Security and Privacy in Smart Grid

by

Asmaa Abdallah

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Smart grid utilizes different communication technologies to enhance the reliability and efficiency of the power grid; it allows bi-directional flow of electricity and information, about grid status and customers requirements, among different parties in the grid, i.e., connect generation, distribution, transmission, and consumption subsystems together. Thus, smart grid reduces the power losses and increases the efficiency of electricity generation and distribution. Although smart grid improves the quality of grid's services, it exposes the grid to the cyber security threats that communication networks suffer from in addition to other novel threats because of power grid's nature. For instance, the electricity consumption messages sent from consumers to the utility company via wireless network may be captured, modified, or replayed by adversaries. As a consequent, security and privacy concerns are significant challenges in smart grid.

Smart grid upgrade creates three main communication architectures: The first one is the communication between electricity customers and utility companies via various networks; i.e., home area networks (HANs), building area networks (BANs), and neighbour area networks (NANs), we refer to these networks as customer-side networks in our thesis. The second architecture is the communication between EVs and grid to charge/discharge their batteries via vehicle-to-grid (V2G) connection. The last network is the grid's connection with measurements units that spread all over the grid to monitor its status and send periodic reports to the main control center (CC) for state estimation and bad data detection purposes.

This thesis addresses the security concerns for the three communication architectures. For customer-side networks, the privacy of consumers is the central concern for these networks; also, the transmitted messages integrity and confidentiality should be guaranteed. While the main security concerns for V2G networks are the privacy of vehicle's owners besides the authenticity of participated parties. In the grid's connection with measurements units, integrity attacks, such as false data injection (FDI) attacks, target the measurements' integrity and consequently mislead the main CC to make the wrong decisions for the grid.

The thesis presents two solutions for the security problems in the first architecture; i.e., the customer-side networks. The first proposed solution is security and privacy-preserving scheme in BAN, which is a cluster of HANs. The proposed scheme is based on forecasting the future electricity demand for the whole BAN cluster. Thus, BAN connects to the electricity provider only if the total demand of the cluster is changed. The proposed scheme employs the lattice-based public key NTRU crypto-system to guarantee the confidentiality and authenticity of the exchanged messages and to further reduce the computation and communication load. The security analysis shows that our proposed scheme can achieve the privacy and security requirements. In addition, it efficiently reduces the communication and computation overhead. According to the second solution, it is lightweight privacy-preserving aggregation scheme that permits the smart household appliances to aggregate their readings without involving the connected smart meter. The scheme deploys a lightweight lattice-based homomorphic crypto-system that depends on simple addition and multiplication operations. Therefore, the proposed scheme guarantees the customers' privacy and message integrity with lightweight overhead.

In addition, the thesis proposes lightweight secure and privacy-preserving V2G connection scheme, in which the power grid assures the confidentiality and integrity of exchanged information during (dis)charging electricity sessions and overcomes EVs' authentication problem. The proposed scheme guarantees the financial profits of the grid and prevents EVs from acting maliciously. Meanwhile, EVs preserve their private information by generating their own

pseudonym identities. In addition, the scheme keeps the accountability for the electricity-exchange trade. Furthermore, the proposed scheme provides these security requirements by lightweight overhead; as it diminishes the number of exchanged messages during (dis)charging sessions. Simulation results demonstrate that the proposed scheme significantly reduces the total communication and computation load for V2G connection especially for EVs.

FDI attack, which is one of the severe attacks that threatens the smart grid's efficiency and reliability, inserts fake measurements among the correct ones to mislead CC to make wrong decisions and consequently impact on the grid's performance. In the thesis, we have proposed an FDI attack prevention technique that protects the integrity and availability of the measurements at measurement units and during their transmission to the CC, even with the existence of compromised units. The proposed scheme alleviates the negative impacts of FDI attack on grid's performance. Security analysis and performance evaluation show that our scheme guarantees the integrity and availability of the measurements with lightweight overhead, especially on the restricted-capabilities measurement units.

The proposed schemes are promising solutions for the security and privacy problems of the three main communication networks in smart grid. The novelty of these proposed schemes does not only because they are robust and efficient security solutions, but also due to their lightweight communication and computation overhead, which qualify them to be applicable on limited-capability devices in the grid. So, this work is considered important progress toward more reliable and authentic smart grid.

Acknowledgements

First and foremost, I am grateful to Almighty God for His countless blessings and for giving me the knowledge and strength to accomplish my research work.

I would like to express my deep gratitude to my supervisor, Prof. Sherman (Xuemin) Shen. Thank you for your continuous guidance, support, and encouragement throughout my PhD study. Without your guidance, advice, and valuable input on my research ideas and writings, this work would have not been possible.

I gratefully acknowledge my PhD committee members, Prof. Hossam Hassanein, Prof. Xinzhi Liu, Prof. Kankar Bhattacharya, and Prof. Liang-liang Xie for their valuable comments and insightful suggestions that helped increase the quality of the thesis.

I sincerely thank my colleagues at the Broadband Communications Research (BBCR) group. Thank you for your friendship, support, and beneficial discussions, especially in the security and smart grid BBCR subgroups.

My deepest appreciation and grateful thanks go to my parents, brother, and sisters for their prayers, support, and care. Your love and encouragement have been and will always be a great source of inspiration in my life.

