\mathbf{B}_k \tag{4.4.12}$$

where $p(\mathbf{B}_k|\mathbf{Y}_k)$ is the perceptual posterior of the Kalman filter and $\mathbb{R}$ denotes the entire space where the state $\mathbf{B}_k$ lies. Assuming that the noise terms in (4.4.3) and (4.4.5) are Gaussian, the posterior $p(\mathbf{B}_k|\mathbf{Y}_k)$ can be simplified to its mean and covariance matrix at each cycle. In [5], it was shown that (4.4.12) can be further simplified to:

$$h_{k|k} = \frac{1}{2}\log(\det\{(2\pi e)\mathbf{P}_{k|k}\}) \tag{4.4.13}$$

where det{.} is the determinant operator. The use of det{.} in (4.4.13), was originally intended to condense the whole information of the matrix into a single number. However in [7], it elaborated that (4.4.13) was not suitable for the SG as it was not sensitive enough to the changes in the environment. Consequently the following expression, that originates from the mentioned paper, will be used instead to calculate the entropic state at time k in this architecture:

$$h_{k|k} = \frac{\mathrm{Tr}\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{Y}_k\}^2)\}}{\mathrm{Tr}\{\mathbf{P}_{k|k-1}\}} \tag{4.4.14}$$

where Tr represents trace operator, diag{.} is the diagonal operator and $h_{k|k}$ is the entropic state. In [7], the efficiency of (4.4.14) for control and cyber-attack detection was proven and illustrated. For this reason, it will be retained for the CDS architecture being elaborated. Mathematically, (4.4.14) simplifies the information between the filtering-error covariance $\mathbf{P}_{k|k-1}$ and the error between the state estimate $\hat{\mathbf{B}}_{k|k-1}$

and current states calculated at cycle k into a single metric. The denominator of (4.4.14) normalizes the equation such that $h_{k|k}$ can only take values ranging from 0 to 1 when the environment is operating in the absence of uncertainty. The degree of disturbance affecting the SG can then be characterized through the entropic state; the lower $h_{k|k}$ is, the greater the amount of disturbance or uncertainty in the system. Since the SG will be facing different situations during its operation such as the normal day to day routine and cyber-attacks, we can further dissociate the entropic state with the two following important properties:

i. When the environment is operating in the absence of uncertainty, $h_{k|k}$ will always be positive because of the probabilistic representation of the uncertainties.

ii. When uncertainties are present, $h_{k|k}$ will fluctuate around values which are less than 1. Thus, to distinguish between normal uncertainties, such as process disturbance, due to the probabilistic nature of the environment, and abnormal uncertainties, such as cyber-attack, a suitable threshold γ can be chosen such that if $h_{k|k}$ is below γ, then this would indicate presence of attack and to switch on CRC.

Consequently these two characteristics will become key for FDI attack detection and mitigation as explained in the next sections.

4.4.3.2 Task-Switch Control

Task-Switch Control (TSC) occupies a central role in the CDS in the activation of the two special set of switches which are responsible for CC and CRC. Only one set of switches can be ON at a time; if the CC switches are ON, then the CDS will be operating under CC and the CRC switches will be OFF. Since the amount of

uncertainty influencing the SG is uncontrollable experimentally, the entropic state allows us to quantify this phenomenon. Hence, this enables us to use the entropic state as an attack detector and risk raiser as follows:

$$\zeta_k = \begin{cases} 0, & \text{if } h_{k|k} > \gamma. \\ 1, & \text{otherwise.} \end{cases} \tag{4.4.15}$$

where ζ_k is the result of detection. Thus when ζ_k is 0, the CC switches are ON and the executive, responsible for CC, will be operating. When ζ_k is 1, the CRC switches are ON and the CRC subsystem, responsible for CRC, will take over. It should be noted that CC switches are OFF during that mode. So far, (4.4.15) explains how the system will transition from CC to CRC. However, in order to revert the system back to CC, another mechanism is required as the entropic state is perturbed and cannot be relied on during the attack. Thus, using the same principles from [8], the TSC is fitted with a watchdog timer (WDT) and internal memory defined as TSC memory. This will be expanded in the section covering the CRC subsystem where the principles of predictive adaptation and risk control theory, discussed in the previously cited paper, will be re-adapted and modified for this CDS structure tailored for the AC model. It is to be highlighted that only one mode of control can be active due to the design of the switches. As a matter of fact, knowing exactly when the attack occurs is critical in order to secure the SG through CRC during the FDI attack.

4.4.4 Executive

From a design perspective, the Executive is the most important entity of the CDS as it is responsible for control of the SG in the absence of uncertainty. With this goal in

mind, it consists of Reinforcement Learning (RL) and Cognitive Control (CC), which can be further subdivided into the action space, planner, working memory and policy.

4.4.4.1 Reinforcement Learning: Bayes-UCB

Asides from its role in the calculation of the entropic state during each PAC, the feedback channel is also involved in the calculation of internal rewards during the planning stages of the RL [47] algorithm in the executive. RL in the CDS is based on the current entropic state at each cycle which is subsequently used to optimize an objective function for optimal control in the network. Before we elaborate on the pivotal role of RL with the other components of the executive, Bayes-UCB [48] RL algorithm will be covered briefly in order to give an overview on how it operates. Bayes-UCB represents the current state of the art from a class of multi-armed bandit algorithms called UCB algorithms [49], which are based on the principle of optimism in the face of uncertainty. In the approach to the multi-armed bandit problem, the algorithm updates the estimate of the reward distribution for each action using a Bayesian method. The action that will be applied is then chosen according to the one that will yield the highest reward. Consequently, Bayes-UCB algorithm is an index policy that uses the prior distribution to pick a dynamic quantile of the posterior estimates for the index for each action. Hence, at each discrete time t, the algorithm will select the action A_t that satisfies the following condition:

$$A_t = \underset{a}{\operatorname{argmax}}\, q_a(t) = Q(1 - \frac{1}{t(\log(t))^c}, \lambda_a^{t-1}) \tag{4.4.16}$$

where $Q(\alpha, \pi)$ refers to the quantile of order α of the distribution π. Moreover, by assuming that the rewards follow a Bernoulli distribution, and when the prior distribution of each action is Beta(1,1), [50] shows that (4.4.16) can be further simplified. To maintain consistency of the used notations in this paper, (4.4.16) can be reduced to:

$$\begin{aligned} A_k &= \underset{a}{\operatorname{argmax}}\, q_a(k) = Q(1 - \frac{1}{k(\log(k))^c}; \\ &\quad \text{Beta}(S_a(k) + 1, N_a(k) - S_a(k) + 1) \end{aligned} \tag{4.4.17}$$

where k is the PAC cycle number, S_a is the cumulative reward for action a, N_a is the number of times action a has been chosen and c is real parameter. As the CDS is a construct that draws its origin from the neuroscience of the brain, it is to be emphasized that Bayes-UCB shares many common traits to the Bayesian approach of decision making in human brains [51]. Following this brief coverage of Bayes-UCB, it will be shown in the next section, pertaining to Cognitive Control, how the RL algorithm integrates the system configuration **H** of the power grid, the generative model of the perceptor and the process model in the Kalman filter together for optimal state estimation.

4.4.4.2 Cognitive Control

CC can be considered in many ways as the heart of the CDS as it brings together all the components, described so far, for goal oriented action on the SG. CC is made up of two important modules namely the *planner* and the *policy*. The planner is involved in the extraction of a set of prospective actions from the action-space A and their evaluation during the planning cycles (i.e., shunt cycles [6] in CDS terminology)

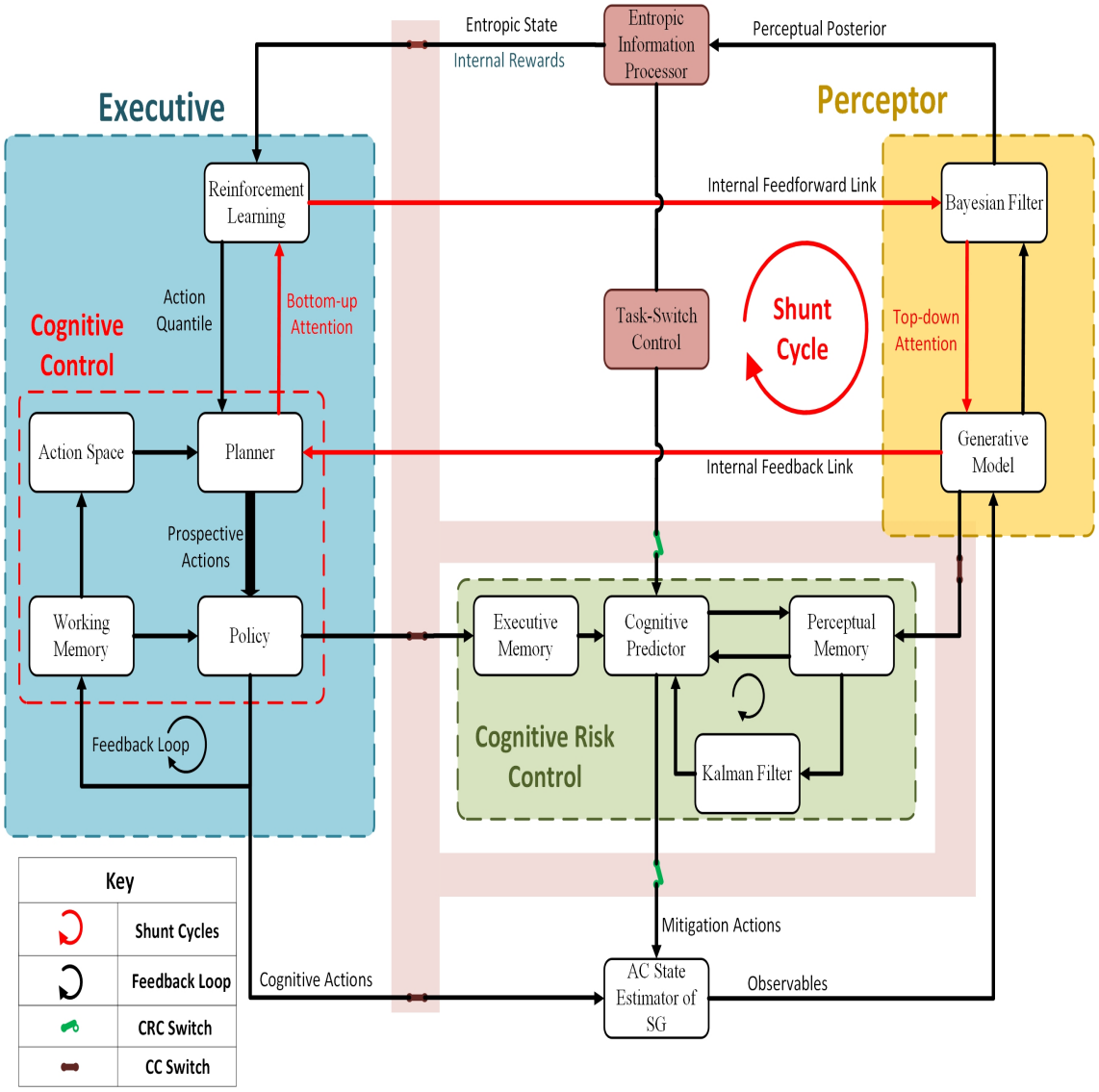

Figure 4.1: Architectural Structure of CDS incorporating CC and CRC for AC State Estimation and Attack Mitigation in SG.

during each PAC. Consequently, under the influence of attention from one PAC to the next, the policy learns the most appropriate actions yielding the maximum rewards to be applied. In the context of the SG, the action space consists of discrete weight values that can be attributed to the different meters. Thus, under the influence of attention,

the CDS will learn the optimal weight values for the different meters for optimal state estimation. Those meters, which are detrimental for the state estimation, will be assigned lower weight values while those, which are crucial, will be given larger weight values as the CDS keeps learning about its environment to better perform its set goal. Planning in CC brings together all the other modules previously discussed. The process starts with the selection of a randomly chosen prospective action $a_k^{i,j}$ which represents weight value a^i for meter j during cycle k. This hypothesized weight value is then applied virtually to the weight matrix $\mathbf{W}$ in (4.3.13) and (4.3.14) to form $\mathbf{W}_k^{i,j}$. $\mathbf{W}_k^{i,j}$ is then used to calculate a new planned state estimate, $\hat{\mathbf{x}}_k^p$, using the same procedures mentioned in the last paragraph of section 4.3.3. Thus, the same preceding calculated state of the AC state estimator, $\mathbf{x}_{k-1}$, is used the initial guess for the current cycle using any of those iterative techniques cited. However, the number of iterations is limited to N_p iterations this time around. Due to the different weight matrices being examined, each iteration of using a $\mathbf{W}_k^{i,j}$ will also involve a different hypothesized gain, $\mathbf{G}_k^p$, during planing. Since state estimation is computationally costly, by doing this process through the steps mentioned previously, this allows the CDS to learn during the planning stages at a lower resource cost. With $\hat{\mathbf{x}}_k^p$ denoting the planned state estimate using the modified weight matrix with the hypothesized weight, the planned cumulative sum involving $\hat{\mathbf{x}}_k^p$ is then calculated:

$$\mathbf{B}_k^p = \sum_{i=k-L}^{k-1} \mathbf{x}_i + \hat{\mathbf{x}}_k^p \tag{4.4.18}$$

where $\mathbf{B}_k^p$ is the planned cumulative sum involving $\hat{\mathbf{x}}_k^p$ instead of $\hat{\mathbf{x}}_k$. Using this new cumulative sum, a planned entropic state, $h_{k|k}^p$, is subsequently calculated as follows:

$$h_{k|k}^{p} = \frac{\det\{\mathbf{P}_{k|k-1} - (\mathrm{diag}\{\hat{\mathbf{B}}_{k|k-1} - \mathbf{B}_{k}^{p}\}^{2})\}}{\det\{\mathbf{P}_{k|k-1}\}} \tag{4.4.19}$$

The presence of uncertainties in the environment, whether stochastic or probabilistic, will cause a deviation in the output of the generative model of the perceptor from the estimated hidden state of the Kalman filter. Hence, the goal of (4.4.19) is to reduce this divergence by finding the best configuration weights for the respective meters. This condition is satisfied whenever the $\mathbf{W}_{k}^{i,j}$ generates $h_{k|k}^{p}$ closer to the optimal value of 1, which implies that the planned estimated state of the AC state estimator reduces the propagated variation in the generative model.

