A QUICKEST DETECTION FRAMEWORK FOR SMART GRID

A Dissertation
Presented to
the Faculty of the Electrical and Computer Engineering Department
University of Houston

in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in Electrical Engineering

by
Yi Huang

May 2013

A QUICKEST DETECTION FRAMEWORK FOR SMART GRID

Yi Huang

Approved:

Chair of the Committee
Dr. Zhu Han, Associate Professor
Electrical and Computer Engineering

Committee Members:

Dr. Haluk Ogmen, Professor
Electrical and Computer Engineering

Dr. Amin Khodaei, Assistant Professor
Electrical and Computer Engineering

Dr. Rong Zheng, Associate Professor
Computer Science

Dr. Lijun Qian, Associate Professor
Electrical and Computer Engineering
Prairie View A&M University

Dr. Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Dr. Badrinath Roysam, Professor and Chairman,
Electrical and Computer Engineering

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Zhu Han, for his excellent guidance and continual support during the course of my degree. Working with him was a wonderful experience and his wise knowledge, constructive advice, and constant encouragement, that he shared during my stay at the University of Houston has been invaluable. He contributed significantly to both my research and my professional development. I would also like to thank my dissertation committee and professors, Dr. Haluk Ogmen, Dr. Leang-San Shieh, Dr. Wei-Chuan Shih, Dr. Rong Zheng, Dr. Lijun Qian, and Dr. Amin Khodaei, for their encouragement, insightful comments, valuable discussions, and accessibility. I would like to show my deep appreciation to my collaborators, Dr. Yu Cheng, Dr. Husheng Li, Dr. Kristy A. Campbell, and Dr. Lifeng Lai. My appreciation also goes to all the staff of the ECE Department.

During my graduate studies at the University of Houston, I have had the pleasure of meeting many students, who have helped me directly or indirectly in completing my studies and have made my Ph.D. a rewarding experience. I owe my thanks to them. In particular, I would like to thank my labmates, Dr. Zhou Yuan, Nam Nguyen, Lanchao Liu, Najmeh Forouzandeh Mehr, Mohammad Esmalifalak and Jia Meng. I appreciate all the helpful discussions that I had with them over the years. I also thank my close friends, who have become an inseparable part of my life.

I am deeply indebted to my parents and my family who have been a constant source of inspiration, support, and love throughout this degree and my life. This thesis is dedicated to them. Thank you all for everything.

A QUICKEST DETECTION FRAMEWORK FOR SMART GRID

An Abstract
of a
Dissertation
Presented to
the Faculty of the Electrical and Computer Engineering Department
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in Electrical Engineering

by
Yi Huang

May 2013

Abstract

The smart grid technology has significantly enhanced the robustness and efficiency of the traditional power grid network. The integration of such smart functionalities into the power grid also poses many risks such as increasing system complexity, network security risk, end-user data privacy issues, uncertainty of the renewable energy generation, and etc. Although the smart grid has been investigated heavily in many directions and aspects when it was raised for the first time, the research on the power system issues and the quickest detection techniques on smart grid networks are still limited.

In this dissertation, we explore specifically in three areas: system status, security issue, and resource management in smart grid networks. First, we propose a CUSUM-based defense strategy against the false data injection attack in smart grid networks. In comparison to classical approaches, the advantages of the proposed CUSUM-based defense mechanism include the low complexity approach of solving unknown parameters in the probability density function of post change distribution, and the development of Markov chain based model for analyzing the proposed approach for performance guarantee.

Second, we propose a quickest estimation scheme to determine the network topology with minimum detection/decision delay while maintaining a given accuracy constraints from the dispersive environment. The conventional topology estimation requires a long process of network status analysis for ensuring the normality. The proposed algorithm helps detect and identify the topological error efficiently and promptly for smart grid state estimation via just using online power measurement, and furthermore, reduce on vulnerability on system failure.

Finally, we investigate the energy profile allocation scheme for end-user that is capable of determining the best choice of energy profiles as few samples as possible for long-term usage under the accuracy constraint while balancing the exploration and exploitation. In other words, an online learning technique is developed to learn the evolution of the power pattern in terms of reliability over time. We derive the close form for the confident interval and obtain an upper bound for the

expected regret for the proposed scheme.

In conclusion, the proposed technologies concerning different aspects of smart grid issues, such as cyber security issues, network topology problem, alternative renewable energy resource allocation, can provide a lot of benefits to a power grid society, and will enhance the grid reliability and stability, utility services, emission control, and end-user experience in enabling better communications access to the grid, which could potentially translate into effective efficient utility operations and better living environment for human beings.

Table of Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction and Background

1.1 What is a Smart Grid?

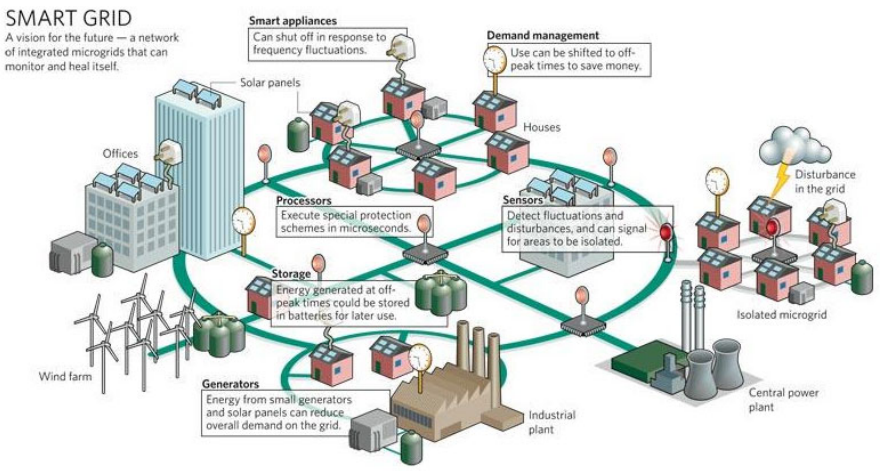


Figure 1.1 *An illustration of a "smarter" power grid [1].*

By tradition, the power grid is referred to as an electricity system, which supports several operations including the electricity control, electricity distribution, electricity transmission, and electricity generation. The traditional power grids are generally used to carry power from a few central generators to a large number of users or customers. Today, a smart grid (SG) - an enhancement of the 20th century power grid, is also known as a smart electrical/power grid, intragrid, intelligrid, or future-grid. In comparison to the traditional power grids, the smart grid uses two-way flows of electricity and information to create an automated and distributed advanced energy delivery network. A brief comparison [5] between the traditional power grid and the smarter power grid is described in Table 1.1.

The development of a smart grid has grown rapidly in recent years because of its promising economic, environmental and social benefits [6]. With the aid of modern communication technologies, a future power grid has the capability of supporting two-way information and electricity

Table 1.1 *A brief comparison between the traditional power grid and the smarter power grid.*

Traditional Power Grid	Smarter Power Grid
Electromechanical	Digital
Few sensors	Sensor throughout
One-way communication	Two-way communication
Centralized generation	Distributed generation
Failures and blackouts	Adaptive and islanding
Manual monitoring	Automated montioring
Utility centralized control	Consumer more choices

flow, resolving power outages efficiently, expediting renewable energy integration into the grid, and empowering people with better tools for optimizing their energy consumption.

The concept of the smart grid is to open a two-way, real-time communications path between utilities and consumers. The current grid is not able to handle real-time events, such as lower prices at certain times of the day when electricity rates are in less demand and therefore lower. According to [7], $787 billion federal stimulus bill passed in 2009, and more than $3.4 billion was targeted at modernizing the US electricity transmission and distribution systems and promoting investments in smart grid technologies.

The benefits for integration of the smart grid are both utility companies and the customers. It can be illustrated in Figure 1.1 [1]. For the utility companies, the smart grid allows them to have better controllability by balancing supply and demand, and to provide greater reliability by making their power purchases more rational. The consumers can save money having far more control over their usage of electricity . Furthermore, another important benefit of the smart grid is the reduction in CO2 emissions for the environment. Because supply will be better fitted to demand, fewer new generating plants need to be built. Less electricity is required because the power is more efficiently delivered and consumed. However, implementing the necessary changes is a long way development [8], because many researchers expect the resistance from regulators and consumers, who summon the complexity of the system as well as concerns about privacy and security.

1.2 Legislations, Programs, and Standards for Smart Grid

Since 2001, U.S. Department of Energy (DOE) have started several power controls and communications workshops, which objects on the integration of distributed energy resources [9]. DOE's GridWise [10, 11] is one of successful programs that influences a broad view of transformation to smart grid. From the legislations and government policy of views, U.S. federal government has proven its policy for smart grid based on two important Acts of Congress: 1) Energy Independence and Security Act of 2007 [12] is to specify studies on the state and security of smart grid; establishes a federal advisory committee and inter-government agency task force; frames technology research, development and demonstration; directs the advancement of inter-operability; and creates a matching fund program to encourage investment in smart grid [9], and 2) American Recovery and Reinvestment Act of 2009 [13] includes 3.4 billion dollars in funding for the smart grid Investment Grant Program and 615 million dollars for the smart grid Demonstration Program.

According to the report [3] from National Institute of Standards and Technology (NIST) , the U.S. Energy Independence and Security Act of 2007 directed NIST to coordinate the research and development of a framework to achieve inter-operability of smart grid systems and devices as well as the anticipated requirements of smart grid are the following:

- Improving power reliability and quality;
- Optimizing facility utilization and averting construction of back-up (peak load) power plants;
- Enhancing capacity and efficiency of existing electric power networks;
- Improving resilience to disruption;
- Enabling predictive maintenance and self-healing responses to system disturbances;
- Facilitating expanded deployment of renewable energy sources;
- Accommodating distributed power sources;
- Automating maintenance and operation;

- Reducing greenhouse gas emissions by enabling electric vehicles and new power sources;
- Reducing oil consumption by reducing the need for inefficient generation during peak usage periods;
- Presenting opportunities to improve grid security;
- Enabling transition to plug-in electric vehicles and new energy storage options;
- Increasing consumer choice;
- Enabling new products, services, and markets.

Furthermore, several major smart grid standardization roadmaps and studies have also developed in different countries and organizations. For instance, NIST IOP Roadmap [3] in U.S.A, Mandate CEN/CENELEC M/441 [14] in E.U., BMWi E-Energy Program [15] and BDI initiative Internet der Energie [16] in Germany, SGCC Framework [17] in China, METI Smart Grid roadmap [18] in Japan, Smart Grid Roadmap 2030 [19] in Korea, P2030 [20] from IEEE, SG 3 Roadmap [21] from IEC SMB, D2.24 [22] from CIGRE, and SERA [23] from Microsoft.

Although those road-maps and standardizations may need to be continuously enhanced and developed to acclimatize the changes from regulatory, political, and technical aspects, IEEE P2030 is considered as an important standard. From [20], IEEE P2030 focuses on a system level approach to the guidance for inter-operability components of communications, power systems, and information technology platforms; smart grid inter-operability provides organizations the ability to communicate effectively and transfer meaningful data, even though they may be using a variety of different information systems over widely different infrastructures, sometimes across different geographic regions and cultures. In other words, IEEE P2030 treats the smart grid network as a large, complex ”system of systems” [24] and provides guidance to navigate the numerous smart grid design pathways throughout the electric power system and end-use applications.

In addition, Smart Grid Information Clearinghouse [2] has created visual overview maps for smart grid projects and programs (i.e., Advanced metering infrastructures, equipment manufac-

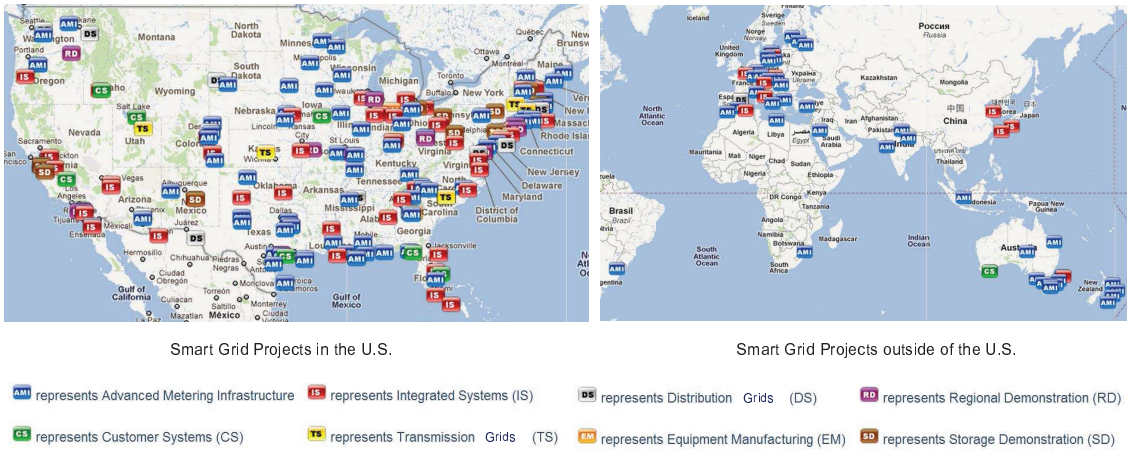


Figure 1.2 *The display map of smart grid projects [2].*

turing, integrated systems, distribution grids, transmission grids, storage demonstration, regional demonstration, customer systems) all over the world. As shown in Figure 1.2 [2], these maps roughly show the location and the objectives of those smart grid projects. Put differently, there already exist several integrated system projects.in the U.S., Europe, and East Asia, although we are just at the beginning of the smart grid transition. In [25], authors explicitly point out that almost every countries, has dedicated a significant amount of investments to projects of addressing the integration of different smart grid technologies and applications; most of the technologies are known and matured, but the integration of the technologies and the grid still has the certain degree of challenge.

So as to realize this new grid archetype, a conceptual model is provided by NIST as shown in Figure 1.3 [3]. Note that, in this figure, the solid line represents the securely bidirectional communication/information flows, the dash line represents the bidirectional electricity flows, and the cloud represents the domain. We can use this conceptual model as a reference for the various parts of the electric system because it is the place where the smart grid standardization process takes. It divides the smart grid into seven domains; each domain encompasses several essential smart grid elements, which includes devices, systems, or programs. These elements are able to make decisions and exchange information necessary for performing applications. Note that NIST propose this model from

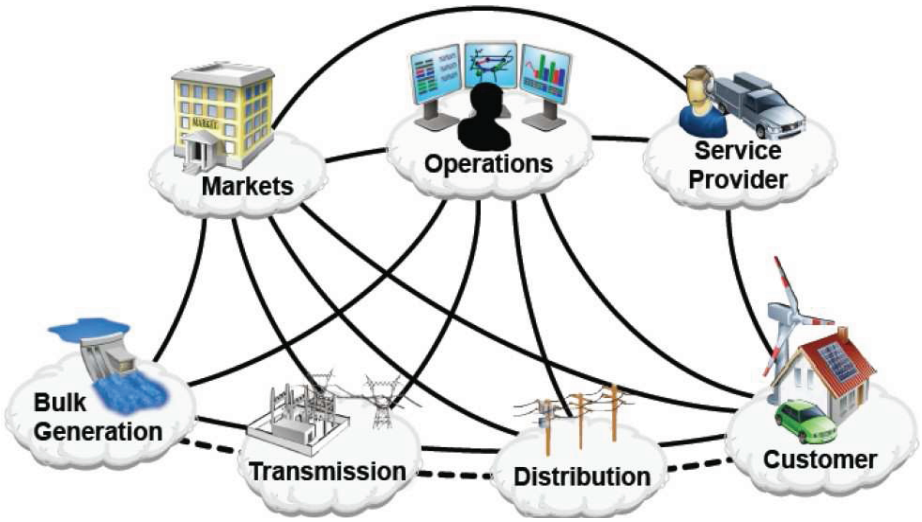


Figure 1.3 *The smart grid conceptual model from NIST [3].*

the perspectives of the different roles involved in the smart grid [3]. From 1.3, the customers domain is about end-users of electricity that includes the generation, storage, and management of the use of energy; the market domain has the operators and participants in electricity markets; the service providers domain contains the organizations, which provide services to electrical customers and utilities; the operations domain incorporates the managers of the movement of electricity; the bulk generation domain is composed by the generators of electricity in bulk quantities, and this domain may also has the ability of storing energy for later distribution; the transmission domain encompasses the carriers of bulk electricity over long distances and may also store/generate electricity; the distribution domain has the distributors of electricity to and from customers, and also store/generate electricity.

1.3 Structure of Smart Grid

Given the vast landscape of the smart grid research [19, 26–36], different researchers may express different visions for the smart grid due to different focuses and perspectives. The authors in [29], [32], and [36] give a general concept of smart grid and possible techniques can be utilized. In [33] and [19], the authors focus on the current smart grid standardizations as well as analyz-

ing thoroughly for possible future smart grid standards. The legal framework governing metering activities and policies in Europe is drawn in [34]. A practical perspective and technologies for the smart grid distribution system is developed in [28]. The authors in [27] and [30] address of the work related to cyber security and explored the privacy issues in smart grid. An overview of hybrid network architecture (i.e., communication networks for electric system automation) is explored in [31]. In [26], authors analyze how, where, and what types of wireless communications are suitable for deployment in the electric power system. A survey in [35] address the communication architectures in the power systems, including the communication network compositions, technologies, functions, requirements, and research challenges is also developed. We can conclude three major systems in smart grid from a technical perspective [24].

1.3.1 Smart Infrastructure System

The smart infrastructure system is the infrastructure of energy, information, and communication underlying of the smart grid. This system sustains three subsystems: 1) the smart energy subsystem for advanced electricity generation, delivery, and consumption; 2) the smart information subsystem for advanced information metering, monitoring, and management; and 3) the smart communication subsystem for communication connectivity and information transmission among systems, devices, and applications.

To understand the smart grid, we concisely discuss the traditional power gird. The traditional power grid is unidirectional in nature [37]. Electricity is often generated at a few central power plants by electromechanical generators. These types of generators are primarily driven by the force of flowing water or by the heat engines, which are fueled with chemical combustion or nuclear power. To fulfill the economic need and safety concern, the generating plants are located away from heavily populated areas due to their large size. From generating plans, the electricity is stepped up to a higher voltage for transmission. On the transmission grid, electricity moves over long distances to substations. Upon arrival at a substation, the electricity will be stepped down from the transmission level voltage to a distribution level voltage into the distribution grid. Finally, upon

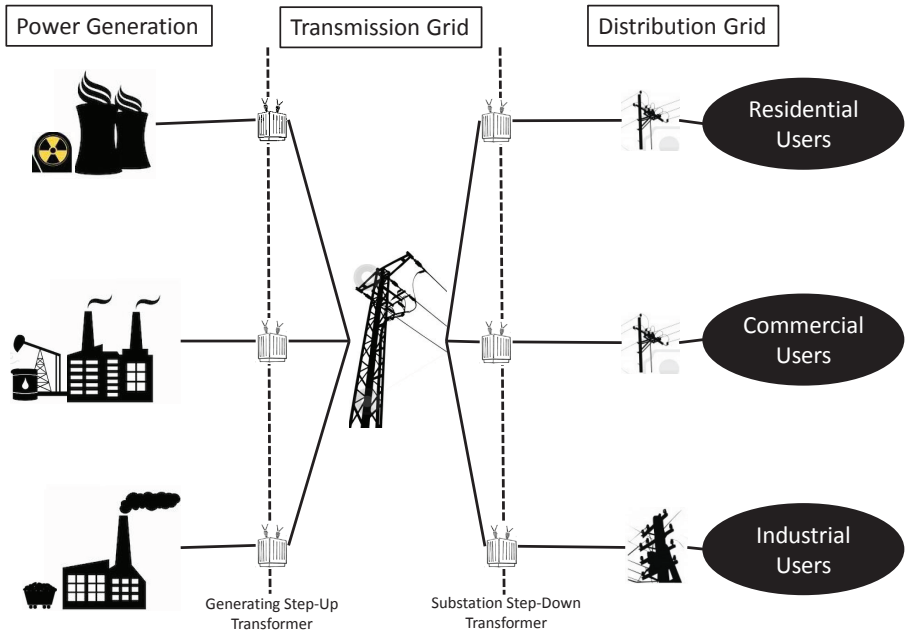


Figure 1.4 *An illustration of classical power grid.*

arrival at the service location, the power is stepped down again from the distribution voltage to the required service voltage, and eventually delivers to end-users. In contrast with the traditional power grid, the electricity generation and the power flow pattern in smart grid are more flexible since the distribution grid may also be capable of generating electricity by using renewable energy resource.

In summary, the smart energy subsystem of smart infrastructure is able to support two-way flow of electricity and information. In other words, the action of transmitting information or delivering electricity is bidirectional. In the existing grid (traditional power system) as shown in Figure 1.4, the utility company sells/delivers the electricity from the power generator, transmission grid, distribution grid, to consumers in one way. In smart grid, the electricity is able to send back to the grid verse vise. Thanks to the easy access of renewable energy, the consumers may use solar plane, wind turbine, or etc to put-back electricity to the grid. It is extremely helpful to send power back to the grid when demand is high (i.e., with additional power supply, the grid will be able to balance loads by peak shaving). In addition, this two-way flow of electricity and information is important because, in near future, each regional power grid can be formed into a macro-grid with several

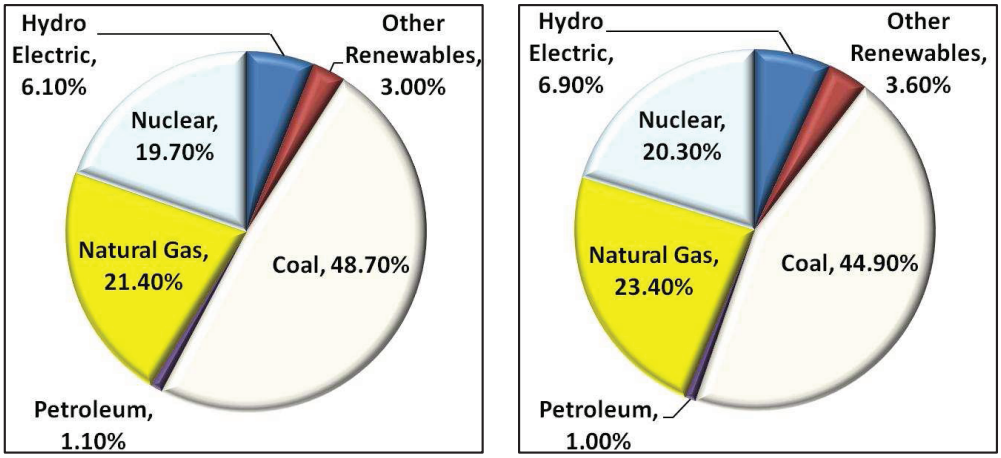


Figure 1.5 *United States electricity generation by sources (Left: 2008, Right: 2009) [4].*

micro-grid [38, 39]. This backward flow can be helpful in a micro-grid that has been islanded due to power failures in the macro-grid. Consequently, the local power supply quality is improved (i.e., Multiple DGs has the same reliability, and lower capacity margin than a system of equally reliable generators) [40]. Improves the grid efficiencies, reliability, high penetration of renewable sources, self-healing, active load control [41]. Note that there are many energy sources used to generate electric power. In Figure 1.5 [4], the statistics show U.S. electricity generation by source in 2008 and 2009. As fossil fuels get depleted and generally get more expensive, it is expected that the renewable energy will play a more important role in the future power generation.

The development of smart grid relies on both advancement of power equipment technology and improvement of sophisticated computer control, monitoring, and analysis from exclusively central utility offices to the distribution and transmission grids. The authors in [42] promote an information technology perspective to address the concerns of distributed automation, such as interoperability of data exchanges and integration with existing and future devices, systems, and applications. Therefore, for smart information subsystem, an advance Information management is need for a large amount data in smart grid. The functionalities of management system are the information, information integration, and the information optimization [43]. It can help to improve information effectiveness and reduce communication burden via soring only useful information [44]. One

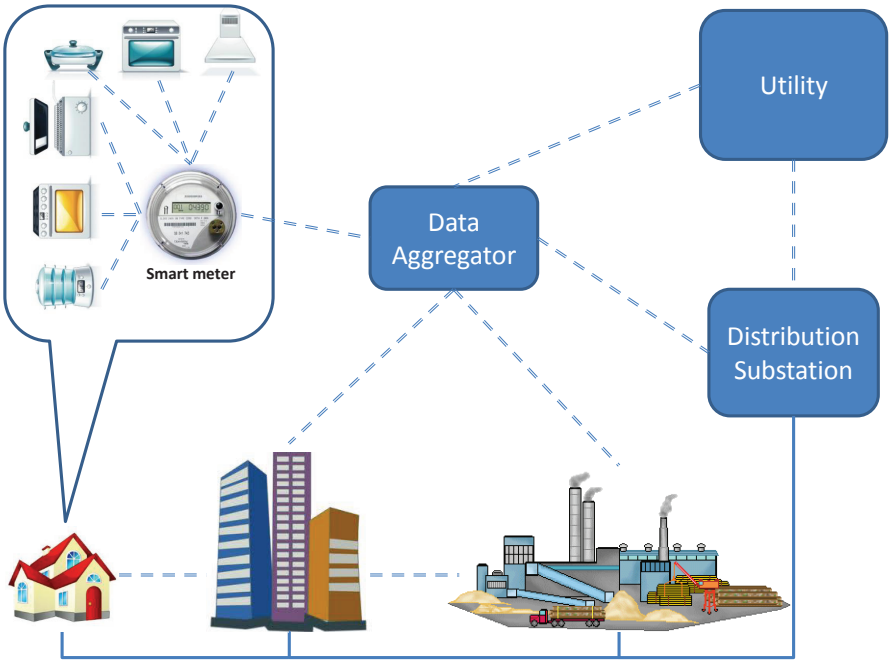


Figure 1.6 *An illustration of the smart metering structure*.

important sensing layer is a network of smart meters, which support two-way communications between the meter and the central system. A smart meter is usually an electrical meter that records consumption in intervals of an hour or less and sends that information at least daily back to the utility for monitoring and billing purposes [45]. Figure 1.6 illustrates an example of the smart metering structure. The smart meter collects the power consumption information of the household applicants such as the refrigerator, television, and dishwasher; smart meter also has the capability of sending the control commands to household applicants if necessary; a data aggregator, which can be an access point or gateway, collects the data that generates by the smart meters from different buildings; finally, all data can be further routed to the electric utility or the distribution substation [24]. Note that the task of information transmission/routing is belong to the smart communication subsystem.

The smart communication subsystem must be able to maintain the quality of service of data (i.e., critical data must delivered promptly), guaranteeing the reliability (i.e., smart grid is a large and heterogeneous network), and be pervasively available and have a high coverage for any event in the grid in time [46]. The smart communication subsystem must fulfill some basic requirements.

First, it supports the quality of service of data is important because the critical data (e.g. the grid status information) must be delivered promptly. Second, it guarantees the reliability of a large and heterogeneous network, because a large number of devices are connected and different devices and communication technologies are used in smart grid. Third, it must be pervasively available and have a high coverage since smart grid can respond to any event in the grid in time. Finally, The communication subsystem must guarantee security and privacy. An example of a communication network used in smart grid is shown in Figure 1.7. End-user devices and smart meters use ZigBee, WiFi, and power-line communications. Wireless mesh networks are used for information exchanges between users. Communities are connected to their electric utility via free-space optical, satellite, microwave, or cellular systems. A substation communicates with an electric utility over the power-line. Therefore, communication and networking technologies of smart grid can include both wireless technologies and wired technologies [24].

In wired technologies, fiber-optic Communications have a long history with power companies to help connect their generation network with their network control facilities. Furthermore, fiber-optic communication has the unique property, which is able to immunize from electromagnetic and radio interference, so that it is ideal for high voltage operating environment [47]. Next one is the power-line communications, which is capable to carry data on a conductor via electric power transmission. Utility companies have been using PLC for remote metering and load control applications around the world [48]. One advantage of using power line communication is that has deployment cost comparable to wireless technologies since the power-lines are already almost everywhere [49]. However, the debate on power-line communications in future smart grid is still open; some researchers claim that power-line communications work well with some applications [49, 50], while others concern on its security issue due to the nature of power-lines [51, 52].

On the other hand, wireless has already been widely used in our daily life and can be deployed anywhere and anytime. The advantages of wireless technologies over wired technologies include the low installation cost, mobility, rapid deployment, etc., and the great suitability of remote-end applications [53]. There are several wireless communication technologies, which are suitable for smart

grid. First, the cellular Communication Systems on smart grid, such as GSM [54] and 3G [55, 56], is a radio network distributed over large land areas (i.e., it is also called the cell). Each cell is served by at least one fixed-location transceiver known as a base station. For several decades, the wireless communication has been proven as a matured technology; the authors in [57] state that the existing cellular communication systems is quick and inexpensive to obtain data communications coverage over a large geographic area. Second, the wireless Communications based on 802.15.4, such as WirelessHART, ISA100.11a, and ZigBee. WirelessHART is a self-organizing, time-synchronized, and self-healing mesh architecture; it supports to operate in 2.4 GHz band using IEEE 802.15.4 standard radios; WirelessHART was defined as a multi-vendor, inter-operable wireless standard for the requirements of process field device networks. ISA100.11a is an open wireless networking technology standard developed by the International Society of Automation; it is officially known as *Wireless Systems for Industrial Automation: Process Control and Related Applications*. Using WirelessHART or ISA100.11a for wireless sensor network applications in smart grid is suggested in [24], such as a substation or a generation plant. ZigBee is one widely used communication technology in the customer home network domain of the smart grid by the U.S. NIST [3]; it is designed to fulfill the specific objectives, which includes a long battery life, low data rate, and secure networking for radio-frequency applications. Because of a standardized platform for exchanging data between smart metering devices and appliances on customer domain, ZigBee is selected by many electric utilities as the communication technology for the smart metering devices [37]; its features include demand response, advanced metering support, real- time pricing, text messaging, and load control [58].

For point-to-point communications, microwave/free-space optical communications technologies are widely used; the distinctive nature of microwave, such as the small wavelength, permits users to utilize conveniently the sized directional antennas so that users can obtain secure information transmission at high bandwidths. The statistics shows that at least half of the worldwide mobile base stations are connected using point-to-point microwave technologies [59]; microwave has been the primary solution for rapidly rolling out cost-effective world mobile back-haul infrastructure [60].