To my Parents.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

EVs . . . . . Electric Vehicles
PLC . . . . . Power Line Carrier
PMUs . . . . Phasor Measurement Units
HANs . . . . Home Area Networks
NANs . . . . Neighbourhood Area Networks
V2G . . . . . Vehicle-to-Grid Networks
WAN . . . . . Wide Area Network
BANs . . . . Building Area Networks
IANs . . . . . Industrial Area Networks
CC . . . . . Control Center
DoS . . . . . Denial-of-Service Attacks
FDI . . . . . False Data Injection Attacks
DGs . . . . . Distributed Generators
QoS . . . . . Quality of Service
WMN . . . . Wireless Mesh Network
WAMS . . . . Wide-Area Measurement System
SCADA . . . Supervisory Control and Data Acquisition Systems
BEVs . . . . . Battery Electric Vehicles
PHEVs . . . . Plug-in Hybrid Vehicles
ICS . . . . . Industrial Control System
RTUs . . . . . Remote Terminal Units
HMI . . . . . Human Machine Interface
MMSE . . . . Minimum Mean Squared Error
LRT . . . . . Likelihood Ratio Test
PKI . . . . . Public Key Infrastructure
KP-ABE . . . Key-Policy Attribute-based Encryption
ABE . . . . . Attribute-based Encryption Scheme
EPPDR . . . Efficient Privacy-Preserving Demand Response Scheme
EAP . . . . . Extensible Authentication Protocol
IBC . . . . . Identity-based Cryptography Scheme
SSS . . . . . Shamir Secret Sharing Scheme
CS . . . . . Cramer-Shoup Cryptosystem
ECC . . . . . Elliptic Curve Cryptography

UBAPV2G . Unique Batch Authentication Protocol for V2G Communications
CA . . . . . Central Authority
TPM . . . . Trusted Platform Module
DSSS . . . . . Direct Sequence Spread Spectrum
GLRT . . . . Generalized Likelihood Ratio Test
CUSUM . . . Cumulative Sum Control Chart Test
LMP . . . . . Locational Marginal Price
LR . . . . . Load Redistribution Attack
TA . . . . . Trusted Authority
SVP . . . . . Shortest Vector Problem
LWE . . . . . Learning With Error Problem
NSS . . . . . NTRU Signature Scheme
BSs . . . . . Base Stations
SMs . . . . . Smart Meters
APs . . . . . Smart household Appliances
LAs . . . . . Local Aggregators
AP . . . . . Access Point
CSs . . . . . Charging Stations
PIDs . . . . . Pseudonym IDs
LS . . . . . Local Substation
SE . . . . . State Estimator
MUs . . . . . Measurement Units

Chapter 1

Introduction

Over many decades, the electricity grid, with approximately the same infrastructure, has generated electrical power from central generator plants resources, such as oil, hydro, and gas. In addition, it has distributed the energy in one direction from generation stations to end-users. In the early twenty-first century, scientists are utilizing the great progress in information and communication technology and merging it with the regular electricity grid to form the smart grid. The main target of this incorporation is to connect the components of the electricity grid via communication networks, such as Internet or sensor networks, to gather data about grid conditions and consumers' requirements.

This integration accordingly enhances the electricity distribution and diminishes the huge waste in the ordinary grid; the grid operator can estimate the peak time, when the demand for electricity is high, and the off time, when the demand is low, and exploits this knowledge to regulate the electricity consumption and reduce the losses by two ways. The first one is to encourage the customers to use their high-consumption electrical appliances during the off period by reducing the electricity price during that period. Second, the electric vehicles (EVs) can work as temporary storages for the power; accordingly, the electricity will be charged from the grid to the vehicle in the off period; in contrast, the electricity will be discharged from the vehicle to the grid in the peak period. In next section, we define the smart grid, its main objectives, the new deployed devices, and the different techniques used in it.

1.1 Smart Grid Definition

Smart grid utilizes the communication technologies to improve the power fault detection, reduce the electrical waste and enhance the self-healing feature of the grid. To fulfill these characteristics, various communication techniques, such as power line carrier (PLC), WiFi, ZigBee, and Internet, are utilized to connect different parts of the grid, i.e., generation, distribution, transmission, and consumption systems, together. Then, smart grid utilizes this connection to exchange information about grid conditions between participated parties; also, it allows the electricity flow between the service provider and customers in both directions.

The main targets of smart grid are to reduce the power losses and maintain the electricity generation-consumption ratio to stabilize the power grid, e.g., the electricity amount in the grid is around a certain level all the time. Consequently, the efficiency of electricity generation increases. This upgrade requires merging sensors and measurement devices, such as smart