4.4.4.3 Internal Rewards

Moving forward with equations that describe the planning steps, the stage is now set to define to the relationship between the previous steps and the calculation of the internal rewards during RL. The hypothesized internal rewards, $r_{k}^{i,j}$, associated with each prospective action $a_{k}^{i,j}$, for cycle k can be written as:

$$r_{k}^{i,j} = h_{k|k}^{p} - h_{k|k} \tag{4.4.20}$$

As it can be seen from (4.4.20), the objective of RL, when operating under CC, is to minimize the amount of uncertainty in the SG by searching for an improved weight configuration during every PAC that will result in a better entropic state than the previous cycle. In other words, RL attempts to restrict to the amount of uncertainty or disturbance during the state estimation process to the range computed by the Kalman filter in the perceptor. Referring back to the steps described so far that led

to (4.4.20), we can see that the CDS, as defined in this specific architecture, learns from the past and present actions to pick the best actions for the future. To assist in this task, after undergoing the shunt cycles during every PAC, the working memory holds temporarily the actions that have achieved the highest quantile from Bayes-UCB in (4.4.17) and applies them to the system before starting the next PAC. Thus, when the next PAC starts and a new set of prospective actions are evaluated according to their quantile values, if any of those actions achieves a higher quantile than the quantile of its respective meter in the working memory, then the higher achieving action will replace that previously considered best action. This way of performing control in the SG can also be viewed from a Bandit perspective, whereby it can be considered as a Contextual Bandit problem where every cycle presents new situations to be faced. According to those conditions, the actions performed on the SG will modify the system configuration to a new set point, from which the RL algorithm will have to adapt. This then continues on until the CDS is brought to rest.

4.5 FDI Attack Mitigation

It was shown in [7; 8] and mentioned earlier that in the presence of deterministic uncertainties, such as the FDI cyber-attack, the attack can be detected by using a threshold γ on the entropic state as risk raiser. When those situations arise, CC must expand its functionality to be able to face those hostile events [6], also referred to as risks. Through the use of CRC, in [8], it was shown for the first time how the principle of predictive adaptation [58] can be engineered to be applicable to cyber-physical systems which are at the threat of cyber-attacks. In [8], this principle led to the introduction of Posterior Executive and its supporting structures to be able

to bring the attack under control once detected. In that paper, that system was evaluated rigorously with computational experiments whereby many states were being targeted by FDI attacks. It was shown that the resulting algorithm was very efficient at dealing with multiple attacks at the same time and bringing them under control in a timely manner. This was mostly possible due to the fact that the DC model was being investigated in that paper. Since the mathematics of state estimation in the DC model is linear and does not involve recursions compared to the ones involved in the AC model, the planning stages involved during the RL in posterior executive was not computationally expensive. However, when it comes to the AC model, where the state estimation is nonlinear and computationally expensive by itself, the method introduced in [8] for bringing the cyber-attack under control is not practical as it would be a burden from a computational viewpoint. Consequently, a new solution that is able to deal with the computation needs to be introduced such that it would still be as efficient as the methodology, we proposed in our earlier work, and can still be a practical solution if applied on the field. To this end, we propose to introduce CRC through the CRC module which will be consisting of a smaller number of supporting structures. It will be shown later in this paper, how the principle of predictive adaptation can still be re-engineered to face this particular problem under different conditions. Before proceeding further, we will cover briefly the principle of predictive adaptation.

4.5.0.1 Predictive Adaptation

Since the CDS is a construct that is founded on the cognitive neuroscience of the prefrontal cortex, the principle of predictive adaptation, which serves as basis for

CRC, also shares a similar origin in the neuroscience of the brain. According to [58], there is an ability that exists in more advanced organisms, especially humans, that expresses itself to anticipate changes within the organism's environment or itself and to adapt to them before they are predicted to happen. Thus, it can be said that the organism is able to preadapt to these expected changes due to this predictive ability. This preadaptation capability is rooted in its ability to restructure past and deep-rooted individual experience to form new adaptive structures of goal-directed behavior according to their temporal relationship. The author also goes one step further, suggesting that this prediction and preparation ability is one of the dominant functions of the prefrontal cortex. Furthermore, similar to primates and the CDS, the internal PAC becomes the neural framework for prediction and cognitive control. Moreover, according to the cognit model [59], this process is assisted by internal memories, more specifically the perceptual memory and the executive memory. After their formation in the two moieties of the left cortical hemisphere, the perceptual memory gains experiences through information from the senses-in posterior (PTO) cortex while executive memory gains experiences through action-in from frontal cortex. These experiences will be referred to as past experiences in the next sections. Consequently, those stored experiences and emotional inputs will allow the PAC to adapt to unexpected changes. Hence, before and during the pursuit of an objective, the prefrontal cortex will selectively choose past experiences to guide and correct the course of action to that objective. As a matter of fact, this use of past experiences is the foundation of prediction and error correction in the cognitive functions of the prefrontal cortex. For a more detailed narration on the CDS and predictive adaptation, the reader is advised to cosult [8]. Infact, when this principle was first inspired

from [58] and presented for engineering applications in [60], it had the form shown in Fig. 4.2. However, it should be noted that at that time, CDS was still in its early days of conception and not as advanced and established as it is today. While the left of the Nearest neighbor classifier [61], in Fig. 4.2, is now what constitutes the Executive of the CDS in Fig. 4.1, the right side of the latter was relatively new for situations involving risk. Accoring to Fig. 4.2, the window of past of experiences, as the name mentions it, is tasked with storing the past actions that were performed by the policy over a recent lapse of time. The role of the MAP (maximum a posteriori) unit, which is also inspired from Bayesian Inference, is then to select the best experience from this window according to the proposed action of the policy with the help of the Nearest neighbor classifier. Nevertheless, it was still lacking some of the concepts mentioned in [58]. Most recently, in [8], we proposed a new form of CRC which is closer to principle of predictive adaptation in [58]. In that paper, it was demonstrated through simulations how this principle can be used to bring risk under control in the DC model of the SG. In the next section, we will show how this principle and the one from [58] can be inspired from to bring CRC forward in a new light for the AC model.

4.5.0.2 Cognitive Risk Control

Conceptually, the CRC module is the second most important component of the CDS, behind the executive. When the FDI attack is present, the latter will be detected with the drop in the entropic state to below γ. At that instant, the TSC will switch off the executive, which is responsible for CC, and turn on the CRC module, which is assigned to dealing with the attack. CRC is now in control of the SG. In order for

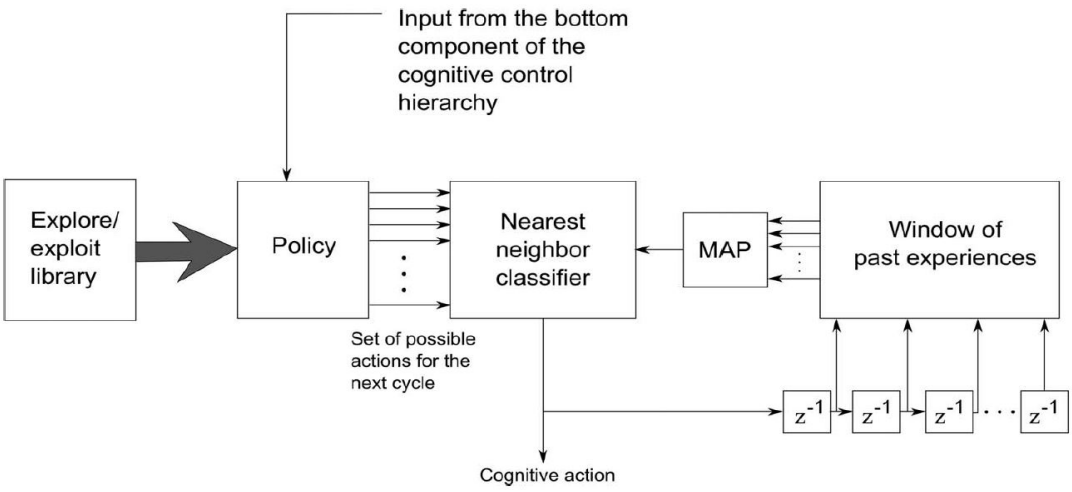

Figure 4.2: First version of preadaptive control mechanism suggested for CDS from [60]

the CRC module to apply CRC, it is fitted with supporting structures that will allow it to embody the principle of predictive adaptation that was discussed previously. Those structures can be expanded as follows:

i. Exective Memory and Perceptual Memory: From a neuroscience perspective, those memories are different from the working memory due to their temporal relationship. While the working memory stores actions that will be performed for the current PAC, the executive and perceptual memories keep a record of the experiences gained by the CDS for a longer period of time. More specifically, the executive memory stores the motor actions that were held in working memory for the past L_{PE} PACs and the perceptual memory stores sensory information of the generative model of the perceptor over the same past L_{PE} PACs. We will show in the coming sections how those two memories, rooted in the neuroscience of the brain, will become the basis for prediction and error correction in the CDS.

Depending on how those two memories are computationally implemented, each index (row or column) of the executive memory relates to the actions applied to the same index of the perceptual memory whereby the latter gives the context under which those actions were applied. However, it should be highlighted that the storage of experiences into those memories only occurs when the CDS is operating under CC. During the cyber-attack, since the environment is no longer viable, the two memories are disconnected from the normal PAC and no longer accommodate experiences.

ii. Cognitive Predictor (Prediction): Building on the roles of the two previous memories, the Cognitive Predictor (CP) expands on the stored past experiences to bring forward prediction and adaptation behavior in the CDS in the presence of FDI attack. At the first instance of attack detection, the CP will use the statistical properties of the signals stored in perceptual memory in order to identify the states being targeted by the FDI attack. Consequently, the CP will calculate an indirect mean of the states, $\tilde{\mathbf{x}}_{PE}$, of the AC state estimator from the sensory information gathered from the generative model of the perceptor over the past L_{PE} PACs as follows:

$$\tilde{\mathbf{x}}_{PE} = \frac{\sum_{j=1}^{L_{PE}} \mathbf{B}_j}{L \times L_{PE}} \tag{4.5.1}$$

where $\mathbf{B}_j$ is the stored accumulator values in the perceptual memory from (4.4.1) and L_{PE} is the length of time over which the executive and perceptual memories extend. Using $\tilde{\mathbf{x}}_{PE}$, the safety limits, $\mathbf{x}_{\text{lim}}$, of the states $\mathbf{x}_k$ is then calculated as the maximum deviation from $\tilde{\mathbf{x}}_{PE}$ for the different relevant states in perceptual

memory. Although more advanced signal properties can be exploited for the same task, this will be left as an open topic for further research. Although in a practical scenario relating to the current operation of the SG where $\mathbf{x}_k$ varies continuously, when it comes to a situation involving risk whereby these same states are perturbed and their real values are unknown, $\mathbf{x}_{\text{lim}}$ serves as a range denoting the highest probability where those states will lie. Moreover, $\tilde{\mathbf{x}}_{PE}$ and $\mathbf{x}_{\text{lim}}$ can be combined to identify the attacked states as follows:

$$\varepsilon_k^x = [(\tau^x \times \mathbf{x}_{\text{lim}}) - |\mathbf{x}_k - \tilde{\mathbf{x}}_{PE}|] < 0 \tag{4.5.2}$$

where ε_k^x is a logical vector that identifies the targeted states and τ^x is a vector that defines the volatility or deviation tendency of the different states. Furthermore, the prior history of the measurements and signals in the SG can be used to determine a more suitable τ^x so as to decrease the probability of false alarms during the attack. This is due to (4.5.2) which tells us that a particular state will not be considered as attacked as long as it lies within the range defined by past experiences and history of the states. When the states being targeted by the FDI attack have been identified, the CP send the index of those states to the Bayesian filter which will assist in predicting the most correct values of the attacked states using the stored past experiences, which is the next important topic to be discussed.

iii. Bayesian filter: Similar to section 4.4.2.2, where the Bayesian filter was explained for the perceptor, the CRC module is also fitted with a Kalman filter

to track the hidden states output from the AC state estimator to the perceptor. Comparable to the perceptor where the filter operates indirectly through the output of the generative model, the Kalman filter for CRC works similarly but operates under a different concept. Nevertheless, the equations defining the CRC Kalman filter will be the same as (4.4.2) to (4.4.11), which was defined earlier in the perceptor. The CRC Kalman filter is only active when the CDS is performing under CRC. More specifically, the CRC filter has a bigger role during CRC for bringing attack under control and the adaptation process of the CDS to the FDI attack. While the adaptation process will be covered in the next paragraph, focus will be laid on bringing the attack under control at this point in the paper. While the perceptor perceives and keeps track of the current perturbed system under attack, the CRC Kalman filter views the system differently. Once the CP has identified the states under attack for the current cycle, ε_k^x is then used to replace those attacked states in $\mathbf{x}_k$ with their respective counterparts in $\tilde{\mathbf{x}}_{PE}$. Consequently we will denote this new vector of modified states as $\tilde{\mathbf{x}}_k$. This vector is then sent to the generative model to follow the same perception process as mentioned previously in the section pertaining to the perceptor. Moreover, in order to keep the generative model under control, ε_k^x is also used to create $\tilde{\mathbf{B}}_k$. As the FDI attack causes a perturbation in the AC state estimator and the generative model subsequently, CP will generate a vector $\tilde{\mathbf{B}}_k$ to keep the latter under control. $\tilde{\mathbf{B}}_k$ is a vector of the states output by the generative model, whereby the targeted states have been replaced by the average of their respective values for the different stored $\mathbf{B}_j$ values in the perceptual memory. $\tilde{\mathbf{B}}_k$ is then sent to the CRC Kalman filter to follow

the perception process as explained in the section pertaining to the perceptor. Thus, the CRC Kalman filter, by virtue of replacing the identified attacked states with the closest match according to past experiences, keeps a different representation of the system. Consequently, by perceiving the system as if the attack was not present, this allows the CP to perform control according to the adaption process that will be explained in the next section. Since the end goal of the FDI attack is to deviate some specified states towards some predetermined values, the replacement of those identified states with those predicted by Kalman filter, rooted in past experiencces, can bring the attack under control. It is to be noted that when such an attack occurs, the real correct state values are unknown to the system and thus the best that we can do is to maintain those values within a range of where they used to lie before the attack occured. Hence, as the CRC Kalman filter is based on the generative model, we can calculate an estimate of the states of the AC state estimator indirectly by dividing the resulting filtered output with L. Those values that match the index of the attacked states will then replace the values reported by the AC state estimator for the current cycle. This vector will be denoted as $\tilde{\mathbf{x}}_k$. As a result, the output of the state estimator, during the attack, is kept under control by not sending those hijacked states further down the SG, rather it reports the output states as if the attack was not present. However, the CDS is still aware of the presence of the attack through its internal representation of the system states. There is a mechanism in place, that will be explained later, that allows the CDS to detect when the attack is absent and to switch over to CC. In the next section, it will be shown how the corrected output of the CRC Kalman filter allows control to

be performed for optimal state estimation.

iv. Cognitive Predictor (Adaptation): Once the attacked states have been identified and the Kalman filter has predicted what their values should be, the stage is now set for the CP to adapt the system to the current perturbed situation by picking the best control actions according to the past experiences stored in executive memory. More specifically, CP will use $\tilde{\mathbf{B}}_k$ to mediate between the perceptual and executive memories. From this point, CP will now find the closest experience in perceptual memory that matches $\tilde{\mathbf{B}}_k$. This can be done in a variety of ways such as minimum sum of squared errors or Euclidean distance. Once the nearest match is found, the indices pertaining to the match in perceptual memory is used to find the counterpart indices in executive memory. It is to be reminded that the executive memory stores the past actions applied to the system during CC. Thus, the relative index in that memory relates to the actions applied in the context of that situation. In some ways, it is a matching of inputs to outputs similar to how a Neural network [54] operate. Therefore, those actions, which are the nearest weight combination for state estimation in that situation, are only aimed at maintaining the state estimation for the unaffected states and attacked states under control.