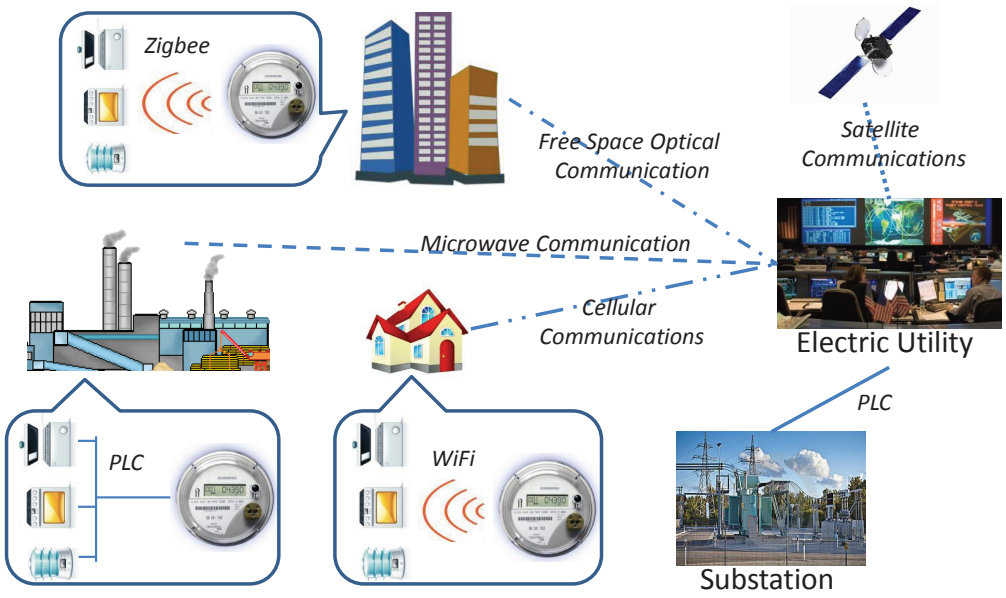


Figure 1.7 *An illustration of a smart grid communication network.*

Another important issue in the smart communication subsystem is the end-to-end communication management. This heterogenous communication subsystem uses various communication technologies, devices, and network structures. Therefore, we need to identify each entity and solve the problem of how to manage end-to-end communications (i.e., perhaps between any pair of entities).

1.3.2 Smart Management System

The smart management system provides advanced management and control services and functionalities for smart grid. One reason for the smarter power grid is the widespread deployment of functionality based on its smart infrastructure. In other words, the grid becomes smarter via leveraging the technology and capability upgrades with the development of new management applications and services. The smart management system takes advantage of the smart infrastructure to pursue various advanced management objectives.

Most of management objectives are related to operation cost reduction, emission control, energy efficiency improvement, utility maximization, and supply and demand balance. In the frame-

work of smart grid, many such goals become achievable in comparison to the traditional power grid in past. In current works for smart management, there are three mainly focused objectives as follows. First, Energy efficiency and demand profile improvement includes the minimizing energy loss and demand profile shaping, which help match the demand to the available supply. Classical methodologies of shape demand profile includes shifting, scheduling, or reducing demand so that management system is able to reshape a demand profile full of peaks to a nicely smoothed demand profile, or reduce the peak-to-average ratio or peak demand of the total energy demand [61–66]. Second, the smart management system focuses on the utility and cost optimization, and price stabilization. The various perspectives of this objective incorporate from an individual user cost, bill, and profit [67,68], an single energy bill and aggregate utility of a group of users [69,70], to the cost and utility of electricity industry and system [68, 71]. Stabilization of prices is also essential, because the realtime wholesale energy market prices to the end consumers can be treated as a closed loop feedback system, which may results the price volatility via unstable or lack robustness. Finally, the another important management objective is the emission control, which has a significant impact on environment protection. However, minimizing emission by utilizing renewable energy is not always the best approach due the cost of power generation from renewable energy source is not always the lowest in general.

1.3.3 Smart Protection System

smart protection system provides security/privacy protection services, failure protection, and advanced grid reliability analysis. Enabling by smarter infrastructure in smart grid, smart grid services are administrated by a smarter management system and the necessarity of a smarter protection system must be considered as well, which includes effectively efficiently tackling cyber security issues, preserving privacy, and supporting failure protection mechanisms. This system must address not only unintentional compromises, such as user errors, equipment failures, and natural disasters, but also deliberate cyber attacks, such as from disgruntled employees, industrial spies, and terrorists.

Cyber Security is a never-ending game. Smart grid security is no exception to this paradigm.

The authors in [3, 72] agree that the smart grid cyber security is one of greatest challenges. For security of smart metering, during recent years of newly installing smart meters, its vulnerabilities can easily be monetized [73]. In other words, the widespread deployment of smart meters leads to a potentially large number of opportunities for cyber attacks, and even can result large-scale disasters. The Security in Monitoring and Measurement devices (i.e., sub automats control stations, phase measurement units, etc.) can lead to another system vulnerabilities. The effective services of smart grid depend on these widely deployed accurate measurement devices; measurements are typically sent to Supervisory Control and Data Acquisition (SCADA) Systems [74], and then state estimators in SCADA estimate the system state through analysis of measurement data and power system models. Therefore, it is very important to guarantee the integrity of the data. For failure prediction and prevention, the system can try to prevent failures via predicting the weak points or the region of stability existence in the energy subsystem of smart infrastructure. Once the failure occurs, the system must be able to execute failure identification, diagnosis, and recovery that the first step must be quickly locating and identifying the failure to avoid cascading events. Moreover, system reliability is another major topic for smart protection system. U.S. annual cost of outages is almost one third of the total electricity revenue in 2002 [75]. While expecting widespread deployment of distributed generation, some fluctuant and intermittent renewables resource power pattern may compromise the stability of the grid [76], Furthermore, in smart grid, the reliability and stability also depends on the reliability of the measurement system, which is used to monitor the reliability and stability of the grid [77].

1.4 Challenges for Smart Grid

Indeed, the smart grid technology has improved the robustness and efficiency of traditional power grid networks by exploiting the modern technological advances in sensing, measurement, and control devices with two-way communications. The information exchange among users, operators, and control devices significantly improves the efficiency in production, transmission, and distribution in the power system. However, the integration of intelligence into the power grid also poses

many challenges such as increasing system complexity, network security risk, end-user data privacy issues, uncertainty of the renewable energy generation, and etc. Here, we list some important challenges and issues, which are worth exploring.

1.4.1 Impact of System Complexity

The advanced infrastructure used in smart grid is a double-edged sword. On one hand, it can easily result to pose the vulnerability of system failures or cyber attacks due to increase system complexity and expand communication paths. On the other hand, it is the main foundation of intelligent power grid in future that can serve end-users better. Note that a fully implemented smart grid may consist of tens of millions of nodes. Such large scale network has the difficulty to anticipate system failures due to many dependent or independent unpredictable factors in the grid or even the manipulation of an unpredictable/intelligent adversary [78]. Many researchers suggest dividing the entire system into many sub-grids so that the system complexity can be reduced. In other words, a safety net between each sub-grid can be established for reducing the impact of system failures and attacks. However, it does not mean that no connection among these sub-grids and electric utilities, and as a result, the failures or attacks cannot be completely isolated. Thus a complete solution needs to consider both autonomy and inter-connectivity.

1.4.2 Establishment of A Large Scale Deployment

In the fully implemented smart gird, there are possible tens of millions of nodes. The leak experience in large-scale distributed control approaches to addressing the complex power system component interactions, and in large-scale deployment of new technologies, such as batteries, thermal storages, DGs, and EVs [79]. How to establish a large amount of devices in a large-scale smart grid is not a trivial problem. Likewise, in other types of network, two approaches may be applicable for smart grid: 1) the bottom-up approach is the piecing together of systems to give rise to grander systems, which are grown up from many individually formed subsystems (i.e., the performance of the whole system may not be good enough), and 2) the top-down approach is allow that a central-

ized powerful grid operator formulates the high-level framework of the system and refines in greater details for the subsystems (i.e., this approach needs a powerful operator to initially design the whole architecture, which is not an easy task). The advantages and disadvantages of both top-down and bottom-up approaches need to be investigated.

1.4.3 Complication of Decision Making Process

As previously mentioned about the increased system complexity and expanded communication paths, processing failures in smart grid is usually required to solve much more complex decision problems, but within a limit time [75] (i.e., smart grid is delay sensitive; it must response the abnormal events promptly to avoid the catastrophic failures and system instability). Note that a fully implemented smart grid has at least millions of nodes. It is not an easy task. Some studies propose a system of using multiple distributed decision-makings. For instance, a large number of failure sensors can be deployed evenly to the system; each sensor has the responsibility to monitor several devices and makes decisions locally. The complexity of the decision making process can be simplified and the failure response time can be decreased. However, a locally optimal decision is not always globally optimal. We need to consider how to balance the response time and the effectiveness of the local decision. The complete solution based on this idea is needed to investigate.

1.4.4 Cryptographic Inter-operability between Different Systems

The smarter power grid is enabled by a vary of different communication protocols and technologies. Due to various aspects of security needs, each of them may take different cryptography requirements. Therefore, understanding the cryptographic inter-operability between systems is an essentials problem. In order to securely issue and exchange cryptographic keys from different systems, one possible approach is to design a public key infrastructure [80], which are similar to the layered approach in communication models. However, as an initial idea [80], a complete solution based on this approach is required to investigate thoroughly.

1.4.5 Confrontation between Privacy Preservation and Information Accessibility

In smart grid, millions of user data have been collected for performance analysis so that the utility company can provide the better power service to end-users. Balancing between privacy preservation and information accessibility is always challenging. Considering an example of power flow optimization, if the users are willing to release more information about the demand patterns, the management system can make smarter decisions on optimizing supply for the demand. However, more accessible information usually means more privacy leaks, which may easily disclose user profiles and behaviors. The authors in [24] promote an idea of defining several privacy preservation levels similar to those in access control, each of which describes a tolerable amount of information leak. In other words, each level based on the accessible information, the management objectives are defined.

1.4.6 Effective Utilization of Intermittent and Fluctuant Renewable Energy Resources

In smart grid, the distributed renewable energy generations are grown and will be widely used as alternative energy resource. The authors in [81] discuss the problems of intermittent and fluctuant nature of wind and solar generation, which requires much more complicated forecasting and scheduling; in addition, understanding and exploring both long-term and short-term renewable source patterns and likely behavior are necessary in future smart grid. In other words, the utilization of the renewable resources, such as hydro, wind and solar, also makes management more difficult due to their fluctuant and intermittent nature. Maintaining reliability and satisfying operational requirements are the first priorities for the management system, meanwhile must taking into account the uncertainty and variability of energy sources. Thus, the utilization of distributed renewable energy resources poses many challenges and opens up many new research topics such as online learning model of renewable resource, optimal deployment of the additional energy reserves, and etc.

Among these challenges, the top three areas we concern most are network stability, network

security, and resource allocation which are required the modern techniques in smart grid to response promptly on irregular events such as grid instability via fault status, network intrusion detection, or distributive renewable energy resource management.

1.5 Contributions of this Dissertation

Although the smart grid have been investigated heavily in many directions and aspects when it was raised for the first time, the research on the power system issues and the quickest detection techniques on smart gird networks is still limited. In this dissertation, we like to explore specifically in three areas: system status, security issue, and resource management in smart grid networks.

In Chapter 3, we propose a CUSUM-based defense strategy against the false data injection attack in smart grid networks. Most cumulative sum (CUSUM)-based tests expectedly have the perfect knowledge of the likelihood functions. The contributions and advantages of the proposed CUSUM-based defense mechanism include: 1) it is able to tackle the unknown parameters in the probability density function of post change distribution via the low complexity approach, 2) the decision-making of the proposed scheme for detecting attack is based on using multiple online samples/observations rather than using a single observation while maintaining a certain level of decision accuracy, and 3) Markov chain based approach is developed to analyze the proposed approach for performance guarantee (i.e., the analytical model of proposed scheme can guide us to configure a detection system based on some detection performance requirement such as false alarm rate, average detection delay, and missed detection ratio under a detection delay constraint). The accuracy of the analytical model and detection with performance guarantee are also discussed. The Simulations are conducted with MATPOWER 4.0 package for different IEEE test systems.

In Chapter 4, we propose a quickest estimation scheme to determine the network topology as quickly as possible with given accuracy constraints from the dispersive environment. Unlike the conventional topology estimation requires a long process of status analysis that the sensor at each bus senses, collects, analyzes, and then finally, sends the status measurement to the control

center; in this chapter, one essential objective is to help detect and identify the topological error efficiently and promptly for smart grid state estimation via just using online power measurement, and furthermore, reduce on vulnerability on system failure. In addition, the proposed algorithm is software-based, which has the ability of avoiding the deployment of additional sensors and the cost of expensive hardware. A Markov chain based analytical model is also constructed to systematically analyze the proposed scheme for the on-line estimation. With the analytical model, we are able to configure the system parameters for the guaranteed performance in terms of the false alarm rate and missed detection ratio under a detection delay constraint. The performance is evaluated through both analytical and numerical simulations with the MATPOWER 4.0 package. It is shown that the proposed scheme achieves the minimum average stopping time, but retains the comparable estimation accuracy and false alarm rate.

In Chapter 5, we investigate the energy profile allocation scheme for end-user that is capable of determining the best choice of energy profiles as few samples as possible for long-term usage under the accuracy constraint while balancing the exploration and exploitation. In other words, a online learning technique is developed to learn evolution of power pattern (i.e., taking into account the uncertainty and variability of energy source) in term of reliability overtime. We derive the close form for the confident interval and obtain an upper bound for the expected regret for the proposed scheme. From the simulation results, we can show that a user can effectively switch and select the best energy profile with the minimum delay while balancing the exploitation and exploration. The great potential of the proposed algorithm, online quickest multiarmed bandit algorithm, includes the solution of online strategizing allocation in randomized environments such as electrical vehicles scheduling, DRER allocation, and etc.

Finally, in Chapter 6, we conclude our work and explore the possible extensions of our proposed framework. We propose some future work, such as the extension for fault detection in the fully-distributed smart grid, the optimality of sequential BDD algorithm in smart grid estimation, an quickest search on profile scheduling and utilization for grid-to-vehicle (G2V) / vehicle-to-grid (V2G) , and real-world implementation in Universal Soft-ware Radio Peripherals 2 (USRP2) hard-

ware (i.e., mimic the power grid network via USRP2).

1.6 Organization of this Dissertation

The remainder of this paper is organized as follows. We briefly discuss the foundation of basic quickest detection and how its frameworks can help the smart grid challenges along with literature research of different fields in Chapter 2. Chapter 3 investigates a false data injection defense strategy in smart grid state estimation while considering unknown attacker model and multi-thread observations. In Chapter 4, an innovative approach employing both quickest estimation and multiuser detection is proposed to identify promptly the network topology error for reducing the vulnerability of system failure. 5 discovers a quickest multiple arm bandit algorithm for end-user to punctually select the most profile (i.e., reliable, efficient) energy resource profile among others for long-term usage under an appropriate decision time, exploration, and exploitation. The conclusion and a brief discussion about the related future works is given in Chapter 6.

Chapter 2

Quickest Detection for Smart Grid

Many applications, such as those in image analysis [82] [83], medical diagnosis [84] [85], or econometrics [86] [87] [88] [89], involve primarily off-line analysis to detect a change in statistical behavior during a per-specified frame of time or space. There are many applications of change detection in which it is of interest to perform on-line (e.g., in real time) detection of such changes in a way that minimizes the delay between the time a change occurs and the time it is detected. This type of problem is known as the quickest detection problem.

2.1 Why is Quickest Detection on Smart grid?

To address these challenges in smart grid, the advance digital signal analysis and detection is important and must be processed in a realtime manner, such that the information can be updated and the correct action can be taken as quickly as possible before major failure. Such data is delay-sensitive, and the unexpected delay of information can cause the instability of smart grid networks.

In other words, the smart grid technology has improved the robustness and efficiency of traditional power grid networks by exploiting the modern technological advances; the information exchange among users, operators, and control devices significantly improves the efficiency in production, transmission, and distribution in the power system. However, the integration of intelligence into the power grid needs to act punctually on abnormal situations (i.e., system fault, attacker, short-cut, etc.) [6].

One interesting approach for this type of problem, the quickest detection (QD) [90], can be employed for decoding on-line or real-time information in a way that minimizes the delay between the time a change occurs and the time it is detected while maintaining a certain level of detection accuracy.

2.2 Basic Quickest Detection

The idea of QD attempts to determine a change as quickly as possible based on real-time observations such that the user-defined condition is satisfied. The strict term of the user-defined condition is known as the decision rules, which optimize the tradeoff between the stopping time and decision accuracy (i.e., pre-defined error probability, initial prior probability of each hypothesis occurring, etc. [90]). The classification of QD includes: (1) Bayesian framework (e.g., the sequence probability ratio test (SPRT)), which detects the distribution changes between two known distribution at random times. It requires the full knowledge about the prior distributions of changing time. (2) non-Bayesian framework (e.g., the cumulative sum (CUSUM) test), which detects a change of distribution to known/unknown distributions at random times. It is executable with unknown distributions of changing time. We now give more detail relating these concepts.

2.2.1 Probability Spaces

We summarize some essential notions from probability theory that will be useful in the sequel. Most of this material can be found in many basic books.The basic notion in a probabilistic model is that of a random experiment, in which outcomes are produced according to some chance mechanism. From a mathematical point of view, this notion is contained in an abstraction a probability space, which is a triple (Ω, $\mathcal{F}$, P) consisting of the following elements [90]:

- a sample space Ω of elemental outcomes of the random experiment;
- an event class $\mathcal{F}$, which is a nonempty collection of subsets of Ω to which we wish to assign probabilities;
- a probability measure (or probability distribution) P, which is a real-valued set function that assigns probabilities to the events in $\mathcal{F}$.

To be able to manipulate probabilities, the event class is not allowed to be arbitrary, but rather be assumed that it is a σ-field (or σ-algebra); the assumption is that $\mathcal{F}$ is closed under complemen-

tation and under countable unions. The usual algebra of set operations then assures that $\mathcal{F}$ is closed under arbitrary countable sequences of the operations: union, intersection, and complementation. Such a class necessarily contains the sample space Ω and the null set. The elements of $\mathcal{F}$ are called events. A pair (Ω, $\mathcal{F}$) consisting of a sample space and event class is called a measurable space or a pre-probability space. The probability measure P is constrained to have the following properties, which axiomatize the intuitive notion of what probability means [90]:

- $P(\Omega) = 1$;
- $P(F) \geq 0, F \in \mathcal{F}$
- $P\left(\cup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} P(F_n)$, where all sequence $\{F_k; k = 1, 2, \cdots\}$ of elements of $\mathcal{F}$ satisfying $F_m \cap F_n = 0, m \neq n$. P is constrained to be non-negative, normalized , and countably additive.

2.2.2 Stopping Rule

We discuss the notions and some general properties relating to stopping time. Consider a probability space (Ω, $\mathcal{F}$, P) and a filtration $\{\mathcal{F}_k; k = 0, 1, \cdots\}$. A stopping time is an extended random variable T taking values in the set $\{0, 1, 2, \cdots\} \cup \{\infty\}$, with the property [90] that

$$\{\omega \in \Omega | T\{\omega\} \leq k\} \in \mathcal{F}_k, \tag{2.1}$$

where a stopping time associated with a filtration is an extended random variable taking values in the time set of the filtration, with the property that it can assume the value k only on events that are measurable with respect to the filtration at k. For a filtration $\{\mathcal{F}_k; k = 0, 1, \cdots\}$ define the σ-field $\mathcal{F}^{\infty}$ as the smallest σ-field containing $\cup_{k=0}^{\infty} \mathcal{F}_k$. A stopping time T associated with $\{\mathcal{F}_k\}$ defines the σ-field $\mathcal{F}_T$ and can be expressed as

$$\mathcal{F}_T = \{F \in \mathcal{F}_{\infty} | F \cap \{T \leq k\} \in \mathcal{F}_k; k = 0, 1, 2, \cdots\}, \tag{2.2}$$

where events in $\mathcal{F}_T$ is prior to T. If the sequence $\{X_k\}$ is a random sequence adapted to $\{\mathcal{F}_k\}$, then X_T is an $\mathcal{F}_T$ measurable random variable. Therefore, if the filtration is the minimal filtration

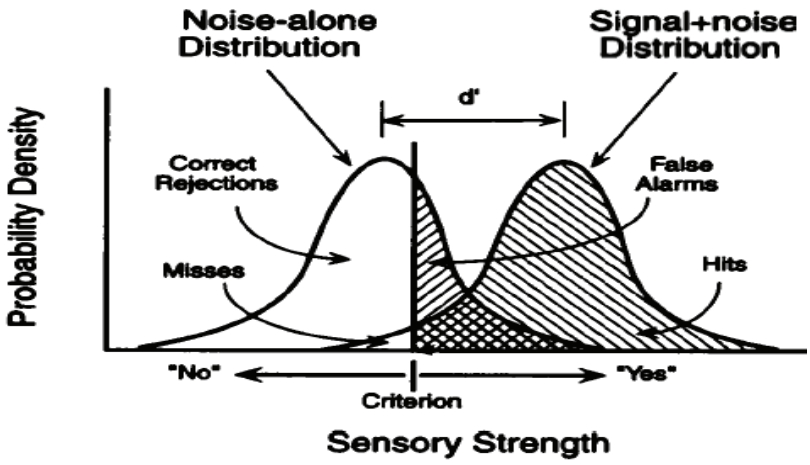


Figure 2.1 *An illustration of statistical hypothesis signal detection paradigm.*

generated by a random sequence $\{X_k\}$, then whether $T = k$ or not can be determined by observing $X_0, X_1, \cdots, X_k$.

2.2.3 Statistical Hypothesis Testing

A statistical hypothesis test (SHT) is a efficient scheme of making decisions using observation data, whether from an observational study (not controlled) or a controlled experiment. One important element is known as the statistical significant; if it is unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significance level [91]. The objective of SHT is to determine what outcomes of an experiment would lead to a rejection of the null hypothesis for a pre-specified level of significance as well as helping to decide whether experimental results contain enough information to cast doubt. One common method is the Bayesian approach to hypothesis testing; it is to base rejection of the hypothesis on the posterior probability. The critical region of a hypothesis test is the set of all outcomes which cause the null hypothesis to be rejected in favor of the alternative hypothesis. For instance in Figure 2.1, we try to determine whether the signal is transmitting, and we need to pick one of two possible responses: 1) "Yes," the signal is transmitting, or 2) "No," there is only noise. We have the null hypothesis H_0 (no signal, noise only)

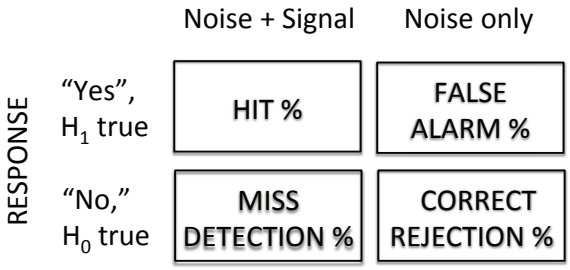


Figure 2.2 *An illustration of statistical hypothesis signal detection paradigm.*

and the target hypothesis H_1 (signal and noise):

$$\begin{aligned} H_0: \quad y(t) &= \quad n(t), \\ H_1: \quad y(t) &= \quad h(t) + n(t), \end{aligned} \tag{2.3}$$

where $y(t)$ denotes the received signal, $h(t)$ is the channel model, and $n(t)$ is the noise model. From Nieman-Person Theorem [90], the decision is given via the likelihood ratio test based on the posterior probabilities of H_1 and H_0:

$$\begin{aligned} &\text{Select } H_1 \quad \text{if } \log \tfrac{P(y|H_1)}{P(y|H_0)} \geq \eta, \\ &\text{otherwise} \qquad\quad \text{select } H_0. \end{aligned} \tag{2.4}$$

In other words, we select H_1 when the likelihood ratio is greater or equal to the predefined threshold η; otherwise, the null hypothesis is selected. We can describe this example in a graphic way.

As shown in Figure 2.1, H_0 is belong to the noise-alone distribution, H_1 is belong to the signal with noise distribution, d' is the distance between the distributions' mean, and the criterion represents as η. A decision-making point is at the criterion, where we need to decision "Yes" or "No." Figure 2.2 shows the relationship between the decision response and the hypotheses. If the decision is "Yes," under the true H_1, then it is a positive hit. If "No," under the true H_0, then it is a positive/correct rejection. However, when the decision is "Yes," under the true H_0, then we call it "false alarm;" when "No," under the true H_1, it is a "missed detection." The false alarm is also known as a Type 1 error, which incorrectly reject of a true null hypothesis. The missed detection is a Type 2 error, which failure to reject a false null hypothesis. To determine the detection performance,

a receiver operating characteristic (ROC) is usually produced.

A ROC is a graphical plot, which illustrates the performance of a binary classifier system as its discrimination threshold is varied [92]. It is created by plotting the fraction of positive hits out of the all hits vs. the fraction of false positives out of all miss, at various criterion or threshold settings. The main functionality of ROC analysis is to help selecting possibly optimal models and to discard suboptimal ones independently. It is related in a direct and natural way to cost/benefit analysis of diagnostic decision making. Figure 2.3 demonstrates the tradeoff between false alarm and detection probability. Note that the x-axis is the false alarm rate, and y-axis is the probability of hit. The sum of probability of hits and missed detection ratio should be 1. The sum of probability of correct rejections and false alarm rate should be 1 as well. Theoretically, the ROC curve, which is closer to the upper left corner, has the better detection performance in comparison to others.

Furthermore, the four results with different criterion/η around the ROC space in Figure 2.3 can incorporate with Figure 2.1,2.2. The case A has the low criterion for detection/decision and results with high positive hits (i.e., H_1 is almost always true and even it is false) as well as a extremely high false alarm ratio. The case B has the equilibrium criterion point by considering fairly both false alarm rate and missed detection ratio but the probability of hits is mediocre. The case C is given a high criterion for decision-making that the missed detection ratio is high (i.e., H_1 is almost always false and even it is true). The case D and the case B have a resemblance in criterion, but with smaller d'. In other words, the distributions of H_1 and H_0 have the higher similarity in the case D that of case B. The main objective is to produce the highest probability of hits while maintaining the minimum false alarm rate. Therefore, balancing between the false alarm rate and missed detection ratio is a critical task for detection/decision schemes.

Therefore, SHT is one essential part for QD technique [93]. In other words, the mechanism of SHT is that the receiver classifies a sequence of observations into one hypothesis among multiple hypotheses with a hypothesis normally representing a type of distributions. There are a lot of recent literature to apply QD with SHT to a variety of networks. The authors in [94] use the CUSUM tests as a collaborative QD for detecting a distribution change in ad hoc networks. The CUSUM test

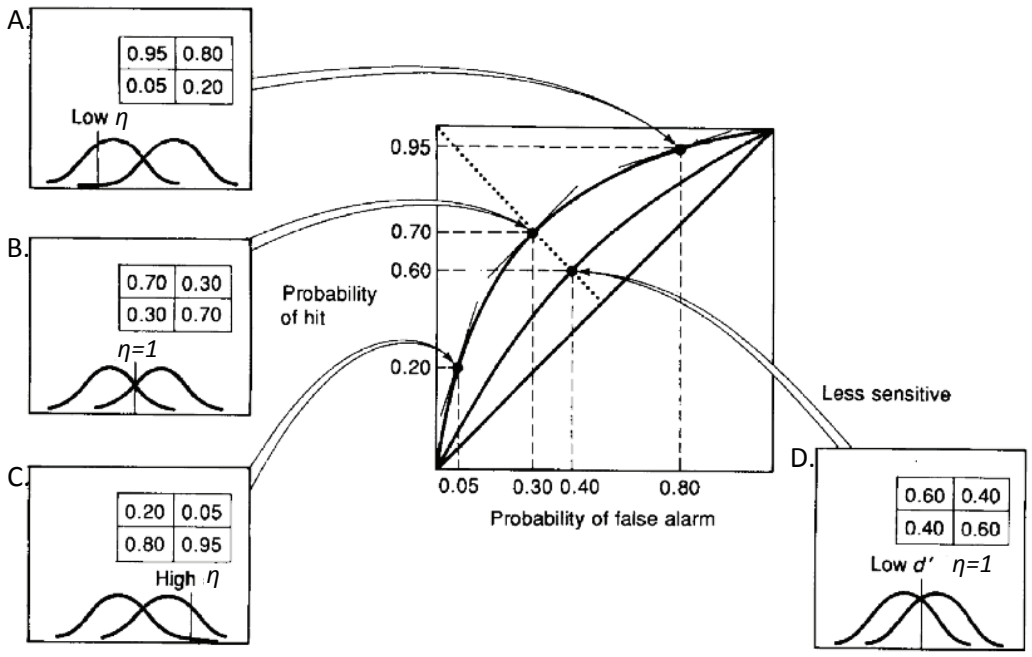


Figure 2.3 *An illustration of receiver operating characteristic curve*.

to address the real-time backoff misbehavior problem in IEEE 802.11 networks is utilized in [95]. The work in [96] applies the modified CUSUM test to detect an abrupt distribution change with an unknown time-varying parameter. In [97], the authors modify the sequential probability ratio test (SPRT) to detect the occupancy of the unknown primary user in a licensed spectrum band for a cognitive radio application. The cognitive radio spectrum sensing with unknown parameters of PU was described in [98]. The QD technique combining with the statistical hypotheses test for detecting abnormal change is practiced in [93]. Similarly, the authors in [99, 100] demonstrate that its multi-hypothesis sequential probability ratio test is asymptotically optimal for the minimum expected sample as the decision risks go to zero. In the generalization [101] of the traditional multi-hypothesis test (MHT), the algorithm first computes a set of upper and lower thresholds for each hypothesis, and then evaluates the input signal against each set to determine the true hypothesis. However, none of the existing work has considered the unique environment of smart grid networks. However, little existing work has considered the unique environment of smart grid networks.

2.3 Motivation on Smart Gird

2.3.1 Network Security

Indeed, smart grid is delay-sensitive and requires the techniques that are able to identify/react on the abnormal change in a very timing manner. If the detection and corresponding response are not made promptly, the grid may become unstable and further cause the catastrophic failures throughout the entire network. For example, in the control center of smart grid, an essential task of the energy management system (EMS) is to estimate the system states by collecting the data from remote meters periodically. If the adversaries are able to inject the malicious data, EMS may produce the false state estimation, which potentially results in wrong decisions on billing, power dispatch, erroneous analysis and even blackout [102]. Thus, the smart grid network must have the protection mechanism, which has the capability of detecting the abnormal change and then making the decision as quickly as possible. Therefore, this strongly motivates us to propose the quick detection based detection scheme.