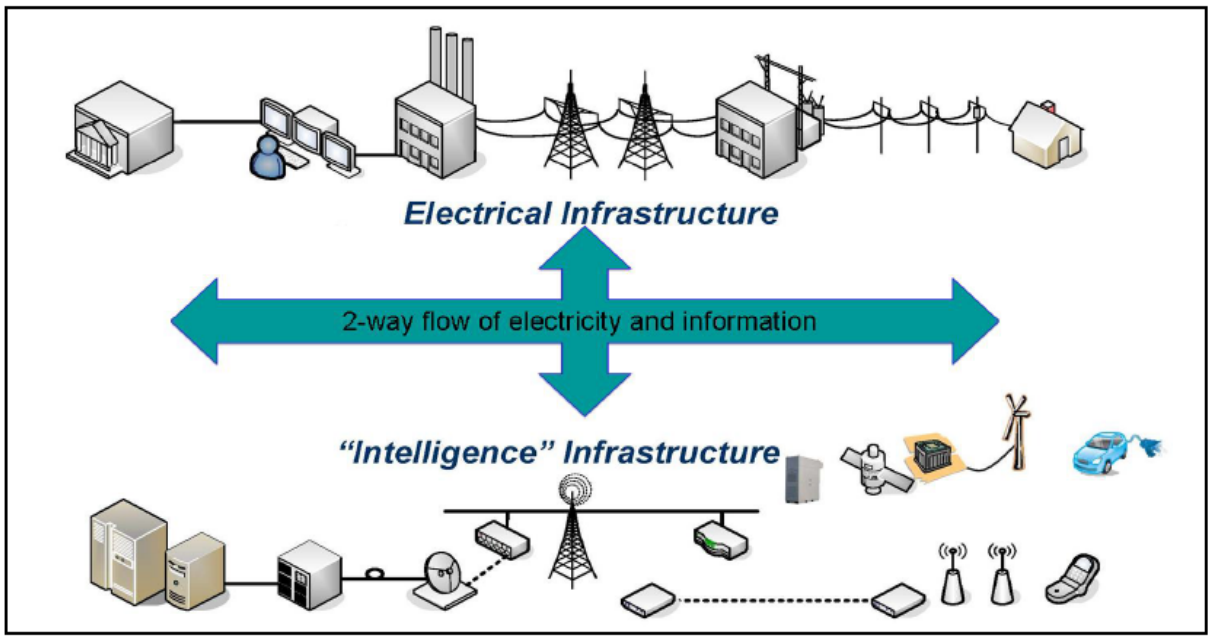

Figure 1.1: Smart grid.[1]

meters and phasor measurement units (PMUs), in the power grid. The main functions for these meters are aggregating information about the power grid status in different zones. According to smart meters, they are deployed in the customer side networks, e.g., houses and industrial buildings to aggregate the electricity consumption for each individual appliance and send the total consumption to the utility. Then, the utility calculates the accurate electricity bill for each customer. While, PMUs are high-speed secured sensors distributed throughout the grid transmission lines. They monitor the grid status and quickly detect any anomaly behaviours and threats that could lead to blackouts.

Smart grid correspondingly introduces new types of networks, such as home area networks (HANs), neighbourhood area networks (NANs), and vehicle-to-grid (V2G) networks, also exploits the existing wide area network (WAN) infrastructure. The HAN is a connection between the house's smart meter and the smart household appliances, in which the smart meter aggregates the electricity readings for the appliances and forwards them to the utility company for billing purposes; the smart meter sends its aggregated readings periodically, e.g., every 15 minutes. The building area networks (BANs) and the industrial area networks (IANs) are types of HANs that connect several HANs in the same residual or industrial area to the local substation. While each NAN forwards the periodic reports for the HANs in its region to the utility control center (CC) via the existing WAN. According to the V2G connection, the grid operator communicates with the EVs to charge/discharge them according to the current electricity demand; the main function of these connections is scheduling the EV charging/discharging operations to avoid the uncoordinated actions that cause negative impacts on the grid. For the CC's connection with the measurement units, it is significant to transfer real-time information about grid status and consumers' requirements to help CC to make the right decision for the grid. Figure 1.1 shows smart grid as integration between power grid and communication architectures [1, 2, 3, 4, 5, 6, 7, 8, 9].

Although these networks provide noteworthy communication services to the different parties

[1]Picture source: http://www.nist.gov/itl/antd/emntg/smartgrid.cfm.

in the grid, they expose the power grid to the cyber security threats that communication networks suffer from. So, smart grid could be threatened by malicious adversaries, who attempt to intercept, modify, or block the exchanged electrical information, or by dishonest consumers, who send fake readings to reduce their payment amount or try to steal the neighbours' electrical shares. In addition, the power grid becomes vulnerable to the thorough attacks that target the grid infrastructure and cause huge blackouts. In next section, we briefly introduce the main security concerns in smart grid.

1.2 Smart Grid Security Concerns

The upgrade of power grid exposes it to the cyber security threats that communication networks suffer from, such as malicious attacks that can forge the electricity consumption readings, extract personal information from the readings, or establish Denial-of-Service (DoS) attacks. In addition, novel threats are introduced to the grid because of its special nature, such as false date injection (FDI) attacks that inject fake information about grid's status to mislead the CC to make wrong decisions that impact negatively on the grid stability and reliability. The security concerns in smart grid can be categorized into three major groups [1, 4, 8, 9, 10, 11, 12, 13, 14, 15, 16]:

- The users privacy and information confidentiality are significant concerns in smart grid especially for customers, i.e., in houses, residential units, and industrial buildings. Personal information and daily habits can be revealed to the outsiders from analyzing the electricity consumption pattern. Therefore, any eavesdropper, who has moderate data analysis tools, can threaten the customers' privacy and extract vital information about the householders, such as when they are in the house, what are the types of electrical appliances they usually use. Thus, the attacker can break the house while the householders are outside. According to the industrial institutes, the adversary may get important data about the institutions production from its consumed power and sell this information to the competitors. In conclusion, the privacy for electricity users and exchanged information confidentiality are fundamental factors that should be considered in any proposed security schemes for smart grid.