This section showed how CC can be brought under risk control using the principle of predictive adaptation, rooted in past experiences. So far we explained how the transition from CC to CRC is carried is out. However, another mechanism is required to return the system back to CC after the FDI attack is over. This is the next topic to be discussed.

4.5.0.3 Task-Switch Control Restoration

In order to restore the CDS from CRC to CC, which is its dominant function under the absence of uncertainty, the TSC is equipped with a memory and watchdog timer (WDT). At the first instance of attack detection, the CP will also calculate $\tilde{\mathbf{B}}_{PE}$ and $\mathbf{B}_{\text{lim}}$, asides from $\tilde{\mathbf{x}}_{PE}$ and $\mathbf{x}_{\text{lim}}$. Here, $\tilde{\mathbf{B}}_{PE}$ is the predicted output of the generative model and $\mathbf{B}_{\text{lim}}$ are the safety limits for these values, similar to $\mathbf{x}_{\text{lim}}$ for $\tilde{\mathbf{x}}_{PE}$. Unlike (4.5.1), $\tilde{\mathbf{B}}_{PE}$ can be calculated directly by averaging out the values stored in perceptual memory for the different respective states. $\mathbf{B}_{\text{lim}}$ can then be calculated using the same procedure as $\mathbf{x}_{\text{lim}}$. When the switch from CC to CRC occurs, the TSC memory will store the $\mathbf{B}_{\text{lim}}$ calculated by the CP. Consequently, during every PAC after CRC has been switched ON, the TSC compares the values of attacked states to the safety limits, $\mathbf{B}_{\text{lim}}$, which define how the system should be operating if the attack was not present. This process can be performed in a similar process to (4.5.2):

$$\varepsilon_k^b = \Big[\big(\tau^b \times \mathbf{B}_{\text{lim}}\big) - |\mathbf{B}_k - \tilde{\mathbf{B}}_{PE}|\Big] < 0 \tag{4.5.3}$$

where ε_k^b is a logical vector that checks if the previously identified attacked states ε_k^x are within a certain safety range according to past experiences, and τ^b is a vector that defines the deviation tendency of the accumulator values as τ^x was for the AC state estimator. When these values are in conformity with the ones calculated by the CP, the TSC WDT will trigger for the next T_{TSC} PACs. When the timer starts, the CDS will still be operating under CRC. However, no mitigation actions will be applied. This allows the CDS to be proactive should an FDI attack occur during this transition. Finally, after the next T_{TSC} PACs have elapsed and if $h_{k|k}$ is greater

than γ, then CRC is switched OFF and CC takes over once again. At this point, the uncertainty present due to the attack is considered to be no longer affecting the CDS.

4.5.0.4 Complete Algorithm

The stage is now set to describe how the CDS algorithm in [7] and [8] has been modified to handle CC and CRC in the nonlinear case of AC state estimation. Tables 4.1, 4.2 and 4.3 summarizes the notations, which were described for CC and CRC, throughout this paper. Those tables extend from [7] and [8], where the CDS algorithm was first introduced for the DC state estimation case. Algorithm 3 gives a detailed description of the CDS routine during every PAC. Lines 15 to 24 shows three different modified applications of the AC state estimation algorithm. Those lines shows the flexibility of the algorithm as it can be easily integrated with the nonlinear state estimation algorithms such Honest or Dishonest Gaussian Newton methods, whereby they all rely on an initial guess of the states followed by iterations until they reach a certain error tolerance. Those techniques still hold prevalence in the CDS. However, they can be made more powerful by bringing along the cognition ability of the CDS, which has its roots in the brain. Assuming that the system is quasi-static, by limiting the number of iterations and having a good initial guess of the states, it will be shown in the computer experiments that it is possible to perform optimal state estimation along with planning performed by the executive. Nevertheless, during an attack, since those conditions no longer hold and for those reasons, we have to shift to the original AC state estimation algorithm as shown on lines 19 to 21. The use of the *a priori* estimated states by the AC state estimator and the limitation on the number of iterations also make the CDS sensitive to changes in those values such that the

entropic state will drop greatly at the first instance of state deviation introduced by the FDI attack. As during the attack, those states will be perturbed, reverting back to the original state estimation algorithm allows the CRC to better identify those attacked states so that the attack can be brought under control consequently. Lines 23 to 24 are meant for the initial initialization of the CDS algorithm, whereby neither CC nor CRC are ON. The system is allowed to run for some cycles until the generative model and the Kalman filter of the perceptor have settled on the track. The CDS takes over when those conditions are met. The variable *BayesReward*, defined in lines 63 to 68, is used to maintain the consistency of the Bayes UCB algorithm, which uses a frequentist approach relying on bounded rewards of the Bernoulli reward distribution. Since the cumulative rewards can become negative during the planning, the threshold, f, is used together with *BayesReward* to decrease the outcomes of negative saturation of rewards. Consequently, consider an action that has been considered as unsuitable for a long period of time. In the event that a situation arises, whereby that action is now the best action, the use of f for bounding the cumulative negative rewards, makes it such that the RL algorithm can increase the quantile for that action at a faster rate so as to be eventually picked by achieving the highest quantile. As a result, f permits those unselected actions to make a come-back should the right situations arise. Moreover, since the rewards will get smaller and smaller over time as the Kalman filter converges, R_n and R_p on lines 55 and 57 respectively have been introduced to improve the sensitivity of picking the right actions. Thus, R_n and R_p, which serves as positive rewards and negative rewards magnifiers, allows the good actions to build up the quantiles faster while also penalizing the negative ones. Furthermore, as Bayes UCB is based on bounded rewards in [0,1], *BayesReward*,

on lines 63 and 65, is used to handle those circumstances. Referring to line 63, where the cumulative rewards have been negatively saturated, *BayesReward* will take a value of 0 and this value is used to update the respective quantile. Similarly, the same procedure is repeated on line 65, where the negative reward saturation has not yet been reached. Lastly, on line 67, in the event that the reward for a specific action is positive during planning, then *BayesReward* will take on those values to update the relative quantile value. Moreover, in order to add stability to the algorithm during startup, the quantiles of the default configuration of the weights are initially biased with values of α. Additionally, cognitive confidence cycles, n_{cc}, are also implemented for this reason. n_{cc}, on line 71, are a pre-set number of cycles during which the cognitive controller in the executive will evaluate the several actions to gain an initial feeling of the different prospective actions based on their calculated quantile values. However, during those cycles, those actions are not performed on the system. Once those n_{cc} cycles have passed, the RL algorithm will start off by applying the best actions learnt during those cycles. Thus, this assists in bringing down the random chaotic behavior of the RL algorithm at startup. Referring back to the SG, this kind of behavior can be harmful if it is not controlled in some way. Building on the CC algorithm presented in [7], where it was demonstrated that the algorithm can overcome one of the critical limitations of RL algorithms by being able to apply many actions over the different meters, the algorithm presented in this paper retains the same essence under a different approach. Hence, as shown in lines 49 and 50, N_p is a hyper-parameter that allows us to save computational resources during the evaluation of the different actions during the planning stages using the state estimation algorithm. Nevertheless, the greater N_p is, the better the planning.

However, this increase in efficiency comes at the cost of increased computational power, which is a parameter that the system designer has decide on depending on the resources available and the system at hand. Since the CC/CRC algorithm presented is founded on [7] and [8], where the entropic state's ability to detect the FDI was validated, it will be shown in the computational experiments that follow that $h_{k|k}$ is still valid under the nonlinear case. The switching mechanism of the TSC, which is a constituent of the feedback channel, to transition from CC to CRC and vice-versa is explained on lines 34 to 45. The two conditional statements on line 38 assists in assuring a smooth transition from CRC to CC. This will be demonstrated in the simulations in the next sections. Resetting the CP on line 45, when a changeover from CC to CRC has occurred, allows the CP to adapt to new different situations of attack using the evolving perceptual and executive memories. Additionally, the CP, on line 79, is executed only during the transition from CC to CRC. Therefore, this saves computational resources and makes the algorithm very efficient. Moreover, this also implies that new values of $\tilde{\mathbf{x}}_{PE}$, $\mathbf{x}_{\text{lim}}$, $\tilde{\mathbf{B}}_{PE}$ and $\mathbf{B}_{\text{lim}}$ are calculated whenever a situation involving the FDI attack arises. Since these attacks can target different states in different instances, CRC can also evolve using the perceptual and executive memories. As a result, this shows CRC's robustness and excellent adaptive properties, similar to how the brain adapts to its continuously changing environment. Lines 89 to 97, gives an overview of how the CRC Kalman filter and CP works reciprocally. During the FDI attack, the perceptor keeps the true track of the system while the CRC Kalman filter views the system as if the attack is absent. Consequently, there are two different interpretations of the current system taking place. The computer simulations, that follow, will provide greater insight on how those two representations

work together. Once the attack appears to be over, during the WDT on lines 84 to 87, CP is still using past experiences to pick the best set of weights for optimal state estimation since the perceptor and the feedback channel are still adapting to the new situation. Nevertheless, the CDS will still operate under CRC as long the T_{TSC} cycles are not over. This serves as a defense mechanism in case a new attack occurs during this period. Lastly, due to the fact that ε_k^x is calculated during every PAC while operating under CRC, the states will not be considered under attack if they are within the safety limits evaluated by the CP. Thus, it can be said that the CP keeps a snapshot of the system parameters, before the attack, as reference in order to perform control during CRC.

4.6 Computational Experiments

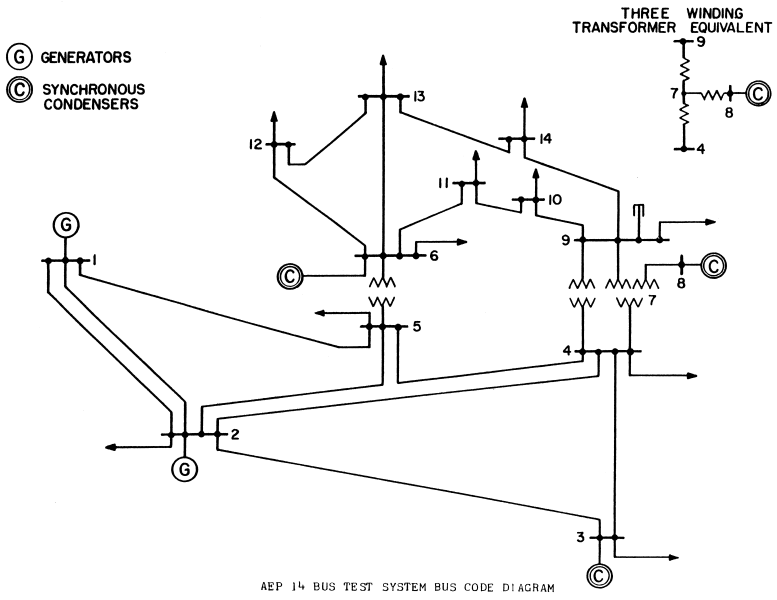

Figure 4.3: Line diagram for the IEEE 14 Bus network.

Table 4.1: Summary of Notations for CC and CRC modified from [8] (Part 1 of 3)

Notation	Definition
M	Total number of PAC cycles
N	Total number of shunt/planning cycles
n_{cc}	Number of cognitive confidence cycles
$\mathbf{z}_k$	Vector of measurements taken at cycle k
$\mathbf{x}_k$	Vector of calculated states by AC state estimator at cycle k
$\mathbf{W}_k$	Weight matrix at cycle k
$\mathbf{B}_k$	Vector retaining the cumulative sum of the states at cycle k
L	Window over which the past states is being accumulated
$\hat{\mathbf{B}}_{k\|k}$	Filtered estimate of the cumulative sum from the generative model at cycle k
$\mathbf{P}_{k\|k}$	Process error covariance matrix at cycle k
$\hat{\mathbf{B}}_{k+1\|k}$	Predicted estimate of the cumulative sum for cycle k+1
$\hat{\mathbf{P}}_{k+1\|k}$	Predicted error covariance matrix for cycle k+1
$\hat{\mathbf{B}}_{k\|k-1}$	Predicted estimate of the cumulative sum for current cycle k which was calculated during the previous cycle $k-1$
$\mathbf{P}_{k\|k-1}$	Predicted error covariance matrix for current cycle k which was calculated during the previous cycle $k-1$
$h_{k\|k}$	Entropic state at cycle k
γ	Threshold for attack detection
A	Set of all possible actions stored in action space
A_1	Set of selected prospective actions for planning
$a_k^{i,j}$	Prospective action involving the virtual application of a^i to meter j during planning
$\mathbf{W}_k^{i,j}$	Modified weight matrix where meter j's weight value has been replaced by a^i during planning
$\mathbf{G}_k^p$	Hypothesized gain during planning
$\hat{\mathbf{x}}_k^p$	Hypothesized state estimate during planning
$\mathbf{B}_k^p$	Hypothesized cumulative sum involving $\hat{\mathbf{x}}_k^p$ during planning
$h_{k\|k}^p$	Hypothesized entropic state during planning

Table 4.2: Summary of Notations for CC and CRC modified from [8] (Part 2 of 3)

Notation	Definition
$r_k^{i,j}$	Internal reward associated with each prospective action $a_k^{i,j}$
a_m	Action stored in working memory
$S_a^{i,j}$	Cumulative reward for action $a^{i,j}$
$N_a^{i,j}$	Number of times action $a^{i,j}$ has been chosen
$Q_a^{i,j}$	Quantile of action $a^{i,j}$
c	real parameter for Bayes-UCB
u	Number of prospective actions to select from action space
f	Negative rewards saturation threshold
α	Quantile initial bias
N_s	No. of maximum iterations for AC state estimation algorithm during CC
N_p	No. of maximum iterations for AC state estimation algorithm during planning
R_p	Positive rewards magnifier during planning
R_n	Negative rewards magnifier during planning
$\tilde{\mathbf{x}}_{PE}$	Predicted states by predictor
L_{PE}	Time span over which the executive and perceptual memories extend.
$\mathbf{B}_j$	j^{th} accumulator value stored in perceptual memory
$\mathbf{x}_{\text{lim}}$	Safety limits of $\tilde{\mathbf{x}}_{PE}$
ε_k^x	Logical vector where the current states, under attack, have been indentified
τ^x	Vector that defines the deviation tendency of the states
$\tilde{\mathbf{B}}_k$	Vector where the index of the attacked states in ε_k^x have been used to replace the relative $\mathbf{B}$ values in $\mathbf{B}_k$ with their closest match in perceptual memory
$\tilde{\mathbf{x}}_k$	Vector where the index of the attacked states in ε_k^x have been used to replace the relative $\mathbf{x}$ values in $\mathbf{x}_k$ with their average output of the CRC Kalman filter
x_{PE}^i	Predicted value of the i^{th} affected state according to past experiences