There have been many researches for smart grid security in literature. A framework for analyzing the cyber-attacks impact in a smart grid is presented in [103]. The work in [104–107] formulates the attacks, which are able to evade from the conventional detection in smart grid. The authors in [108] address the problem of false data injections in the case of cyber-attacks of the power system. The micro-grid vulnerability is discovered in the smarter power system via using the false data injection attack in [109]. The novel false data injection attacks against the energy routing process is investigated in [110]. In this paper, we like to focus on studying in the non-stealthy context with the proposed detection scheme that can be an interesting practical contribution for smart grid networks.

To address this type of security problem in smart grid, EMS in the control center needs to be equipped with the capability of real-time detection of malicious attacks by analyzing the statistical behavior of the state estimation process to timely prevent further damage to the entire network. From the quickest detection (QD) framework [90], the cumulative sum (CUSUM) based approach fits well

to this type of detection problems because of its non-Bayesian properties. Such framework aims to determine a change of the observed statistics as quickly as possible based on online observations, the user-defined decision rules, and the requirement of detection accuracy. The decision rules should be properly designed to optimize the tradeoff between the stopping time and decision accuracy.

2.3.2 Network Status

The state estimation plays a major role in building such real-time models of power grid networks. Two types of measurement data are collected for state estimation in modem energy management systems (EMS) of smart grid networks, namely: i) the status data of switches and breakers, and ii) the analog data of bus voltage, power injection, power flow, and reactance. The status data is used to determine the real-time topology of the network. The analog data is used to determine the loading/voltage profile of line and transformer. However, both status and analog data are distortive because of the missing data, communication errors, or measurement errors. Errors in status data will show up as errors in the network topology, which will also cause the state estimation errors. In practice, the tree search algorithm [111] [112] to detect the erroneous data for the network topology processor is widely implemented. The authors in [113] applied the sequential search method through the network graph. In [114] [115] [116], the authors propose the methods via using state estimation results for the topology error detection. Although the techniques are improved, the computational complexity of most existing approaches for the determination of topology error is high in practice.

The EMS needs to efficiently combat the topological error in a realtime manner to timely prevent further damage to the entire network [117,118]. In other words, the network topology should be determined as quickly as possible so that one can detect/identify the erroneous data to maintain a reliable database for the state estimator; otherwise, erroneous data can result in topology errors that invalidate the whole real-time modeling process on smart grid networks. This type of estimation problem can be interested via applying the quickest detection (QD) concept - SPRT from [90]. SPRT aims to determine a change of the observed statistics/distributions as quickly as possible based on

online observations, the user-defined decision rules, and the requirement of detection accuracy. The decision rules need to be properly designed to optimize the tradeoff between the stopping time and decision accuracy.

2.3.3 Resource Management

The smart grid (SG) technology has improved robustness and efficiency of traditional power grid networks by exploiting the modern technological advances in sensing, measurement, and control devices with two-way communications. The information exchange among users, operators, and control devices significantly improves the efficiency in production, transmission, and distribution in the power system A key paradigm of smarter power generation is the distributive generation (DG), which can take the advantage of distributive renewable energy resource (DRER) system. The concept of DG and DRER further promotes the development of new grid concept,"the micro-grid," which has the great potential for the future smart grid . For example, the micro grid (a localized group of DRER generators and loads) can independently handle demand and supply for this specified region without the help of the macro grid; meanwhile, it can avoid the disturbance in the macro grids so that power reliability and quality can be improved. By intentionally isolating the grid, the study in has found that it provides a higher local power stability via DRERs than that of the power system as a whole.

One important challenge for utility companies is how to deliver energy to serve the end-users better [24]. On the end-user side, users like to determine one among many profiles that produces the most reliably and efficiently [119], as its electricity source. A profile of DRER system can include the combination of geothermal heat, sunlight, wind, tides, and hydro energy, which are often small-scale power generators (i.e., they provide alternative resources and enhance the power quality and stability for the macro grid). To find the best profile that can give most profit, there are two main problems. First, the completive environment in energy market, and therefore, the utility companies unlikely publish such sensitive data; otherwise, all consumers will use the company who has the best energy profile and the other companies get nothing in business. We face another problem that

the power generation pattern of the renewable energy source is often unknown, stochastic, and hard to predicted. Therefore, it is difficult for users to make a decision on which profile they should be connected with for returning most profit.

The question can be formulated as a multiarmed bandit (MAB) problem, where each arm is for one profile in our case. The MAB problem considers the tradeoff between the exploitation (to utilize the existing profile) and exploration (to test the new profile). In [120], reinforcement learning policies are developed facing the exploration versus exploitation dilemma. An elegant algorithm is proposed in [121] to make exploitation-versus-exploration decisions based on uncertain information provided by a random process. In the stochastic multi-armed bandit problem, a modification of the upper confident bound algorithm is considered for an improved bound on the regret with respect to the optimal reward [122] [123]. Beyond the MAB problem, an user also faces to make a decision to select a DRER as quickly as possible based on uncertain information in realtime. This type of problem is known as the quickest detection (QD) problem [90]. The classic cumulative sum test in [124] becomes one of the powerful tools for non-Bayesian framework in quickest detection. It aims to determine a change of the observed statistics as quickly as possible based on online observations, the user-defined decision rules, and the requirement of detection accuracy. The decision rules should be properly designed to optimize the tradeoff between the stopping time and decision accuracy.

Chapter 3

Real-time Detection of False Data Injection in Smart Grid Networks: An Adaptive CUSUM Method and Analysis

In this chapter, a counter-measurement strategy of the false data injection attack is considered in the form of adversary detection. The problem formulation of detecting the false data injection is based on the bad data detection (BDD) on smart grid state estimation. The proposed scheme is able to determine the possible existence of adversary as quickly as possible without violating the given constraints such as a certain level of detection accuracy such as false alarm. In [125], some preliminary results are studied. The highlights of our contributions are follows: chapter

- We have developed a novel defense strategy for smart grid via an online statistical analysis using a sequence of data meanwhile controlling the detection delay and error probability under the desired levels. The application of the proposed algorithm to smart grid is very natural due to its delay-sensitive property. The conventional state estimation [126,127] for bad data detection uses measurements to balance false alarm rate or missing detection rate, while our approach aims to minimize the detection delay under the error probability constraint. In addition, the conventional approach makes decisions via only snapshot measurements, but the proposed sequential framework would lead to more reliable decisions over time.

- The proposed algorithm is able to detect the presence of false data attack in that the probability density function of the post-change is unknown due to the unknown parameters. Unlikely, the classical CUSUM test expectedly has the perfect knowledge of the likelihood functions. To solving the unknown, instead of using the existing generalized likelihood ratio test (GLRT) approach, which has high complexity, the chapter proposes a new low complexity approach with shorter decision delay and more accurate decision that is asymptotically equivalent to the GLRT test.

- An analytical model for the proposed algorithm is developed, which provides the theoreti-

cal guidance for quantitative performance analysis. With the analytical model, it gives the insight on system parameters configuration for the on-line detection of false data injection attack. System parameters can also be computed for guaranteed performance in terms of three fundamental performance metrics: the false alarm rate, average detection delay, and missed detection ratio under a detection delay constraint. In other words, our analytical model can guide us to configure a detection system based on some detection performance requirement.

- The performance of the proposed algorithm is evaluated by both mathematic analysis and numerical simulation. Notes that simulations are conducted under MATPOWER 4.0 package [128] for different IEEE test systems to ensure the experiment accuracy and proficiency.

The remainder of this chapter is organized as follows. The system model is given in Section 3.1. The proposed scheme, *adaptive CUSUM algorithm*, is given and analyzed mathematically in Section 3.2. The analytical model based on the Markov chain with the theoretical performance analysis under three fundamental performance metrics is described in Section 3.3. The performance analysis of analytical and numerical simulation results are provided in Section 3.4. The conclusion is given in Section 3.5. Table 3.1 includes some important notations throughout this chapter.

3.1 System Model and Problem Formulation

Figure 3.1 illustrates the IEEE 4-buses test system with 2 generators; each bus has its corresponding voltage (V_q) and phase angle (θ_q); the control center sends the power measurement data (z_{qr}), and then the state estimator defines the states of power system that can be used in the different functions such as the automatic generation control (AGC), optimal power flow (OPF), or EMS. The operator makes the final decision for controlling generators and managing load. Notes that, in this figure, "G" represents as the generators, the black dot represents available active power-flow measurements, and the triangular on the bus represents the load of the region or the city

As an essential role in the power system, the state estimator uses the steady system model and evaluate measurement data to estimate the system state (i.e., the voltages at all buses over the

time) [129]. Speaking in general, the realistic power system state estimation with a total of B active buses can be described as

$$\mathbf{Z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}, \tag{3.1}$$

where $\mathbf{Z}$ denote the measurement data, $\mathbf{x}$ represents the unknown state including the voltage level V_q and phase angel θ_q of each bus $q \in B$, and $\mathbf{e}$ is the Gaussian measurement noise with zero mean and covariance matrix $\mathbf{\Sigma}_e$. Noticing that a nonlinear $\mathbf{h}(\mathbf{x})$ is determined by the network topology, the real power flow from bus q to bus r can be expressed as

$$\begin{aligned} M_{qr} = &\quad V_q^2(g_{sq} + g_{qr}) - V_q V_r(g_{qr}\cos\theta_{qr} + b_{qr}\sin\theta_{qr}), \\ \tilde{M}_{qr} = &\quad -V_q^2(b_{sq} + b_{qr}) - V_q V_r(g_{qr}\cos\theta_{qr} - b_{qr}\sin\theta_{qr}), \end{aligned} \tag{3.2}$$

where the admittance of the series branch between buses q and r is $(g_{qr} + jb_{qr})$, and the admittance of the shunt branch at bus q is $(g_{sq} + jb_{sq})$. The formulations of real and reactive power injection can be constructed in the similar way such as described in (3.2).

For simplicity, the linear state estimation model is applied in this chapter. Notice that all shunt elements, bus, branch and reactive power flow are neglected and the bus voltage magnitude is known [126]. The power flow and power injection can be linearized and described as

$$\begin{aligned} M_{qr} = &\quad \tfrac{\theta_{qr}}{X_{qr}}, \\ M_q = &\quad \textstyle\sum_{r\in B_q} M_{qr}, \end{aligned} \tag{3.3}$$

where M_q is denoted to the power injection, B_q is the set of bus numbers that are directly connected to bus q, X_{qr} is the reactance between bus q and bus r. Further, we can simplify[1] (3.1) to

$$\mathbf{Z}_n = \mathbf{H}\mathbf{x} + \mathbf{e}_n, \tag{3.4}$$

where $\mathbf{H}$ is the constant Jacobian matrix, $\mathbf{Z}_n = [Z_{n,1}, \cdots, Z_{n,m}]^T$ with m measurements at the observation index $n \in 1, 2, 3, \cdots$, and $\mathbf{x} = [\theta_2, \ldots, \theta_B]^T$. Notices that phase angle for bus 0

[1]The reason we utilize the linear model is that, practically, for security constraint unit commitment (SCUC) and market operation purposes, most of control centers use linear state estimation because of two reasons. First, the phase differences are relatively small so that linear model can be employed. Second, due to the complexity of computing AC model, the linear model is used for the time limitations in the power system operation [130].

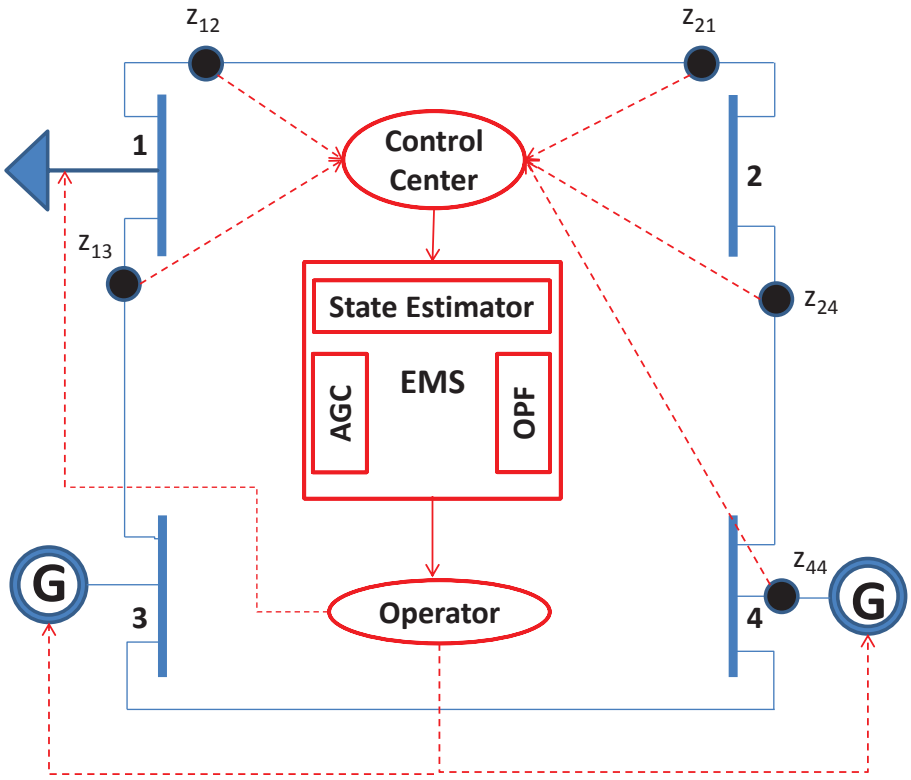


Figure 3.1 *An illustration of the 4 bus power network, control center, a few main functions (AGC, OPF, EMS), and the operator.*

θ_0 is assumed known as reference angle, and the size of $\mathbf{Z}_n$ is normally greater than that of $\mathbf{x}$. In [126, 131], One objective of (4.4) is to determine $\hat{\mathbf{x}}$ so that minimize

$$(\mathbf{Z}_n - \mathbf{H}\hat{\mathbf{x}})^T \mathbf{\Sigma}_e^{-1} (\mathbf{Z}_n - \mathbf{H}\hat{\mathbf{x}}). \tag{3.5}$$

By applying the weighted least square, the estimated system state $\hat{\mathbf{x}}$ is

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{Z}_n. \tag{3.6}$$

For BDD system, we compare the power-flow measurements $\mathbf{Z}_n$ with the estimated active power-flow $\hat{\mathbf{Z}}_n$ by the phase angle estimate $\hat{\mathbf{x}}$. $\hat{\mathbf{Z}}_n$ can be written as

$$\hat{\mathbf{Z}}_n = \mathbf{H}\hat{\mathbf{x}} = \mathbf{H}(\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{Z}_n = \Im \mathbf{Z}_n, \tag{3.7}$$

where $\Im$ is known as the *hat matrix*. Define the residue vector as

$$\mathbf{R}_n = \mathbf{Z}_n - \hat{\mathbf{Z}}_n. \tag{3.8}$$

The expected value and the covariance of residual $\mathbf{R}_n$ are

$$E(\mathbf{R}_n) = \mathbf{0}, \tag{3.9}$$

and

$$\mathbf{\Sigma_R} = [\mathbf{I} - \mathbf{H}(\mathbf{H}^T\mathbf{\Sigma}_e^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{\Sigma}_e^{-1}]\mathbf{\Sigma}_e, \tag{3.10}$$

respectively. The system can perform BDD by analyzing $\mathbf{R}_n$ [126].

In summary, the conventional state estimation for false data injection detection uses only snapshot measurements, and therefore, we like to apply the online quickest detection technique using a sequence of measurements that would lead to more reliable decisions.

3.2 CUSUM-based Defending Mechanism

In this chapter, we propose an adaptive CUSUM algorithm on quickest change detection for defending false data attack in smart grid state estimation. The proposed scheme evaluates the measurements before the potential bad data is removed by BDD. The detection formulations such as presented in [90] [124] is no longer useful, because the unknown exists in the post-change distribution and may change over the detection process. Our main motivation is to derive a detection model with considering the existence of the unknown, and then develop an analytical model that can guide us configure the detection system for performance guarantee based on the fundamental detection requirements. The proposed scheme does not require the Maximum Likelihood (ML) estimate of the unknown, thereby making the computation process much simpler.

Under false data injection, the false data $\mathbf{b}_n$ is maliciously injected into the power flow measurement vector as

$$\mathbf{Z}_n = \mathbf{H}\mathbf{x} + \mathbf{b}_n + \mathbf{e}_n. \tag{3.11}$$

Residual vector $\mathbf{R}_n$ can be well approximated by a Gaussian random variable because of Gaussian thermal measurement noise $\mathbf{e}_n$ [104] [132]. When there is no attack, the residual vector

$\mathbf{R}_n$ follows Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma_R})$. Under attack, $\mathbf{R}_n$ follows $\mathcal{N}(\mathbf{a}_n, \mathbf{\Sigma_R})$, where

$$\mathbf{a}_n = \mathbf{K}\mathbf{b}_n, \tag{3.12}$$

where $\mathbf{K} = (\mathbf{I} - \Im)$. Notice that $\mathbf{a}_n = [a_{n,1}, a_{n,2}, \cdots, a_{n,m}]^T, \in \mathbb{R}^m$ is the unknown (i.e., no one knows the adversary's statistical model, attack patterns, or mathematical distributions. This issue will be addressed later in this section). Then, we have the binary hypothesis as

$$\begin{cases} \mathcal{H}_0 : & \mathbf{R}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_R}), \\ \mathcal{H}_1 : & \mathbf{R}_n \sim \mathcal{N}(\mathbf{a}_n, \mathbf{\Sigma_R}), \end{cases} \tag{3.13}$$

and assumes the false data injection becomes active at random-time moment τ. In other words, a change of the distribution from $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma_R})$ to $\mathcal{N}(\mathbf{a}_n, \mathbf{\Sigma_R})$ at τ. Notes that we process the measurement data before a BDD removes the potential residual.

We denote T_h as the stopping time for declaring the best arm under current observation. τ is a change time. In other words, it is the switch point from one distribution belongs to the normal state to another distribution under the attack. Based on the Lorden's formulation [90], we minimize the worst case of detection delay, which can be described as

$$T_D = \inf_{T_h \in \mathcal{T}} \sup \operatorname{esssup} E_\tau[(T_h - \tau + 1)^+ | \mathcal{F}_{\tau-1}], \tag{3.14}$$

which $\tau > 1$, $\mathcal{F}_\tau$ denote the smallest α-field with respect to the observations, $\mathcal{T}$ is the set of all stopping time with respect to $\mathcal{F}_\tau$, and E_τ is the expectation that the change time is τ. However, most CUSUM-based models assume the perfect knowledge of the likelihood functions [124]. In the scenario of intrusion detection in smart grid state estimation, the variable from the $\mathcal{H}_1$ distribution cannot be completely defined because of the unknown. We also face that multiple measurements are correlated each together in a single online observation. Thus, we need to employ the technique to solve the issues for real-time detection of false data injection in smart grid networks.

The proposed quickest detection algorithm, *adaptive CUSUM algorithm*, is recursive in nature, and each recursion comprises two interleaved steps: i) unknown variable solver based on Rao Test and ii) multi-thread CUSUM test. The proposed CUSUM algorithm considers and cooperates

the likelihood ratio term of m active power measurements in order to determine the stopping time T_h, which can be described as follows

$$T_h = \inf\{n \geq 1 | S_n > h\}, \tag{3.15}$$

where the detection threshold h is a function of false alarm rate (FAR), and its value is determined numerically. At n, the cumulative statistic S_n can be solved recursively and described as

$$S_n = \max\left[0, S_{n-1} + L_n\right], \tag{3.16}$$

where the S_n returns to zero for statistical accuracy if its value is negative, $S_0 = 0$ initially, and

$$L_n = \log \frac{f_1(\mathbf{R}_n)}{f_0(\mathbf{R}_n)}, \tag{3.17}$$

being the likelihood ratio function based on measurements at n denoted as the observation vector $\mathbf{R}_n$ $(R_{n,l}, l \in 1, 2, \cdots, m)$. In (3.17), $f_1(\mathbf{R}_n)$ is the distribution associated with the hypothesis $\mathcal{H}_1$ with false data injection, and $f_0(\mathbf{R}_n)$ is the distribution associated with the hypothesis $\mathcal{H}_0$ in the normal state. Therefore, the control center is able to declare the alarm when the accumulation crosses a certain threshold h, the cumulative process is terminated, and ARL is equivalent to T_h.

Due to the unknown $\mathbf{a}_n$ in (3.12), the authors in [133] propose to implement the generalized likelihood ratio test (GLRT) in the Page's CUSUM algorithm with the unknown. The idea is to apply likelihood ratio test (LRT) by replacing the unknown with the ML estimation. The GLRT approach is asymptotically minimax, and can be written as

$$S_n = \max_{1 \leq n \leq T_h} \max_{a_n} \sum_{i=n}^{T_h} \log \frac{f_1(\mathbf{R}_i | \mathbf{a}_i)}{f_0(\mathbf{R}_i)}. \tag{3.18}$$

In other words, we minimize the effect of the unknown while considering the worst case situation (i.e., the second maximization in (3.18)). Thus, by applying GLRT in the CUSUM algorithm, we can ensure a certain level of detection accuracy for quickest detection, while minimizing the potential effect from the unknown in the system. However, the recursive expression of (3.18) for the CUSUM test is no longer available as shown in (3.16). It is because GLRT needs to compute

every unknown element of $\mathbf{a}_n$ by estimating up to the current observation n. In other words, GLRT approach requires storing the estimated data and ML-estimating the unknown at every point. Thus, in practice, the GLRT is too difficult from the view points of hardware and software implementation. Moreover, the work in [134] states that Rao test might be more robust than the GLRT in the presence of real operating situations and require a smaller computational complexity than the GLRT for their implementation. In [135], it shows that performance of Rao test based detectors works better than that of GLRT in terms of the parameter estimation and handling training-free scenario.

For the multi-thread CUSUM algorithm with the unknown, the desired approach is to solve recursively, avoiding ML estimation. Thus, we consider the Rao test [136], which is the asymptotically equivalent test model of GLRT. The derivation of Rao test is similar to the locally most powerful (LMP) test but much simpler. The Rao test has the straight-forward calculation by taking derivative of L_n with respect to the unknown evaluated around the region of interests. In our case, we analyze the case where the region is around zero due to the hypothesis $\mathcal{H}_0$ has zero mean. The Rao test also doesn't involve the complex computation like the ML estimation does. The statistics in [136] of the Rao test for detection can be modified and rewritten at observation n

$$\mathcal{K}(\mathbf{R}_n) = \frac{\partial L_n}{\partial \mathbf{a}_n}\bigg|^T_{\mathbf{a}_n=\mathbf{0}} \left[\mathbf{J}^{-1}(\mathbf{a}_n)\bigg|_{\mathbf{a}_n=\mathbf{0}}\right] \frac{\partial L_n}{\partial \mathbf{a}_n}\bigg|_{\mathbf{a}_n=\mathbf{0}}, \tag{3.19}$$

where $\mathbf{J}$ is the Fisher information matrix [137]. By inspecting (5.16) and evaluating (3.12)-(3.13), we notice that the computation of the inverse Fisher information matrix can be simplified and equivalent to the covariance of residual.

Based on (3.13), we can write the binary hypothesis $\{\mathcal{H}_0, \mathcal{H}_1\}$ by expanding the multivariate normal distributions. Next, we apply (5.16) to (3.17) by taking its derivative with respect to $\mathbf{a}_n$ evaluated at $\mathbf{a}_n = \mathbf{0}$. Finally, by recursion, the multi-thread CUSUM-based statistic can be described

$$S_n = \max\left\{0, S_{n-1} + \mathcal{I}_n\right\}, \tag{3.20}$$

where $\mathcal{I}_n = \left[(\mathbf{R}_n^T\boldsymbol{\Sigma}_{\mathbf{R}}^{-1})^T + \boldsymbol{\Sigma}_{\mathbf{R}}^{-1}\mathbf{R}_n\right]^T\boldsymbol{\Sigma}_{\mathbf{R}}\left[(\mathbf{R}_n^T\boldsymbol{\Sigma}_{\mathbf{R}}^{-1})^T + \boldsymbol{\Sigma}_{\mathbf{R}}^{-1}\mathbf{R}_n\right]$. Notices that the cumulative statistic is now independent from the unknown variable, and (4.31) becomes a scalar quantity

once it is computed. In summary, the control center observes actual power-flow measurements and generates vector $\mathbf{R}_n$ of residual for m active power measurements at observation index n. We construct the proposed scheme, *adaptive CUSUM algorithm*, with two interleaved steps: the unknown variable solver and multi-thread CUSUM test. The control center is able to tackle with false data injection by examining adaptive CUSUM statistic in (4.31) against the threshold h at n. Remember that h is the detection threshold to be set according to the desired value of the error probability. The alarm rises when the CUSUM statistic S_n exceeds the threshold h. The framework of the adaptive CUSUM algorithm of the proposed scheme is shown in Algorithm 4.1.

Algorithm 3.1 *Adaptive CUSUM algorithm*

$n \leftarrow (1, 2, 3 \cdots)$
$\mathbf{R}_n \leftarrow$ compute the difference between $\hat{\mathbf{Z}}$ and $\mathbf{Z}$.
repeat
  Update of: $n \leftarrow n + 1$
  continues the observation
  Unknown Solver based on Rao Test:
  eliminate $\mathbf{a}_n$ by taking derivative of L_n with respect $\mathbf{a}_n$ evaluated at $\mathbf{0}$
  Multi-thread CUSUM test:
  compute recursively S_n for all m measurements at current n as shown in (4.31)
until $T_h = \inf\{n \geq 1 | S_n > h\}$ is determined
Terminate the adaptive CUSUM process
Report the determined hypothesis and ARL

3.3 Markov Chain based Analytical Model

In this section, we develop the Markov chain based analytical model to systematically examine the proposed scheme for the false data injection attack. The Markov chain based analytical model produces the quantitative performance analysis and theoretical guidance on the proposed scheme's parameter configuration for performance guarantee under three fundamental performance metrics: the expectation of false-alarm rate, the expectation of missing-detection rate, and the expectation of detection delay.

3.3.1 Analysis Model

For analysis purpose, we discretize $\mathbb{R}^+ \bigcup 0$ into the finite sets $\{U_1, \cdots, U_{F-1}, U_F\}$, where $U_1 = 0$, and U_F is the set whose value is greater than or equal to h. In other words, F is the total

number of transition from 0 to the state that has the value greater than or equal to h. There are several approaches for discretization [138] [139]. In this chapter, we employ uniform sampling for simplicity[2]. Moreover, from (4.31), we know that the sequence exhibits the Markov property, which the current state $j = S_n$ at observation n only depends on the previous state $i = S_{n-1}$ at $n-1$, but not on the past history [140].

The transition probabilities of the Markov chain for the proposed scheme from state i at $(n-1)$ to state j at n can be described as

$$\begin{aligned} P_{ij} &= P(S_n = j | S_{n-1} = i), \text{ under } \mathcal{H}_0, \\ \hat{P}_{ij} &= P(S_n = j | S_{n-1} = i), \text{ under } \mathcal{H}_1. \end{aligned} \tag{3.21}$$

Note that The Markov chain based analytical model for the proposed scheme involves two different transition probabilities matrix (TPM): one is under the normal state environment; and the other one is under the false data attack. The normal TPM can help determining the initial state as well as false alarm rate. With the initial states, the average detection delay and detection delay can be analyzed by using the TPM under attack. We can calculate TPMs: $\mathbf{P}$ and $\hat{\mathbf{P}}$ with the size of $(F+1)\times(F+1)$, under the hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ according to $f_0(\mathbf{R}_n)$ and $f_1(\mathbf{R}_n)$, respectively. Here, we assume that the attacker's strategy is stationary for the illustration purpose[3].

The initial steady state probability of the Markov chain, which the process starts from an unalarmed state, can be determined

$$\pi_j^0 = \frac{\pi_j}{\sum_{i=0}^{F-1} \pi_j}, \quad \text{given} \quad j \in \{0, U_1, \cdots, U_{F-1}\}, \tag{3.22}$$

and the steady-state probability can be determined as

$$\pi_j = \sum_{i=0}^{F} P_{ij} \pi_i, \tag{3.23}$$

[2]Other discretization methods can be employed like the μ-law or A-law in PCM.

[3]Notices that, for the non-stationary case, $\hat{P}_{ij}$ cannot be determined in advance because $\mathbf{R}_n$ is Gaussian distributed with time-varying mean (i.e., $\mathbf{R}_n$ depends on $\mathbf{b}_n$ under $\mathcal{H}_1$). To address this issue, $\mathbf{b}_n$ from attackers can be bounded in a range. For instance, $\mathbf{b}_n$ follows the uniform distribution in $[\mathbf{b}_n^{\min}, \mathbf{b}_n^{\max}]$, and then $\hat{P}_{ij}$ can be estimated using the conditional probability technique such that $\hat{P}_{ij} = E_{\mathbf{b}_n}\left[(P[S_n = j-1|\mathbf{b}_n)\right]$. However, the conditional probability calculation is usually not tractable due to complexity. Here, we like to discuss with a simplified situation for illustration purpose.

where $j \in \{0, U_1, \cdots, U_F\}$ and $\sum_{j=0}^{F} \pi_j = 1$.

Next, based on the Markov chain model, we study the theoretical performance analysis of detection delay, false alarm rate and missed detection ratio expectations, respectively, in the following subsections.

3.3.2 Expectation of Detection Delay

To determine the expectation ($E_{\hat{\mathbf{P}}}[T_D]$) of detection delay, we utilize the weighted average of the expected number of transitions from every initial state ($\pi_0^0, \pi_1^0, \cdots, \pi_{F-2}^0, \pi_{F-1}^0$) to state U_F based on $\hat{\mathbf{P}}$. We set Ω_{gF}, $g \in \{0, U_1, \cdots, U_{F-1}\}$ as the expected number of transitions for state g to state U_F. Following the derivation from [140], the numerical value of Ω_{iF} can be determined

$$\Omega_{iF} = 1 + \sum_{g \neq F} \hat{P}_{ig} \Omega_{gF}, \tag{3.24}$$

where the transition probability $\hat{P}_{ig} \in \hat{\mathbf{P}}$ is from state i to state g. The expectation of detection delay can be obtained from the results of (4.34) and (3.24) as

$$E_{\hat{\mathbf{P}}}[T_D] = \sum_{i=0}^{F-1} \pi_i^0 \Omega_{iF}. \tag{3.25}$$

3.3.3 Expectation of False Alarm Rate

The expectation ($E_{\mathbf{P}}[\text{FAR}]$) of false alarm rate is the probability that the proposed CUSUM statistic S_n reaches to the state U_F when there is no attacker in the network. As described in [140], $E_{\mathbf{P}}[\text{FAR}]$ is equivalent to the probability that S_n stays at state U_F (i.e., exceeding threshold h) under hypothesis $\mathcal{H}_0$.