- Second, the data integrity is a serious concern. As adversaries may attempt to alter or fabricate the exchanged messages between different parties in the grid. There are two obvious examples of integrity attacks in smart grid. First one is when a malicious consumer attempts to forge his/her electricity reading by compromising the smart meter to reduce the electricity bill. Second attack is the FDI attack, in which the adversary compromises several measurement units and exploits them to inject false information about the grid conditions; this attack misleads the CC to make improper decisions for the grid. The consequences of this attack extremely affect all the involved parties in the grid.

- Third concern is the grid's resources availability. Malicious adversaries can target the network resources, e.g., by DoS attacks. They attempt to block, delay or corrupt the transmitted information or make it unavailable to the different parties in the grid. For instance, an attacker can falsify a large number of electricity request messages via a group of compromised smart meters and ask for huge amount of energy trying to make

the utility CC unavailable to the legitimate users. Consequently, smart grid networks should be robust to network availability attacks, because the network unavailability could lead to severe consequences, such as loss the real-time monitoring of the critical power infrastructures, which subsequently lead to large-scale power system disasters, i.e., huge blackouts.

In summary, merging communication technologies in power grid exposes the grid to unfamiliar security problems that imported from the communication networks in addition to new threats due to the power grid nature. In general, the main security risks for smart grid are network resources and information availability, data integrity, and customers privacy and data confidentiality. Therefore, we study the security threats regarding the three different scenarios for smart grid networks and propose efficient security techniques to solve these problems or alleviate their impacts on the grid's performance. The succeeding section presents our objectives and the thesis motivations.

1.3 Motivation and Objectives

The thesis addresses the security threats in the three different communication architectures of smart grid and proposes efficient security and privacy-preserving schemes:

- The first part of the thesis addresses the security threats during the electricity readings aggregation at the customer side at HANs, BANs, and IAN, e.g., in houses, residential buildings, or industrial buildings, additionally, the problems during transmitting these aggregated readings via NANs to the utility for billing purposes. Mainly, the privacy of consumers is the central concern for these networks; also, the transmitted messages integrity and confidentiality should be guaranteed. Furthermore, the authenticity of different connected parties must be assured.

  Many research works are proposed to addresses these security concerns especially preserving-privacy mechanisms; the suggested approaches can be divided into three classes. First class solutions are deploying hardware devices, such as EVs or temper-resistance devices, to conceal the real consumption of the place. However, these solutions are expensive to deploy with every device in the grid. While second class studies suggest applying signal distortion function on the electrical reading signal at the transmitter, then reconstructing the original signal again at the receiver. But this operation requires difficult computations and complex data mining techniques. Finally, cryptographic security schemes are used to guarantee the security and privacy of the connected parties and the exchanged messages. Some privacy-preserving public key systems are proposed to conserve the customers' privacy in addition to the confidentiality and integrity of the exchanged messages. These privacy schemes are based on two main cryptographic ideas: homomorphic encryption and anonymization techniques. However, most of the proposed schemes are high computation and communication complexity, and the deployed devices in customer-side networks smart grid, i.e., smart meters and smart appliances, are limited computation capabilities and restricted power abilities [45, 48, 50, 61, 70].

  Consequently, we have proposed two lightweight security frameworks to guarantee the privacy and security requirements for customer side networks. The first proposed scheme is based on forecasting the electricity demand for a cluster of HANs in a residential area.

Utilizing the NTRU lattice-based crypto-system, the proposed scheme guarantees the confidentiality and authenticity of exchanged messages and to further reduce computation and communication load. While the second proposed solution is a lightweight security and privacy-preserving scheme that deploys lightweight lattice-based homomorphic crypto-system to secure the readings aggregation process inside HANs. Thus, the proposed scheme guarantees the customers privacy and messages integrity with lightweight overhead.

- The second part of the thesis analyzes the security threats for V2G connections. Generally, the main security concerns for V2G networks are the privacy of vehicle's owners besides the authenticity of different parties.

  Several research studies are proposed to conserve the demanded security level and bound the threats impact on V2G connection. These existing solutions are based on four different security procedures, such as encryption schemes, authentication mechanisms, physical security methods, and anonymization techniques. First group of proposed schemes utilizes public key schemes, such as elliptic curve cryptography (ECC) and symmetric key schemes, as advanced encryption standard (AES) to create a security framework that guarantees V2G security requirements. According to the second category, several researchers use different authentication mechanisms, such as Diffie-Hellman key exchange protocol, to assure the validity of vehicles and other related devices, i.e., aggregators. Third category is based on anonymizing the vehicles utilizing several anonymization mechanisms, such as partially-blind signature scheme, so that no one can like between the vehicles' identities and their charging behaviours or locations. According to physical layer protection mechanisms, various physical security techniques are utilized, e.g., using a channel-based key management to setup a symmetric key between two remote nodes to resist DoS and jamming attacks [72, 75, 77, 79, 83]. However, the proposed solutions in the literature provide high communication and computation overhead and require usage of special hardware devices.

  Thus, we have proposed lightweight secure and privacy-preserving V2G connection scheme, where the grid assures confidentiality and integrity of exchanged information during (dis)charging electricity sessions and overcomes EVs' authentication problem. The proposed scheme guarantees the financial profits of the grid and prevents EVs from acting maliciously. Meanwhile, EVs preserve their private information by generating their own pseudonym identities. In addition, the scheme keeps the accountability for electricity-exchange trade. Furthermore, the proposed scheme provides these security requirements by lightweight overhead; as it diminishes the number of exchanged messages during (dis)charging sessions.