Algorithm 3: Complete Algorithm for implementation of CC and CRC in the defined structure

Initialization:
memstate := short-term memory for storing L past outputs from the AC state estimator
memexec := Executive memory for storing L_{PE} past contents of CC working memory
mempercept := Perceptual memory for storing the respective L_{PE} past outputs of generative model for the different states
CC and CRC Switches := CC switches for CC and CRC switches for CRC
A_1 := set of selected unique cognitive actions for planning
a_m := load working memory with default configuration of weights for all meters
BayesReward := short-term memory for storing cummulative entropic reward
Set action quantiles of the default configuration to α
$\mathbf{B}_0$, $\hat{\mathbf{B}}_{1|0}$, $\mathbf{P}_{1|0}$, $\mathbf{W}_0$, n_{cc}, f, γ, τ^x, τ^b, T_{TSC}, N_s, N_p
$a_m \leftarrow a_0$
$k \leftarrow 1$

Begin **while** $k \leq M$ **:**
    AC State Estimation:
    if *CC Mode is ON* **:**
        Use $\mathbf{x}_{k-1}$ as initial guess for calculating $\mathbf{x}_k$ with N_s iterations using the incoming $\mathbf{z}_k$
        Caculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
        memstate $\leftarrow \mathbf{x}_k$
    else if *CRC Mode is ON* **:**
        Use flat start to calculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
        memstate $\leftarrow \mathbf{x}_k$
    else:
        Use flat start to calculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
        memstate $\leftarrow \mathbf{x}_k$
    Caculate $\mathbf{x}_k$ using the incoming $\mathbf{z}_k$
    memstate $\leftarrow \mathbf{z}_k$
    Generative Model:
    Calculate $\mathbf{B}_k$ using the past L outputs from *memstate*

(Continued on next page)

(*Continuing as from line 28 onwards*)
Bayesian Filter/Kalman Filter:
Update Steps:
Calculate $\mathbf{K}_k$, Calculate $\hat{\mathbf{B}}_{k|k}$, $\mathbf{P}_{k|k}$
Prediction Steps:
Calculate $\hat{\mathbf{B}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$

Feedback Channel:
Calcualte $h_{k|k}$
if *CRC Switches are ON* **:**
    Caluculate ε_k^b, raise $flag$ if it conforms with past experiences
    if *flag is raised and* $h_{k|k} > \gamma$ **:**
        Increament WDT
        if *WDT is equal to* T_{TSC} **:**
            Switch OFF CRC *Switches*, Turn ON CC *Switches*
            Reset WDT
if $h_{k|k} < \gamma$ *and CC Switches are ON* **:**
    Switch OFF CC *Switches*, Switch ON CRC *Switches*
    Reset Cognitive Predictor

Executive:
if *CC Switches are ON* **:**
    Planning:
    for *all cognitive actions* $a_k^{i,j} \in A_1$
        Calculate $\hat{\mathbf{x}}_k^p$ using N_p iterations and corresponding $\mathbf{B}_k^p$
        Calculate corresponding $h_{k|k}^p$

        Internal reward:
        Calculate $r_k^{i,j}$
        if $r_k^{i,j} < 0$ **:**
            $r_k^{i,j} = r_k^{i,j} \times R_n$
        else if $r_k^{i,j} > 0$ **:**
            $r_k^{i,j} = r_k^{i,j} \times R_p$

        Learning: (Continued on next page)

(*Continuing as from line 58 onwards*)
Learning:
Update $S_a^{i,j}$ for $a_k^{i,j}$
Update $N_a^{i,j}$ for $a_k^{i,j}$
if $S_a^{i,j} < f$ **:**
  $S_a^{i,j} = f$
  $BayesReward = 0$
else if $S_a^{i,j} < 0$ **:**
  $BayesReward = 0$
else:
  $BayesReward = S_a^{i,j}$
Update $Q_a^{i,j}$ for $a_k^{i,j}$ using $BayesReward$
if $Q_a^{i,j} > Q_{a_m}^{i,j}$ **:**
  $a_m^{i,j} \leftarrow a_k^{i,j}$
if $k > n_{cc}$ **:**
  Policy:
  Apply a_m to the AC state estimator
  $mempercept \leftarrow \mathbf{B}_k$
  $memexec \leftarrow a_m$

CRC Module:
if *CRC Switches are ON* **:**
  if *CC transitioned to CRC* **:**
    Cognitive Predictor:
    Calculate $\tilde{\mathbf{x}}_{PE}$, $\mathbf{x}_{\text{lim}}$, $\tilde{\mathbf{B}}_{PE}$, $\mathbf{B}_{\text{lim}}$ using *mempercept*
    Calculate ε_k^x , $\tilde{\mathbf{B}}_k$
    Use $\mathbf{P}_{k|k-1}$ and $\tilde{\mathbf{B}}_k$ for initialising the CRC Kalman filter
  if $WDT > 0$ **:**
    Cognitive Predictor:
    Find best set of actions in *memexec* that matches $\tilde{\mathbf{B}}_k$ and apply them
  else:
    (Continued on next page)

(*Continuing as from line 86 onwards*)
else:
    CRC Kalman Filter:
    Update Steps:
    Calculate $\mathbf{K}_k$, Calculate $\hat{\mathbf{B}}_{k|k}$, $\mathbf{P}_{k|k}$
    Prediction Steps:
    Calculate $\hat{\mathbf{B}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$
    AC State Estimator Correction:
    Use $\hat{\mathbf{B}}_{k|k}$ and ε_k^x to modify $\mathbf{x}_k$ into $\tilde{\mathbf{x}}_k$
    Cognitive Predictor:
    Find best set of actions in *memexec* that matches $\tilde{\mathbf{B}}_k$ and apply them
$k = k + 1$

Table 4.3: Summary of Notations for CC and CRC modified from [8] (Part 3 of 3)

Notation	Definition
$\hat{x}_k^{p,i}$	Hypothesized value of state x_k^i if action a_k^i is applied
$\tilde{\mathbf{B}}_{PE}$	Predicted output of generative model
$\mathbf{B}_{\text{lim}}$	Safety limits of $\tilde{\mathbf{B}}_{PE}$
ε_k^b	Logical vector that checks if the accumalor values in $\mathbf{B}_k$, relevant to the identified states in ε_k^x, are within a certain safety range
τ^b	Vector that defines the deviation tendency of the accumalor values, $\mathbf{B}_k$
T_{TSC}	Maximum recorded elapsed time for the WDT

In this section, two different experiments are carried out to show the capability of CC and CRC respectively. The first experiment shows CC's potential for optimal state estimation for the AC state estimator by using the optimization of the entropic state as objective function. In the second experiment, it is demonstrated how this same entropic state can be used for FDI attack detection and consequently causes the CDS to switch over to CRC, which then brings the attack under control. As IEEE bus networks have generally been used as benchmarks for evaluation in the other papers referenced in this paper and relating to this topic, the IEEE 14-bus network, as shown in Fig. 4.3, will be used for assessing the architecture proposed in this paper. Since this particular network comprises of a large number of measurements and states, the results for the two different experiments will focus on certain aspects of the network that are relevant to the actual simulation. For both experiments, the data used to simulate the network configuration comes from the 14-bus case file in *MATPOWER* [53] which is an Electric Power System Simulation and Optimization Tools for MATLAB and Octave. Moreover, in order to bring about the modification for the AC state estimation algorithm, the doSE function of *MATPOWER* was modified for the requirements of the architecture. Originally, the algorithm uses Honest Gaussian Newton method with a maximum number of iterations of 100 and error tolerance of 10^{-5}. It also uses a *Flat Start* initialisation each time the function is called. During the *Flat Start*, all the values of the different states for the initial guess are set to 1 unit.

4.6.1 Cognitive Control for BDD

In the first experiment, the measurement signals relating to the state values were available from the case data in *MATPOWER* [53]. For this simulation, a noisy version of those signals were then generated with a signal-to-noise ratio (SNR) of 20 dB to create $\mathbf{z}$. From the case data, 39 measurement signals are used to calculate the 29 state values of the IEEE 14-bus network, half of which are the voltage magnitudes and the other are the voltage angles for the different buses involved. The total duration of this experiment is 2000s. The parameter L of the generative model of the perceptor was set to 20. In regards to the initialization of the Kalman filter, the initial estimates of the values to be received from the generative model are assigned a value of 0 and the diagonal elements of $\mathbf{Q}$ were set to 0.0324. Those of $\mathbf{R}$ were assgined a value of 0.01. On the executive side of the CDS for CC, the action space is made up of 156 actions, whereby each meter can be assigned a weight value from the following: 25 50 100 200. f was set to -0.5, $\alpha = 0.5$ and the number of planning/shunt cycles to be evaluated during each PAC was set to 20. R_p and R_n were set to 4 and 3 respectively. The goal of this experiment is to highlight this architecture's properties in terms of adaptability and robustness towards optimal state estimation to changing conditions. Consequently, in order to create a perturbation in the system, the SNR of the following meters is changed to 5 dB at the mentioned times: $\mathbf{t = 500s}$ for meter 2, $\mathbf{t = 700s}$ for meter 15 and $\mathbf{t = 1000s}$ for meter 20. This simulated context can be viewed as meter malfunction or a random attack, where the attacker only has limited access to meters to perform his task. In this simulation, CC is started at $\mathbf{t = 200s}$ with 10 cognitive confidence cycles. As mentioned in the earlier sections, CC is not started at $\mathbf{t = 0s}$ as some time (cycles) have to be allowed so that the Kalman filter

can settle on the track in order for the algorithm to be operate effectively.

Referring to Fig. 4.4, 4.5 and 4.6, it can be seen that CC makes the whole network dynamic, whereby the executive of the CDS is assigning the best weight values for the meters for optimal state estimation on a cycle to cycle basis. Consequently, the cognitive controller shows it ability to learn from the current and past cycles to choose the best actions for future. Moreover, the constant modification of the weight values adds another level of nonlinearity on top of the already very complex and nonlinear AC state estimator. While this may appear to be over-complicated at first, the results show that this is not only feasible but it also makes the SG more powerful. As it can be seen in Fig. 4.5, at the first instance of meter malfunction for meter 2 at $\mathbf{t = 500s}$, this has virtually no effect on this system at all as the CDS has assigned a lower weight value to that meter compared from the rest. While Fig. 4.5 shows the graphs of weight values for the some of the meters pertinent to this simulation, it is left to reader to realize that all the meters are undergoing weight reconfigurations every cycle. Thus, the different respective weight values for the meters are not all the same since the cognitive controller is adapting to the probabilistic nature of the noisy signals continuously. It is also shown that the algorithm is able to apply more than one action during each PAC under a stable manner. At $\mathbf{t = 700s}$, when meter 15 starts malfunctioning, we can now really see the capability of the architecture. As shown in Fig. 4.5, it takes only a couple of cycles for the cognitive controller to learn and adapt to the new situation by lowering the weight assigned to meter 15 and compensating for it by boosting the other meters. Thus, we can see that state 16 is the most affected and state 25 is also afflicted to a lower extent. Compared to

the traditional AC state estimator, the CC algorithm is able to keep this perturbation under control as demonstrated in the referenced figure. Adding to that, Fig.4.6, which shows the sum of squared errors for the states during this experiment compared to the absence of noise, provides more insight on the situation. From that figure, it can be seen that CC is performing better at that instance. Through the learning process of the cognitive controller together with the perceptor, the CDS is able to restore the state signal to how it was before the malfunction occurred. At $\mathbf{t = 1000s}$, when the meter malfunctions occurs for meter 20, we can see that state 25 is more affected that state 15 this time around. Similar to the first situation, we can see that this has almost no impact as far as state estimation involving the CC algorithm is concerned. In fact, there is a weight reconfiguration that takes place as from that malfunction, whereby other meters, such as meters 26, 30 and 39, have the weight values increased to keep state estimation under control. Consequently, this shows the robustness of the algorithm to adapt and act according to the evolving situations. Although some of those weight values are changed at a later point in time, this is due to the frequentist approach of the Bayes UCB coupled with the probabilistic origin of the noise. As a result of those reconfigurations in earlier situations, this highlights the cognitive ability of the controller to trust certain meters more than the others. This simulation demonstrated CC's ability to pick the best set of meters for state estimation on the go. The optimization of the entropic state plays a pivotal role in the CDS for control. As mentioned earlier in this paper, $h_{k|k}$ can be used to assess the state of the grid. Thus $h_{k|k}$ allows us to quantify the disturbances in this grid. For example, when the meter malfunction occurred at $\mathbf{t = 700s}$, we can see a decrease in $h_{k|k}$ in the subsequent cycles. However, the cognitive controller makes up for it by

taking necessary actions to bring the entropic state as close as possible to its optimal value of 1. Referring back to Fig.4.6, it can be seen the CC has performed better than the traditional algorithm. Lastly, the Chi-squares test was not implemented in this experimented as it is based on statistical properties of the signals while the approach proposed is rooted on the principle of cognition of the brain.

As voltage fluctuations are common occurrence disturbances in power systems, this simulation was designed to provide the reader a greater intuition on how the algorithm is able to distinguish between what constitutes a perturbation and the normal condition. When the states of the AC state estimator is experiencing important fluctuations, this is propagated to the generative model and therefore affects the entropic state as a result. Since $h_{k|k}$ serves as an embodiment of the grid's performance, it was illustrated in the earlier simulation how those perturbation would cause a decline in the entropic state. Since the objective function of CC is to always bring $h_{k|k}$ as close as possible to 1, the optimization of $h_{k|k}$ allows CC to reduce fluctuations in the system and keep state estimation under control. Additionally, it was shown in Fig. 4.4 that when critical meter malfunction occurred, this caused the estimated states to experience greater deviation. This was then propagated to the generative model and the Kalman filter as result, thereby lowering $h_{k|k}$ for a number of cycles. Nevertheless, CC is then quickly able to evaluate the situation and apply the best actions to increase $h_{k|k}$ and keep state estimation under control. Consequently, this experiment showcases the importance of each of the individual roles of the different components of the CDS and how they work together for goal oriented action on the SG.