According to [140], it states the transition probability matrix $\mathbf{P}$ always has a special eigenvector with only one eigenvalue $\lambda = 1$ and the rest is zero. Thus, we can obtain the solution by

re-elaborating the equation (4.35) into the matrix form

$$\begin{bmatrix} P_{00}-1 & P_{01} & \cdots & P_{0F} \\ P_{10} & P_{11}-1 & \cdots & P_{0F} \\ \vdots & \vdots & \ddots & \vdots \\ P_{F0} & P_{F1} & \cdots & P_{FF}-1 \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \pi_F \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{3.26}$$

Then, the expectation of false alarm rate can be determined by

$$E_{\mathbf{P}}[\text{FAR}] = \pi_F. \tag{3.27}$$

3.3.4 Expectation of Missed Detection Ratio

We define the missing detection probability as the probability that the detection delay is greater than or equal to a detection delay constraint C. The expectation ($E_{\hat{\mathbf{P}}}[\text{MDR}]$) of the missing detection probability is, starting from the initial state, the summation of probabilities that S_n stays at a state other than state U_F at time C. Let $p_i(s)$ denote the probability of the state variable at time s and at state i. We set the initial condition for the transition probabilities

$$p_i(0) = \pi_i^0, \tag{3.28}$$

where $i \in \{0, U_1, \cdots, U_{F-1}\}$ and $p_F(0) = 0$. By the iteration, at each s, the state probability vector is updated by the previous state probability vector in a matrix form as

$$\begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \\ p_{F-1}(s) \\ p_F(s) \end{bmatrix}^T = \begin{bmatrix} p_0(s-1) \\ p_1(s-1) \\ \vdots \\ p_{F-1}(s-1) \\ p_F(s-1) \end{bmatrix}^T \hat{\mathbf{P}}, \tag{3.29}$$

and

$$p_F(s) = 0, \quad s \in \{0, C-1\}. \tag{3.30}$$

Here the $p_F(s)$ at every s of state U_F is reset to zero for the next iteration since we only concern the missing detection case only. The expectation of missed detection ratio under the given delay

constraint C can be obtained as

$$E_{\hat{\mathbf{P}}}[\text{MDR}] = \sum_{i=0}^{F-1} p_i(C). \tag{3.31}$$

3.4 Performance Analysis

In this section, we present the analytical and numerical simulations to demonstrate the performance of the proposed scheme. This section is composed by two main sub-sections. The first sub-section demonstrates the performance of the proposed scheme from the simulated data. In other words, we heuristically configure the parameter and analyze the detection performance. The second sub-section involves both analytical and numerical results under the realistic power test systems by MAPOWER 4.0 package [128]. For the experiment setup of the second subsection, we first apply the analytical model to theoretically analyze the performance of the detection system for guiding the system parameter configuration. Then, we use the parameter from the theoretical analysis to confirm the accuracy of the analysis in the first half of the subsection, and then demonstrate the performance of the detection system in the second half of the subsection. Assumes that the adversary is able to inject the false power flow measurement at the random time and the sample rate is normalized[4].

3.4.1 Simulation Results with Simulated Data

Figure 3.2 illustrates the relation between the detection parameters (S_n, h) and performance metrics (FAR, T_D). Note that, in this figure, the x-axis is the observation index (n), and y-axis is the recursive CUSUM statistic (S_n); Case 1 with FAR of 1% corresponds to h_1, and Case 2 with FAR of 0.1% corresponds to h_2. The proposed algorithm signals the alarm and then terminates the process at $T_h = 7$ and 8, respectively. The number of measurements $m = 4$. On the detector side, the detector has no information about the adversary statistical model, distribution, or any unknown.

[4]Since the measured noise is white Gaussian (independent over time), the performance of the quickest detection is depended on the number of observations. In other words, the decision time is related to the sampling rate, and the decision time is equivalent to the number of observation divided by the sampling rate.

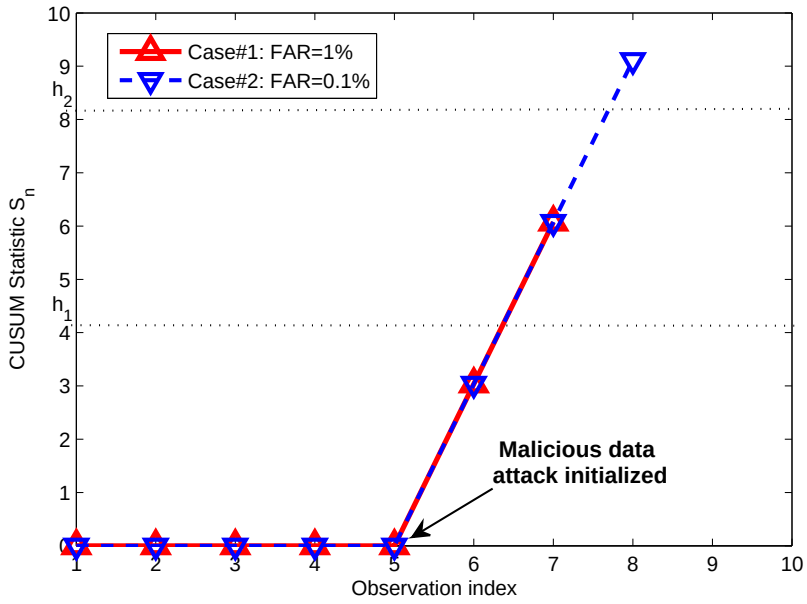


Figure 3.2 *The simulation of the adaptive CUSUM algorithm.*

The adversary manipulates and injects the false data into the system at the random time. As shown in Figure 3.2, we consider that Case 1 with FAR of 0.01 and Case 2 with FAR of 0.0001. The adversary becomes active and injects the false data at $n = 6$. In other words, a change distribution is at $\tau = 6$ from $\mathcal{N}(0, \mathbf{\Sigma_R})$ to $\mathcal{N}(\mathbf{a}_n, \mathbf{\Sigma_R})$, where $\mathbf{a}_n$ is unknown. For both cases, the curve of adaptive CUSUM statistic (S_n) shows the sudden increase right after a change of distributions. The proposed algorithm quickly responses the abnormal event by signaling an alarm when S_n passing the threshold. At the observation index 7, the threshold parameters h_1 and h_2 are corresponded to Case 1 and Case 2, respectively. As a result, h_1 is less than h_2, because of the different FARs. For the smaller FAR, the stricter constraint that causes increasing the threshold; the higher requirement for system to declare the decision. ARL (T_h) of the adaptive CUSUM algorithm is 7 and 8 at S_n of 6.07 (Case 1) with $h = 5.97$ and 9.11 (Case 2) with $h = 8.19$, respectively. ARL (T_D) of detection delay is 1 for cases 1 and 2 for the Case 2 in this simulation. The proposed algorithm is able to signal the alarm and terminates the process after the active false date attack.

Figure 3.3 shows the characteristics of the proposed algorithm by varying FAR for the accu-

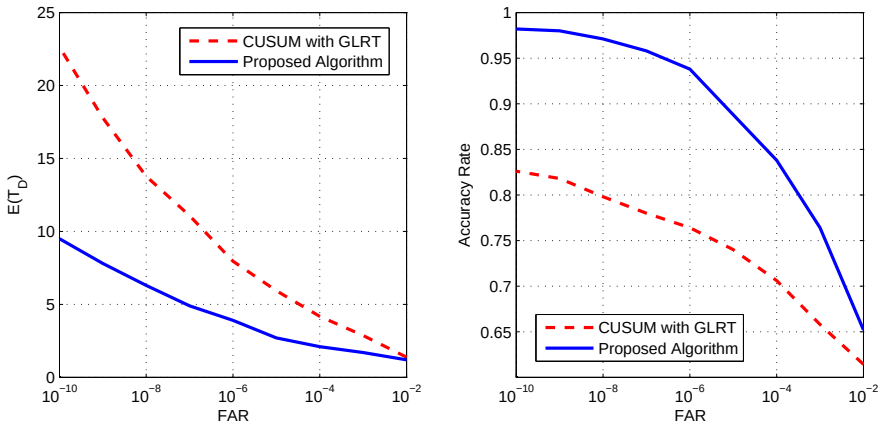


Figure 3.3 *The performance analysis of the adaptive CUSUM algorithm in comparison with CUSUM GLRT.*

racy rate and expected ($E[T_D]$) of detection delay in comparison to that of the CUSUM GLRT. We run 5000 realizations for the simulation. FAR is vary from 10^{-10} to 10^{-2}. The false data injection is begun at the 6th observation index. The accuracy rate in Figure 3.3(Right) represents the ratio of successful detection that the algorithm terminates the process and declares the existence of adversary after the 6th observation index (the actual attack index). As shown in the figure for both proposed scheme and the CUSUM GLRT, the stricter FAR is, the greater expected detection delay and higher detection accuracy we have. The expected detection delay of CUSUM GLRT seems to increase exponentially while that of proposed scheme steady raises as FAR decreases. $E[T_D]$ of the proposed scheme has the average 50% less than that of CUSUM GLRT. We also obtain the better accuracy rate as FAR decreases. By giving the sufficiently low FAR, the proposed scheme is able to reach the accuracy above 95% while CUSUM GLRT struggles it below 83%. Therefore, the proposed scheme outperforms the CUSUM GLRT in terms of shorter decision time and higher detection accuracy. The simulation result also shows the tradeoff between the detection delay, false alarm and accuracy rate. The smaller FAR causes higher delay but better accuracy, i.e., the system needs to spend more observations for making a decision.

3.4.2 Simulation Results with MATPOWER 4.0

3.4.2.1 Accuracy of the analytical model

In this section, the power flow data for all simulations are generated by MATPOWER 4.0 instead of random independent variables in the previous subsection. MATPOWER 4.0 is a Matlab simulation tool for solving power flow and optimal power flow problems. It provides realistic power flow data and test systems that uses widely in research-oriented study as well as in practice. We consider 4 popular IEEE test systems from the MATPOWER 4.0 package. Case 1 is the IEEE 4-bus test system, which has 2 generators for 4 measurements; Case 2 is the IEEE 57-bus test system, which has 7 generators for 80 measurements; Case 3 is the IEEE 118-bus test system, which has 54 generators for 186 measurements; and Case 4 is the IEEE 2383-bus test system, which has 326 generators for 2896 measurements. The analytical performance measures and the simulation results are compared under same setting and input data to examine. Hence, by using power flow data sets with 4 different study cases from MATPOWER 4.0, the performance indices (E[FAR], E[MDR], $E[T_D]$) comparisons between the analytical and simulation result can be conducted. With the parameter from the theoretical analysis, the performance indices are simulated so that we can properly configure the proposed algorithm for the guaranteed performance. Notice that both theoretical analysis and simulation are plotted together to confirm accuracy of analysis and demonstrate the performance.

Figure 3.4 gives us an insight of the relationship between the system parameters h and the detection delay $E[T_D]$ of the proposed scheme. The higher the threshold, the larger the delay. Also shown in Figure 3.4, both analytical and simulation results are matched closely in all IEEE 4-bus, 57-bus, 118-bus test systems. The maximum difference between the analysis and simulation is around 2% in the case of IEEE 2383-bus test system.

The numerically examination is presented for understanding the impact of the fundamental performance metric FAR on system parameters h of the proposed scheme. As shown in Figure 3.5, the analytical and simulation result are close. Notes that the logarithmic scale is used in the figure

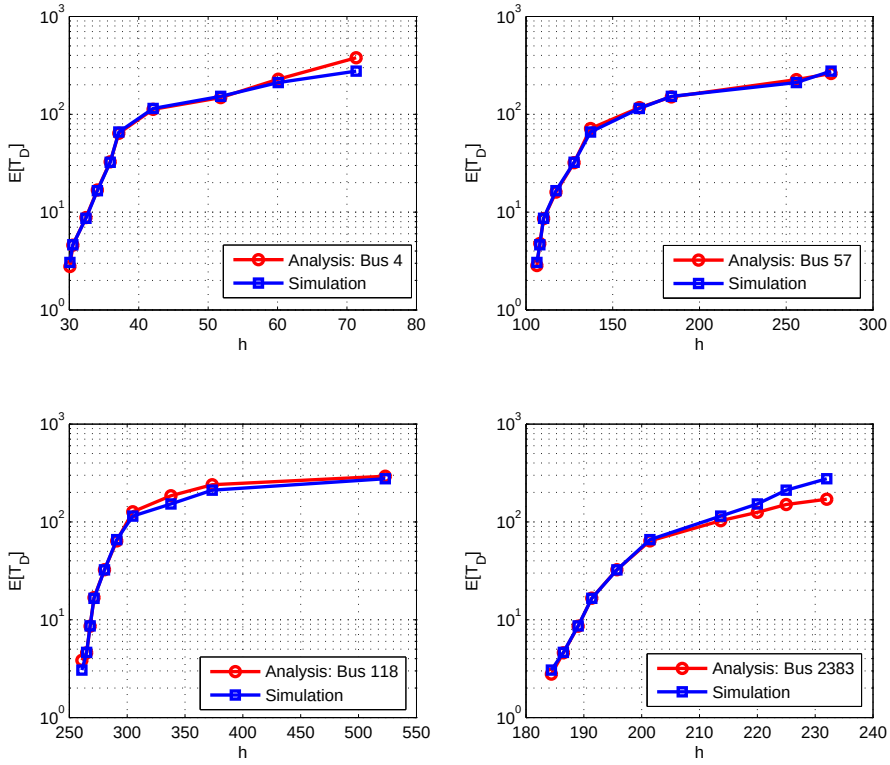


Figure 3.4 *The expectation of detection delay for different IEEE Bus test systems.*

for the vertical axis. In cases of IEEE 4-bus and 57-bus test systems, the difference percentage between the analysis and simulation is very small and near zero. However, as the number of bus increases (total number of active power flow measurement increases, too), the maximum difference percentage is about 8% in IEEE 2383-bus test system. More measurements can cause the larger variance when we try to calculate the covariance for computing $\mathbf{R}$. From the figure, we also can observe that a larger h yields a smaller false alarm rate as expected.

The analytical result of $E[\text{MDR}]$ is demonstrated under 2 scenarios of the delay constraints, in which $C = 7$ and $C = 18$. The result is shown in Figure 3.6 that helps us study the impact of the missed detection ratio on h of the proposed scheme. The logarithmic scale is used in the figure for the vertical axis. From the figure, the larger constraint C results smaller expectation of missed detection ratio as expected. In other words, the probability of detection rises if we allow to

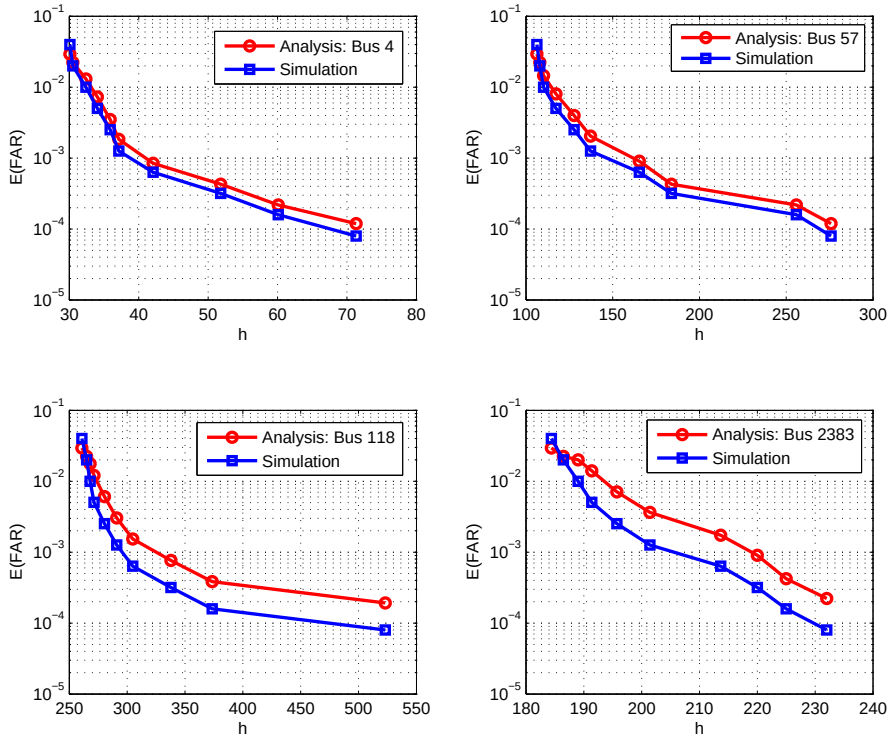


Figure 3.5 *The expectation of false alarm rate for different IEEE Bus test systems.*

increase the cost of longer delay. We also compute the mean of expected missed detection ratio as the base line, in comparison with the analytical results for 4 different IEEE test systems. The trend of analysis follows the base line closely. However, as the number of active power flow measurement increases, the gap between them becomes obvious, especially, in case of IEEE 2383-bus test system, the maximum difference percentage is obtained around 10%. More measurements can cause the larger variance when we try to calculate the covariance for computing $\mathbf{R}$. In addition, the smaller h is, the better the expectation of missed detection ratio that corresponds to the result of expected false alarm rate in Figure 3.5 as the tradeoff.

3.4.2.2 Detection with performance guarantee

From Figure 3.4-3.6, we demonstrate the performance metrics with different h. It also helps us to configure the system parameter h for guaranteed performance under three fundamental met-

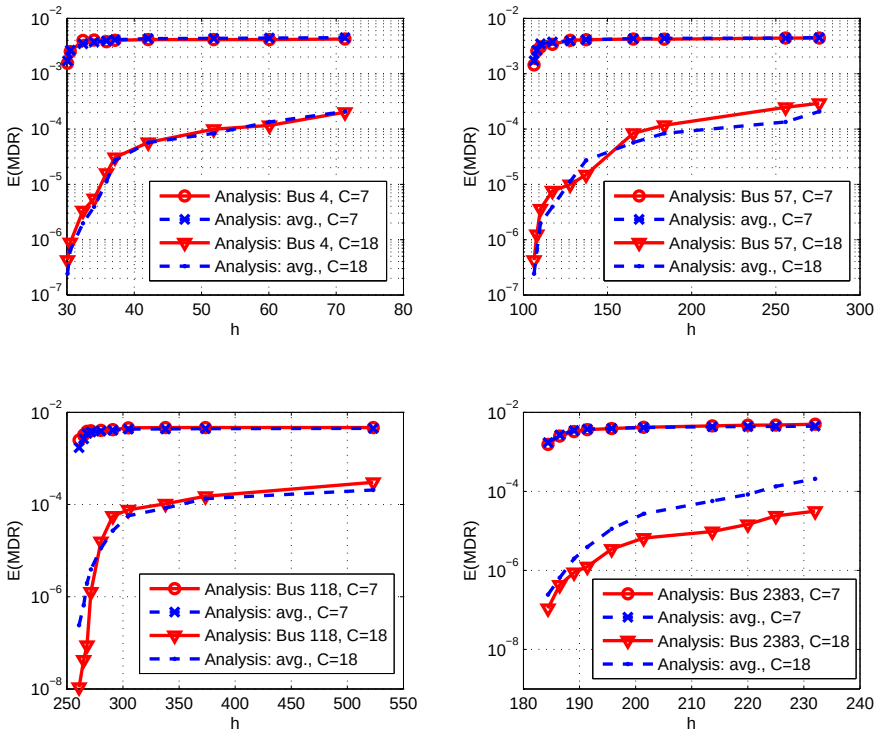


Figure 3.6 *The expectation of missed detection ratio for different IEEE Bus test systems.*

rics. For each different IEEE test system, we can select the proper configuration of h from the reasonable range to satisfy the desired performance constraints. For examples, the configuration of h is set to 135 for IEEE 57-bus test system; the analytical model of the proposed scheme shows that the expectation of the false alarm rate of 0.001, the expectation of detection delay of 20, and the expectation of missed detection ratio of 0.00005 under the delay constraint $C = 18$. In addition, if we wish to have a certain level of detection probability, we can compute the numerical value of detection probability from Figure 3.4; with its corresponding h, we can explicitly determine the cost of detection delay from Figure 3.4 and the tradeoff for the false alarm rate from Figure 3.5. The above analysis can be extended to other IEEE power systems in a similar way.

In Figure 3.7, we show the CUSUM statistics S_n over observation index n for the IEEE 4-bus, 57-bus, and 118-bus test systems. For the simulation setup, we considers that the false alarm rate of 0.01 is presented, and the active false data injection attack is initialized after the observation

index 15. Note that The x-axis is the observation (n), and y-axis is the recursive CUSUM statistic (S_n). The proposed algorithm signals the alarm and then terminates the process at $T_h = 24$, 37, and 45, respectively. For the simulation results, in the IEEE 4-bus test system, the system is alarmed the after 24 observations with the corresponding detection threshold of 34.51; the detection delay is 9. In the IEEE 57-bus test system, the system is alarmed the after 37 observations with the corresponding detection threshold of 133.52 and the detection delay of 22. In the IEEE 118-bus test system, the system is alarmed the after 45 observations with the corresponding detection threshold of 283.14; the detection delay in this test system is 30. As expected, the simulation also shows that the detector need more observations to make the decision, when the number of the power-flow measurements and buses increases. Notices that the numerical results of each IEEE test system in Figure 3.7 is corresponded to our analytical results, which are presented in Figure 3.4-3.6.

3.5 Conclusion

In this chapter, we propose the adaptive CUSUM algorithm for defending false data injection attack in smart grid networks. We successfully derive a detection model with considering the existence of the unknown, and then develop an analytical model that can guide us configure the detection system for performance guarantee based on the fundamental detection requirements. Our proposed scheme for smart grid state estimation composes two interleaved steps: i) introduces the unknown variable solver technique based on Rao Test, and ii) applies the multi-thread CUSUM algorithm for determining the possible existence of adversary as quickly as possible without violating the given constraints. Furthermore, we develop the Markov chain based analytical model to characterize the behavior of our proposed scheme. We can quantitatively study the system parameters to achieve the guaranteed detection performance in term of three fundamental metrics (E[FAR], E[MDR], and $E[T_D]$). The analytical and numerical simulation results have shown that the proposed scheme is efficient in terms of detection accuracy and minimum detection delay. Overall, the proposed scheme is able to achieve the important objectives of smart grid security in terms of real-time operation and security requirement.

Table 3.1 *The description of some important symbols and abbreviations*.

Notation	Description
EMS	energy management system
QD	quickest detection
CUSUM	cumulative sum
SHT	statistical hypothesis test
BDD	bad data detection
AGC	automatic generation control
OPF	optimal power flow
ARL	average run length
GLRT	generalized likelihood ratio test
TPM	transition probabilities matrix
B	number of buses in power system
V_q	voltage measurement at the bus q
θ_q	phase measurement at the bus q
X_{qr}	reactance between bus q and r
M_{qr}	power flow measurement from bus q to r
M_q	power injection measurement at bus q
n	observation index
m	total number of active power measurement
$\mathbf{Z}$	a vector of power measurement (M_qr,M_q, or both)
$\mathbf{x}$	the unknown state vector for state estimation
$\mathbf{e}$	a vector of measurement noise
$\mathbf{H}$	Jacobian matrix
T_D	detection delay for the proposed algorithm
T_h	the moment when detector raises the alarm
τ	the moment when adversary initializes the attack
S_n	CUSUM statistic at observation index n
$\mathbf{P}$	the transition probability matrix for Markov chain
π_i^0	the steady state probability that a detector starts from a normal state i
π_i	the steady state probability that a detector is at state i

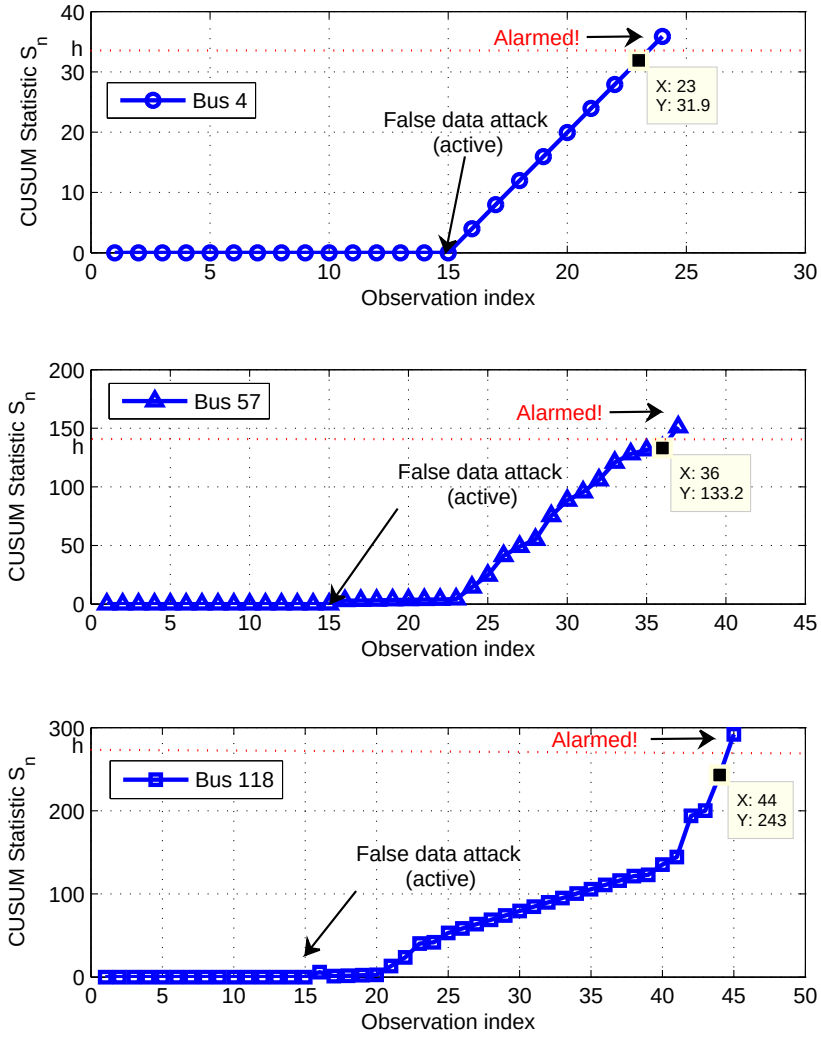


Figure 3.7 *The detection simulation of the adaptive CUSUM algorithm with MATPOWER 4.0 power-flow measurements for the IEEE 4-bus test system, IEEE 57-bus test system, and IEEE 118-bus test system.*

Chapter 4

Adaptive Quickest Estimation Algorithm for Smart Grid Network Topology Error

In this chapter, we employ the *adaptive estimation algorithm* to help detect and identify the topological error efficiently for smart grid estimation. The proposed scheme is able to *i)* execute the demodulation of mutually interfering streams of information that produces by all buses in the power network, and then *ii)* determine the current network topology as quickly as possible without violating the given constraints such as a certain level of estimation accuracy includes false alarm. With knowledge of the present network topology, one can explicitly determine and identify the topology error in an efficient way. Furthermore, we develop an analytical model for the proposed algorithm that provides theoretical guidance for quantitative performance analysis. With the analytical model, it provides us the insight on system parameters configuration for the on-line quickest estimation. System parameters can also be computed for guaranteed performance in terms of fundamental performance metrics: the false alarm rate and missed detection ratio under a detection delay constraint. In other words, our analytical model can guide us to configure a detection system based on some detection performance requirement. The Markov chain based analytical model for the proposed scheme involves two different transition probabilities matrix (TMP): one is under the normal state environment; and the other one is under the malicious data attack. The normal TMP can help determining the initial state as well as false alarm rate. With the initial states, the missed detection ratio can be analyzed by chapter using the TMP under attack. The performance of the proposed algorithm is evaluated by both mathematic analysis and numerical simulation. It is demonstrated in term of fundamental metrics (e.g., false alarm rate, missed detection ratio, average sample number).

The remainder of this chapter is organized as follows. The system model is given in *Section* 4.1. The proposed scheme is given and analyzed mathematically with the Markov chain based analytical model in *Section* 4.2. The performance analysis are provided in *Section* 4.3, and the conclusion is drawn in *Section* 4.4. The table below is included some important symbols throughout

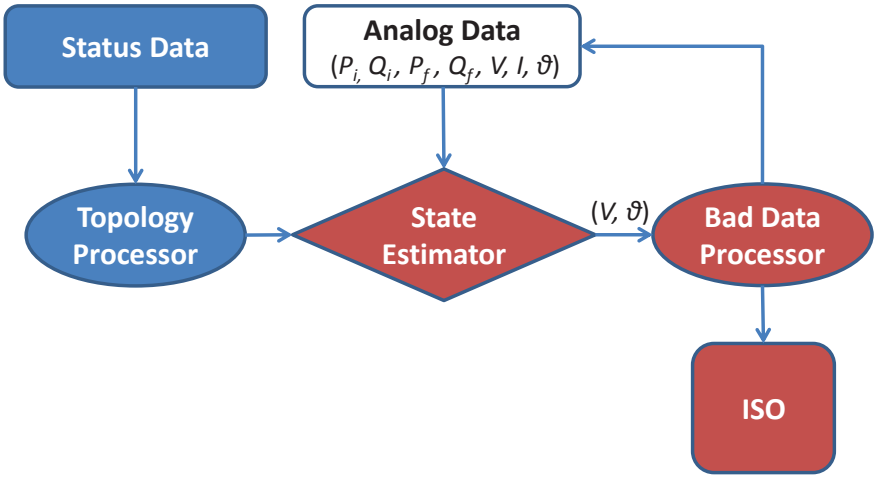


Figure 4.1 *The illustration of status data effect on the state estimation processor and further on the ISO.*

the entire chapter.