- Finally, the third part of the thesis studies the FDI attacks and analyze the previously proposed detection techniques. FDI attacks target the measurements units that spread in the power grid to monitor its status. The objective of FDI attacks is to mislead the main CC to make the wrong decisions for the grid. Several solutions are proposed to mainly detect the presence of any malicious measurements; some solutions utilize more powerful state estimation tools for accurate bad data detection. While, many studies utilize various optimization techniques to detect the attacks and provide protection models. Certain cryptographic schemes, such as watermarking, are employed to detect FDI attacks [87, 88, 99, 106, 111, 122, 123, 129, 134]. However, most of the research works focus on detecting FDI attacks not resisting them.

In the thesis, we have proposed an FDI attack prevention technique that protects the integrity and availability of the measurements at measurement units and during their transmission to the CC, even with the existence of compromised units. The proposed scheme alleviates the negative impacts of FDI attack on grid's performance. Security analysis and performance evaluation show that our scheme guarantees the integrity and availability of the measurements with lightweight overhead.

1.4 Outlines of the Thesis

The thesis is organized as follows. Chapter 2 presents a detailed background about smart grid and its security concerns. Chapter 3 defines the main security and privacy problems in the three main communication architectures in smart grid and the related work in the literature. Chapter 4 presents the first proposed solution for security and privacy problems in customer-side networks and demonstrates its security analysis and performance evaluation. Chapter 5 describes the second proposed scheme, investigates its security features, and evaluates its performance. Chapter 6 introduces the proposed solution for security and privacy threats in V2G connections. Chapter 7 proposes the efficient prevention technique for FDI attacks. Finally, Chapter 8 concludes the thesis and indicates the future directions.

Chapter 2

Background and Literature Survey

Smart grid can be defined as the incorporation between the communication and information technology and the traditional power grid. It utilizes the networking techniques to exchange information about the grid conditions and customers' demands. The main target of this integration is to improve the power generation process and reduce the electricity losses. In addition, smart grid merges the renewable power resources with the traditional power generators to cover the increased electricity demand. Another benefit for smart grid is assisting in CO2 emissions reduction and environment protection. Additionally, more distributed generators (DGs) are inserted in smart grid to satisfy the high electricity demands; they mostly are renewable resources-based generators, such as wind turbines, and solar panels. Furthermore, original techniques, such as micro-grids and V2G connection, are utilized in smart grid. The micro-grid offers electrical self-sufficiency for a specific area using one or more DGs and storage units and allows the area to be isolated or connected to the main grid according to the current status of the grid; this feature protects the micro-grid in case of blackout and assists the self-healing of the grid. In addition, smart grid utilizes the EVs' batteries as temporary storage units for the extra generated power during low demand periods; V2G networks organize the charging/discharging operations of the EVs' batteries to guarantee balanced electricity level in the grid [1, 2, 3, 4, 5, 6]. This chapter defines the smart grid importance, describes its architecture, and briefly introduces its main security concerns.

2.1 Smart Grid Benefits

According to the service provider, i.e., utility companies, smart grid technology can significantly enhance the reliability and efficiency of the power grid. The grid's reliability means reducing the probability of blackouts and guaranteeing the required level of electricity supply to all customers. The electricity company is responsible for providing a specific electricity demand to each customer according to its type, i.e., residential or industrial. In case of electricity shortage, there will be huge financial and economic losses for the customer especially industrial ones and as consequence for the electricity company, which is obligated to pay a fine to the affected customer.

On the other hand, the grid's efficiency is to reduce the power losses; the efficiency can be satisfied by rearranging the electricity consumption patterns for users. For instance, the electricity company can encourage the residential customers to use their high-consumption appliances at the low peak load period by decreasing the power price at that period. In

addition, applying security scheme in smart grid can assist to decrease the power theft, which is an effective reason for electricity losses in many countries. As a result, the electricity generation will be organized and even reduced. Moreover, the insertion of renewable generation resources in the new smart grid can lessen the burden on the traditional plant generators.

Smart grid can improve the efficiency of the maintenance and replacement operations for the involved devices in the grid. For example, there are many deployed sensors in smart grid for monitoring purposes; they monitor the performance of the different devices and send an alarm message to the control center in case of error. Finally, smart grid is a friend to the environment, as it organizes the electricity production and uses renewable generation resources. Accordingly, smart grid plays a significant role in CO2 emission reduction. To conclude, the utility companies are interested in smart grid to assure the optimal usage of the electrical power and provide more luxury services to the customers, and consequently, increase their financial profits [1, 4, 5, 7, 8].

2.2 Smart Grid Architecture

Smart grid introduces new components and protocols in the power grid to achieve the smart grid's functions. This section introduces the smart grid's reference model, its different layers and their functions, and then the smart grid's systems.