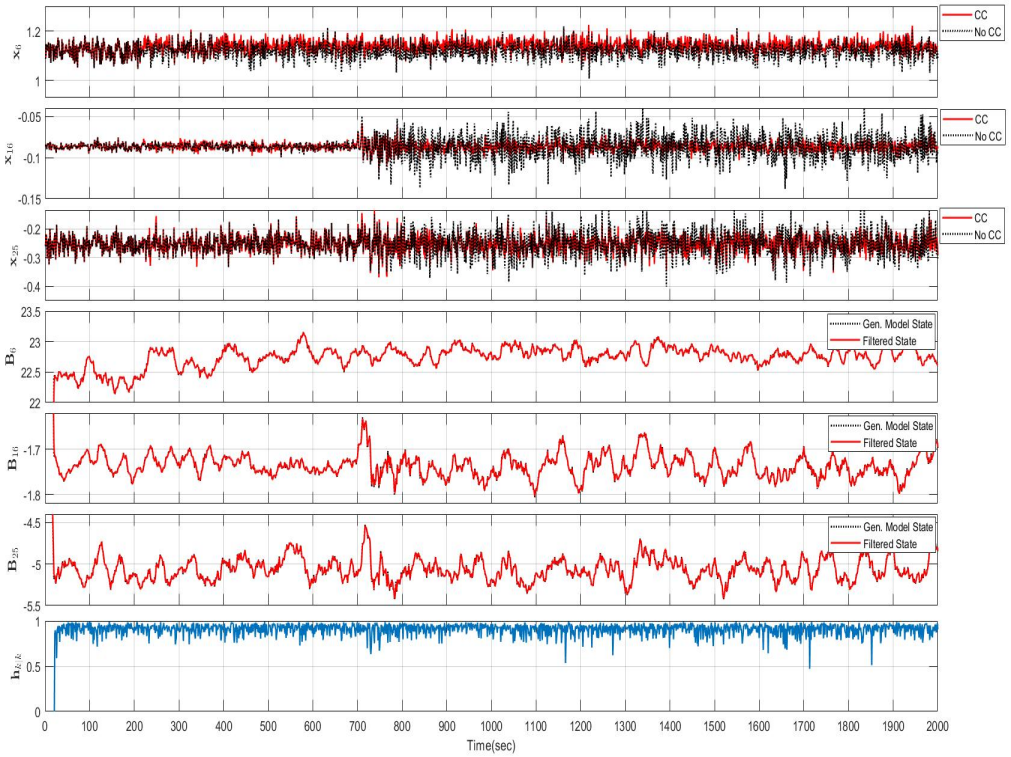

Figure 4.4: Graphs of States, Generative models and Entropic State

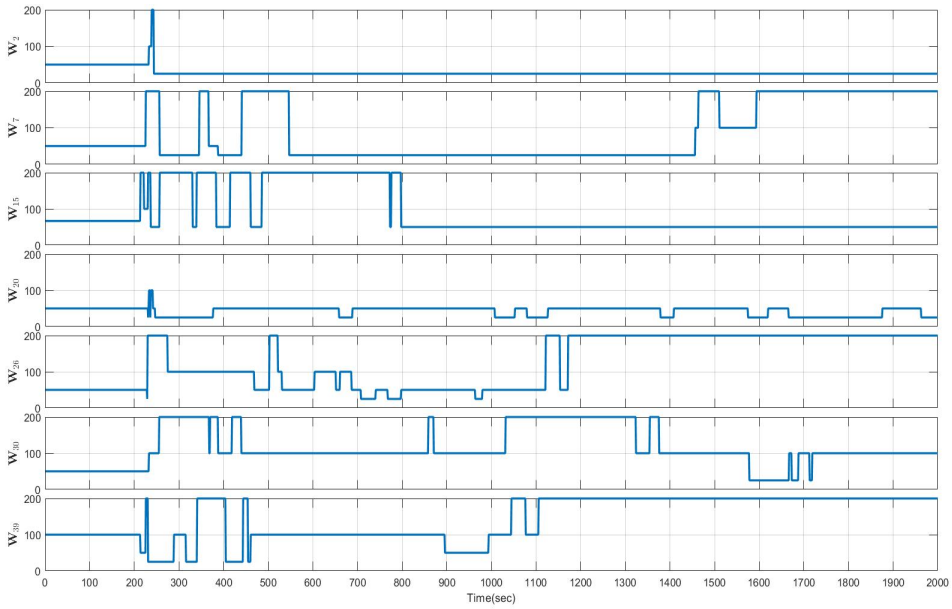

Figure 4.5: Graphs of weight values of the meters with time

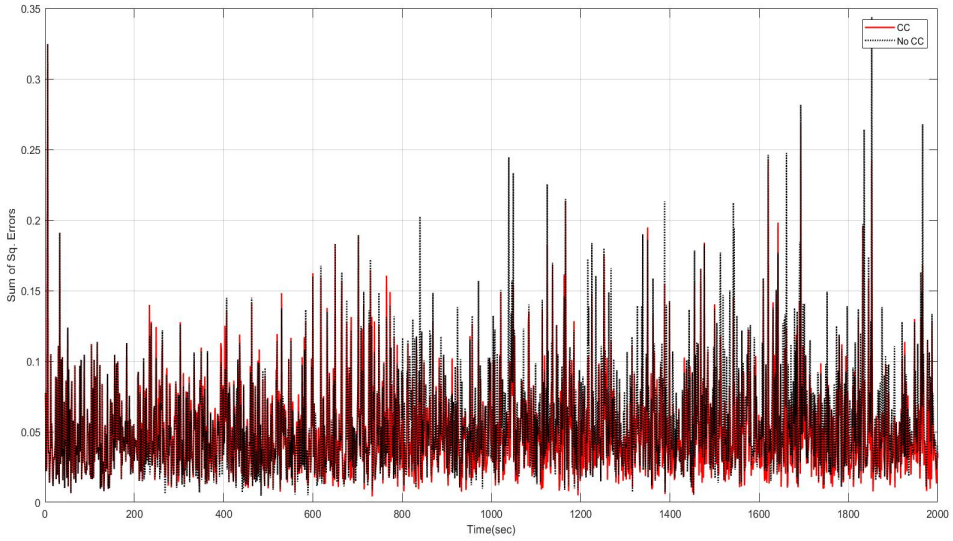

Figure 4.6: Graphs of the SSE of the states compared to flat mean without noise

4.6.2 Cyber-Attack Mitigation

In this experiment, the same IEEE 14-bus network, from the earlier simulation is put to the test in a case involving perfect FDI attack. The same conditions and parameters are used except that this time, the meter malfunctions are not present. The parameters relating to CRC are set as follows: $\gamma = 0.4$, L_{PE} =60, T_{TSC} = 40 cycles, τ^x and τ^b are both initialized with a value of 5. In this experiment, the attacker applies the FDI attack at three different buses, whereby at the first bus, it increases the voltage angle, $\mathbf{x}_3$, by 0.7 radians while at the other two buses, it decreases the voltage magnitude, $\mathbf{x}_{13}$ by 0.5 units for one of them and increases the voltage magnitude, $\mathbf{x}_{16}$ by 0.3 units at the other. In this experiment, CRC is started at $\mathbf{t = 500s}$ and the attack takes place at $\mathbf{t = 1000s}$ lasting up to $\mathbf{t = 1700s}$. Figures 4.7, 4.8 and 4.9 display the results. Referring to Fig. 4.7 and 4.9, it can be seen that the CDS is still operating under CC until $\mathbf{t = 1000s}$ despite the fact CRC was started at

$\mathbf{t = 500s}$. Thus, the system is operating as intended and the switch over will not take place until the entropic state falls below γ. As soon as the FDI attack is initiated at $\mathbf{t = 1000s}$, there is an immediate drop in $h_{k|k}$ to below that threshold. Furthermore, it was demonstrated in [7] that the greater the attack, the greater will be the dip in the entropic state. Moreover, the effects of the FDI attack also propagates throughout the generative model, causing a bigger drop in $h_{k|k}$ following the first instance of intrusion as shown in Fig. 4.7, where $h_{k|k}$ decreases further for a few cycles after $\mathbf{t = 1000s}$. Therefore, we can see in Fig. 4.9 that this further drop in $h_{k|k}$ causes the CDS to detect the attack and to switch over to CRC as from $\mathbf{t = 1001s}$ During CRC, the CDS operates entirely based on the past experiences acquired in the past L_{PE} cycles which are stored in the perceptual and executive memories. Using those experiences, CDS is able to identify those states under attack using attention. At the same time, the other states, which are not under attack, are under the control a different version of CC rooted in those past experiences and the principle of predictive adaptation. Referring to Fig. 4.10, which shows some of the other states that are not under attack, we can see that the weight configuration is being applied optimally for state estimation under a stable approach. Thus, the CDS is able to correctly identify that $\mathbf{x}_3$, $\mathbf{x}_{13}$ and $\mathbf{x}_{16}$ are under attack as shown in Fig. 4.7. Through the actions of CRC, the risk attached to the intentional deviation of those states is rapidly brought under control. However, in a realistic situation involving those FDI attacks, the real values of those hijacked states are unknown to the control center. Consequently, the best way to keep those states under control is to maintain them within a tolerable threshold of how they were operating before the attack occurred. Hence, in Fig. 4.7, we can see that the CRC module brings the AC state estimator under control

in very few PACs by using their stored reference set point of the generative model before the attack and then using it as input to the CRC Kalman filter to calculate an indirect estimate of the state values. By correctly modifying those state values, it ensures that the domino effect of cascaded incorrect control decisions based on those attacked states will not occur. Nevertheless, during CRC, the CDS is keeping track of two different representations of the SG; one which perceives the AC state estimator under attack through the perceptor and another one whereby the attack is not present. They both play critical roles in this architecture, whereby one of them is concerned with mitigating cyber-attack and the other is concerned in determining when the attack is over and switch back to CC. Moreover, referring to Fig. 4.8, it can be seen that CRC is still able to apply the optimal weights for the different meters during the attack through the use of past experiences. For example, it can be seen that $\mathbf{x}_7$ and $\mathbf{x}_{20}$ are experiencing less deviation than the traditional algorithm. At $\mathbf{t = 1700s}$, although the cyber-attack is over, we can see from Fig. 4.9 that the CDS is still operating under CRC for an additional 58 cycles before reverting to CC. As mentioned earlier in the section pertaining to task-switch control restoration, the feedback channel will switch back to CC until certain conditions, such as the running out of the watchdog timer, matching of the current and past experiences and the entropic state being above γ, are met. We can also infer from Fig. 4.7 that when the FDI has stopped as from $\mathbf{t = 1700s}$, this perturbation in the states, received from the state estimator, propagated through the generative and caused a deflection in $h_{k|k}$ once again. Nevertheless, we can also see from the same figure that once the states are no longer being perceived under attack by the CRC, their real values are now

being used as the real output of the AC state estimator to be sent to the other critical modules of the SG such as contingency analysis. Those conditions ensure smooth transition from CRC to CC, as shown from the results. Thus, this shows the key importance of the entropic state for control and cyber-attack detection in the SG in the architecture described. As a result, this simulation illustrated the potential of the CDS, though CC and CRC, for optimal state estimation and FDI attack mitigation. Since the CDS is an architecture inspired by the brain, it was shown that by mimicking some of its important fundamentals, such as the principles of cognition and predictive adaptation, bringing the risk, associated with cyber-attack, under control becomes a reality.

The two computational experiments carried out in this section were performed on a system running Windows 10 with an Intel i7-8750H processor. The computational running time of the first experiment was **39s** and **26.5s** for the second experiment. The architecture presented in this paper is based on [7] and [8], where CC and CRC was presented respectively for the first time for the DC state estimator of the SG. In this paper, we were able to re-evaluate the capacity of the entropic state for detecting the FDI attack in the AC case. In fact, in [7], it was argued and compared how the use of $h_{k|k}$ for cyber-attack has very desirable properties compared to other detection methods. The architecture proposed in this paper is revolutionary as it is able to cater for control, cyber-attack detection and cyber-attack mitigation at the same time. If the architecture is to be applied to bigger networks than the one used for the simulation, there are some modifications than need to be performed so that the architecture can be scaled up. First of all, the number of cognitive confidence cycles and shunt cycles will have to be increased accordingly since more meters will have

to be evaluated by the executive so that it can develop its own sense of the starting optimized configuration. Moreover, it is encouraged to keep the action space relatively small and distinct so that the planned rewards, evaluated during planning, are distinguishable from each other. Furthermore, N_s and N_p should also be small since the computational cost of one iteration of the state estimation algorithm is greater as the network becomes bigger. Furthermore, the diagonal elements of $\mathbf{Q}$ will have to be adjusted accordingly. Unlike many applications where $\mathbf{Q}$ is supported by a mathematical formula, the ideal values of $\mathbf{Q}$ for this architecture have to be defined by the designer since it correlates to both the sensitivity of control and cyber-attack detection. Additionally, past history of the data of the SG can be used to fine tune those parameters. A more detailed discussion on setting those values can be found in [7]. In both simulations carried out earlier, only one core was being used for the computations. Nevertheless, in order to speed the computing for bigger networks, High Performance Computing (HPC) and Parallel Computing can be employed.

In the previous computational experiment, the CDS was very fast at detecting the FDI attack within a few cycles due to the category of attack applied. However, in the presence of other types of attack, such as the slow ramp attack, the detection time can vary slightly. Thus, the sampling frequency of the AC state estimator is going to be critical for the efficiency of the architecture in regards to those scenarios. Consider a situation whereby the state estimator performs state estimation every 5 seconds. Thus, it can take up to 50 seconds for the CDS to detect the attack depending on their intensity. Nevertheless, in order for the cyber-attack to be really devastating, many states will have to be compromised. In those conditions, the detection property