4.1 System Model

Before reviewing the classical formulation of the state estimation for the power network, we like to present the scope of smart grid state estimation as described in Figure 4.1. Note that, in this figure, analog data are included the measurement data of power-injection P_i, reactance-injection Q_i, power-flow P_f, reactance-flow Q_f, bus voltage V, bus current I, and bus phase ϑ. The network topology processor uses the telemetered data of breaker and switch status to determine the present network topology of the system. Then, the state estimator processes both sets of measurement data (status and analog data) globally and takes advantage of its redundancy to detect any data error. If data error exists, Bad data processor will notify the state estimator, and then the estimated state of power system will be discarded and re-estimated again. Otherwise, the independent system operator (ISO) makes the decision for controlling generators and managing load by applying the current state into the different functions such as the automatic generation control, optimal power flow, or energy management system. If the accident occurs, erroneous data can magnify the negative impact on these smart grid operations. Therefore, the efficient online error detection on measurement data is

essential.

While the errors for analog data will cause the state estimation errors, the effect of topology error shall be analyzed to understand how it can be used for determining topology error. First, we consider the state estimation problem as estimating phase angles θ_q, by observing the real-time measurements of active power-flow. The initial phase angle θ_0 is known as reference angle, and therefore only N angles have to be estimated. In other words, we have the total of N active angles (buses) in the system. The voltage level of each bus and reactance of each transmission line are assumed to be known.

4.1.1 Traditional Bad Data Detection

We first review the classical formulation of the state estimation using the normal equations. At the observation index $n \in \{1, 2, 3, \cdots\}$, the control center observes a vector $\mathbf{Z}_n$ of M actual power measurements. The non-linear equation relating the state vector $\mathbf{x}$ is

$$\mathbf{Z}_n = \mathbf{h}(\mathbf{x}) + \mathbf{e}_n, \tag{4.1}$$

where $\mathbf{Z}_n = [Z_{n,1}, \cdots, Z_{n,M}]^T$, and $\mathbf{e}_n$ is the Gaussian measurement noise with zero mean and covariance matrix $\mathbf{\Sigma}_e$. By applying the Gauss-Newton method [141], the unknown state $\mathbf{x}$ can be estimated iteratively as

$$\hat{\mathbf{x}}^{s+1} = \hat{\mathbf{x}}^s + (\mathbf{H}_s^T \mathbf{\Sigma}_e^{-1} \mathbf{H}_s)^{-1} \mathbf{H}_s^T \mathbf{\Sigma}_e^{-1} [\mathbf{Z}_n - \mathbf{h}(\hat{\mathbf{x}}_s)], \tag{4.2}$$

where the estimated system state $\hat{\mathbf{x}}^s \in \mathbb{R}^M$, s is iteration number, and $\mathbf{H}_s \in \mathbb{R}^{M \times N}$ is the Jacobian evaluated at $\hat{\mathbf{x}}^s$,

$$\mathbf{H}_s = \left. \frac{\partial \mathbf{h}(\hat{\mathbf{x}}^s)}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}^s}. \tag{4.3}$$

By decoupling the real and reactive part of measurements and state vectors, we will assume the phase differences between two buses in the power network are all small. Then, a linear approximation of (4.1) is accurate, and we obtain

$$\mathbf{Z}_n = \mathbf{H}\mathbf{x} + \mathbf{e}_n, \tag{4.4}$$

where $\mathbf{Z}_n$ is the set of power measurements[1] (i.e., the power-flow, power-injection, or voltage), $\mathbf{x}$ is the set of real part of $[\theta_1, \theta_2, \ldots, \theta_{N-1}, \theta_N]^T$ (bus angles), and $\mathbf{H} \in \mathbb{R}^M$ is the measurement Jacobian matrix with respect to phase angles. As a result, the estimated state $\hat{\mathbf{x}}$ is

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{Z}_n. \tag{4.5}$$

For the bad-data detection (BDD) system, we compare the power-flow measurements $\mathbf{Z}_n$ with the estimated active power-flow $\hat{\mathbf{Z}}_n$ by the phase angle estimate $\hat{\mathbf{x}}$. $\hat{\mathbf{Z}}_n$ can be written as

$$\hat{\mathbf{Z}}_n = \mathbf{H}\hat{\mathbf{x}} = \mathbf{H}(\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{Z}_n = \Im \mathbf{Z}_n, \tag{4.6}$$

where $\Im$ is known as the *hat matrix*. Define the residue vector as

$$\mathbf{R}_n = \mathbf{Z}_n - \hat{\mathbf{Z}}_n. \tag{4.7}$$

The expected value and the covariance of residual $\mathbf{R}_n$ are

$$E(\mathbf{R}_n) = \mathbf{0} \tag{4.8}$$

and

$$\mathbf{\Sigma}_{\mathbf{R}} = [\mathbf{I} - \mathbf{H}(\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1}] \mathbf{\Sigma}_e, \tag{4.9}$$

respectively. Generally, the bad $\mathbf{Z}_n$ such as corrupted data, missed data, or topology error can typically trigger a BDD alarm since the measurement residual $\mathbf{R}_n$ in 4.7 increases.

4.1.2 Topology Error

The effect of topology error can be presented in $\mathbf{H}$ and effected on $E(\mathbf{R}_n)$, which is no longer to be zero. As mentioned in the introduction section, the system can perform bad data detection (BDD) to determine the error by some comprehensive algorithms such as the sequential search algorithm, Chi-squares test with WLS state estimation, or the largest normalized residuals test. Nevertheless, some of known computational issues and shortcoming for these algorithms are

[1] Each measurement has different dimension, and the following study is based on power-flow estimation.

explicitly described in [141]. From (4.4), the measurement of state estimation under the network topology error can be modeled in the following manner [142]

$$\mathbf{H} = \mathbf{H}_e + \mathbf{B}, \tag{4.10}$$

where $\mathbf{H}$ is the true Jacobian measurement, $\mathbf{H}_e$ is the incorrect Jacobian measurement due to topology errors, and $\mathbf{B}$ is the Jacobian error matrix. Next, we substitute (4.10) into the linear approximation model (4.4) that yields

$$\mathbf{Z}_n = \mathbf{H}_e\mathbf{x} + \mathbf{B}\mathbf{x} + \mathbf{e}_n, \tag{4.11}$$

Next, the statistical characteristics of the new residual vector can be derived as below. The residual under error with its covariance is

$$\mathbf{R}_n = [\mathbf{I} - \mathbf{H}_e(\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}\mathbf{H}_e)^{-1}\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}][\mathbf{B}\mathbf{x} + \mathbf{e}_n]; \tag{4.12}$$

$$\mathbf{\Sigma}_R = [\mathbf{I} - \mathbf{H}_e(\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}\mathbf{H}_e)^{-1}\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}]\mathbf{\Sigma}_e. \tag{4.13}$$

and the expected value of residual $\mathbf{R}_n$ is

$$E(\mathbf{R}_n) = [\mathbf{I} - \mathbf{H}_e(\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}\mathbf{H}_e)^{-1}\mathbf{H}_e^T\mathbf{\Sigma}_e^{-1}][\mathbf{B}\mathbf{x}]. \tag{4.14}$$

In other words, the ISO receives an scrabble measurement data and need to identify/determine the erroneous status data that are generated by multiple buses in the power grid network. This type of problem can also be seen as the detection problem for dealing with the demodulation of the mutually interfering digital streams of information.

4.1.3 Problem Formulation

In fact, we can consider each bus in smart grid networks as a single transmitter through a common communication channel that carries the information such as the network topology. Let's first consider the expanded version of (4.4) that yields

$$\begin{pmatrix} Z_{n,1} \\ Z_{n,2} \\ \vdots \\ Z_{n,M} \end{pmatrix} = \begin{pmatrix} H_{1,1} & \cdots & H_{1,N} \\ H_{2,1} & \cdots & H_{2,N} \\ \vdots & \ddots & \vdots \\ H_{M,1} & \cdots & H_{M,N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} e_{n,1} \\ e_{n,2} \\ \vdots \\ e_{n,M} \end{pmatrix}. \tag{4.15}$$

In contrast, we can also describe the power measurement matrix of (4.15) in the formulation of sequential process; for the representation of entire power grid network, each power measurement information $\mathbf{Z}_n$ can be expressed as

$$\begin{array}{rcl} Z_{n,1} & = & \sum_{i=1}^{N} H_{1,i}\hat{x}_i + e_{n,1}, \\ Z_{n,2} & = & \sum_{i=1}^{N} H_{2,i}\hat{x}_i + e_{n,2}, \\ & \vdots & \\ Z_{n,r} & = & \sum_{i=1}^{N} H_{r,i}\hat{x}_i + e_{n,r}, \\ & \vdots & \\ Z_{n,M-1} & = & \sum_{i=1}^{N} H_{M-1,i}\hat{x}_i + e_{n,M-1}, \\ Z_{n,M} & = & \sum_{i=1}^{N} H_{M,i}\hat{x}_i + e_{n,M}, \end{array} \tag{4.16}$$

where the row $r \in \{1, 2, \cdots, M\}$, and the measurement noise is $e_{n,r}$ at r.

The idea is to enable several buses to send information simultaneously through a communication channel; on the control center, the noisy version of the superposition of signals is obtained from a crowd of the active buses; next, the operator works an efficient way to decode and estimate the useful information that is sent by the individual buses. We like to online estimate $\mathbf{H}$ with minimal delay in order to help one computes $\mathbf{B}$ and $\mathbf{H}_e$ further identify the topology error efficiently. Thus, the main task is that how to estimate the each element of $\mathbf{H}$ as quickly as possible with a certain level of error probability, i.e.,

$$\min \quad k, \quad \text{s.t. } P_r(\mathbf{H} \neq \hat{\mathbf{H}}) \leq \eta, \tag{4.17}$$

where k represents the index of each repetition, $\hat{\mathbf{H}}$ is the estimated measurement Jacobian matrix, and η is a certain threshold of error probability. We assume the errors in status data of breakers and switches result in erroneous assertion of network topology in term of branch outage, bus-split, or shunt capacitor/reactor switching. At ISO, the true measurement Jacobian matrix is already determined. After a short period of time, the topology error is occurred in the network. Our objective is to determine the present measurement Jacobian matrix under topology errors as little delay as possible, so that help to compute the resulting error matrix $\mathbf{B}$ and $\mathbf{H}_e$ efficiently to identify the problem. In other words, the present network topology is sequentially estimated with the minimal delay.

4.2 Proposed Detection/Decision Scheme

In this section, we investigate one type of the detection/estimation mechanisms against topology error in smart grid. In brief, we have developed a novel estimation strategy via an online statistical analysis of a sequence of data, which can control the detection delay and error probability under the desired levels. The conventional state estimation methods in [141] [143] for bad data detection uses measurements to balance the false alarm rate and the missing detection rate, while our approach aims to minimize the detection delay under the error probability constraint.

The proposed scheme, the *adaptive quickest estimation algorithm*, is the modification of the classic sequential probability ratio test [90]. To estimate the network topology efficiently based on the observed measurements, the proposed scheme contains two interleaved steps that is computed iteratively until the completion of $\hat{\mathbf{H}}$. With such information, ISO can explicitly determine topology error matrix $\mathbf{B}$ via comparing $\mathbf{H}$. Notes that we also develop an analytical model based on the Markov chain, which provide a guidance to configure the detection system for performance guarantee. In this section, we first give the overview of the proposed quickest estimation algorithm. Details of the proposed algorithm are then described and analyzed with a Markov chain based model.

4.2.1 Overview

There are many ways to express the network topology [143] [141]. The network topology is also known as the measurement Jacobian matrix $\mathbf{H}$. In the matrix, an element $H_{r,i}$ and can take the values from set $\{1, -1, 0\}$ respectively, with 1 indicating the power-flow from bus a to bus b, -1 from the bus b to bus a, or the "off" switch status (0) between the two buses. If the erroneous status data occurs in the network, then on-off statuses may are exchanged. Notes that the "off" status under the error may become the active power measurement with either positive or negative direction.

In a general step, we assume that $\vec{\mathbf{H}}_1, \cdots, \vec{\mathbf{H}}_{r-1}$ have been already estimated. The proposed scheme works as follows:

1. The true $\mathbf{H}$ with true $\vec{\mathbf{H}}_r$ removed is used to estimate $\hat{\mathbf{x}}$ by (4.5) beforehand, which is then

further used to estimate the Jacobian matrix under possible topology errors, i.e., $\hat{\mathbf{H}}$. This step is to ensure the calculation accuracy of $\hat{\mathbf{H}}_r$ in later steps. As the number of measurements $M >> N$, removing one measurement will not impact the accuracy in estimating $\hat{\mathbf{x}}$.

2. Consider $\mathbf{Z}_n$ in the multiple access format as in (4.16). Given $\vec{\mathbf{H}}_1, \cdots, \vec{\mathbf{H}}_{r-1}$, $Z_{n,r}$ will be modified as $Z'_{n,r} = Z_{n,r} - \sum_{i=0}^{r-1} \hat{x}_i \hat{H}_{r,i}$. $Z'_{n,r}$ will be used to estimate $\vec{\mathbf{H}}_r$.

3. The elements of $\vec{\mathbf{H}}_r$ will be estimated with an iterative algorithm, to be discussed in details.

4. With $\vec{\mathbf{H}}_r$ obtained, go back to *Stage* 1, and update $r = r + 1$. The algorithm ends when the $\hat{\mathbf{H}}$ is fully constructed.

4.2.2 Methodology

After removing previous measurements related to $\vec{\mathbf{H}}_1, \cdots, \vec{\mathbf{H}}_{r-1}$ in *Stage* 1, $Z'_{n,r}$ of *Stage* 2 are then used to estimate $\vec{\mathbf{H}}_r$ with two interleaved steps in *Stage* 3 in Section 4.2.1.

4.2.2.1 Step 1: Statistical Hypothesis Test

Next, to solve each $H_{r,i}$, we expand each power measurement to the continues-time sequence as shown (4.16). With $Z_{n,r}$ updated to $Z'_{n,r}$ as described in *Stage* 2 Section 4.2.1, the present measurement can be implicitly described as

$$Z'_{n,r} = \hat{x}_i H_{r,i} + e_{n,r}, \tag{4.18}$$

which items related with $i \in \{r+1, \cdots, N\}$ merge into $e_{n,r}$ as the background noise (sum of interference and measurement noise). Without loss of generality, $Z'_{n,r}$ can be composed by three possible statistical hypotheses ($\mathcal{H}_0$, $\mathcal{H}_1$, $\mathcal{H}_2$) of

$$\begin{cases} \mathcal{H}_0 : Z'_{n,r} \sim \mathcal{N}(-\hat{x}_r, \sigma_r^2), \\ \mathcal{H}_1 : Z'_{n,r} \sim \mathcal{N}(0, \sigma_r^2), \\ \mathcal{H}_2 : Z'_{n,r} \sim \mathcal{N}(\hat{x}_r, \sigma_r^2), \end{cases} \tag{4.19}$$

where σ_e^{22} plus the interference of rest items has power of

$$\sigma_r^2 = \sigma_e^2 + \sum_{j=r+1}^{N} (\hat{x}_j)^2. \tag{4.20}$$

To sequentially estimate each $H_{r,i}$ in (4.18) correctly, three possible combinations of binary hypothesis test (BHT) are considered, namely,

$$\hat{H}_{r,i} = \begin{cases} \textit{Test 1:} & \mathcal{H}_0 & \text{vs.} & \mathcal{H}_1, \\ \textit{Test 2:} & \mathcal{H}_1 & \text{vs.} & \mathcal{H}_2, \\ \textit{Test 3:} & \mathcal{H}_2 & \text{vs.} & \mathcal{H}_1, \end{cases} \tag{4.21}$$

where $\hat{H}_{r,i}$ is the estimated value of $H_{r,i}$. Now, we can formulate a quickest estimation problem for each test in (4.21), and the procedure is described below in generalized term of the hypothesis test: $\tilde{\mathcal{H}}_0$ vs. $\tilde{\mathcal{H}}_1$. Therefore, we like to make a decision between these hypotheses in a way of minimizing an appropriate measure of error probability and cost.

4.2.2.2 Step 2: Sequential Decision Problem

Let $(1 - \hbar)$ represents the prior probability of $\tilde{\mathcal{H}}_0$, and $\hbar$ represent the prior probability of $\tilde{\mathcal{H}}_1$. Based on [90], we can formulate the posterior probability P_k of $\tilde{\mathcal{H}}_1$ with a sequence of $Z_{n,r}^k$ in present

$$P_k = \frac{\hbar \Pi_{i=1} f_1(Z_{n,r}^i)}{\hbar \Pi_{i=1} f_1(Z_{n,r}^i) + (1 - \hbar) \Pi_{i=1} f_0(Z_{n,r}^i)}, \tag{4.22}$$

where $f_1(\cdot)$ is the probability density function of $\tilde{\mathcal{H}}_1$, $f_0(\cdot)$ is the probability density function of $\tilde{\mathcal{H}}_0$. By the recursion, the posterior probability P_k of $\tilde{\mathcal{H}}_1$ can be rewritten as

$$P_k = \frac{P_{k-1} f_1(Z_{n,r}^k)}{P_{k-1} f_1(Z_{n,r}^k) + (1 - P_{k-1}) f_0(Z_{n,r}^k)}, \tag{4.23}$$

where the initial value P_1 is equal to $\hbar$. We now can recursively determine P_k in realtime. The constraint function for the decision threshold shall be defined as well, so that the system can know whether $\tilde{\mathcal{H}}_1$ or $\tilde{\mathcal{H}}_0$ is the true hypothesis.

[2] In the case of items with $H_{r,i} = 0$, the σ_e^2 is equivalent to σ_r^2.

The estimation constraint c_e with the error probability cost for the decision rule is formulated as

$$c_e = c_0(1-\hbar)\alpha + c_1\hbar\beta, \tag{4.24}$$

and the terminal decision is

$$\mathcal{D}_k = \begin{cases} 1, & P_k \geq \frac{c_0}{c_0+c_1}, \\ 0, & P_k < \frac{c_0}{c_0+c_1}, \end{cases} \tag{4.25}$$

where $c_j > 0, j \in (0,1)$, is the cost of falsely rejecting $\hat{H}_j$, α is $P_0(\mathcal{D}_k = 1)$ known as the false alarm rate (FAR), and β is $P_1(\mathcal{D}_k = 0)$ known as the missed detection ratio (MDR). If the posterior probability is greater than or equal to the threshold, $\tilde{\mathcal{H}}_1$ is declared ($\mathcal{D}_k = 1$) as the true hypothesis; otherwise, $\tilde{\mathcal{H}}_0$ is declared ($\mathcal{D}_k = 0$). For the better estimation accuracy, we are able to reevaluate c_e by incorporating the Bayes optimal sequential decision rule based on [90] and can be rewrite as

$$C_v(P_v) = \min\left\{\min\left[c_1 P_{v-1}, c_0(1-P_{v-1})\right], C_{v-1}(P_{v-1})\right\}, \tag{4.26}$$

where $C_0(P_0) = \min\left[c_1\hbar, c_0(1-\hbar)\right]$ and $v \in \{1, 2, \cdots, V\}$ (V is the length of the training sequence). Note that the training sequence is needed for the computation of the new cost function C_v before the likelihood test can be executed. By incorporating both cost function vector $\mathbf{C}$ and posterior vector $\mathbf{P}$, we can determine the minimum upper-bound posterior probability $P_U^{\min}$ and maximum lower-bound posterior probability $P_L^{\max}$ of P_k, as below

$$\begin{cases} P_L^{\max} = \max\left[0 \leq \mathbf{P} \leq \hbar | \mathbf{C} = c_1\mathbf{P}\right], \\ P_U^{\min} = \min\left[\hbar \leq \mathbf{P} \leq 1 | \mathbf{C} = c_0(1-\mathbf{P})\right]. \end{cases} \tag{4.27}$$

With the similar formulation of (4.25), the threshold of accepting $\tilde{\mathcal{H}}_0$ in terms of $P_U^{\min}$ and $P_L^{\max}$ can be described as

$$A = \frac{1-\hbar}{\hbar}\frac{P_L^{\max}}{1-P_L^{\max}}, \tag{4.28}$$

and the threshold of accepting $\tilde{\mathcal{H}}_1$ is

$$B = \frac{1-\hbar}{\hbar}\frac{P_U^{\min}}{1-P_U^{\min}}. \tag{4.29}$$

Finally, the minimal stopping time T of our proposed scheme with the Bayesian optimal sequential decision rule, can be written as

$$T = \inf \{k \geq 1 | \Lambda_k \ni (A, B)\}, \tag{4.30}$$

where Λ_k is the sequence of the likelihood ratios

$$\Lambda_k = \frac{f_1(Z_{n,r}^k)}{f_0(Z_{n,r}^k)} \Lambda_{k-1}, \tag{4.31}$$

with the initial value of $\Lambda_0 = 1$ at time interval $k = 0$. A decision is made at each interval k for whether continue sampling, or terminate the test and declare the true hypothesis. If either Λ_k excesses the threshold B (to declare the true hypothesis is $\tilde{\mathcal{H}}_1$) or is less than the threshold A (to declare the true hypothesis is $\tilde{\mathcal{H}}_0$) , the hypothesis decision is made and the process is terminated. Now, we want to decode the element of $\mathbf{H}$ of bus $i + 1$. By the source separation method, $\hat{x}_i \hat{H}_{r,i}$ is now eliminated from the sequence of the observation. The newly updated sequence of observation is

$$Z_{n,r}^k = Z_{n,r}^k - \hat{x}_i \hat{H}_{r,i}, \tag{4.32}$$

where $\hat{H}_{r,i}$ denotes the estimation of $H_{r,i}$. *Stage* 3 is terminated until the completion of estimating $\vec{\mathbf{H}}_r$.

At *Stage* 4, we update the next available measurement ($Z_{n,r+1}$) and $\vec{\mathbf{H}}_r = \vec{\mathbf{H}}_{r+1}$. Then, we return *Stage* 1 and repeat the stages from the beginning of this section until recovering the last element $\hat{H}_{M,N}$ of $\mathbf{H}$. In other words, the $\hat{\mathbf{H}}$ is fully constructed. Notes that a summary of proposed scheme is shown in Algorithm 4.1.

4.2.3 Complete Algorithm

4.2.4 Mathematical Analysis

In this subsection, we develop the Markov chain based analytical model to systematically examine the proposed scheme. The Markov chain based analytical model produces the quantitative

Algorithm 4.1 *Adaptive Estimation for* $\hat{\mathbf{H}}$

known $\hat{\mathbf{x}}$
repeat
  unknown $\vec{\mathbf{H}}_r$
  repeat
    unknown $H_{r,i}$ in (4.18)
    repeat
      compute the posterior probability P_k
      the training sequence:
      for $v = 1$ to V **do**
        compute the cost function, $C_v(P_v)$, and $v \leftarrow v + 1$
      end for
      Threshold calculation:
      compute the boundary posterior probabilities $\{P_L^{\max}, P_U^{\min}\}$, respectively in (4.27)
      compute the boundary thresholds A and B
      Likelihood test:
      compute Λ_k
      Update of: $k \leftarrow k + 1$
      continue the observation
    until $T = \inf\{k \geq 1 | \Lambda_k \ni (A, B)\}$ in (4.31)
    report the true hypothesis and store $\hat{H}_{r,i}$
    update $Z_{n,r}$ via removing $\hat{x}_i \hat{H}_{r,i}$ in (4.32)
    update $H_{r,i}$=$H_{r,i+1}$,
  until completion of estimating $\vec{\mathbf{H}}_r$
  update $Z_{n,r}$=$Z_{n,r+1}$, and $\vec{\mathbf{H}}_r$=$\vec{\mathbf{H}}_{r+1}$
until completion of estimating $\hat{\mathbf{H}}$

performance analysis and theoretical guidance on the proposed scheme's parameter configuration for performance guarantee under fundamental performance metrics: the expectation of false-alarm rate and the expectation of missed detection ratio.

4.2.4.1 Analysis Model

For analysis purpose, we discretize $\mathbb{R}^+ \bigcup 0$ into the finite sets $\{U_1, \cdots, U_{F-1}, U_F\}$, where $U_1 = 0$, and U_F is the set whose value is greater than or equal to h. In other words, F is the total number of transition from 0 to the state that has the value greater than or equal to h. There are several approaches for discretization [138] [139]. In this chapter, we employ uniform sampling for simplicity[3]. Moreover, from (4.31), we know that the sequence exhibits the Markov property, which the current state $j = \Lambda_k$ at observation k only depends on the previous state $i = \Lambda_{k-1}$ at $k - 1$, but not on the past history [140].

[3]Other discretization methods can be employed like the μ-law or A-law in PCM.

The transition probabilities of the Markov chain for the proposed scheme from state i at $(k-1)$ to state j at k can be described as

$$\begin{aligned} P_{ij} &= P[\Lambda_k = j | \Lambda_{k-1} = i], \text{ under } \tilde{H}_0, \\ \hat{P}_{ij} &= P[\Lambda_k = j | \Lambda_{k-1} = i], \text{ under } \tilde{H}_1. \end{aligned} \tag{4.33}$$

We can calculate the transition probability matrices $\mathbf{P}$ and $\hat{\mathbf{P}}$ with the size of $(F+1) \times (F+1)$, under the hypothesis $\tilde{H}_0$ and $\tilde{H}_1$ according to $f_0(Z_{n,r}^k)$ and $f_1(Z_{n,r}^k)$, respectively.

The initial steady state probability of the Markov chain, which the process starts from an normal state, can be determined as

$$\pi_j^0 = \frac{\pi_j}{\sum_{i=0}^{F-1} \pi_j}, \quad \text{given} \quad j \in \{0, U_1, \cdots, U_{F-1}\}, \tag{4.34}$$

and the steady-state probability can be determined as

$$\pi_j = \sum_{i=0}^{F} P_{ij} \pi_i, \tag{4.35}$$

where $j \in \{0, U_1, \cdots, U_F\}$ and $\sum_{j=0}^{F} \pi_j = 1$.

4.2.4.2 The Expectation of False Alarm Rate

Next, based on the Markov chain model, we study the theoretical performance analysis of false alarm rate and missed detection ratio expectations, respectively. The expectation ($E_{\mathbf{P}}[\text{FAR}]$) of the false alarm rate is the probability that the proposed statistic Λ_k reaches to the state U_F when the hypothesis $\tilde{H}_0$ is true. According to [140], it states the transition probability matrix $\mathbf{P}$ always has a special eigenvector with only one eigenvalue $\lambda = 1$ and the rest is zero. Thus, we can obtain the solution by re-elaborating the equation (4.35) into the matrix form as

$$\begin{bmatrix} P_{00}-1 & P_{01} & \cdots & P_{0F} \\ P_{10} & P_{11}-1 & \cdots & P_{0F} \\ \vdots & \vdots & \ddots & \vdots \\ P_{F0} & P_{F1} & \cdots & P_{FF}-1 \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \\ \vdots \\ \pi_F \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{4.36}$$

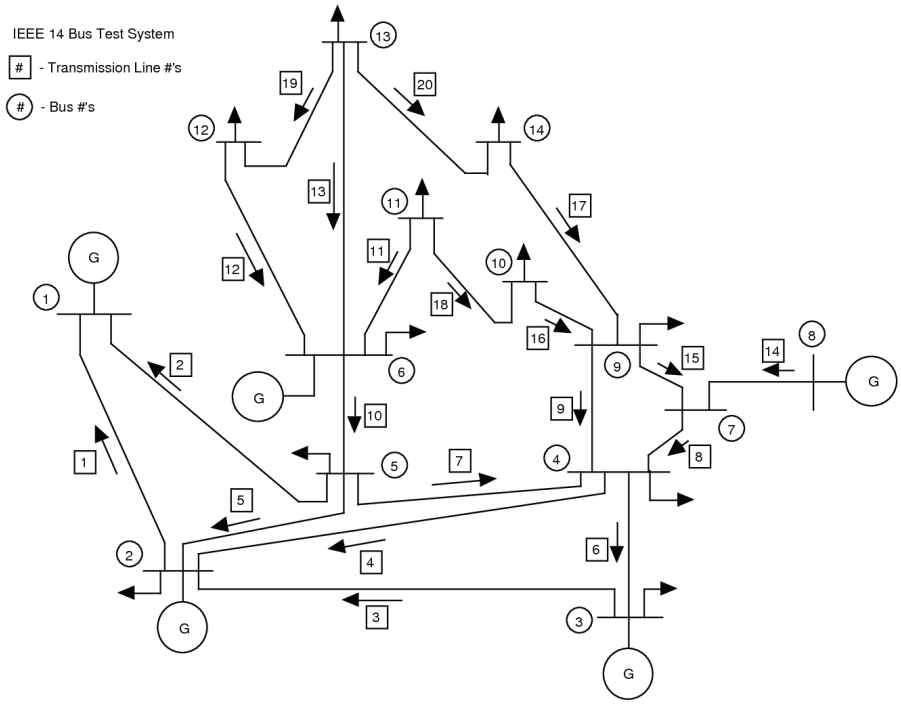


Figure 4.2 *The schematic graph of IEEE 14-Bus test system with 5 generators (G).*

Then, the expectation of false alarm rate can be determined by

$$E_{\mathbf{P}}[\text{FAR}] = \pi_F. \tag{4.37}$$

4.2.4.3 The Expectation of missed detection ratio

We define the missed detection probability as the probability that the detection delay is greater than or equal to a detection delay constraint C. The expectation ($E_{\hat{\mathbf{P}}}[\text{MDR}]$) of the missed detection probability (MDP) is, starting from the initial state, the summation of probabilities that Λ_k stays at a state other than state U_F at time C. Let $p_i(s)$ denote the probability of the state variable at time s and at state i. We set the initial condition for the transition probabilities as

$$p_i(0) = \pi_i^0, \tag{4.38}$$

where $i \in \{0, U_1, \cdots, U_{F-1}\}$ and $p_F(0) = 0$. By the iteration, at each s, the state probability vector is updated by the previous state probability vector in a matrix form as

$$\begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \\ p_{F-1}(s) \\ p_F(s) \end{bmatrix}^T = \begin{bmatrix} p_0(s-1) \\ p_1(s-1) \\ \vdots \\ p_{F-1}(s-1) \\ p_F(s-1) \end{bmatrix}^T \hat{\mathbf{P}}, \tag{4.39}$$

and

$$p_F(s) = 0, \quad s \in \{0, C-1\}. \tag{4.40}$$

Here the $p_F(s)$ at every s of state U_F is reset to zero for the next iteration since we only concern the missed detection case only. The expectation of missed detection ratio under the given delay constraint C can be obtained as

$$E_{\hat{\mathbf{P}}}[\text{MDR}] = \sum_{i=0}^{F-1} p_i(C). \tag{4.41}$$

4.3 Performance Analysis and Simulation

In this section, we use both analytical and numerical simulations to demonstrate the performance of the proposed scheme by MATPOWER 4.0 package [128]. MATPOWER 4.0 provides realistic power flow data and test systems that uses widely in research-oriented study. All simulations are performed under IEEE 14-bus test system as shown in Figure 4.2, which has 5 generators for 20 measurements; the arrow represents the power-flow directions, and the triangular (attached on the bus) is the load. Notes that we first apply the analytical model to theoretically analyze the performance of the detection system for guiding the system parameter configuration. Next, we use the parameter from the theoretical analysis to confirm the accuracy of the analysis, and then demonstrate the performance of the detection.