2.2.1 Smart Grid Reference Model

There are many proposed frameworks to identify the structure of smart grid. According to [2], smart grid reference model composes of seven functional domains:

- *Bulk Generation:* Electricity is usually generated from non-renewable resources, such as coal and gas generators. In smart grid, the renewable sources, e.g., wind turbines and solar panels, are merged with the traditional ones to satisfy the increased demands and reduce CO2 emissions.
- *Transmission:* Several substations and transmission lines are utilized to transmit the produced power to consumers.
- *Distribution:* Distribution domain spreads the electricity to individual customers and communicates suppliers and users via communication infrastructure.
- *Operation:* This domain controls and monitors the transmission and distribution domains to obtain information about the power system's activities.
- *Market:* This domain contains all the parties involved in the electricity-trade operation to sustain the balance between supply and demand.
- *Customer:* Customers in smart grid not only consume electricity but also generate it by distributed generators and store the extra power in rechargeable batteries.
- *Service Provider:* The electricity is provided to customers via service provider that is responsible for services, such as billing and customer accounts management.

2.2.2 Smart Grid Layers

According to [3], smart grid is composed of five layers that arrange the involved parties in the grid:

- *Application layer:* provides smart grid applications for both customers and utilities.
- *Security layer:* satisfies the security requirements for all involved parties in smart grid.
- *Communication layer:* provides a two-way reliable and secure data transmission.
- *Power control layer:* monitors and controls the power transmission operation using PMUs, sensors, transformers, meters and storage devices.
- *Power system layer:* delivers the electricity to customers through power generation, transmission and distribution systems.

2.2.3 Smart Grid Systems

Smart grid differs from traditional power grid in several ways. The most important difference is that smart grid can exchange electricity and information about the grid conditions between suppliers and end users in both directions. The main purposes of this communication are to decrease the total consumption of electricity, preserve the demands for electricity approximately at the same level all the time, and consequently reduce the overall cost of this service. According to [1], smart grid is divided into smart infrastructure system, smart management system, and smart protection system.

2.2.3.1 Smart Infrastructure System

Smart infrastructure system supports the bidirectional flow of data and power; it consists of three subsystems: smart energy subsystem, smart information subsystem, and smart communication subsystem.

- *Smart energy subsystem* controls generation, delivery, and consumption of electricity. In traditional grids, the electricity is generated in few central huge-size power generators. Then, generated power is transmitted via transmission grid to substations, finally distributed to customers through distribution grid. So, it's a unidirectional process. In contrast, smart grid is bidirectional; it utilizes DGs, such as solar panels and wind turbines, to enhance the grid consistency. This expansion leads to two new concepts:

  Micro-grid, which is a small-scale grid with its own DGs and loads, has self-power sufficiency. Accordingly, micro-grid is able to disconnect from the original grid if any failure happened; micro-grid considers a small independent grid producing its own power. However, the communication with the grid did not fully disconnected; micro-grids still exchange information with the whole grid to decide when to reconnect with it.

  The other concept is V2G connection. EVs' batteries are charged from the grid at low-demand times and work as electrical storage. Still, the charging operation requires efficient scheduling techniques for coordinated charging to conserve the optimal power system

performance and keep the peak power demand at minimum level. The grid restores the power back from EVs in high-demand periods, i.e., EVs act as DGs and supply electricity back to the grid.

- *Smart information subsystem* is responsible for information measurement, grid status monitoring, and user appliances control. It deploys certain metering and measurement devices, such as smart meters devices that are part of the automatic metering infrastructure, which is a technology to collect the energy metering data for analysis and electricity billing purposes. The main function of smart meter is to compute the total amount of consumed electricity for a unit, e.g., a house, every specific interval and send the data to the central control for monitoring and billing purposes. Moreover, smart meter can control the appliances, i.e., connect or disconnect them, and manage the loads and future demands to reduce the electricity bill.

  In addition, smart monitoring and measurement devices, i.e., sensors and PMUs, are utilized. First, sensors are used to monitor the real-time mechanical and electrical conditions for power system in addition to analyze the failures if happened. Wireless sensor networks are strongly recommended to accomplish this mission because of its effective cost; however, sensors are low power nodes and vulnerable to attacks or severe environmental conditions. Second, PMUs are secure measurement devices that are based on measuring the phase angle of the power model to determine the power system's state. PMUs are utilized to forecast any failure before happening. Huge amount of information is generated from the metering devices; this data should be stored and analyzed to extract best benefit. Cloud computing is a good candidate for that huge information storage. However, cloud computing suffer from certain security and privacy threats in addition to the expensive cost of that service.

- *Smart communication subsystem* is responsible for exchanging the collected information among different devices in the grid utilizing both wired and wireless networks. Smart grid utilizes a combination of networks to guarantee the required QoS and support reliability, availability, and security and privacy requirements with low installation cost. Several communication technologies are suggested for smart grid, such as wireless mesh network, cellular communication system, IEEE 802.15.4-based technologies, satellite communications, fibre optics communications, and PLC. TCP/IP protocol is strongly candidate for managing the smart grid's communication subsystem; however, it should overcome many challenges related to the used heterogeneous communication networks and find low cost methods to smoothly upgrade the existing grid to become smart grid.

2.2.3.2 Smart Management System

Smart management system mainly concerns about management and control mechanisms for the novel applications and services in smart grid. The main functions of this system are energy efficient usage and cost optimization. Mainly, smart management system aims to smooth the demand profile shape by shifting and rescheduling the loads. Consequently, energy losses are minimized, overall generation cost is decreased, and system's reliability is increased. To satisfy these objectives, various optimization techniques can be exploited.