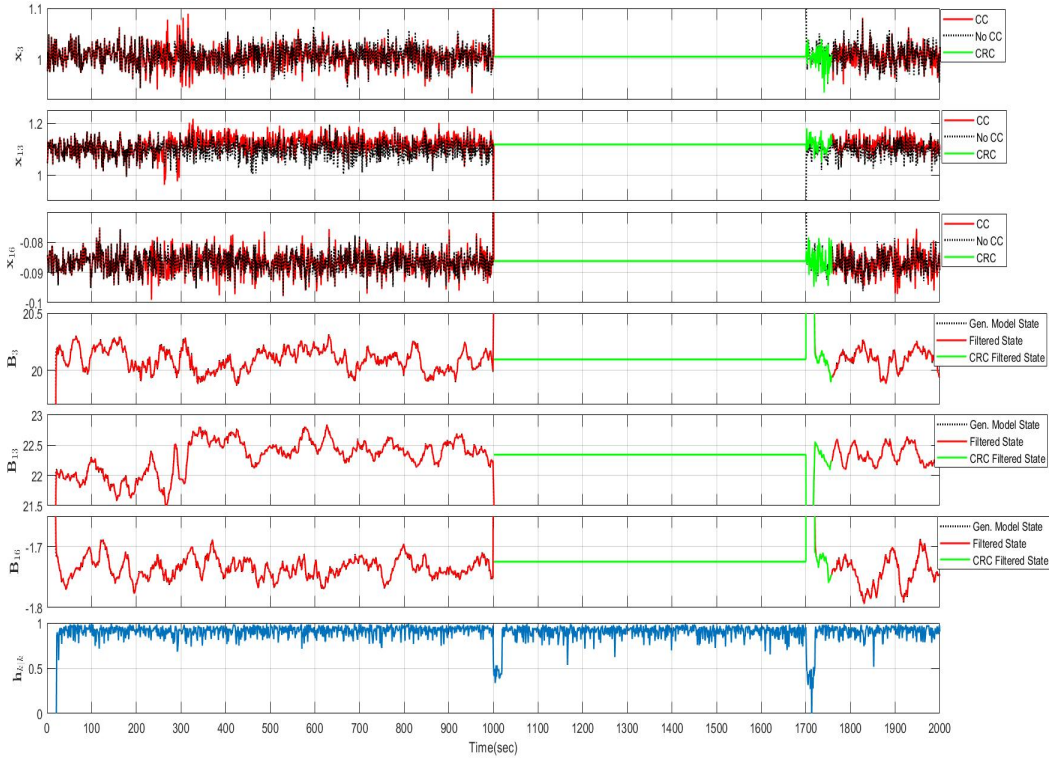

Figure 4.7: Graphs of attacked States, Generative models and Entropic State

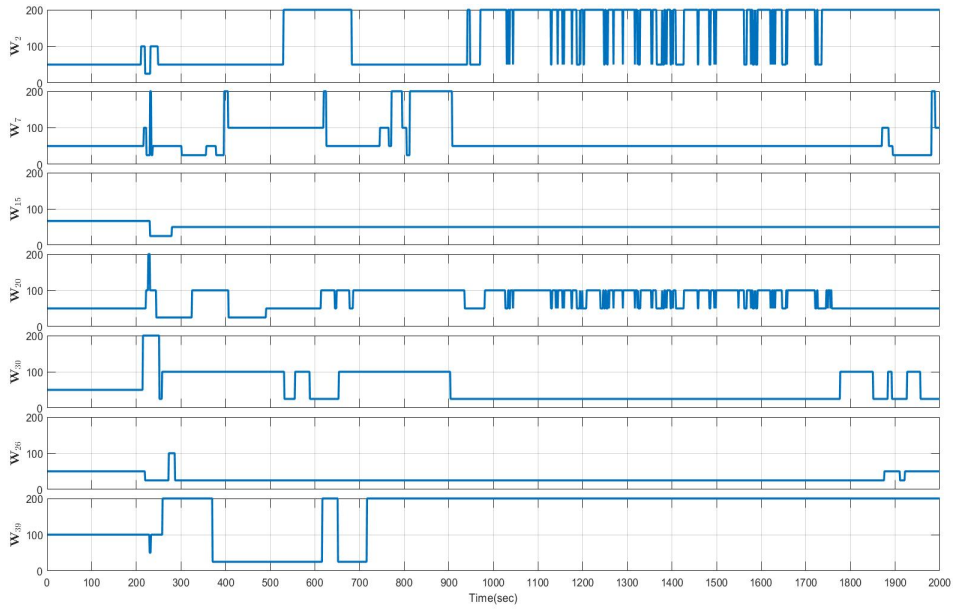

Figure 4.8: Graphs of weight values of the meters with time

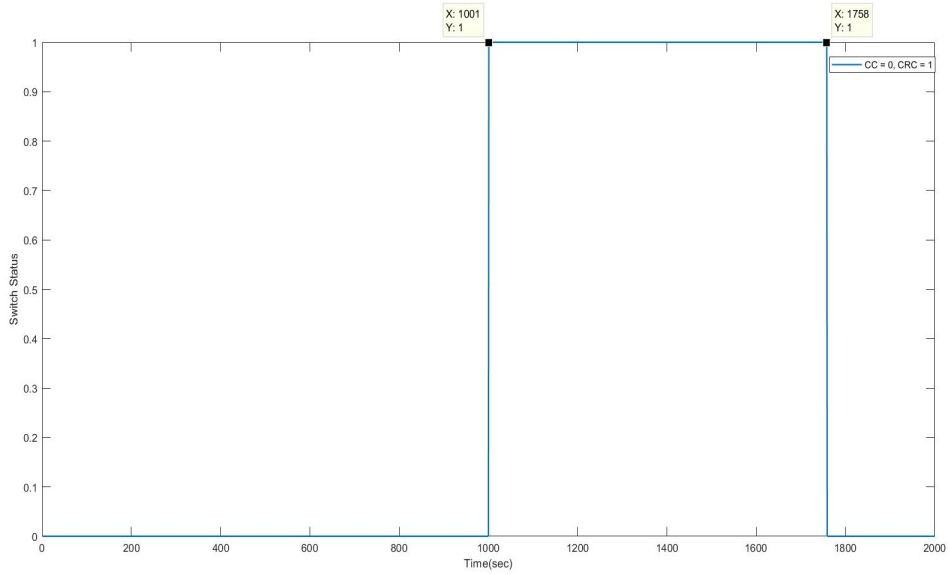

Figure 4.9: Graph of status of Switches

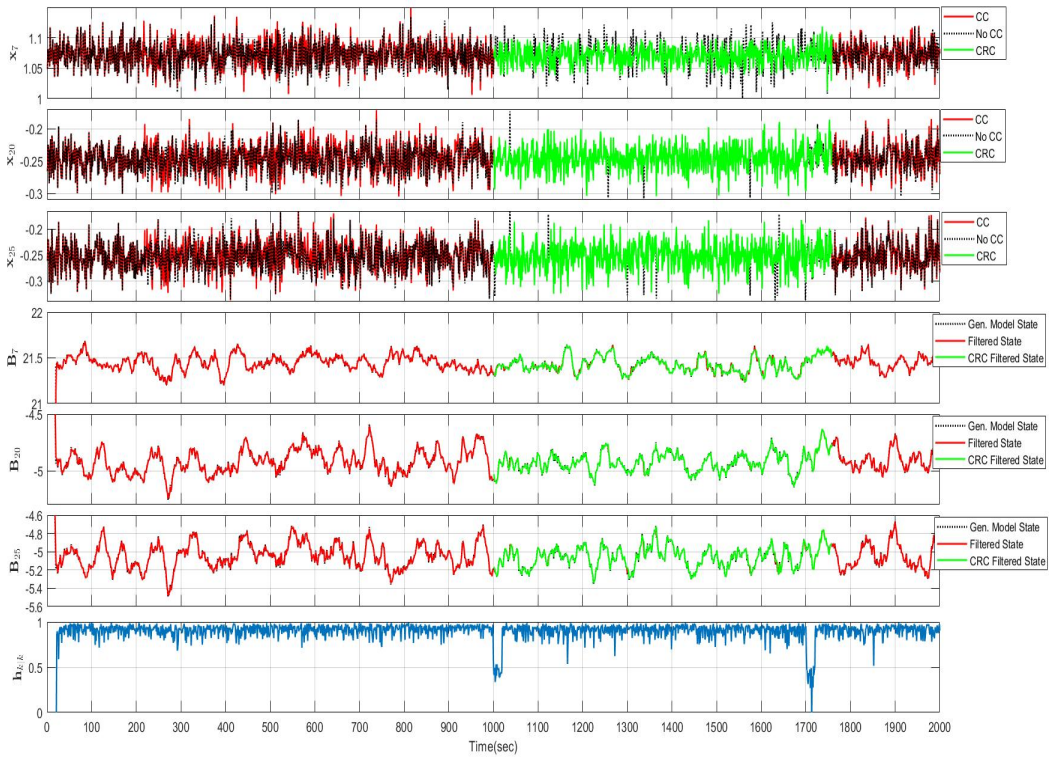

Figure 4.10: Graphs of some non-attacked States, Generative models and Entropic State

of $h_{k|k}$ shows its real robustness for practical applications; $h_{k|k}$ experiences greater deviation when the attack is larger and more powerful. On the other hand, a higher sampling frequency will also increase the detection accuracy. This accuracy can also be improved by tuning the parameters of $\mathbf{Q}$ or modifying the threshold for cyber-attack detection accordingly. Lastly, τ^x and τ^b can be estimated more accurately by using past historical data of the grid. This will consequently decrease the rate of false alarms if applied in a real life scenario.

4.6.3 Current Existing Approaches

In this section, the approach presented in this paper in the context of nonlinear state estimation, cyber-attack detection and mitigation will be contrasted with some other techniques applied for the same problems. Current literature on the topic has been mostly focused on the FDI attack detection aspect using machine learning approaches [62; 63; 64] while some have looked into the AC state estimation facet of the grid [65]. In fact, very few have tried to approach both problems at the same time [66] but none, to our knowledge, also presented a solution to mitigate the undesirable effects of the attack once detected, unlike the CDS. In [62], the authors proposed a deep learning approach to detect sparse cyber-attacks in the AC smart grid using a stacked auto-encoder in an interval state based estimation defense mechanism. Using the extracted features from their training dataset, these are then used to improve the accuracy of electric load forecasting in an attempt to decrease the uncertainty in the state variables. However, their method was not tested with cases involving perfect FDI attack unlike the simulations that were presented earlier. Moreover, their technique can exhibit false negatives over time as deep learning cannot adapt dynamically

to evolving situations compared to reinforcement learning. From a technical point of view, all machine learning techniques require supervised parameter tuning and their method is no exception. Additionally, the stacked auto-encoder's performance is only conditional on the training data used. Nevertheless, in a real world scenario where datasets may or may not be available or may be huge, fine-tuning these parameters can be very difficult. It also lacks to the ability to adapt to the new data without re-training. Compared to their approach, the CDS does not rely on such assumptions, such as requiring prior training data, as it is made specifically to learn from current data and optimize for the future. Moreover, the parameter tuning is also easier since the cognitive action is applied from a multi-armed bandit perspective whereby the parameters regulate the rate of exploration and exploitation. The approach in [62] can be reduced to an algorithm that calculates an upper bound and lower bound for electric load forecast. Those bounds are then used to identify the normal state variables. The weakness of this method is that by narrowing those variables to a certain threshold level, an attacker, with prior knowledge of this threshold, can craft an attack to remain undetected. Furthermore, the performance of their technique is also conditional on the update rate, which could be hard to find in a practical setting. The CDS, as shown and demonstrated in this paper, does not rely on so many assumptions as the ones used in [62]. Compared to the approach in [62], the algorithm presented in this paper is not only less computationally intensive but also provides a solution to deal with the cyber-attack once detected. Similarly, in [63], the authors use another machine learning technique known isolation forest to detect those FDI attacks based on the hypothesis that these kind of attacks will have the shortest average path length in a constructed random forest. They also make use of principal

component analysis to handle the dimensionality problem that arises with growing power systems. However, their approach was only evaluated on the linear state estimation problem and not the AC state estimation problem discussed in this paper. Similar to the technique in [62], they share the same weaknesses such as requiring prior data set to be trained on, parameter tuning and inability to adapt to evolving power systems on the fly without retraining and data pre-processing. The method in [63] also does not cover the problem of optimizing state estimation nor providing a cyber-attack mitigation plan once the attack is detected, compared to the solution discussed in this paper. In [64], the authors propose a machine learning approach for cyber-attack detection in large-scale smart grids using statistical correlation between measurements. In their method, symbolic dynamic filtering is used to extract features from patterns of changes in FDI attacks, which is used with Dynamic Bayesian networks to discover causal interactions between the smart grids sub-systems. Dynamic Bayesian networks is then used in conjunction with other learning algorithms to detect FDI attacks using free energy as anomaly index. Compared to the attack model in this manuscript where the most dangerous kind of FDI attack was simulated, the attack model in [64] was more relaxed in the sense it assumed that the attacker was limited on the number of measurements he could manipulate. Thus, the efficiency of their algorithm might not be the same in the face of a more dangerous attack such the one simulated previously. Moreover, similar to the two previous techniques in [62] and [63], the performance of the proposed machine learning approach is conditional on the training data and parameter tuning. Furthermore, in [64], many assumptions were made on the training data and operating conditions of the grid. Thus, in a real world situation, where such assumptions no longer hold, the efficiency of the

proposed approach might degrade and be too hard to implement to be a practical solution. Furthermore, a predefined threshold is also employed to detect the cyber-attack. Consequently, similar to the solution proposed in [62], an attacker armed with such prior knowledge can still craft an attack to remain undetected from their approach. In contrast, the threshold employed in the CDS is the entropic state, which is dynamic nature, and easier to employ since the optimization scheme in the executive is always trying to bring that value to the optimal value of one. As a result, it does not set a predefined threshold on each of the state variables but the entropic state is really an agglomeration of all the information throughout the whole grid on a cycle-to-cycle basis. Any kind of attacks or disturbance will propagate through the generative model to influence the entropic state negatively as shown in the simulations. As mentioned previously when comparing the CDS to other machine learning approaches mentioned, the technique proposed in this paper is not reliant on prior training data and is less costly from a computational perspective. Additionally, the solution in [64] is a stand-alone method in the sense that it can be used only for detection and it does not cater to state estimation or cyber-attack mitigation, compared to the technique, discussed in this manuscript, where a solution is presented to take care of these three important problems at the same time in a united architecture. Besides the machine learning detection approaches discussed in [62; 63; 64], the authors of [65] proposed a machine learning technique for improving state estimation. Using relevance vector machine, they present a nodal load estimation process that can be utilized for pseudo-measurement generation to improve the accuracy of distributed system state estimation. Although the method did achieve good results, the performance is heavily reliant on the training dataset and the preprocessing performed on

it. Hence, for a large scale evolving grid, prior data might data might not be available or additional pre-processing steps, asides the ones performed in that paper, might be required. Additionally, the pruning process of their algorithm requires some predefined threshold, which could get more complicated in a larger and more elaborate dataset. Compared to their solution, the CDS does not require prior training data or any of pre-processing carried out in [65]. In fact, the CDS construct presented in this paper is a sophisticated solution that brings together AC state estimation, modelling filtering and RL for online decision making with the result being optimal state estimation. Furthermore, the optimization of the objective function of the CDS also allows the architecture to tackle to the other fundamental problems of FDI attack detection and mitigation. Compared to the method in [65] where parallel programming had to be employed to cut the computational cost, the algorithm presented in this paper is less computationally intensive. In fact, if parallel programming was used in the simulations carried out earlier, then the performance of the algorithm would be even better. Consequently, this highlights the flexibility of the CDS as a solution for different engineering problems where resources could be costly. On the other hand, in [66], the authors make use of traditional methods rather than machine learning to improve state estimation. Their technique is reliant on additional hardware in the form of a special graphical processing unit (GPU) designed for processing large datasets. In their design, the GPU serves as platform for hosting a lateral two level dynamic state estimator based on the extended Kalman filter model. However, the algorithm is inherently complex and computationally intensive. Firstly, the process of generating the uniform set of measurements for their simulation is already very computationally costly. Adding to that, the outcome of this process also affects the

efficiency of the method. Moreover, the performance of the technique is also dependent on the programming style for implementing the algorithm. Furthermore, there is some form of parameter tuning involved through the application of Holt's exponential smoothing technique for the determination of the parameters of the extended Kalman filter. In comparison, the solution proposed in the paper is not reliant on prior data sets and additional hardware. The computational complexity of the CDS is also small compared to their technique. Lastly, the architecture presented in this paper is closer to being a practical solution compared to all the previous techniques discussed in this section as it is a novel approach, inspired by a new way of thinking, to handle the essential problems of nonlinear state estimation, FDI attack detection and mitigation together. While the methods discussed provide possible solutions to certain aspects of the problems elaborated, they fail at tackling all the avenues of these problems at the same time.