Figure 4.3 shows an illustrative example of decoding/estimating $H_{r,i}$ of bus i. The dot represents Λ_k, $\hbar$ is set to 0.5 with the maximum cost constraint. The simulation result shows that the

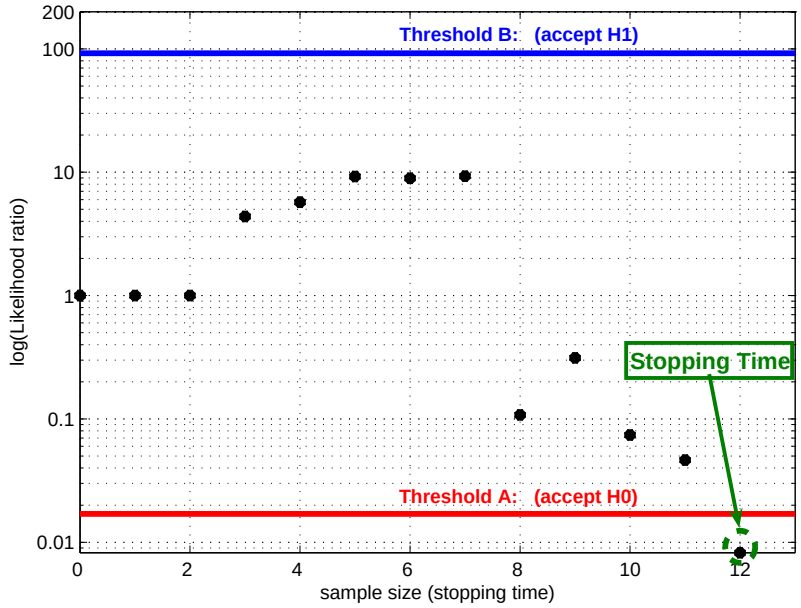


Figure 4.3 *Simulation of the proposed scheme in quickest detection for determining $H_{r,i}$ of the bus i under an IEEE 14-bus test system.*

threshold A is 0.017 and the threshold B is 92. From Figure 4.3, the decision is declared at minimum stopping time T is 12 after falling under the threshold A, and the value of $\hat{H}_{r,i}$ is -1 in this case.

The numerically examination is presented in Figure 4.4 for understanding the impact of ASN on system parameters c_0 and c_1 of the proposed scheme. Notes that x-axis is the bus index, and y-axis is the corresponding ASN for each bus. There are three settings: i) the low cost with $c_0 = 1$ and $c_1 = 2$; ii) the median cost with $c_0 = 3$ and $c_1 = 4$; iii) the high cost with $c_0 = 6$ and $c_1 = 8$. As shown in the figure, the higher cost of falsely rejecting the hypothesis causes larger ASN (i.e., the system needs to spend more observations for making a decision). In other words, the higher cost setting may results the better estimation accuracy, but have the longer decision delay.

As shown in both Figure 4.5 and in Figure 4.6, the analytical performance measures and the simulation results are compared under same setting and input data for the examination. By using power flow data sets from MATPOWER 4.0, the performance indices (E[FAR], E[MDR]) comparisons between the analytical and simulation results can be conducted. With the parameter from the

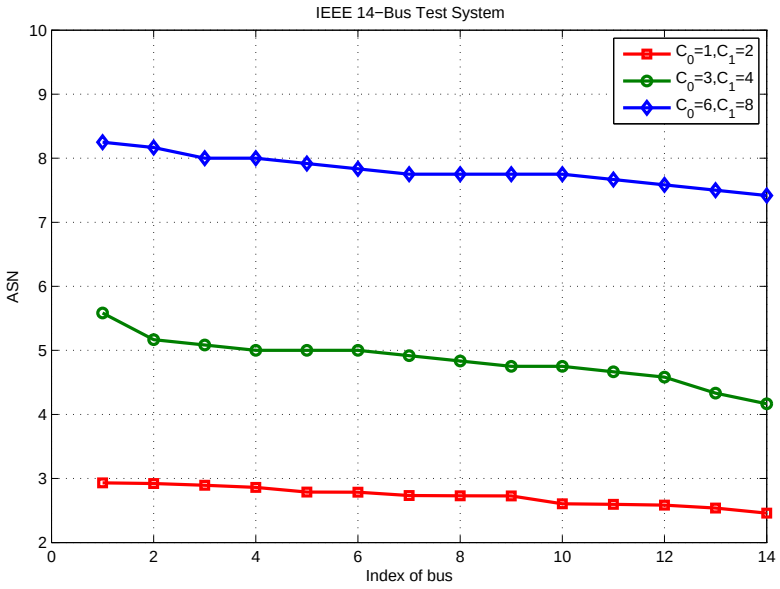


Figure 4.4 *The ASN under 3 sets of error cost scenarios for IEEE 14-Bus test system.*

theoretical analysis, the performance indices are simulated (both theoretical analysis and simulation are plotted respectively to confirm accuracy of analysis and demonstrate the performance), so that we can properly configure the proposed algorithm for the guaranteed performance.

Figure 4.5 illustrates the relation between the system parameter c_0 and performance metrics (FAR). Notes that the logarithmic scale is used in the figure for the vertical axis. FAR is computed via applying c_0 from 1 to 10 in the ascending order, while retaining the value of c_1 to 1 (i.e., c_0 is the cost of falsely rejecting $\hat{H}_0$). As shown in Figure 4.5, the analytical and simulation results are fairly close. Figure 4.5(a) is the result of the Markov chain based analytical model of the proposed scheme. In Figure 4.5(b), the central mark is the median, and the edges of the box are the 25^{th} and 75^{th} percentiles. The whiskers extend to the most extreme data points (the maximum and minimum value of FAR). The maximum value of FAR is close to the edge of box; it means that the worst case for each set is not far away from the majority value of FAR. The difference percentage between the median and the maximum FAR is about 8%, but smaller FAR is always desired in this simulation. To compare with Figure 4.5(a) and 4.5(b), the analytical FAR usually falls right-on or slightly-below the 25^{th} percentile edge of the box but it never exits the minimum FAR of numerical result. In other

words, the analytical model gives us more ideal/theorical value of FAR for the proposed algorithm that is less than the median FAR of numerical result. From Figure 4.5, we also observe that a larger c_0 yields a smaller false alarm rate as expected.

The result is shown in Figure 4.6 that helps us study the impact of the missed detection ratio on c_1 of the proposed scheme (i.e., c_1 is the cost of falsely rejecting $\hat{H}_0$). Notes that the logarithmic scale is used in the figure for the vertical axis. MDR is computed via applying c_1 from 1 to 10 in the ascending order, and the value of c_0 is equal to 1 through the process. Figure 4.6(a) is the expectation of missed detection ratio under detection delay constraint of the Markov chain based model. As presented in Figure 4.6(b), the center mark is the median of MDR, and the edges of rectangular are the 25^{th} and 75^{th} percentiles. The whiskers extend to the maximum and minimum value of MDR, respectively. The maximum MDR of each set is much close to the majority of MDR in comparison to the minimum MDR; the smaller the worst case of MDR, the better performance. In Figure 4.6(a) and 4.6(b), we observe that the value of analytical MDRs falls on the region between the 25^{th} percentile and median MDR of numerical simulation; in other words, the median MDR of numerical result is fairly close to the analytical results. The detection delay constraint, which is introduced in the formulation of calculating analytical MDR, may give additional reenforcement on the simulation accuracy. From Figure 4.6, both analytical and numerical simulations show that the larger constraint c_1 results smaller missed detection ratio as expected. In other words, the probability of true estimation rises if we allow to increase the cost of longer delay.

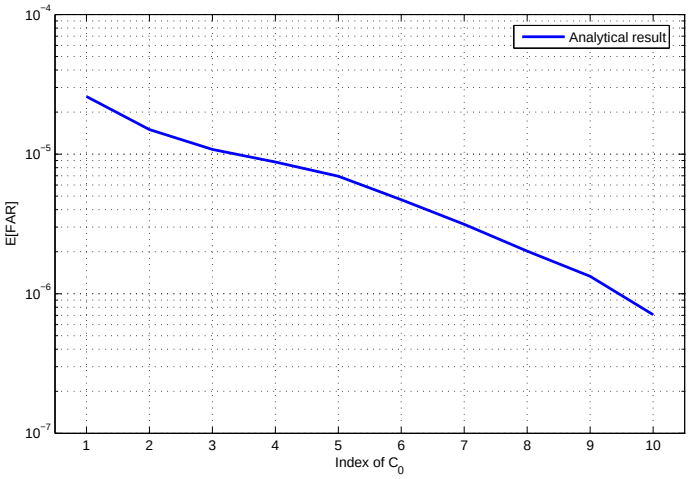
From Figure 4.5 and 4.6, we demonstrate the performance metrics with different h. It also helps us configure the system parameter c_1 and c_0 for the guaranteed performance under fundamental metrics. We can select the proper configuration of $[c_1, c_0]$ from the reasonable range to satisfy the desired performance constraints. For examples, in the low-cost configuration (i.e., the low cost setting: both $c_0 = 1$ and $c_1 = 1$), one can explicitly determine the expectational FAR of 0.0005 and expectational MDR of 0.007. In other words, the guarantee performance of proposed algorithm is able to estimate the $\mathbf{H}$ with minimal delay while maintaining a comparable low error rates.

Finally, we consider that the performance of proposed scheme in term of the computational

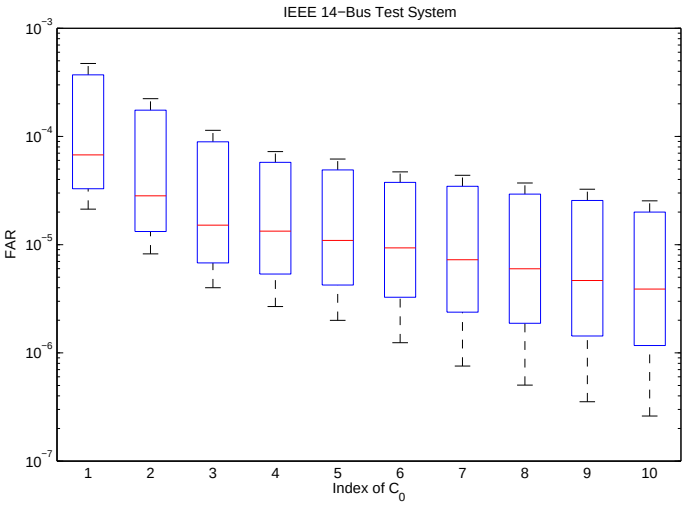
complexity. Since the proposed scheme heavily involves with the hypothesis testing, we like to understand the difference of computation time between the proposed scheme and the conventional hypothesis test algorithm, which are used in [101] [100] [99]. The *size of an instance* is set to the number of cycles, which is used to compute ASN. The computational time is simulated by increasing the number of cycle. From Figure 4.7, the computational time of the proposed scheme is much smaller than that of the conventional algorithm. Although the performance of both algorithm seems linearly increases as the number of cycle for calculating ASN, the proposed method runs on the 70% less of computational time than that of the conventional algorithm.

4.4 Conclusion

In this chapter, the main objective is to perform online estimation for the present network topology with minimal delay, in order to help detect and identify the topological error efficiently in smart grid networks. The proposed scheme, the *adaptive quickest estimation algorithm*, successfully determines the current network topology as quickly as possible without violating the given constraints such as a certain level of estimation accuracy. As the present topology is solved, the operator can quickly determine and identify the network topology error timely. Furthermore, we are able to develop the Markov chain based analytical model to characterize the behavior of our proposed scheme; one can quantitatively study the system parameters to achieve the guaranteed detection performance in term of fundamental metrics. With the aid of MATPOWER 4.0 package, the both analytical and numerical simulations are conducted under the simulated power test system to ensure the its accuracy and proficiency; the simulation results have shown that the proposed scheme is efficient in terms of detection accuracy and minimum detection delay. The guaranteed performance of the proposed scheme under the low-cost scenario can be explicitly determined in terms of FAR and MDR, and the computational complexity of the quickest estimation of proposed algorithm is much efficient than the conventional scheme.

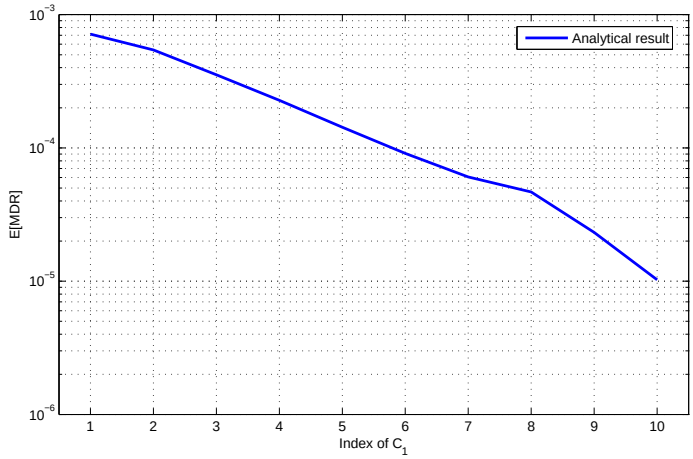


(a) The expectation of false alarm rate via Markov chain based analytical model

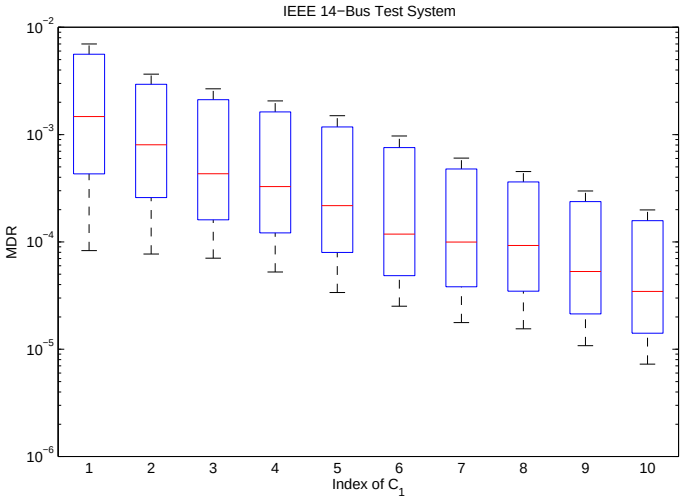


(b) The false alarm rate via numerical simulation

Figure 4.5 *The comparison between the analytical and numerical results under* $c_1 = 1$ *for the IEEE 14-Bus test system.*



(a) The expectation of missed detection ratio via Markov chain based analytical model



(b) The missed detection ratio via numerical simulation

Figure 4.6 *The comparison between the analytical and numerical results under* $c_0 = 1$ *for the IEEE 14-Bus test system.*

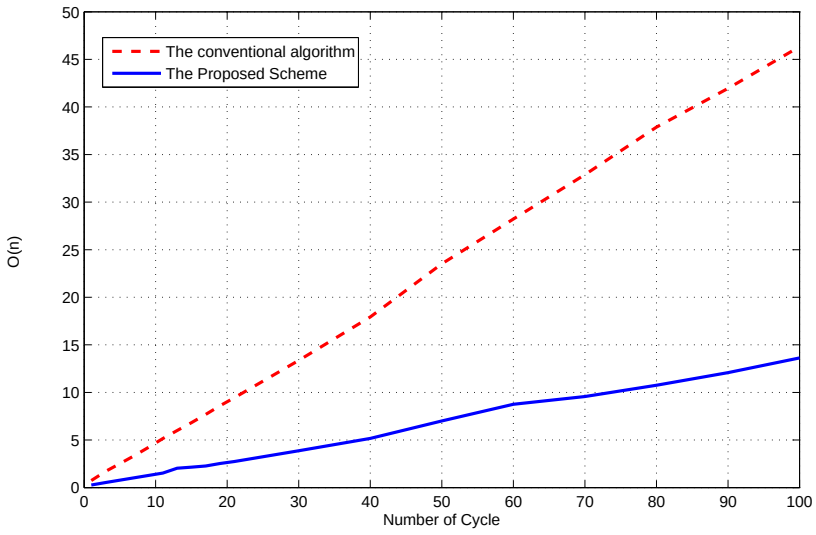


Figure 4.7 *Performance comparison of computational complexity by varying the number of ASN calculation cycle*.

Chapter 5

Online Quickest Multiarmed Bandit Algorithm for Distributive Renewable Energy Resources

In this chapter, we develop an online quickest MAB algorithm to determine the best choice of energy profile as few samples as possible, under the constraint of accuracy. For each energy profile containing various sources of distributive renewbles, such as solar, wind, hydro, etc, the utility is the sum of the cumulate log likelihood ratio plus a confident interval. The cumulate logarithm likelihood ratio represents how well this energy profile performs, which is the exploitation. The confident interval adds weight to find the new profile, which is the exploration. We derive the close form for the confident interval and obtain an upper bound for the expected regret. From the simulation results, we can show that a user can effectively switch to profiles and, at the end, select the best energy profile. In order to make our proposed algorithm in more realistic simulation environment, the preliminary of the proposed algorithm is that 1) this is an application intend to give an end-user more dynamic tool to select the better service, 2) the energy market is competitive as always and therefore the profile information is revealed, only if be tested (i.e., The utility companies unlikely publish such sensitive data; otherwise, all consumers use the best one and the others get zero), and 3) during the algorithm execution process, the system is able to remain in a short period steady or quasi-steady state; to tackle the fluctuation of renewable energy pattern over time, the first step is that a customer applies our proposed scheme to find the best profile, the newt step is that a customer uses and trades with this profile, the last step is that after a certain period of time, the renewable energy distribution is changing sufficiently, and then our algorithm is triggered to find the new best profile and then trade again. Note that, in general speaking, the power pattern of wind generation has better stability in comparison to the solar/thermal or hydro generations as shown in 5.1. Note that, in this figure, the power in Million Watt (MW) of BPA balancing authority load and total wind, hydro, and thermal generations were recorded last seven days from July 16, 2011 to July 23, 2011; the measurement is obtained based on five minute reading form BPA SCADA system for

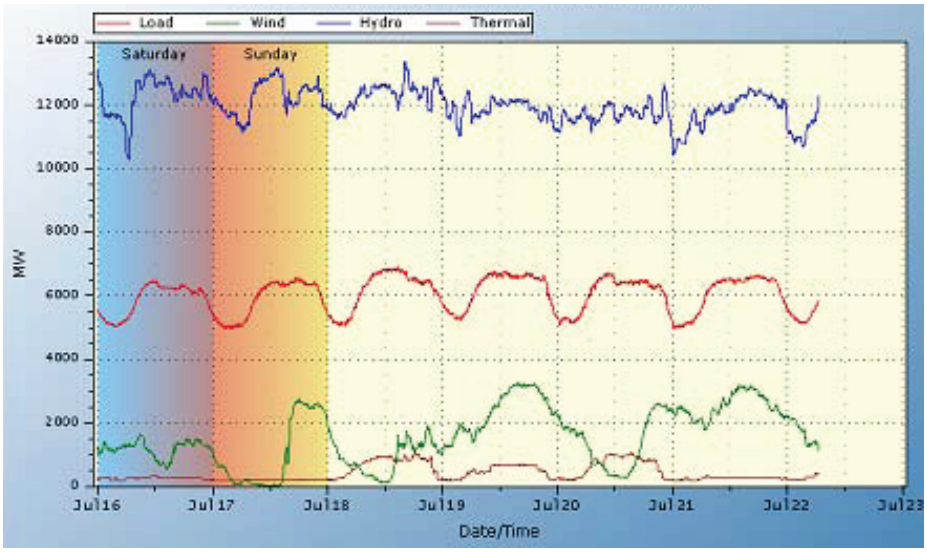


Figure 5.1 *The statistical curve of technical operation of British Pipeline Agency (BPA).*

point 45583, 79687, 79682, and 79685; balancing authority load is red, wind generation in green, hydro generation in blue, and thermal generation in brown.

5.1 System Model

We illustrate an example for our system model in Figure 5.2. Note that, in this figure, *Profile A* composes one hydro generator and two wind generators; *Profile A* composes one hydro generator and two wind generators; *Profile B* composes one hydro generator, one wind generator, and one geo-thermal generator; *Profile C* composes two solar generators and one hydro generator; *Profile D* composes two hydro generators and one wind generator. As shown in Figure 5.2, the user decides to buy *Profile C* as the main renewable power resource from the 4th user (i.e., the payment from the user is credited into the 4th user trading account). However, *Profile A* has the most stable and effective power pattern among others in the grid since it has two wind generators. To have the highest profit, obviously, the user should pick *Profile A* instead. Therefore, the key point is how the user knows that which profile is the best energy resource, while the renewable energy power patterns might not have the prior statistics.

With a total of K profiles, user u needs an efficient algorithm to find as quickly as possible

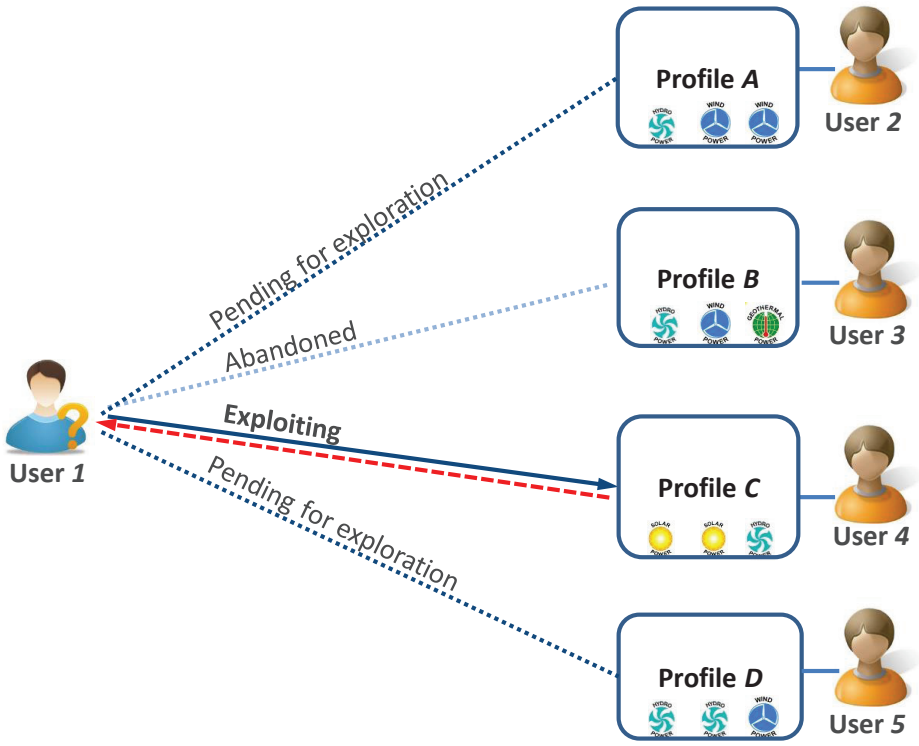


Figure 5.2 *The illustration of distributive renewable energy resource allocation for finding the most profit one among multiple energy profiles.*

the best profiles $i \in \{1, K\}$, which can deliver the highest profit. Thus, the one of main tasks is that how to minimize the decision time T for making the decision as quickly as possible while maintaining a certain level of error probability, i.e.,

$$\begin{aligned} \min \quad & T, \\ \text{s.t.} \quad & P_r \leq \alpha, \end{aligned} \tag{5.1}$$

where T is also called the minimum stopping time, and P_r is the detection error (which might be false alarm or missing probability). In other words, we want to minimize the number of observation under the constraint that the error probability is less than a certain thresholds α.

At each t by accessing profiles i, user u can obtain a utility $\mathcal{U}_i^u(t) > 0$, while paying the cost $\mathcal{C}_i^u(t) > 0$, which is charged by another user equipped with this profile. Consequently, user u can collect a profit from profile i

$$\mathcal{P}_i^u(t) = \mathcal{U}_i^u(t) - \mathcal{C}_i^u(t). \tag{5.2}$$

Notes that $\mathcal{P}_i^u(t)$ can be obtained at the end of t if user u chooses the i^{th} profile. The rest of $\mathcal{P}_k^u(t)$ for $k \neq i$ is temporarily unknown. In other words, user u cannot access the information until the decision is made. We also have no idea about the distributions for profiles power patterns. When the user chooses a sequence $\{i^u(1), i^u(2), \cdots, i^u(T)\}$, the total profit after a period of time can be represented as

$$\mathbb{P}^u(\hbar) = \sum_{t=1}^{\hbar} \mathcal{P}^u_{i^u(t),t}, \tag{5.3}$$

where $\hbar$ can be any time such as the switch time, the decision time, etc.

On the other hand, the regret until $\hbar$ can be determined by

$$\mathbb{R}^u(\hbar) = \mathbb{P}^u_{\max}(\hbar) - \mathbb{P}^u(\hbar), \tag{5.4}$$

which $\mathbb{P}^u_{\max}(\hbar) = \max_{i\in[1,K]} \sum_{t=1}^{\hbar} \mathcal{P}^u_{i,t}$ is the maximum total profit from time 1 to $\hbar$ if user u always chooses the true optimal profile.

Our goal is to propose an online quickest multiarmed bandit algorithm, which involves the quickest change detection and multiple arm bandit frameworks. For our problem, the traditional detection formulations such as presented in [90] [124] are no longer valid, because the unknown exists in the process and may also change over the process. Our main motivation is to derive a detection model with considering existence of the unknown, and then develop a new confident interval that can help us configure the detection system for performance guarantee based on the detection requirements. This confidence interval can be used to balance between exploitation and exploration. With the proposed scheme, the user is able to find the most efficient, effective and profit profile. We assume that each user can only observe one profile at a time and omit notation u of specifying user for simplicity.

5.2 Proposed Profile Selection Scheme

5.2.1 Overview

At the beginning, user u runs each arm/profile once (we also call it as "the initial test-drive") for determining the likelihood ratio of each arm at $t = 0$. Based on the observation, the decision whether to select arm i is determined at each t, and we can start cumulating the statistical measurement $S_{i,t}$ in a recursive way, namely

$$S_{i,t} = S_{i,t-1} + \left[\frac{\ln(L_{i,t}) + \Im_t}{n_i}\right] \tag{5.5}$$

where n_i is the number played on arm i so far, $L_{i,t}$ is the likelihood ratio function, and $\Im_t$ denotes the confident interval. $L_{i,t}$ is the probability ratio of two distributions (Q_0 over Q_1). Let Q_1 denote the distribution of the maximum profit arm and Q_0 be the distribution of noise only. Notes that Q_0 can be assumed known, but Q_1 is unknown. The problem we face here is how to make a best choice after a certain time. In other words, our goal is to find the best arm in a way of minimizing the decision delay while maintaining an appropriate measure of the error probability.

We denote T as the stopping time for declaring the best arm under current observation. τ is a change time (e.g., it is the switch point from one distribution belong to arm i to another distribution belong to arm j, $j \neq i$). Based on the Lorden's formulation [90], we minimize the worst case of detection delay, which can be described as

$$T_D = \inf_{T \in \mathcal{T}} \sup \operatorname{esssup} E_\tau[(T - \tau + 1)^+ | \mathcal{F}_{\tau-1}], \tag{5.6}$$

which $\tau > 1$, $\mathcal{F}_\tau$ denote the smallest α-field with respect to the observations belong to arm i, $\mathcal{T}$ is the set of all stopping time with respect to $\mathcal{F}_\tau$, and E_τ is the expectation that the change time is τ. To solve the minimum T_D, the conventional CUSUM algorithm is the best-known technique to tackle this type of problem [124]. However, most CUSUM-based models assume the perfect knowledge of the likelihood functions. However, in our case, the distribution of Q_1 cannot be identified before the user makes final decision (we will address this matter in the later section). T can be determined

in terms of statistical measurement and user-defined threshold from

$$T = \inf\{t \geq 1 | (S_{i,t} > \eta)\}, \tag{5.7}$$

where η denotes the user-defined threshold[1].

The proposed algorithm is able to do following decision at each t: 1) Let $\max_{m \neq i} P_{m,0}$ denote the next highest likelihood ratio of initial test-drive at $t = 0$. If $S_{i,t} < \max_{m \neq i} P_{m,0}$, it abandons the current observation of arm i and switch the next arm with the highest ratio for starting new observation to the next arm[2]; 2) when $S_{i,t} > \eta$, it claims the current arm (5.7), which is the best choice among K arms; 3) after the decision is make, $S_{i,t}$ is continually cumulated and any arm still has the probability to be tested, so as to avoid the case that the decision is wrongly made.

5.2.2 Definition of Confident Interval

In this section, we define the confident interval in (5.5) by the following lemma.

Lemma 5.1. *The confident interval for the proposed scheme can be defined as follows for any* $K > 0$ *and any* $t > 0$

$$\Im_t = \sqrt{\frac{8 \ln(n) ||x_{i,t}||^2}{n_i}}, \tag{5.8}$$

where $x_{i,t}$ *is the online observation of arm* i, n_i *is the number of times that arm* i *has been played so far, and* n *is the overall number of plays done so far. The proof of Lemma 5.1 is provided in Appendix A.*

It is able to deal with situation of exploitation-exploration tradeoff among K arms. For instance, for exploitation, if user u keeps playing arm i for a while, $\Im_t$ of the current arm decreases. As a result, the chance is increased for switching and exploring to the others. However, for exploration, if the current observation of arm i has the significantly large log likelihood, $S_{i,t}$ can be quickly accumulated.

[1] It is a function of error detection, and its value is determined numerically.

[2] This corresponds to a reset event similar in classic CUSUM test, which is to reset statistical measurement to zero. In our case, we switch to another potential candidate, if it is less than a certain value.