2.2.3.3 Smart Protection System

Smart protection system protects the grid from threats, which could be user errors, equipment failures, natural disasters, and cyber-attacks, by offering defense mechanisms and preserving security and privacy of the grid. In smart grid, DGs with their fluctuant and intermittent renewables resources could threaten the reliability and stability of the grid. Thus, smart grid uses micro-grids. Micro-grids deployment leads to less power flow within the entire grid, as loads are served locally within every micro-grid, which consequently reduces the possibility of cascading failures. Another way to guarantee the grid's reliability is to assure the consistency of measurement system by depending on powerful secured measurement devices.

To predict any failure occurrence in the grid, the PMUs' data is utilized to identify the stability region and predicates the weak points in the grid to identify the probability of failures and where they could happen. If the failure happened, the system's knowledge about topology and PMUs measurements helps to quickly identify and fix it and prevent cascading events. In other words, the grid satisfies the self-healing feature, which is the ability to prevent the spread of failures and quick recovery of the grid. For micro-grid protection, micro-grid can work in two modes: normal and island modes. In the normal mode, the micro-grid connects and exchanges electricity with the main grid. However, if any abnormal conditions, such as power failures, occur, the micro-grid switches to island mode, which isolates the micro-grid and stops the electrical flow with the main grid to protect the micro-grid customers and prevent cascade failure. So, micro-grid with the isolation capability improves the self-healing and increases the grid's ability to work well during normal times and outages as well.

In addition, cyber security is one of the serious challenges in smart grid. The cyber adversaries can compromise the power grid via communication systems to perform malicious actions, such as obtain user private information, gain access to CC, and alter load conditions to destabilize the grid in addition to new security and privacy issues due to the deployment of smart meters, sensors, and PMUs. For smart meters security, smart meter is the most vulnerable part in smart grid; it suffers from many security threats that can falsify the consumed electricity amount; for example, malicious users can compromise their smart meters to reduce their energy meter readings and pay lower electricity bills. Moreover, the extensive deployment of smart meters increases the opportunity for adversaries to inject bad data in the grid. The adversaries' fabricated readings mislead the electric utility and result wrong decisions about local or regional usage and capacity. The adversaries also can launch an effective DoS attack by forging many demand requests for a smart meter that are requesting for a large amount of energy. One of the severe attacks is to target the electricity supply for a country. In traditional grids, this attack is very difficult, as it involves various attacks on generation, transmission, and distribution assets, which are well protected. However, the emergence of millions of smart meters controlled by few central controllers in the grid will simplify this attack. The adversary only compromises these controllers and sends the combination of commands to cause supply interruption. As a result, efficient security techniques to guarantee the confidentiality and integrity of the smart meters' readings are essential.

According to smart meters privacy, the major benefit of smart grid is collecting huge amount of readings data for various appliances in the household. However, this advantage could turn to a privacy concern, as the information about the house energy usage can reveal personal habits and daily activities for householders. To address the privacy of smart meters, several protection approaches have been proposed, such as employing homomorphic encryption during data reading aggregation process, compressing the readings and adding random sequences, or deploying anonymization schemes to conceal the real identity of smart meters.