4.7 Conclusion

This paper is novel for the following reasons:

i. This is the first time that a CDS structure has been proposed for handling the nonlinear state estimation model of the SG. While our earlier two papers, which were focused on bringing the CDS and the SG together, were based on the DC model, the AC model is a more realistic approach to the SG. Consequently, the new construct, which was described in the paper, shows a lot of potential at tackling the future problems that the grid will face in the coming years as it becomes increasingly interconnected with the other aspects of IT such as IoT.

ii. Using our earlier work on the DC model, we were able to extend that research to bring forward new algorithms for CC and CRC for the AC model of the grid, although they were both combined into one algorithm in this paper. While there are some tradeoffs to be made due to the already inherent computational complexity of the AC state estimation algorithms, it was shown that the CC algorithm is revolutionary in the sense that it allows the application of multiple actions during every PAC while still maintaining the stability of state estimation. The algorithm is also very flexible as it allows tuning of certain parameters such as the number of different iterations, where precision is essential, at the cost of computational power. Learning using RL, in this architecture, is carried out in a Bayesian approach using the information from the current and past cycles for future optimal actions.

iii. The CRC algorithm introduced in this paper is also novel as compared to our previous work in [8], where CRC was presented for the first time for the DC model using the principles of risk control and predictive adaptation, the CRC module showed in this paper shows a different approach at bringing risk, associated with cyber-attacks under control, using the same principles. While the previous CRC algorithm for the DC model can be extended for the AC model, the computational cost will be too high for a practical application. Nevertheless, the CRC module, introduced in this paper, has lower computational complexity and can serve as a general solution for both the DC and AC model. Moreover, it was shown through the computational experiments that the algorithm is still robust at bringing the attack under control once detected. Hence, this shows once again the flexibility of the CDS as a whole to tailor for different objectives

and different applications.

iv. The CDS tailored for the AC model of the SG, proposed in this paper, is a unique architecture that is able to make the SG more powerful by providing a new kind of control and cyber-attack mitigation, that are both based on cognition from the brain's perspective, and cyber-attack detection.

In this paper, a new CDS based architecture was united with the SG in order to tackle the issues of nonlinear state estimation and cyber-attack mitigation through CC and CRC. Using the same principles of neuroscience behind the CDS and the principle of predictive adaptation, a new approach to CRC was unveiled such that it can cooperate with the computational cost of AC state estimation during an FDI attack to bring it under control. Computational experiments were carried out to show the individual benefits of CC for optimal state estimation and CRC for bringing the risk, associated with cyber-attack, under control respectively. Moreover, it was also discussed how the algorithm and the parameters can be adjusted so that it can be scaled up to work with bigger networks. In those bigger networks comprising of a large number of meters, a function approximator such as a Neural Network [54] can employed to simplify some of the computations involved. Although this paper focused on the problems of control on state estimation and cyber-attack in the SG, this construct, that was covered in this paper, can also be formulated to work for other similar applications where state estimation is critical such as Vehicular Radar Systems. In order, to adapt the CDS for other applications, the mathematics involving the perceptor and the executive will have to be adjusted accordingly depending on the final goal of the different intended systems. One last benefit of CRC, which was not covered in this paper was the identification of the attacked meters. Although

the past experiences allowed us to determine the attacked states, this can be further extended to isolate exactly which sensors are being compromised. During CRC, the estimated measurements can be calculated using the modified state vector, rooted in past experiences, output by the AC state estimator as follows:

$$\hat{\mathbf{z}} = \mathbf{h}(\hat{\mathbf{x}}_{\mathbf{PE}}) \tag{4.7.1}$$

This equation is similar to (4.3.16) except that in this case $\hat{\mathbf{x}}_{\mathbf{PE}}$ refers to modified output state vector where the attacked states are replaced with the ones mentioned earlier in the CRC section. Using this result, the absolute estimated errors, $\hat{\mathbf{e}}$, associated with all the measurements can be calculated using the following formula:

$$\hat{\mathbf{e}} = |\mathbf{z} - \hat{\mathbf{z}}| \tag{4.7.2}$$

The meters that are under attack will have a larger error associated with them. Hence, a suitable threshold can then be defined to isolate those compromised sensors. Those RTUs that have been an error above the employed threshold will then be considered as risks. As a result, corrective measures can be applied by the operator accordingly after undergoing this identification process. Consequently, further research on this topic can be oriented on similar identification processes when the CDS is applied for other key applications which are at risk of cyber-attack.

Bibliography

[1] S. Haykin, "Cognitive dynamic systems [Point of view]," Proc. IEEE, vol. 94, no. 11, pp. 1910–1911, Nov. 2006.

[2] S. Haykin, "Cognitive Dynamic Systems: Radar, Control, and Radio", Proc. IEEE, Point of View Article, vol. 100, no. 7, pp. 2095-2103, July 2012.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Sel. Areas Commun., vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] S. Haykin, "Cognitive radar: a way of the future." IEEE signal processing magazine 23.1 (2006): 30-40.

[5] M. Fatemi and S. Haykin, "Cognitive control: Theory and application," IEEE Access, vol. 2, pp. 698–710, Jun. 2014.

[6] S. Haykin, J. M. Fuster, D. Findlay, and S. Feng, "Cognitive risk control for physical systems," IEEE Access, vol. 5, pp. 14 664–14 679, Jul. 2017.

[7] M. I. Oozeer and S. Haykin, "Cognitive Dynamic System for Control and Cyber-Attack Detection in Smart Grid," in IEEE Access, vol. 7, pp. 78320- 78335, 2019. doi: 10.1109/ACCESS.2019.2922410

[8] M. I. Oozeer and S. Haykin, "Cognitive Risk Control for Mitigating Cyber-Attack in Smart Grid," in IEEE Access, vol. 7, pp. 125806-125826, 2019. doi: 10.1109/ACCESS.2019.2939089

[9] J. M. Fuster, "Cortex and Mind: Unifying Cognition", Oxford University Press, 2003.

[10] Y. Wang, M. Amin, J. Fu, H. Moussa, "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids", IEEE Access 2017.

[11] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-Physical Systems Security – A Survey", IEEE Internet of Things Journal, vol. PP, no. 99, pp. 1-1, 2017.

[12] J. Hao, R.J Piechocki, D. Kaleshi, et al: 'Sparse malicious false data injection attacks and defense mechanisms in smart grids', IEEE Trans. Ind. Inf.s, 2015, 11, (5), pp. 1–12 (doi: 10.1109/TII.2015.2475695).

[13] X. Fang, S. Misra, G. Xue, D. Yang, "Smart grid - the new and improved power grid: A survey", IEEE Commun. Surveys Tutorials 2012.

[14] S. Sridhar, A. Hahn, M. Govindarasu, "Cyber-physical system security for the electric power grid", Proc. IEEE vol. 99 no. 1 pp. 1-15 Jan. 2012.

[15] F. C. Scheweppe and J. Wildes, "Power system static-state estimation, Part I: Exact model," IEEE Trans. Power App. Syst. , vol. PAS-89, no. 1, pp. 120–125, Jan. 1970.

[16] K. P. V. Priya, J. Bapat, "Bad Data Detection in Smart Grid for AC model", IEEE Indicon 2014.

[17] J J. Grainger and W D. Stevenson JR.,"Power System Analysis 1st Edition", McGraw-Hill Series in Electrical and Computer Engineering, 1994.

[18] Y. Liu, P. Ning, M. Reiter, "False data injection attacks against state estimation in electric power grids", ACM CCS pp. 21-32 2009.

[19] A. Abur and A. Gómez-Expósito,"Power System State Estimation Theory and Implementation", 2004.

[20] A. Monticelli, "State Estimation in Electric Power System A Generalized Approach", Springer Science+Business Media New York, 1999.

[21] Md A. Rahman, and H. Mohsenian-Rad, "False data injection attacks against nonlinear state estimation in smart power grids." In 2013 IEEE Power and Energy Society General Meeting, pp. 1-5. IEEE, 2013.

[25] H. Zhu and G. B. Giannakis, "Robust power system state estimation for the nonlinear AC flow model," in Proc. IEEE North Amer. Power Symp., 2012, pp. 1–6.

[23] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," IEEE Trans. Smart Grid, vol. 3, no. 3, pp. 1362–1370, Sep. 2012.

[24] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid ac-based state estimation: Vulnerability analysis against cyber attacks." IEEE Transactions on Automatic Control 64.5 (2018): 1784-1799.

[25] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," Comput. Netw., vol. 57, no. 5, pp. 1344–1371, 2013.

[26] Z. Yu, W. Chin, "Blind false data injection attack using PCA approximation method in smart grid", IEEE Trans. Smart Grid vol. 6 no. 3 pp. 1219-1226 May 2015.

[27] J. Kim, L. Tong, and R. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," IEEE Transactions on Signal Processing, vol. 63, no. 5, pp. 1102–1114, March 2015.

[28] L. Liu, M. Esmalifalak, Q. Ding, V. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," IEEE Transactions on Smart Grid, vol. 5, no. 2, pp. 612–621, March 2014.

[29] A. Anwar, A. N. Mahmood, M. Pickering, "Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements", J. Comput. Syst. Sci. vol. 83 no. 1 pp. 58-72 2016.

[30] J. Jiang, Y. Qian, "Defense mechanisms against data injection attacks in smart grid networks", IEEE Commun. Mag. vol. 55 no. 10 pp. 76-82 Oct. 2017.

[31] P. McDaniel, S. McLaughlin, "Security and privacy challenges in the smart grid", IEEE Security Privacy vol. 7 no. 3 pp. 75-77 May/Jun. 2009.

[32] R. Deng, G. Xiao, R. Lu, H. Liang, A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks impacts and defense: A survey", IEEE Trans. Ind. Informat. vol. 13 no. 2 pp. 411-423 Apr. 2017.

[33] D. Wang, X. Guan, T. Liu, Y. Gu, Y. Sun, Y. Liu, "A survey on bad data injection attack in smart grid", Proc. IEEE PES Asia-Pac. Power Energy Eng. Conf. pp. 1-6 2013.

[34] K. Manandhar, X. J. Cao, F. Hu, Y. Liu, "Combating false data injection attacks in smart grid using kalman filter", Proceedings of International Conference on Computing Networking and Communications Communications and Information Security Symposium pp. 16-20 2014.

[35] K. Manandhar, X. Cao, F. Hu, Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter", IEEE Trans. Control Netw. Syst. vol. 1 no. 4 pp. 370-379 Dec. 2014.

[36] P.Y. Chen, S. Yang, J. A. McCann, J. Lin, X. Yang, "Detection of false data injection attacks in smart-grid systems", IEEE Commun. Mag. vol. 53 no. 2 pp. 206-213 Feb. 2015.

[37] Y. Liu, L. Yan, J. Ren, D. Su, "Research on efficient detection methods for false data injection in smart grid", International Conference on Wireless Communication and Sensor Network (WCSN) pp. 188-192 December 2014.

[38] D. B. Rawat, C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems", IEEE Signal Process. Lett. vol. 22 no. 10 pp. 1652-1656 Oct. 2015.

[39] Y. Gu, T. Liu, D. Wang, X. Guan, Z. Xu, "Bad data detection method for smart grids based on distributed state estimation", Proc. IEEE Int. Conf. Commun. pp. 4483-4487 2013.

[40] Y. Huang et al., "Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis", IEEE Syst. J. vol. 10 no. 2 pp. 532-543 Jun. 2016.

[41] S. Li, Y. Yilmaz, X. Wang, "Quickest detection of false data injection attack in wide-area smart grids", IEEE Trans. Smart Grid vol. 6 no. 6 pp. 2715-2735 Nov. 2015.

[42] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems,". Journal of Basic Engineering, 82: 34–45, 1960

[43] A. S. Debs, R. E. Larson "A dynamic estimator for tracking the state of a power system", IEEE Trans. PAS vol. PAS-89 pp. 1670-1673 September/October 1970.

[44] E. A. Blood, M. D. Ilic, J. Ilic, B. H. Krogh, "A Kalman filter approach to quasi-static state estimation in electric power systems", 38th North American Power Symposium pp. 417-422 2006 2006.

[45] A Saikia, RK Mehta, "Power system static state estimation using Kalman filter algorithm", EDP Sciences. 2016; 7: 1-7.

[46] C. E. Shannon, "A mathematical theory of communication", Bell Syst. Tech. J., vol. 27, no. 3, pp. 379-423, Jul./Oct. 1948.

[47] R. S. Sutton and A. G. Barto, "Reinforcement Learning", Cambridge, MA, USA: MIT Press, 1998.

[48] E. Kaufmann, O. Cappé and A. Garivier, "On Bayesian upper confidence bounds for bandit problems" Proc. Int. Conf. Artif. Intell. Stat. pp. 592-600 2012.

[49] G. Burtini, J. Loeppky, and R. Lawrence,"A survey of online experiment design with the stochastic multi-armed bandit", CoRR, abs/1510.00757, 2015.

[50] E. Kaufmann, “Analysis of bayesian and frequentist strategies for sequential resource allocation”, Machine Learning [cs.LG]. Télécom ParisTech, 2014. English. ¡NNT : 2014ENST0056¿. ¡tel-01413183¿

[51] P. Reverdy, V. Srivastava ,N. E. Leonard, ”Modeling human decision-making in generalized Gaussian multi-armed bandits”, Proc. IEEE vol. 102 no. 4 pp. 544-571 Apr. 2014.

[52] I. Arasaratnam and S. Haykin, “Cubature Kalman Filters,” IEEE Trans. Autom. Control, vol. 54, no. 6, pp. 1254-1269, Jun. 2009.

[53] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, ”MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education,” Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12-19, Feb. 2011.(Digital Object Identifier: 10.1109/TPWRS.2010.2051168)

[54] S. Haykin, ”Neural Networks and Learning Machines”, 3rd ed. Prentice-Hall, 2009.

[55] R. Xu, R. Wang, Z. Guan, et al. “Achieving Efficient Detection Against False Data Injection Attacks in Smart Grid”, IEEE Access, vol. 5, pp.13787-13798, 2017.

[56] P. Gao, M. Wang, J. Chow, et al., ”Identification of Successive ”Unobservable” Cyber Data Attacks in Power Systems”, IEEE Transactions on Signal Processing, 2016, 64 (21): 5557-5570.

[57] D. J. Kershaw and R. J. Evans, “Optimal waveform selection for tracking systems,” IEEE Trans. Inf. Theory, vol. 40, no. 5, pp. 1536-1550, Sep. 1994.