5.2.3 Online Quickest Multiarmed Bandit Algorithm

For the initial test-drive, let $P_{i,0}$ denote the initial ratio of chosen arm i in term of likelihood ratio test functions

$$P_{i,0} = \frac{L_{i,0}}{\Lambda_0}, \tag{5.9}$$

where $\Lambda_0 = \sum_{j=1}^{K} L_{j,0}$ and

$$L_{i,0} = \frac{q_1(x_{i,0})}{q_0(x_{i,0})}. \tag{5.10}$$

We know that q_0 and q_1 are the probability density function of distribution Q_0 and Q_1, respectively. Here, without loss of generality, we assume the distribution is complex Gaussian.

After playing each arm once for determining a set of $\{P_{1,0}, P_{2,0}, \cdots, P_{K,0}\}$ at $t = 0$, the user can select and play arm i with the highest ratio. The cumulative statistic at observation t can be described as same as (5.5)

$$\begin{aligned} S_{i,1} &= S_{i,0} + \frac{\ln(L_{i,1}) + \mathfrak{I}_1}{n_i}, \\ &\vdots \\ S_{i,t+1} &= S_{i,t} + \frac{\ln(L_{i,t+1}) + \mathfrak{I}_{t+1}}{n_i}, \end{aligned} \tag{5.11}$$

where $S_{i,0} = 0, i \in [1, 2, 3, \cdots, K]$. Reminds that $q_1(x_{i,t})$ of $L_{i,t}$ belongs to the distribution associated with the maximum mean u_1 (the unknown variable), and $q_0(x_{i,t})$ is the distribution associated with zero mean. The technique to solve the unknown is described in detail of Appendix A. For the online detection, $S_{i,t}$ can be expressed in a recursive way via (5.11). As soon as the condition of (5.7) is satisfied, user u abandons the current observation of arm i and switch to another arm with the next highest ratio, if $S_{i,t} < \max_{m \neq i} P_{m,0}$; On the other hand, if $S_{i,t} > \eta$, user u declares the arm under current examination is the best (most profit) one; then, $S_{i,t}$ is continually cumulated for ensuring the correctness of final decision (i.e., if the wrongness is found, it will notify the user to reconsider). In summary, the proposed scheme is illustrated in Algorithm 1.

Algorithm 5.1 *Quickest Search Algorithm*

1: Known a total of active K arms
2: Run the initial test-drive for $\{P_{1,0}, P_{2,0}, \cdots, P_{K,0}\}$
3: Pick arm i with the highest ratio $P_{i,0}$
4: $t \leftarrow (1, 2, 3, \cdots)$
5: **repeat**
6: Cumulate $S_{i,t}$ in terms of $S_{i,t-1}$ and $\mathfrak{I}_t$
7: **if** $S_{i,t} < \max_{m \neq i} P_{m,0}$ **then**
8: Abandon the current $S_{i,t}$
9: Reset and switch to the next potential arm
10: **end if**
11: $t = t + 1$
12: **until** $T = \inf\{t > 1 | S_{i,t} > \eta]\}$
13: Declare the current arm is the best one
14: Continue step 5 to 13, just in case of wrong decision.

5.2.4 Property of Regret

Once T is determined, we can compute $\mathbb{R}^u(T)$ as mentioned in (5.4). The regret is another common measure for the quality of a bandit algorithm. We have the following lemma:

Lemma 5.2. *For any $K > 0$ and any $T > 0$, the regret of any user for the online quickest search can be defined as*

$$E\left[\mathbb{R}^u(T)\right] \leq \frac{32 \ln(T) ||x_{i,t}||^2}{\sqrt{\ln(K)}} \sqrt{\frac{T}{K}}. \tag{5.12}$$

Notes that the upper bound of expected regret is asymptotical for any T. The proof of Lemma 5.2 is provided in Appendix B.

5.3 Simulation

In this section, we verify the performance of the proposed online quickest multiarmed bandit algorithm by the numerical simulation. We assume that the user is able to discover all active profiles in the network. Notes that the profile information is a set of $\{$ identification number (ID), maximum supply power (MSP), minimum sale cost (MSC)$\}$. At the beginning, the user collects profiles information once in order to initiate the test-drive. Then, the user makes a decision on purchasing energy from one profile with the highest ratio. The owner with this particular profile starts sharing/updating both MSP and MSC to the user (i.e., the user transfers credit to the owner's trading account as well

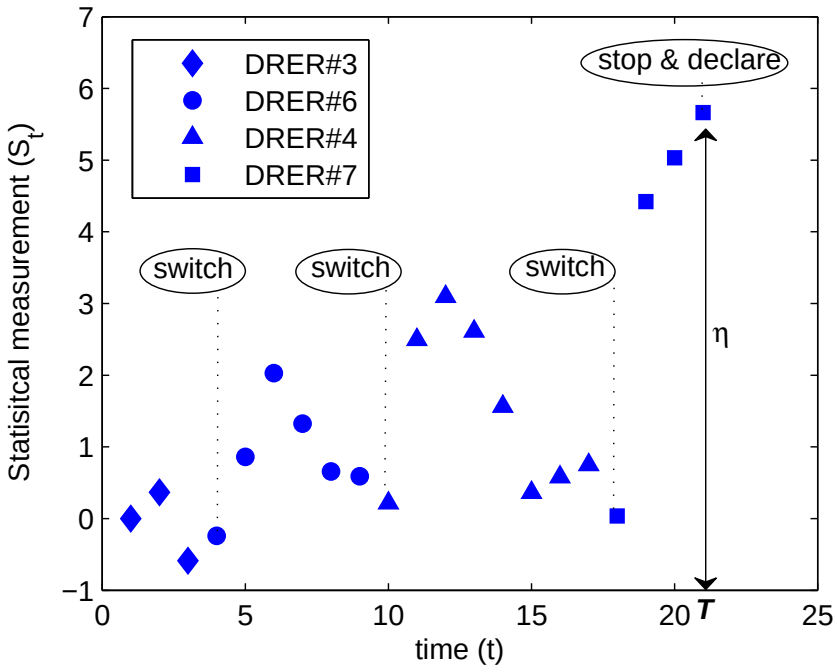


Figure 5.3 *The performance analysis for the proposed scheme.*

at each observation). In simulation, we have 10 profiles and set the7th profile has the unit profit. The other profiles have the profit with uniformly distributed over $(0, 1)$.

Figure 5.3 presents the quickest search for the best one among all profiles. The false positive rate is set to 0.1%. As indicated in this figure, there are 3 switch points; the proposed scheme picks a certain profile according the result of the initial test-drive, exploits it for a couple times, and finally decide to abandons and switches to the another potential profile at time 4, 10, and 18, respectively (from the 3rd profile, the 6th profile, the 4th profile, to the 7th profile). On the other hand, the proposed scheme is stopped at time 21, which is also known as the stopping time. In other words, the proposed scheme declares the 7th profile that is the best choice among others.

The numerically examination is presented for understanding the impact of the fundamental performance metric the false alarm P_f on detection delay T_D of the proposed scheme as shown in Figure 5.4. Notes that the logarithmic scale is used in the figure for the horizontal axis. We run 500 realizations for the simulation. P_f is vary from 10^{-8} to 10^{-2}. We consider three cases; the 1st case

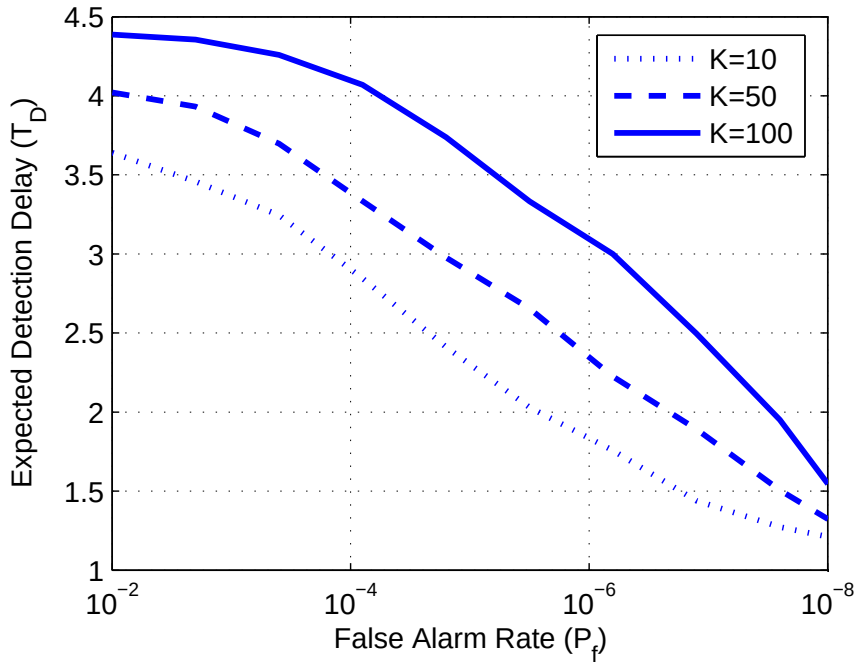


Figure 5.4 *The performance analysis for the proposed algorithm. The y-axis is average run length of detection delay, and x-axis is the probability of the false positive.*

has total of 10 active profiles; the 2nd case has 50 profiles; and the 3rd case has 100 profiles. The simulation result shows the tradeoff between the false positive rate and expected detection delay. As the total number of active profiles increases, the expected detection delay rises slightly. Once the K reaches to 100, the expected T_D reduces much slower than that of the 1st and 2nd cases, which is shown linearly decreased with smaller P_f. we also can observe that a larger T_D yields a greater P_f in each case. From the figure, another conclusion can also be drawn that smaller P_f causes higher T as expected, i.e., the system needs to spend more observations for making a decision.

Finally, we consider that the performance of proposed scheme in term of the expected regret. Figure 5.5 shows how the proposed scheme suffers by varying the size of K and the length of T. Notes that the logarithmic scale is used in the figure for the horizontal axis, the y-axis is the time horizon in logarithm, x-axis is total number of active profiles, and z-axis is the normalized upper bound regret From this figure, we can obtain that the normalized $E\left[\mathbb{R}^u(T)\right]$ decrease gradually as T increases, because the proposed scheme gradually learn the profiles that delivers the most profits.

In addition, as the number of available profiles increases, the normalized expected regret becomes larger initially; however, as the time horizon cumulates, the expected regret decreases gradually. In other words, according to both Figure 5.4 and Figure 5.5, the stricter constraint of the detection error probability is given, the larger T and smaller regret user will have. Thus, the user will have the smarter decision, if more information is collected.

5.4 Conclusion

In this chapter, we proposed the online quickest MAB algorithm for optimal profile allocation. The proposed scheme allows the user to make the best choice among multiple profiles as little delay as possible in realtime based on the observation under no assumption of what distribution the power patterns of profiles follow while maintaining a certain level of detection accuracy. We also mathematically derive a new confident interval and the upper bound of expected regret for the proposed scheme. The results of simulations are successful to demonstrate the methodology of proposed scheme, the detection accuracy of quickest search, and tradeoff between the expected regret, number of active profiles, and constraint of detection accuracy. The proposed algorithm can also have great potential impacts in the analysis of the renewable resource utilization, electrical vehicle charging scheduling, or online strategizing allocation in stochastic environments.

5.5 Appendix A

In (5.10), let q_e depend on the magnitude but not on its variance; we can describe it as

$$\begin{aligned} q_e(x_{i,t-1}) = \frac{1}{\pi \det\left(E[x_{i,t-1}\bar{x}_i]\right)} \times \\ \exp\left[-(\bar{x}_{i,t-1} - \bar{\mu}_e)'(E[x_{i,t-1}\bar{x}_{i,t-1}])^{-1}(x_{i,t-1} - \mu_e)\right], \end{aligned} \tag{5.13}$$

where $\bar{x}_{i,t-1}$ is the complex conjugate. Notes that the mean μ_0 of q_0 holds zero value, and in contrast, μ_1 for q_1 is unknown. By subtracting (5.13) into (5.10), $L_{i,t-1}$ can be rewritten as

$$L_{i,t-1} = \exp\left[-A(x_{i,t-1})\right], \tag{5.14}$$

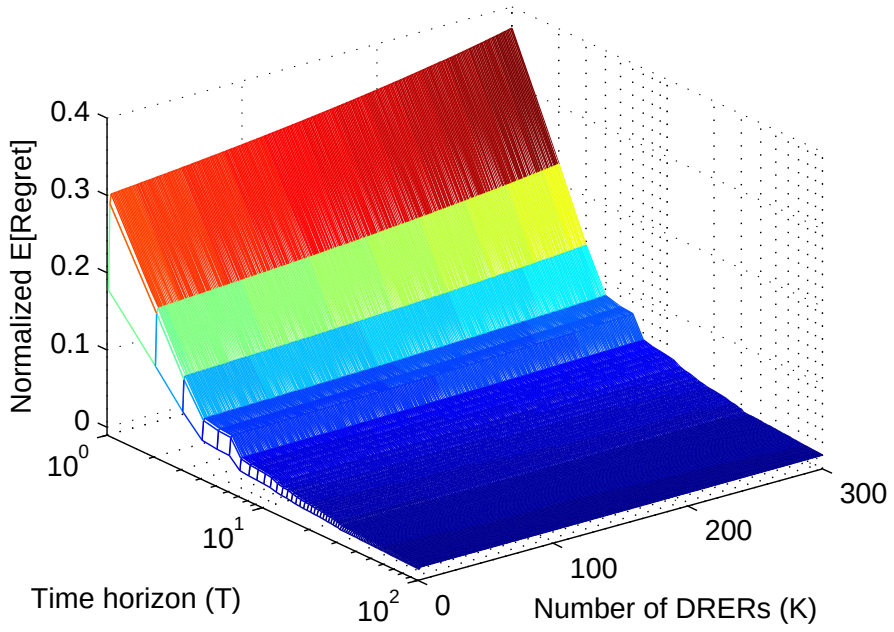


Figure 5.5 *The performance analysis for the proposed algorithm.*

where

$$A(x_{i,t-1}) = \Big((\bar{x}_{i,t-1} - \bar{\mu}_1)'(E[x_{i,t-1}\bar{x}'_{i,t-1}])^{-1}(x_{i,t-1} - \mu_1) - (\bar{x}_{i,t-1})'(E[x_{i,t-1}\bar{x}'_{i,t-1}])^{-1}(x_{i,t-1}) \Big). \tag{5.15}$$

Recall that the unknown complex variable μ_1 is still in (5.15).

One of common approaches for solving the unknowns is the generalized likelihood ratio test (GLRT), which requires storing the estimated data and ML-estimating the unknown at every point. Thus, in practice, the GLRT is too difficult from the view points of hardware and software implementation. Moreover, the work in [134] states that Rao test might be more robust than the GLRT in the presence of real operating situations and require a smaller computational complexity than the GLRT for their implementation. In [135], the authors show that performance of the Rao test based detectors works better than that of GLRT in terms of the parameter estimation and handling training-free scenario.

The Rao test [136] is the asymptotically equivalent the model of GLRT. The Rao test also does not involve the complex computation like the ML estimation does. The derivation of Rao test is similar to the locally most powerful (LMP) test but much simpler. The Rao test has the straightforward calculation by taking derivative with respect to the unknown evaluated around the region of interests. The solution of $A(x_{i,t})$ can be described as

$$\begin{aligned} &\frac{\partial[A(x_{i,t})|\mu_1]}{\partial \mu_1}\bigg|_{\mu_1=\mathbf{0}}^{T} \times \\ &\left[\mathbf{J}^{-1}(\mu_1)\bigg|_{\mu_1=\mathbf{0}}\right]\frac{\partial[A(x_{i,t})|\mu_1]}{\partial \mu_1}\bigg|_{\mu_1=\mathbf{0}}, \end{aligned} \tag{5.16}$$

where $\mathbf{J}$ is the Fisher information matrix [137] of $[A(x_{i,t})|\mu_1]$, and the solution is a two-side test in a quadratic form. We analyze the case in which the region is around zero due to q_0 has zero mean. In addition, we are able to consider one-side Rao Test (μ_1 is always greater than zero) and neglect the inverse Fisher information matrix (it is usually demanded in the multi-parameter environment). The solution of (5.16) leads to the derivative of a function of a complex variable and its conjugate. Given our complex variables x_i and μ_1 are formulated as $(x_r + jx_g)$ and $(\mu_r + j\mu_g)$, respectively.

However, $A(x_{i,t})$ not in terms of μ_r and μ_g, but in terms of μ and $\bar{\mu}$. By knowing that $\mu_r = (\mu + \bar{\mu})/2$ and $\mu_g = (\mu - \bar{\mu})/2j$, we use the chain rule to take the derivative of the function with respect to $\bar{\mu}$ while treating μ as a constant as shown below

$$\begin{aligned} \frac{\partial A(x_{i,t})}{\partial \bar{\mu}}\bigg|_{\mu} &= \frac{\partial \mu_r}{\partial \bar{\mu}}\bigg|_{\mu}\frac{\partial A(x_{i,t})}{\partial \mu_r}\bigg|_{\mu_g} + \frac{\partial \mu_g}{\partial \bar{\mu}}\bigg|_{\mu}\frac{\partial A(x_{i,t})}{\partial \mu_g}\bigg|_{\mu_r} \\ &= \frac{1}{2}\left(\frac{\partial A(x_{i,t})}{\partial \mu_r} + j\frac{\partial A(x_{i,t})}{\partial \mu_g}\right). \end{aligned} \tag{5.17}$$

By integrating the knowledge of (5.16) and (5.17), we can solve the unknown μ_1 in $A(x_{i,t})$ and denote $B(x_{i,t})$ to this function as

$$B(x_{i,t}) = -\frac{x_{i,t}}{2||x_{i,t}||}C(x_{i,t}), \tag{5.18}$$

where $C(x_{i,t})$ is $(2x_{r,t} + x_{g,t})/(x_{i,t}||x_{i,t}||)$.

On the one hand, we first define the upper bound

$$\begin{aligned}\ln \frac{\Lambda_T}{\Lambda_0} &\geq \ln \frac{\max_{j=1,2,\cdots,K} L_{j,T}}{K} \\ &= \min_{j=1,2,\cdots,K} \frac{x_{j,t} C(x_{j,T})}{2||x_{j,t}||} - \ln(K). \end{aligned} \tag{5.19}$$

On the other hand, we define the lower bound

$$\begin{aligned}\ln \frac{\Lambda_t}{\Lambda_{t-1}} &= \ln \frac{\sum_{i=1}^{K} \exp\left(-B(x_{i,t})\right) L_{i,t-1}}{\sum_{j=1}^{K} L_{j,t-1}} \\ &= \ln \sum_{i=1}^{K} P_{i,t} \exp\left(-B(x_{i,t})\right) \\ &\leq \sum_{i=1}^{K} P_{i,t} \left(1 - B(x_{i,t}) + \frac{B(x_{i,t})^2}{2}\right), \end{aligned}$$

$$\leq \frac{x_{i,t}}{2||x_{i,t}||} \sum_{i=1}^{K} P_{i,t} C(x_{i,t}) - \frac{x_{i,t}^2}{8||x_{i,t}||^2} \sum_{i=1}^{K} P_{i,t} C(x_{i,t})^2. \tag{5.20}$$

where the exponential function can be approximated in series of $1 + \frac{C}{1!} + \frac{C^2}{2!} + \cdots$ (we take first three terms for approximating exponential functions). Now, while start sampling, we can sum the upper bound over $t = 1, 2, \cdots, T$ and combine with the lower bound

$$\begin{aligned}&\sum_{t=1}^{T} \sum_{i=1}^{K} P_{i,t} C(x_{i,t}) - \min_{j=1,2,\cdots,K} \sum_{t=1}^{T} C(x_{j,t}) \\ &\leq \frac{2\ln(K)||x_{i,t}||}{x_{i,t}} - \frac{x_{i,t}}{4||x_{i,t}||} \sum_{t=1}^{T} \sum_{i=1}^{K} P_{i,t} C(x_{i,t})^2. \end{aligned} \tag{5.21}$$

Recalls that if the user selects same strategy i as the best strategy at the minimum stopping time T, then we will like to compensate $C(x_{j,t})$ which is unlikely be chosen via dividing the actual $C(x_{j,t})$ over P_i^u of choosing this strategy. It can be described as

$$\hat{C}_{x_{j,t}} = \frac{C(x_{j,t})}{P_{j,t}} \mathbf{1}_{i_t = j}, \tag{5.22}$$

and (5.21) can be elaborated in term of (5.22) to

$$\begin{aligned}&\sum_{t=1}^{T} \sum_{i=1}^{K} P_{i,t} \hat{C}_{x_{i,t}} - \min_{j=1,2,\cdots,K} \sum_{t=1}^{T} \hat{C}_{x_{i,t}} \\ &\leq \frac{2\ln(K)||x_{i,t}||}{x_{i,t}} + \frac{x_{i,t}}{4||x_{i,t}||} \sum_{t=1}^{T} \sum_{i=1}^{K} P_{i,t} \hat{C}_{x_{i,t}}^2. \end{aligned} \tag{5.23}$$

Taking the expectations of both sides of (5.23) and noting that

$$\sum_{i=1}^{K} P_{i,t}\hat{C}_{x_{i,t}} = C_{x_{i,t}}, \tag{5.24}$$

and

$$E_t\left[\sum_{i=1}^{K} P_{i,t}\hat{C}^2_{x_{i,t}}\right] = P_{i,t}\left(\frac{C(x_{i,t})}{P_{i,t}}\right) E_t\left[\mathbf{1}_{i_t=i}\right] = C^2_{x_{i,t}}, \tag{5.25}$$

yields

$$E_t\left[\hat{C}_{x_{i,t}}\right] - \min_{j=1,2,\cdots,K} E_t\left[C_{x_{j,t}}\right] \leq \frac{2\ln(K)||x_{i,t}||}{x_{i,t}} - \frac{x_{i,t}KT}{4||x_{i,t}||}. \tag{5.26}$$

Finally, we treat the condition of (5.26) near zero; notes that the discretionary of $\{T, K\}$ implicitly imply to $\{n_i, n\}$ [120] [121] [122], which n is the number of played and n_i is the number of played on arm i so far. By solving $x_{i,t}$ on the left-hand side of (5.26), the confident interval $\mathfrak{I}_t$ can be formulated.

5.6 Appendix B

First, we denote L^*_t as the true optimal likelihood ratio measurement, which is $\max_{i\in[1,K]} \sum_{t=1}^{T} L^u_{i,t}$ in (5.4). Given $\hat{L}_{i,t}$ to the average likelihood ratio measurement. Consequently, as soon as suboptimal i has been exploited sufficiently, the size of $\mathfrak{I}_t$ in Lemma 5.1 becomes small enough to guarantee that

$$\hat{L}_{i,t} + \mathfrak{I}_t < L^*_t, \tag{5.27}$$

and arm i will be abandoned soon. On the other hand, the condition also holds that

$$\hat{L}_{i,t} < L_{i,t} + \mathfrak{I}_t. \tag{5.28}$$

By combining (5.27) and (5.28), arm i is not played as soon as

$$2\mathfrak{I}_t < L^*_t - L_{i,t}, \tag{5.29}$$

and it implicitly implies that arm i will be ruled out as soon as the number of played on arm i reaches

$$\frac{32\ln(T)||x_{i,t}||^2}{(L^*_t - L_{i,t})^2}. \tag{5.30}$$

In the view of (5.30), the expected regret $E\left[\mathbb{R}(T)\right]$ can be upper bounded by

$$\begin{aligned} E\left[\mathbb{R}^{u}(T)\right] &= L_t^* T - \sum_i L_{i,t} E[t_i], \\ &\leq \sum_{i:L_t^* > L_{i,t}} \frac{32\ln(T)||x_{i,t}||^2}{L_t^* - L_{i,t}}, \\ &\leq \frac{32\ln(T)||x_{i,t}||^2}{\sqrt{\ln(K)}}\sqrt{\frac{T}{K}}. \end{aligned} \tag{5.31}$$

Noticing the sum of denominator can be approximated to $\sqrt{K\ln(K)/T}$ for all suboptimal arms i [120] [122].

Chapter 6

Conclusion and Future Work

In this chapter we summarize the research that has been accomplished in this dissertation and discuss the future research that can be pursued. We describe our findings on energy management systems and significant impact that they can have on the efficiency of smart grid operation. We further discuss problems and solution methods that can be explored from the current point of research.

6.1 Summary and Conclusion

In this dissertation, efforts have been made to explore specifically in three areas: system status, security issue, and resource management in smart grid networks. The techniques that have been developed in this dissertation are listed as follows:

- A CUSUM-based defense strategy is proposed against the false data injection attack in smart grid networks. In comparison to classical approach, the advantages of the proposed CUSUM-based defense mechanism includes: it is able to tackle the unknown parameters in the probability density function of post change distribution via the low complexity approach; the decision-making of the proposed scheme for detecting attack is based on using multiple online samples/observations rather than using a single observation while maintaining a certain level of decision accuracy; and a Markov chain based approach is developed to analyze the proposed approach for performance guarantee. The accuracy of the analytical model and detection with performance guarantee are also discussed.

- A quickest estimation scheme is developed to determine the network topology as quickly as possible with given accuracy constraints from the dispersive environment. Unlike the conventional topology estimation requires a long process of status analysis that the sensor at each bus senses, collects, analyzes, and then finally, sends the status measurement to the control

center. The proposed algorithm helps detect and identify the topological error efficiently and promptly for smart grid state estimation via just using online power measurement, and furthermore, reduce on vulnerability on system failure. In addition, the proposed algorithm is software-based, which has the ability of avoiding the deployment of additional sensors and the cost of expensive hardware. A Markov chain based analytical model is also constructed to systematically analyze the proposed scheme for the on-line estimation.

- An energy profile allocation scheme for end-user is investigated that is capable of determining the best choice of energy profiles as few samples as possible for long-term usage under the accuracy constraint while balancing the exploration and exploitation. In other words, an online learning technique is developed to learn evolution of power pattern (i.e., taking into account the uncertainty and variability of energy source) in term of reliability overtime. We derive the close form for the confident interval and obtain an upper bound for the expected regret for the proposed scheme.

The proposed techniques and algorithms in [144] [145] [146] [147] can surely benefit the smart grid society. Modern power grid depends on smarter infrastructure, smarter management, and smarter protection systems in order to facilitate ubiquitous operations to utilities, end-users, and both essential services and modern conveniences. Therefore, conducting research into smart grid networks, which could be employed in many emerging applications and facilitate the usage of utility operations. The proposed technologies in this disseration concerning different aspects of smart grid issues, such as the cyber security issues, network topology problem, alternative renewable energy resource allocation can provide a lot of benefits to power grid society, and will enhance the grid reliability and stability, utility services, emission control, and end-user experience in enabling better communications access to the grid, which could potentially translate into effective efficient utility operations and better living environment for human being.

6.2 Future Work

This section includes the overview of future research and possible progressions for smart grid networks. Each chapter in this dissertation can be considered as a starting point for further investigation. We describe some of the potential topics and open problems that can be chosen for further advance research in smart grid network as well as other research fields.

6.2.1 Fault Detection for Fully-Distributed State Estimation in Smart Grid

The smart grid is enabled by the advances in sensing, communication, and actuation, power system estimation and control are likely to involve fast information gathering and processing device. Institutionally, electricity industry deregulation has led to the creation of many regional transmission organizations within a large interconnected power system. Hence, the distributed estimation and control becomes another important topic in power system operations. The transition towards more distributed estimation has motivated research for distributed observability analysis and bad date detection as well. Early work for centralized observability and bad data analysis can be found in many literatures. The recent study proposed that the distributed state estimator can perform observability analysis and bad data detection in a distributed manner. Common to all the aforementioned distributed state estimators is the fact that each local area operates independently without sharing any information among its neighbor areas, when processing observability, state estimation and bad data.

As number of measurements and sampling rate of measurements increase, the conventional approach for state estimation suffers from communication bottleneck and reliability problems inherent in systems with one coordination center. Another challenge is that the increasing need for wide area monitoring and control requires the state of the entire interconnection available to all the regional transmission organizations (e.g., for entire eastern interconnection). These challenge create the need for a more distributed approach to state estimation.

We can apply the fault detection algorithm in fully distributed state estimation that removes

the necessarity of the central coordinator. The objective is to show that it is possible to design fully-distributed schemes so that each node converges almost surely to the centralized sufficient statistic. By iteratively exchanging information with neighboring control areas, all local control areas can achieve an unbiased consensus of system-wide state estimation. The framework for performing distributed detection-estimation of $\mathbf{b}$, which is the useful information (e.g., error vector, attacker vector, or etc.). The power system in general, may be viewed as a collection of N substations/nodes. The number of N may be quite large depending on the region of interest. The observation of $\mathbf{Y}$ is a collection of local observation vectors, $\mathbf{Y} = [\mathbf{Y}_1^T, \ldots, \mathbf{Y}_N^T]^T$, where $\mathbf{Y}_n$ is the observation of the n^{th} node and is generated as

$$\mathbf{Y}_n = \mathbf{H}_n\mathbf{X} + \mathbf{Z}_n + \mathbf{b}_n, \tag{6.1}$$

where $\mathbf{H}_n$ corresponds local Jacobian matrix, $\mathbf{Z}_n$,. Notices that $\mathbf{b}_n$ is portions of noise and attack vector respectively influencing the measurements at node n.

6.2.2 Optimality of A Joint Attack Detection and State Estimation Algorithm in Smart Grid

As we mention in the previous chapter, when a attacker occurs in the power network, the ultimate objective of the network operator is beyond a reliable detection of the attack. In fact, detecting the attack will be used as an intermediate step towards obtaining a reliable estimate about the injected false data, which in turn facilitates eliminating the disruptive effects of the false data. By serving the purpose necessitates that the operator obtains the best estimate about the false data injected. Therefore, assuring good estimation performance is the core of estimation and detection problem in the smart grid networks.

To account for the significance of estimation quality, we can define an estimation performance measure and seek to the optimize it while ensuring satisfactory of the detection performance. We proposes to minimize the estimation-related cost subject to appropriate constraints on the tolerable levels of detection errors (MSR and FAR). This approach can provide the operator with the freedom to strike any desired balance between estimation and detection qualities. Once we decide that the

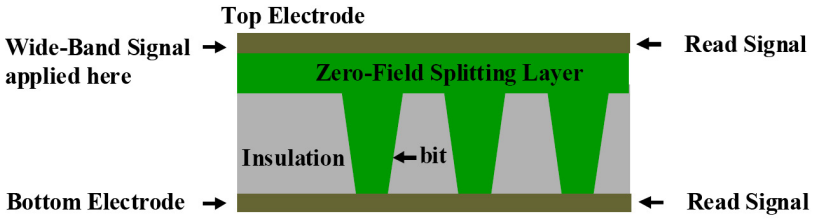


Figure 6.1 *The illustration of the frequency , which accesses the memory.*

observation $\mathbf{y}$ is drawn from hypothesis, an estimated $\hat{\mathbf{b}}(\mathbf{y})$ for $\mathbf{b}$. We capture the quality of the estimate by defining a cost function $C(\mathbf{b}, \hat{\mathbf{b}}(\mathbf{y}))$. By applying the minimum-square error (MSE), the estimation criteria can be described as

$$\text{MSE} : C(\mathbf{b}, \hat{\mathbf{b}}(\mathbf{y})) = ||(\mathbf{b} - \hat{\mathbf{b}}(\mathbf{y})||^2, \tag{6.2}$$

and we can formulate the estimation and detection problem for smart grid state estimation as following optimization problem

$$\begin{aligned} &\min_{\delta_0, \delta_1, \hat{\mathbf{b}}} \quad J(\delta_0, \delta_1, \hat{\mathbf{b}}), \\ &s.t. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.3) \\ &\qquad\qquad P_f \leq \alpha, \\ &\qquad\qquad P_m \leq \beta, \end{aligned}$$

where $J(\delta_0, \delta_1, \hat{\mathbf{b}}) = \frac{\mathrm{E}_{1,\mathbf{b}}[C(\mathbf{b}, \hat{\mathbf{b}}(\mathbf{y}))|\mathbf{1}_{\delta=H_1}]}{P_D} + c_1 P_f + c_2 P_m$ is the performance measurement that involves both estimation $\{\hat{\mathbf{b}}\}$ and detection qualities $\{\delta_0(\mathbf{y}), \delta_1(\mathbf{y})\}$. δ_e is the randomization probability for deciding in favor of the hypothesis H_e. $c_e > 0, e \in (0, 1)$, is the cost of falsely rejecting $\hat{H}_e$, P_D is the probability of detection of H_1 when H_1 is true,P_F is the probability of false-alarm, P_M is the probability of missing-detection.

6.2.3 Other Aspects of Quickest Detection Framework

In addition to the applications in the smart grid, a number of diverse research areas exist for the future work related to what has been presented in this dissertation. Applying the proposed quickest detection framework can be ranged from a single sensing device, modern biomedical exploration, to semiconductor industries.

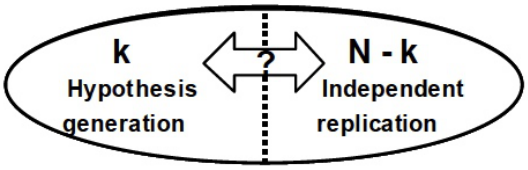


Figure 6.2 *The illustration of the genome model for replication*.

Multi-bits Detection in Memory: Users need to determine whether the memory cell absorbs the stored bits, which are produced by divergent frequencies (e.g., 8 GHz, 10 GHz, 12 GHz). Fig. 6.1 illustrates the reading the input frequency from the memory. Hence, the detection tool must determine one among several frequencies, who accesses the memory, and detect the stored bits as quickly as possible, since the obsolete information might not be useful to the users. The number of frequency bands are more than two. Two possible bits are carried by each individual frequency band; bit "0" contains no information, and bit "1" has the user data. Facing with the challenges discussed beforehand, we are aiming as the quickest detection on the stored bits with high accuracy and shortest observation time. Therefore, we can combine the multiuser detection and CUSUM algorithm together to achieve the goal of this problem; the multiuser detection technique is that decode the input data in the efficient and accurate way so that we know the specific frequency writes on the memory cell; and then, once the particular frequency is determined, CUSUM algorithm can make quickest detection on deciding the existing bit.

Quickest Genome Scan: Human genome is big, lonely place about 3.2 billion base pairs and each pair is 0.5 inches. If combining all pairs in human genome, it can laid end-to-end circle earth. Therefore, the data in the gene sequence is uncountable. The genome of an organism is a complete DNA sequence of one set of chromosomes; for example, one of the two sets that a diploid individual carries in every somatic cell; this includes both the genes and the non-coding sequences. Genome scan can help determining type of genome, replicate the genes in the genome, or etc. From the exclusive health news in *Businessweek* in the April 2010 edition, the genome scan gives human insight into future health risks; the genes information enabled doctors to deliver personalized health care like never before; patients at risk for certain diseases will be able to receive closer monitoring

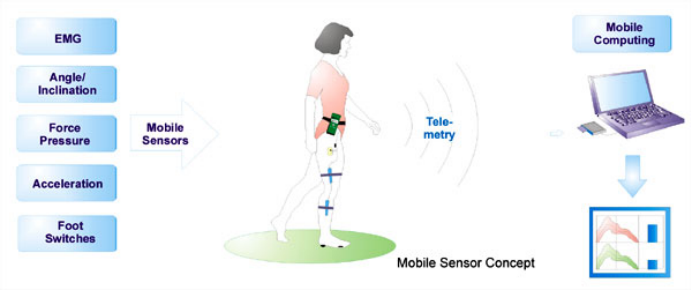


Figure 6.3 *The illustration of the smart sensor device on human.*

and more frequent testing, while those who are at lower risk would be spared unnecessary tests; this will have important economic benefits as well, because it improves the efficiency of medicine. However, the major problem in the traditional genome scan "multiple comparisons" can easily get faux signals by chance alone, and the problem causes the process inaccuracy, massive sample usage on the first stage and less sample left for the replication on the second stage as shown in Fig. 6.2. If k is too small, the replication process is never reached since it is low power to detect the strong signals of genes; if k is too large, it is low power to replicate the signals of genes. Therefore, k samples need to be carefully and accurately determined as quickly as possible for the next stage: genes replication. To solve the problem, we can use some characteristics of CUSUM; CUSUM algorithm is a type of non-Bayesian framework, which doesn't require the completed information from the input sequence; CUSUM algorithm also have the ability to detect the distribution change as little delay as possible at the random time. It means k is minimized and save as much $N - k$ for replication, at same time, without loss the constraints, and resulting the fairly high detection accuracy by maintaining both pre-specified type I and II errors over time.

Ubiquitous Computing on Biomedical Signal Monitoring: The goal of ubiquitous computing in biomedical signal processing is that builds computing systems for supporting and facilitating the daily lives of users, but being the least intrusive possible. The main objective of this application is to present a biomedical assistance environment for the elderly, with a device as a smart object. As shown in Fig. 6.3, the device includes embedded sensors to measure physiological parameters such as heart rate, respiratory rate, and body movement quantities such acceleration,

stand still, up and down. If the assistance is detected by the device such as falling on the ground, heart attack, or any emergency alarm, the device is able to communicate with personal cell-phone via blue-tooth for auto-dialing 911 or to the medic. There are many technological advances reported in literature, but the current scenario is still far away from an everyday life fulfilled with ubiquitous systems. The device requires to detect the sudden change based on the analysis of multi-parameters, so that the first-response medicine can come to help the patient as soon as possible. By applying the adaptive CUSUM algorithm with constraints optimization, which make the quick criterion for deciding the corrected action (i.e., report 911 or not), the detection ability can be accurate, precise, and quick to prevent the further serious injury on the patient.

Bibliography

[1] C. E. Report, "http://www.consumerenergyreport.com/smart-grid/," 2010.

[2] S. G. I. Clearinghouse, "http://www.sgiclearinghouse.org," 2010.

[3] N. I. of Standards and Technology, "Nist framework and roadmap for smart grid interoperability standards," *National Institute of Standards and Technology Report, release 1.0*, no. 1, Jan 2010.

[4] W. W. Fund, "The energy report," in *Energy solutions: Renewable energy and sustainable energy report series*, 2010.

[5] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jul. 2010.

[6] U. D. of Energy (DOE), "The smart grid: an introducation," in *SmartGrid by Department of Energy*, 2010.

[7] K. Bullis, "The big smart grid challenges: regulations, privacy and security concerns, and other issues could hold back developments," *Technology Review Published by Massachusetts Institute of Technology*, vol. 1, pp. 123–130, Jul. 2009.

[8] ——, "How to hack the power grid for fun and profit?" *Technology Review Published by Massachusetts Institute of Technology*, vol. 1, pp. 112–120, 2010.

[9] E. M. Lightner and S. E. Widergren, "An orderly transition to a transformed electricity system," *IEEE Transcations on Smart Grid*, vol. 1, no. 1, pp. 3–10, Jun. 2010.

[10] M. P. D. Coll-Mayor and E. Lightner, "Future intelligent power grids: Analysis of the vision in the european union and the united states," in *Energy Policy*, 2007.

[11] R. G. Pratt, "Transforming the u.s. electric system," in *IEEE Conference on Power Electronics Specialists*, 2004.

[12] D. C. J. Taneja and P. Dutta, "Towards cooperative grids: Sensor/ actuator networks for renewables integration," in *IEEE Conference on Smart Grid Communication*, 2010.

[13] T. C. of the United States, "American recovery and reinvestment act of 2009," in *An Act of the Congress of the United States of America Publ. L. No. 111-5*, 2009.

[14] E. C. for Electrotechnical Standardization (CENELEC)., "Smart meters coordination group: Report of the second meeting held on 2009-09-28 and approval of sm-cg work program for ec submission," 2009.

[15] S. C. OFFIS and M. management coaching., "Untersuchung des normungsumfeldes zum bmwi-foerderschwerpunkt e-energy - ikt-basiertes energiesystem der zukunft," 2009.

[16] B. Initiativ, "Internet of energy - ict for energy markets of the future," 2008.

[17] S. G. C. of China, "Sgcc framework and roadmap for strong and smart grid standards," in *whitepaper*, 2010.

[18] Japan, "Japan roadmap to international standardization for smart grid and collaborations with other countries," 2010.

[19] R. B. J. G. M. S. T. S. M. Uslar, S. Rohjansand and T. Weidelt, "Survey of smart grid standardization studies and recommendations - part 2," in *IEEE Conference on Power Electronics Specialists*, 2010.

[20] IEEE, "P2030/d7.0 draft guide for smart grid interoperability of energy technology and information technology operation with the electric power system (eps), and end-use applications and loads," 2011.

[21] S. S. G. S. Group, "Iec smart grid standardization roadmap," 2010.

[22] I. C. on Large Electric Systems (CIGRE), "Cigre d2.24 ems architectures for the 21st century," 2009.

[23] Microsoft, "Smart energy reference architecture sera," 2009.

[24] G. X. X. Fang, S. Misra and D. Yang, "Smart grid - the new and improved power grid: A survey," *IEEE Journal on Communications Surveys and Tutorials*, vol. 1, no. 1, pp. 110–119, accepted for publication, 2012.

[25] G. F. M. S. J. I. O. A. C. I. P. A. M. C. A. T. O. V. Giordano, F. Gangale and I. Maschio, "Smart grid projects in europe: lessons learned and current developments," in *JRC Reference Reports, Publications Office of the European Union*, 2011.

[26] S. C. B. Akyol, H. Kirkham and M. Hadley, "A survey of wireless communications for the electric power system," in *Prepared for the U.S. Department of Energy*, 2010.

[27] T. Baumeister, "Literature review on smart grid cyber security," in *Technical Report*, 2010.

[28] H. E. Brown and S. Suryanarayanan, "A survey seeking a definition of a smart distribution system," in *North American Power Symposium*, 2009.

[29] L. L. S. Chen, S. Song and J. Shen, "Survey on smart grid technology," in *Chinese Power System Technology*, 2009.

[30] T. M. Chen, "Survey of cyber security issues in smart grids," in *Cyber Security, Situation Management, and Impact Assessment II; and Visual Analytics for Homeland Defense and Security II*, 2010.

[31] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," in *Computer Networks*, 2006.

[32] R. Hassan and G. Radman, "Survey on smart grid," in *IEEE Conference on Southeast Conference*, 2010.

[33] R. B. J. G. M. S. T. S. S. Rohjansand, M. Uslar and T. Weidelt, "Survey of smart grid standardization studies and recommendations," in *IEEE Conference on Smart Grid Communications*, 2010.

[34] D. P. Varodayan and G. X. Gao, "Redundant metering for integrity with information-theoretic confidentiality," in *IEEE Conference on Smart Grid Communication*, 2010.

[35] Y. X. W. Wang and M. Khanna, "A survey on the communication architectures in smart grid," in *IEEE Conference on Computer Networks*, 2011.

[36] Y. Yu and W. Luan, "Smart grid and its implementations," in *CSEE*, 2009.

[37] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jul. 2010.

[38] E. S. T. Platform, "Vision and strategy for europe electricity networks of the future,," in *European SmartGrids*, 2006.

[39] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18–28, Jul. 2010.

[40] S. M. Kaplan and F. Sissine, "Smart grid: Modernizing electric power transmission and distribution; energy independence, storage and security; energy independence and security act and resiliency; integra," in *U.S.A. Government series*, 2009.

[41] R. H. Lasseter and P. Paigi, "Microgrid: A conceptual solution," in *IEEE Conference on Power Electronics Specialists*, 2004.

[42] IEEE, "P2030/d7.0 draft guide for smart grid interoperability of energy technology and information technology operation with the electric power system (eps), and end-use applications and loads," 2011.

[43] J. Roncero, "Integration is key to smart grid management," in *IET-CIRED Seminar*, 2008.

[44] A. Ipakchi, "Implementing the smart grid: Enteriprise information integration," in *GridWise Grid-Interop Forum*, 2007.

[45] F. energy regulatory commission, "Assessment of demand response and advanced metering," in *Federal energy regulatory commission staff report*, 2010.

[46] D. G. D. C. M. Souryal, C. Gentile and N. Golmie, "methodology to evaluate wireless technologies for the smart grid," in *IEEE Conference on Smart Grid Communication*, 2010.

[47] M. McGranaghan and F. Goodman, "Technical and system requirements for advanced distribution automation," in *18th International Conference and Exhibition on Electricity Distribution*, 2005.

[48] J. N. H. Ferreira, L. Lampe and T. Swart, *Power line com- munications: Theory and applications for narrowband and broadband communications over power lines*. John Wiley and Sons, 2010.

[49] A. S. S. Galli and Z. Wang, "S. galli, a. scaglione, and z. wang. power line communications and the smart grid," in *IEEE Conference on Smart Grid Communications*, 2010.

[50] D. W. Rieken and M. R. Walker, "Ultra low frequency power-line communications using a resonator circuit," *IEEE Transcations on Smart Grid*, vol. 2, no. 1, pp. 41–50, Aug. 2011.

[51] C. L. J. Li and K. P. Schneider, "Controlled partitioning of a power network considering real and reactive power balance," *IEEE Transcations on Smart Grid*, vol. 1, no. 3, pp. 262–269, Jun. 2010.

[52] J. Y. N. Pavlidou, A. H. Vinck and B. Honary, "Power line communications: state of the art and future trends," *IEEE Communications Magazine*, vol. 4, no. 4, pp. 34–40, Aug. 2003.

[53] M. G. K. P. P. Parikh and T. S. Sidhu, "Opportunities and challenges of wireless communication technologies for smart grid applications," in *IEEE Power and Energy Society General Meeting*, 2010.

[54] M. Mouly and M. B. Pautet, "The gsm system for mobile communications," in *Telecom Publishing*, 1992.

[55] H. Holma and A. Toskala, *WCDMA for UMTS: Radio access for third generation mobile communications*. John Wiley & Sons, 2005.

[56] C. T. S. You and B. Poulsen, "A market-based virtual power plant," in *IEEE ICCEP*, 2009.

[57] S. Y. Hui and K. H. Yeung, "Challenges in the migration to 4g mobile systems," *IEEE Communications Magazine*, vol. 41, no. 12, pp. 54–59, Feb. 2003.

[58] D. P. S. McLaughlin and P. McDaniel, "Energy theft in the ad- vanced metering infrastructure," in *4th Workshop on Critical Information Infrastructures Security*, 2009.

[59] P. Donegan, "Ethernet backhaul: Mobile operator strategies & market opportunities," *Heavy Reading*, vol. 5, no. 8, Feb. 2007.

[60] S. Little, "Is microwave backhaul up to the 4g task," *IEEE microwave magazine*, vol. 10, no. 5, pp. 67–74, Feb. 2009.

[61] A. M. J. H. V. Bakker, M. Bosman and G. Smit, "Demand side load management using a three step optimization methodology," in *IEEE Conference on Smart Grid Communications*, 2010.

[62] S. Caron and G. Kesidis, "Incentive-based energy consumption scheduling algorithms for the smart grid." in *IEEE Conference on Smart Grid Communication*, 2010.

[63] M. N. C. Ibars and L. Giupponi, "Distributed demand manage- ment in smart grid with a congestion game," in *IEEE Conference on Smart Grid Communication*, 2010.

[64] A. H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120–133, Aug. 2010.

[65] A. G. D. O'Neill, M. Levorato and U. Mitra, "Residential demand response using reinforcement learning," in *IEEE Conference on Smart Grid Communication*, 2010.

[66] D. C. J. Taneja and P. Dutta, "Towards cooperative grids: Sensor/actuator networks for renewables integration," in *IEEE Conference on Smart Grid Communication*, 2010.

[67] F. R. Y. S. Bu and P. X. Liu, "Stochastic unit commitment in smart grid communications," in *IEEE INFOCOM 2011 Workshop on Green Communications and Networking*, 2011.

[68] S. H. L. L. Chen, N. Li and J. C. Doyle., "Two market models for demand response in power networks," in *IEEE Conference on Smart Grid Communication*, 2010.

[69] S. Hatami and M. Pedram, "Minimizing the electricity bill of cooper- ative users under a quasi-dynamic pricing model," in *IEEE Conference on Smart Grid Communication*, 2010.

[70] R. S. V. W. W. P. Samadi, A.-H. Mohsenian-Rad and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *IEEE Conference on Smart Grid Communications*, 2010.

[71] D. K. M. S. X. Z. S. Ghosh, J. Kalagnanam and E. Feinberg., "Incentive design for lowest cost aggregate energy demand reduction," in *IEEE Conference on Smart Grid Communications*, 2010.

[72] A. R. Metke and R. L. Ekl, "Security technology for smart grid networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 99–107, Aug. 2010.

[73] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE transactions on Security and Privacy*, vol. 7, no. 3, pp. 75–77, Feb. 2009.

[74] N. C. System, "Technical information bulletin 04-1," in *Supervisory control and data acquisition (SCADA) systems*, 2004.

[75] K. Moslehi and R. Kumar., "A reliability perspective of the smart grid." *IEEE Transcations on Smart Grid*, vol. 1, no. 1, pp. 57–64, Aug. 2010.

[76] J. Driesen and F. Katiraei, "Design for distributed energy resources," *IEEE Power and Energy Magazine*, vol. 6, no. 3, pp. 30–40, Feb. 2008.

[77] G. K. J. T. W. B. T. K. B. Kroposki, R. Margolis and D. Ton, "Renewable systems interconnection: Executive summary," in *Technical Report NREL/TP-581-42292, U.S. Department of Energy*, 2008.

[78] O. o. E. D. Department of Energy and E. Reliability, "Current cyber security issues," in *Study of security attributes of smart grid systems*, 2009.

[79] E. M. Lightner and S. E. Widergren, "An orderly transition to a transformed electricity system," *IEEE Transcations on Smart Grid*, vol. 1, no. 1, pp. 3–10, Jun. 2010.

[80] T. Baumeister, "Literature review on smart grid cyber security," in *Technical Report, Department of Computer Sicence, University of Hawaii*, 2010.

[81] A. A. C. W. Potter and K. Westric, "Building a smarter smart grid through better renewable energy information." in *IEEE Conference on PSCE*, 2009.

[82] M. Basseville, "Edge detection using sequentialmethods for change level in part ii: Sequential detection of a change in mean," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 1, pp. 29–40, 1981.

[83] R. Trivedi and R. Chandramouli, "Secret key estimation in sequential steganography," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 746–757, 2005.

[84] A. Cohen, *Biomedical Signal Processing*. CRC Press, 1987.

[85] R. G. M. I. M. M. C. Yang, Y.Y. Namgung and W. B. Clark, "Change detection on longitudinal data in periodontal research," *Journal of Periodontal Research*, vol. 28, no. 2, pp. 152–160, 1993.

[86] D. B. E. Andersson and M. Frisen, "Detection of turning points in business cycles," *Journal of Business Cycle Measurement and Analysis*, vol. 1, no. 1, pp. 93–108, 2004.

[87] D. B. E. Andersson and M. Frisyen, "Some statistical aspects of methods for detection of turning points in business cycles," *Journal of Applied Statistics*, vol. 33, no. 3, pp. 257–278, 2006.

[88] E. Andreou and E. Ghysels, "The impact of sampling frequency and volatility estimators on change-point tests," *Journal of Financial Econometrics*, vol. 2, no. 2, pp. 209–318, 2004.

[89] I. L. D.W. K. Andrews and W. Ploberger, "Optimal changepoint tests for normal linear regression," *Journal of Econometrics*, vol. 70, no. 1, pp. 9–38, 1996.

[90] H. V. Poor and Q. Hadjiliadis, *Quickest detection*. Cambridge University Press, 2008.

[91] R. A. Fisher, *Statistical methods for research workers*. Edinburgh: Oliver and Boyd, 1925.

[92] D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press, 2004.

[93] M. Basseville and I. Nikiforov, *Detection of abrupt changes: theory and applications*. Englewood Cliffs: Prentice-Hall, 1993.

[94] C. L. H. Li and H. Dai, "Collaborative quickest detection in ad hoc networks with delay constraint - part i: Two-node network," in *IEEE Conference on Information Sciences and Systems*, 2008.

[95] Y. C. J. Tang and W. Zhuang, "An analytical approach to real-time misbehavior detection in ieee 802.11 based wireless networks," in *IEEE International Conference on Computer Communications*, 2011.

[96] H. D. C. Li and H. Li, "Adaptive quickest detection with unknown parameters," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2009.

[97] H. Z. Y. Xin and S. Rangarajan, "Ssct: A simple sequential spectrum sensing scheme for cognitive radio," in *IEEE Global Communications Conference*, 2009.

[98] C. L. H. Li and H. Dai, "Quickest spectrum sensing in cognitive radio," in *IEEE Conference on Information Sciences and Systems*, 2008.

[99] A. T. V. Dragalin and V. Veeravalli, "Multihypothesis sequential probability ratio tests - part ii: Accurate asymptotic expansions for the expected sample size," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1366–1383, Jul. 2000.

[100] S. Fleisher and E. Shwedyk, "A sequential multiple hypotheses test for the unknown parameters of a gaussian distribution," *IEEE Transactions on Information Theory*, vol. 26, no. 2, pp. 255–259, Mar. 1980.

[101] A. T. V. Dragalin and V. Veeravalli, "Multihypothesis sequential probability ratio tests - part i: Aasymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, Nov. 1999.

[102] Z. H. E. Hossain and V. Poor, *Smart grid communications and networking*. Cambridge University Press, UK, 2012.

[103] S. L. T. Z. D. Kundur, X. Feng and K. L. Butler-Purry, "Towards a framework for cyber attack impact analysis of the electric smart grid," in *IEEE Conference on Smart Grid Communications*, 2010.

[104] R. J. T. O. Kosut, L. Jia and L. Tong, "Malicious data attacks on smart grid state estimation: attack strategies and countermeasures," in *IEEE Conference on Smart Grid Communications*, 2010.

[105] Y. M. L. Xie and B. Sinopoli, "False data injection attacks in electricity markets," in *IEEE Conference on Smart Grid Communications*, 2010.

[106] Q. W. R. B. Bobba, K. M. Rogers and H. Khurana, "Detecting false data injection attacks on dc state estimation," in *the First Workshop on Secure Control Systems*, 2010.

[107] R. J. T. O. Kosut, L. Jia and L. Tong, "On malicious data attacks on power system state estimation," in *In Proceedings of the 45th International Universities' Power Engineering Conference*, 2010.

[108] J. S. B. I. Matei and V. Srinivasan, "Trust-based multi-agent filtering for increased smart grid security," in *20th Mediterranean Conference on Control and Automation*, 2012.

[109] X. Y. G. X. J. Lin, W. Yu and W. Zhao, "On false data injection attacks against distributed energy routing in smart grid," in *IEEE/ACM 3rd International Conference on Cyber-Physical Systems*, 2012.

[110] C. L. M. Talebi and Z. Qu, "Enhanced protection against false data injection by dynamically changing information structure of microgrids," in *IEEE 7th Sensor Array and Multichannel Signal Processing Workshop*, 2012.

[111] P. L. A. M. Sasson, S. T. Ehrmann and L. S. V. Slyck, "Automatic power system network topology determination," *IEEE Transactions on Power Apparatus and Systems*, vol. 92, no. 1, pp. 610–618, Mar. 1973.

[112] K. A. R. T. E. DyLiacco and A. W. Weiner, "Network status analysis for real-time systems," in *The 8th PICA Conference*, 1973.

[113] P. Bonanomi and G. Gramberg, "Power system data validation and state calculation by network search techniques," *IEEE Transactions on Power Apparatus and Systems*, vol. 102, no. 1, pp. 238–249, Mar. 1983.

[114] K. C. L. R. L. Lugtu, D. F. Hacketl and D. D. Might, "Power system state estimation : Detection of topological errors," *IEEE Transactions on Power Apparatus and Systems*, vol. 99, no. 6, pp. 2406–2412, Nov. 1980.

[115] J. Lin and H. Pan, "A static state estimation approach including bad data detection and identification in power systems," in *IEEE Power Engineering Society General Meeting*, 2007.

[116] J. Chen and A. Abur, "Placement of pmus to enable bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 21, no. 8, pp. 53–61, Aug. 2006.

[117] S. C. Y. Z. J. B. D. He, C. Chen and M. Guizani, "Secure service provision in smart grid communication," *IEEE Communications Magazine*, vol. 50, no. 8, pp. 53–61, Aug. 2012.

[118] M. N. Y. L. S. X. Y. Zhang, R. Yu and S. Gjessing, "Cognitive machine-to-machine communications: Visions and potentials for the smart grid," *IEEE Network Magazine*, vol. 26, no. Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. Xie, and S. Gjessing3, pp. 6–13, June 2012.

[119] R. H. Lasseter and P. Piagi, "Microgrid: A conceptual solution," in *IEEE Conference on Power Electronics Specialists*, 2004.

[120] N. C.-B. P. Auer and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Journal on Machine Learning*, vol. 47, no. 2, pp. 235–422, May 2002.

[121] P. Auer, "Using confidence bounds for exploitation-exploration trade-offs," *ournal on Machine Learning*, vol. 3, no. 3, pp. 397–422, Mar. 2003.

[122] P. Auer and R. Ortner, "Ucb revisited: Improved regret bounds for the stochastic multi-armed bandit problem," *Periodica Mathematica Hungarica*, vol. 61, no. 1, pp. 1–11, May 2010.

[123] Y. F. P. Auer, N. Cesa-Bianchi and R. E. Schapire, "The non-stochastic multi-armed bandit problem," *SIAM Journal on Computing*, vol. 32, no. 1, pp. 47–77, May 2003.

[124] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1, pp. 100–115, Jul. 1954.

[125] K. A. C. Y. Huang, H. Li and Z. Han, "Defending false data injection attack on smart grid network using adapative cusum test," in *IEEE Conference on Information Sciences and Systems*, 2011.

[126] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Marcel Dekker, Inc., 2004.

[127] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. Wiley New York et al., 1996.

[128] C. E. M.-S. R. D. Zimmerman and R. J. Thomas, "Matpower steady-state operations, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[129] J. W. F. C. Schweppe and D. B. Rom, "Power system static state estimation," *IEEE Transactions on Power Apparatus and systems*, vol. 89, no. 1, pp. 120–135, Jan. 1970.

[130] A. L. Ott, "Experience with pjm market operation, system design, and implementation," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 528–534, May 2003.

[131] J. Casazza and F. Delea, *Understanding Electric Power Systems*. IEEE Press Understanding Science and Technology Series, A John Wiley and Sons, Inc., 2010.

[132] M. He and J. Zhang, "Fault detection and localization in smart grid: A probabilistic dependence graph approach," in *IEEE Conference on Smart Grid Communication*, 2010.

[133] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, Jul. 1971.

[134] A. D. Maio and S. Iommelli, "Coincidence of the rao test, wald test, and glrt in partially homogeneous environment," *IEEE Letter on Signal processing*, vol. 15, no. 1, pp. 385–388, Apr. 2008.

[135] K. J. Sohn, "Parametric tests for multichannel adaptive signal detection," *Ph.D Dissertation from Stevens Institute of Technology*, Dec. 2007.

[136] A. D. Maio, "Rao test for adaptive detection in gaussian interference with unknown covariance matrix," *IEEE transactions on signal processing*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[137] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[138] M. R. Chmielewski and J. W. Grzymala-BusseJ, "Global discretization of continuous attributes as preprocessing for machine learning," *International Journal of Approximate Reasoning*, vol. 15, no. 4, pp. 319–331, Nov. 1996.

[139] Y. Q. Chen and K. L. Moore, "Discretization schemes for fractional-order differentiators and integrators," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 3, pp. 363–367, Mar. 2002.

[140] D. Gamerman and H. F. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Boca Raton, FL: Chapman and Hall/CRC, 2006.

[141] A. Abur and A. G. Exposito, *Power system state estimation: Theory and Implementation*. Marcel Dekker Inc., 2004.

[142] F. F. Wu and W. R. Liu, "Detection of topology errors by state estimation," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 176–183, Aug. 1989.

[143] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach (Power Electronics and Power Systems)*. Springer, 1999.

[144] H. N.-R. Z. Z. H. H. L. Y. Huang, M. Esmalifalak and L. Song, "Bad data injection in smart grid: Attack and defense mechanisms," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 27–33, 2013.

[145] Y. C.-H. L. K. A. C. Y. Huang, J. Tang and Z. Han, "Real-time detection of false data injection in smart grid networks: An adaptive cusum method and analysis," *IEEE Transactions on Smart Grid: Cyber and Physical Security Systems*, vol. to appear, 2013.

[146] Y. C.-H. L. K. A. C. Y. Huang, M. Esmalifalak and Z. Han, "Adaptive quickest estimation algorithm for smart grid network topology error," *IEEE Journal on Systems: Special Issue on Smart Grid Communication*, vol. to appear, 2013.

[147] H. L.-W. C. Y. Huang, L. Lai and Z. Han, "Online quickest multiarmed bandit algorithm for distributive renewable energy resources," in *IEEE Conference on Smart Grid Communication*, 2012.