[58] J. M. Fuster, "The prefrontal cortex makes the brain a preadaptive system", Proc. IEEE, vol. 102, no. 4, pp. 417-426, Apr. 2014.

[59] J. M. Fuster, "Cortex and memory: Emergence of a new paradigm", J. Cogn. Neurosci., vol. 21, pp. 2047–2072, 2009.

[60] S. Haykin, J. M. Fuster, "On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other", Proceedings of the IEEE. 2014 Mar 21;102(4):608-28.

[61] A. Clark, "Whatever next? Predictive brains, situated agents, and the future of cognitive science," Behav. Brain Sci., vol. 36, no. 3, pp. 181–204, 2013.

[62] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, J. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," IEEE Transactions on Industrial Informatics. 2018 Feb 9;14(11):4766-78.

[63] S. Ahmed, Y. Lee, S. H. Hyun, I. Koo, "Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest," IEEE Transactions on Information Forensics and Security. 2019 Mar 5;14(10):2765-77.

[64] H. Karimipour, A. Dehghantanha, R. M. Parizi, K. K Choo, H. Leung, "A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids," IEEE Access. 2019 May 31;7:80778-88.

[65] K. Dehghanpour, Y. Yuan, Z. Wang, F. Bu, "A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation," IEEE Transactions on Smart Grid. 2019 Jan 17;10(6):5942-51.

[66] H. Karimipour, V. Dinavahi, "Extended Kalman filter-based parallel dynamic state estimation," IEEE transactions on smart grid. 2015 Jan 20;6(3):1539-49.

Chapter 5

Conclusion

5.1 Contributions

This thesis is unique as it introduces a new way of thinking that manifests itself by uniting the SG and the CDS with the end goal being improved control and cyber-security against FDI attacks. All the research work presented in this thesis not only focused on the SG but also brought improvements to the fundamental concepts of the earlier versions of CDS whereby different optimized approaches to performing CC and CRC were presented. Hence, the main contributions made in each of the respective chapters of this thesis are summarized in the next sub-section.

5.1.1 Contributions From Chapter 2

The main contributions of Chapter 2 are listed as follows:

(i) The architectural structure of CDS tailored for the SG is presented for the first time. This is the first known application where the generative was incorporated

in the perceptor to perform CC. In this chapter, the DC model of the SG is used. To get closer to the cognitive neuroscience roots of the CDS, a new way of performing RL in the proposed construct is presented. The entropic state of the CDS tailored for the SG is used as FDI attack detector. Moreover, with the CDS acting as the supervisor of the network, its objective function is to optimize the entropic state in order to improve the state estimation process.

(ii) An algorithm for performing CC in the new achitecture is proposed. In this architecture, the calculated entropic state embodies the essence of the whole perceptor and the environment, which in turn the executive uses via the shunt cycles to find the best weight configuration that get the entropic state closer to its optimal value of 1. The algorithm uses the principles of perceptual and executive attention to increase weight values of meters that it learns to trust over different elapsed PACs. Increasing the weight values is synonym to placing more trust in those specific meters. Furthermore, meters that exhibit characteristics, such as faults, that will be detrimental to the state estimation process are assigned lower weight values.

(iii) Two different experiments are carried out on a 4-bus network and the IEEE 14-bus network respectively. In the first experiment, it is shown that the proposed algorithm performs better than the traditional state estimation process. Moreover, it is also illustrated that the algorithm already takes care of bad detection through its internal optimization process. Thus, in the presence of introduced faults in the simulation, the CDS is able to identify quickly the malfunctioning meters and consequently decrease their impact to the estimated states by

lowering their weight values. After the faults are removed, the algorithm automatically readjusts their weights in a timely manner. In the second experiment where the IEEE 14-bus network is involved, it is shown that the entropic state is able to detect FDI attacks under different contexts whereby the hackers had full to limited knowledge of the network and access to the meters.

To the best of the author's knowledge, the scholarly work presented herein is the first experimental work where the CDS and SG have been combined together for control and FDI attack detection in SG.

5.1.2 Contributions From Chapter 3

The main contributions of Chapter 3 are listed as follows:

(i) Based on the previous published work, the architecture of the CDS for the DC model of the SG is expanded to include CRC to mitigate cyber-attacks. A detailed elaboration on the principle of predictive adaptation is provided from a neuroscience perspective. The important role of past experiences is highlighted on how it assists in this process for goal oriented behavior and correction in the brain. It is further shown how this concept can be applied in the CDS, especially for mitigating the effects of cyber-attacks in the SG. The new architecture presented brings together two executive, whereby one is concerned with CC and the other with CRC is revolutionary as it is a more powerful entity than the original paper from which it was first presented.

(ii) A novel algorithm showing the application of CC and CRC is introduced. In this algorithm, the perceptual memory stores experiences from the senses, which

in the CDS comes from the perceptor, while the executive memory stores experiences related to the actions performed in the executive in the absence of uncertainty. In the presence of uncertainty, for example the FDI attack, the entropic state will detect the attack and trigger task-switch control to switch the CDS to CRC mode. During CRC, experiences stored in perceptual memory is used to identify the attacked states. Moreover, the combination of experiences from the executive and perceptual memories are then used to mitigate the effects of those attacks in the context of risk control. Furthermore, the contents of perceptual memory and entropic state are used to identify when the attacks have subsided and to revert the system back to CC mode.

(iii) The proposed algorithm is then validated through two experiments performed on a 4-bus network and the IEEE 14-bus network respectively. In the 4-bus network simulation, it is shown that this novel algorithm is able to detect the presence of the FDI attack quickly and to switch the CDS to CRC. During CRC mode in the experiment, CRC was able to correctly identify the attacked states and deal with all the targeted simultaneously in a timely manner. In order to show the robustness of the algorithm, it was tested in the IEEE 14-bus network whereby almost half of its estimated states was under attack. Nevertheless, even in a larger network, the algorithm performed very well being able to correctly identify the targeted states and bring the undesired effects under control. Lastly, in both cases, the algorithm was able to identify when the attack was over quickly and to switch the CDS back to CC.

To the best of the author's knowledge, the scholarly work presented herein is the first experimental work where this novel CRC structure was implemented in the case

of the DC model of the SG for the mitigation of FDI attacks.

5.1.3 Contributions From Chapter 4

The main contributions of Chapter 4 are listed as follows:

(i) The brain inspired CDS is now applied to the AC model of the SG, which is a more realistic depiction of the SG. The concepts from chapters 2 and 3 are re-engineered to improve state estimation, detect and mitigate FDI attacks. The various challenges regarding the AC model are discussed and reasons why few published works concerning this model and cyber-attacks have been published are provided. Since this model is more computationally intensive, whereby the nonlinear state estimation involve recursions, the proposed algorithm will have to be able to take those details into consideration. Similar to chapter 2, the entropic state is used to detect FDI attacks and help in optimizing the state estimation process. Moreover, using the concepts of predictive adaptation presented in chapter 3, a novel CRC module is introduced to mitigate the effects of cyber-attacks. The flexibility of this principle is highlighted in the sense that it can take different forms but still rely of some of the supporting structures such as the perceptual and executive memories. The Cognitive Predictor is the new added component to CRC that identifies the attacked states and bring the risk associated with those attacks under control. Finally, the entropic state together with the past experiences stored in perceptual memory help in identifying when the attack is over and switch the system back to CC mode.

(ii) A revolutionary algorithm pertaining to CC and CRC in this construct involving the AC model of the CDS is proposed. Rooted in the concepts from chapter

2, where CC was the emphasis, and chapter 3, where CRC was the main goal, this algorithm attempts to derive a new way of thinking in improving nonlinear state estimation and mitigating the cyber-attacks. From a state estimation perspective, the algorithm is very flexible as it can adapt to the different existing nonlinear state estimation algorithms since it uses the previous initial estimated states as its initial guess to drive and optimize the state estimation process in the subsequent PAC in a Bayesian fashion. At the cost of computational power, the parameters can also be further tuned to improve state estimation. Asides from CC, the CRC aspect of the algorithm makes use of some of the supporting infrastructures from chapter 3 in the context of risk control. Here, the perceptual and executive memories assist the new Cognitive Predictor in identifying the states under attack and mitigating the undesirable effects. When the attack has subsided, task-switch control will revert the CDS back to CC mode using entropic state and perceptual memory as indicators.

(iii) Two different experiments relating to CC and CRC are performed on the IEEE 14-bus network. In the first experiment, it was shown that the algorithm was able to learn quickly which meters were more useful for the state estimation process and attribute higher weight values to those meters. Similarly, meters detrimental to the state estimation process will be assigned lower weight values as a virtue of the trust earned by the RL algorithm. Meter malfunctions were also simulated, similar to chapter 2. In the case of those malfunctions, the algorithm was quick at identifying those and decrease their impacts to the state estimation process by lowering their weight values. On the overall, the proposed algorithm performed better than the traditional AC state estimator at

the cost of slightly higher computational power. However, the parameters of the algorithms are very flexible, allowing the designer to choose the desired computational power to be used. In the second experiment, multiple states were being targeted by FDI attacks. Nevertheless, CRC was able to detect this situation through the entropic state and correctly identify those states with the help of perceptual memory. Perceptual memory together with executive memory were then able to assist the Cognitive Predictor in bringing the risk associated with those attacks under control. Finally, when the attack was no longer present, task-switch control was able to identify that situation and revert the CDS back to CC mode.

To the best of the author's knowledge, the scholarly work presented herein is the first experimental work where this novel construct embodying CC and CRC was implemented for the AC model of the SG in the context of control and mitigation of FDI attacks.

5.2 Limitations

Due to the different scenarios and conditions during which the proposed algorithms were experimentally validated, there are some foreseeable limitations that deserve extra attention if the algorithms were to be implemented in a real-life context. While the following sub-sections covers some of those limitations, the next section will present ways of addressing those.

5.2.1 Network Size

A 4-bus network was used to validate the algorithms in chapters 2 to 3 and the IEEE 14-bus network was used to test the algorithms presented. However, in a practical scenario, the networks of the SG is much bigger than the ones used in our research work. Thus, the state estimation process will be more computationally intensive, especially in the AC model. As a result, the time taken to perform planning steps will be longer and consequently this may lower down the performance of the CDS. Moreover, as there are more states to be estimated, more shunt/planning cycles will need to be implemented in order to maintain the efficiency of the CDS. Furthermore, the covariance matrices of generative model of the perceptor will have to be adjusted accordingly with the size of the network. Hence, prior simulations will have to be performed in order to determine the optimal set of parameters for those components as well as those of the RL algorithm.

5.2.2 Simulation Conditions

All the simulations presented in chapters 2 to 4 were carried out under quasi-static conditions. However, the real SG does not operate fully under quasi-static conditions as the supply and demand for power changes over time. Moreover, the jacobian matrix of the SG remained unchanged throughout the simulations. In a practical scenario, the jacobian matrix will also experience changes when faced with changing supply and demand for power. Consequently, real-life conditions is vastly different from the conditions used in the simulations. Nevertheless, simulating those conditions can be hard from a research perspective as there is no open data available to create those scenarios. In our research, Matpower was used as the package to simulate the SG as

other cyber-security researchers also used it to validate their techniques. Since there is no research tool that can simulate those contexts, this poses a great limitation to validating research for practical scenarios. As a result, how the algorithms presented in this thesis will perform in a real-life scenario remains an open-ended question.

5.3 Future Directions

In the previous chapters, some avenues of future research have already been proposed. Referring to the previous limitations discussed and the entire thesis, the following material provide directions to address those issues.

5.3.1 Larger Networks and Simulation Condtions

In order to get closer to the real SG, larger networks need to be used for simulations. Moreover, more effort should be made so as to re-create more real life conditions. This is can be done in many ways such as adding changing noise and non-quasi-static conditions, evolving jacobian matrix and sudden increase or decrease in load. Although the simulation of those conditions is already a huge challenge in itself and might still be a far cry from a real network grid, it will still be a step in the right direction for future research in this field. Furthermore, there is a huge need for research packages, such as Matpower, that can be used to create those settings for validating new methods involving the SG. Thus, effort should be made to applying the algorithms discussed in this thesis to closer representations of the actual SG.

5.3.2 Recent Advances in Artificial Intelligence

Over the past recent years, there has been tremendous progress in the AI field. As the size of the grid is scaled up, the inverse operation during state estimation and planning can become costly due to the large amount of measurements involved. In order to accelerate the computation, a function approximator such as a neural network [1] can be used. There is also an issue in regards to the way that Bayes UCB is applied in the CDS; once the RL algorithm has chosen a certain weight configuration for the meters, those weights can no longer be evaluated during planning unless they are replaced with new values. Consequently, we can look at the way the advantage function is used in actor-critic methods [2] as inspiration to get around this problem and make the RL algorithm more optimal. Furthermore, this will make the algorithm more stable in the selection of the optimal weight configuration. Additionally, we can also look at recent advances in machine learning and deep learning [3] in order to improve the models employed in the perceptor. Those improvements will not only make the CDS more powerful for the SG, but for other applications, such as Vehicular Radar Systems [4].

5.3.3 Multi-Layered Hierarchical CDS

Currently, the CDS is a single-layered structure which is a very simplified model of the neocortical structure of the mammalian brain where a six-layer laminar structure exists within every region of the cortex [5]. This includes both the perceptor and executive sections. Nevertheless, the single layered CDS that exists currently has already has shown tremendous potential in research areas such as tracking in form of the Cognitive Radar or communication in the form of Cognitive Radio. In the

context of this thesis, it has paved the way for a new way of thinking for uniting the CDS and the SG for control and cyber-security applications. Being a model rooted in the neuroscience of the brain, the CDS has to get closer to real model of the brain by expanding itself into hierarchical structures. This will in turn open the doors to new generations of multi-layered hierarchical CDS with enhanced principles of cognition and higher level capability that can accommodate more challenging engineering applications including the SG.

Bibliography

[1] S. Haykin, "Neural Networks and Learning Machines", 3rd ed. Prentice-Hall, 2009.

[2] Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T.,... and Kavukcuoglu, K. (2016, June). Asynchronous methods for deep reinforcement learning. In International Conference on Machine Learning (pp. 1928-1937).

[3] Geluvaraj, B., Satwik, P. M., and Kumar, T. A. (2019). The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace. In International Conference on Computer Networks and Communication Technologies (pp. 739-747). Springer, Singapore.

[4] S. Feng and S. Hakin, "Cognitive Risk Control for Transmit-Waveform Selection in Vehicular Radar Systems." IEEE Transactions on Vehicular Technology 67.10 (2018): 9542-9556.

[5] Xuyu Qian, Ha Nam Nguyen, Mingxi M Song, Christopher Hadiono, Sarah C Ogden, Christy Hammack, Bing Yao, Gregory R Hamersky, Fadi Jacob, Chun Zhong, et al. "Brain-region specific organoids using mini-bioreactors for modeling ZIKV exposure". In: Cell 165.5 (2016-05), pp. 1238–1254 (cit. on p. 133).