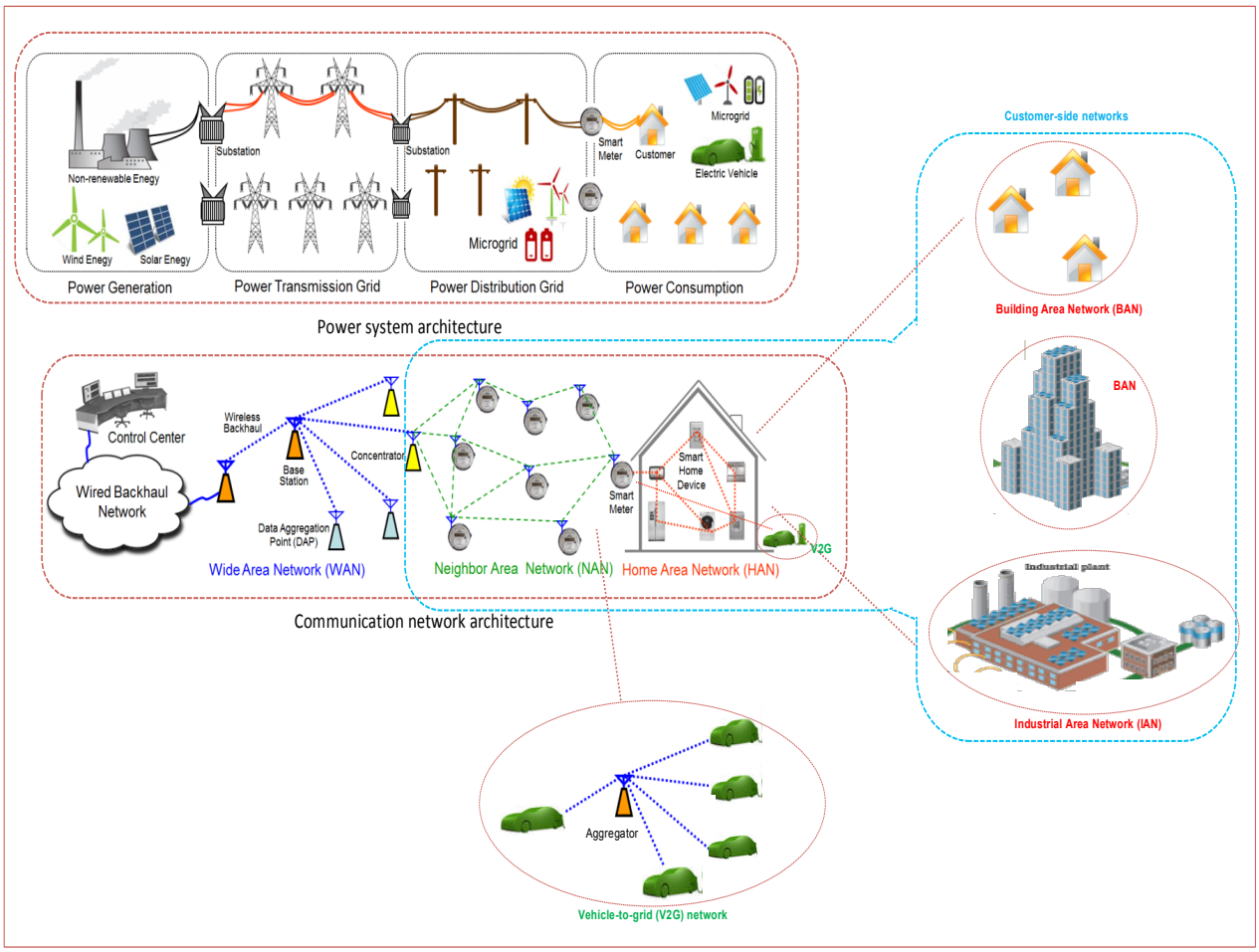

Figure 2.1: Smart grid architecture.[1]

For monitoring and measurement units security, the efficiency of smart grid depends on the accuracy of the deployed measurement units and PMUs. The precision of the measurements is checked by state estimators, which are located in the main CC and used to estimate the power grid status through analyzing the values of the measurements. Therefore, the evaluation of grid status is directly affected by the integrity of the measurements. A typical attack to compromise measurement data integrity is the FDI attack. The invader in this attack can compromise a bunch of PMUs and measurement units in the grid and exploit them to manipulate the state estimate without triggering bad-data alarms. To resist these attacks, many researchers focus on detecting the attacks using efficient state estimators and optimization techniques.

In general, deploying communication networks in power grid not only exposes the grid to security and privacy issues exist in these networks but also adds new threats due to the power grid's nature [1, 10, 11, 12, 13, 14, 15, 16].

2.3 Smart Grid Networks

Four types of communication networks are used in smart grid, which are HANs, NANs, V2G connections, and WANs; each one of them has different data rate, coverage range and con-

[1]Original picture Source: Q. Ho and T. Le-Ngoc, Handbook on Green Information and Communication Systems, ECE Dept., McGill University, Montreal, Canada.

sequently requires different communication technology. Smart grid utilizes these networks to exchange information about grid conditions and customers' demands. Three networks of them are employed to connect the electricity customers with power utility. First network is HAN, which connects smart household appliances to the smart meter inside the house. The second network utilized is NANs, which is responsible for forwarding the electricity consumption reports for all HANs in the region to the utility company. In this thesis, we use the term customer-side networks to refer to HANs, BANs, IANs, and NANs networks. For WAN, it is utilized by NANs to forward the electricity reports to the main utility center. According to V2G network, it is utilized to schedule the charging/discharging operations between EVs and grid. Figure 2.1 shows the power system versus the communication architecture in smart grid.

2.3.1 Home Area Networks (HANs)

HAN is a sort of local area networks; it facilitates the connection between smart devices inside or close to house; HAN represents the communication network between smart appliances, EVs, and smart meter. There are two other networks that can be considered HANs: Building Area Networks (BANs), and Industrial Area Networks (IANs). BAN is a connection between several HANs within the same residential area while IAN connects some HANs in the same industrial area [3].

Smart home appliances, such as refrigerators, washing machines, and ovens, are varied in their communication requirements. For instance, the light bulb sends much less data to smart meter than the air conditioner (AC) so that ACs require more communication infrastructure than bulbs. According to their communication needs, smart appliances can be divided into four categories:

Group 1 consists of small-load appliances, such as light bulbs and phone chargers, where an appliance does not significantly impact the total electricity load, and only needs to inform CC whether the appliance is currently connected or disconnected to the grid. *Group 2* consists of large uncontrollable-load appliances, e.g., stoves, which operate according to the consumer needs, and their usage cannot be delayed to a later time. The appliances in that group need to send only its power consumption and expected duration of usage to CC. *Group 3* consists of controllable large-load appliances, such as ACs and clothes washers. Before any of these appliances is switched on, it should send a request to CC via smart meter, including the appliance's expected electricity requirement, duration of usage, and possible usage times in a day. Based on this information, CC can accept or reject the request according to the dynamic electricity pricing, as well as the agreement between the householder and the utility company. Finally, *Group 4* consists only of EVs, which require extensive exchange of information with CC in order to schedule the charging/discharging processes [9].

While, smart meter is an improved electrical meter that primarily aggregates the readings of electricity consumption for a house every specific time interval and forwards the result at least daily to the power service provider for controlling and billing purposes. Smart meter supports the two-way communication feature with CC; whether this CC is a local control unit or the main CC for the utility. HAN/BAN/IANs' applications, such as industrial energy management or computing total electricity costs, require small data rate to 100 kbps with short coverage distance up to 100 m. Thus, technologies, such as ZigBee, PLC, Ethernet, and WiFi, which are low power, low cost, and secure communication, are widely used [1, 3, 4, 5, 9]. Figure 2.2 shows an example of HAN.

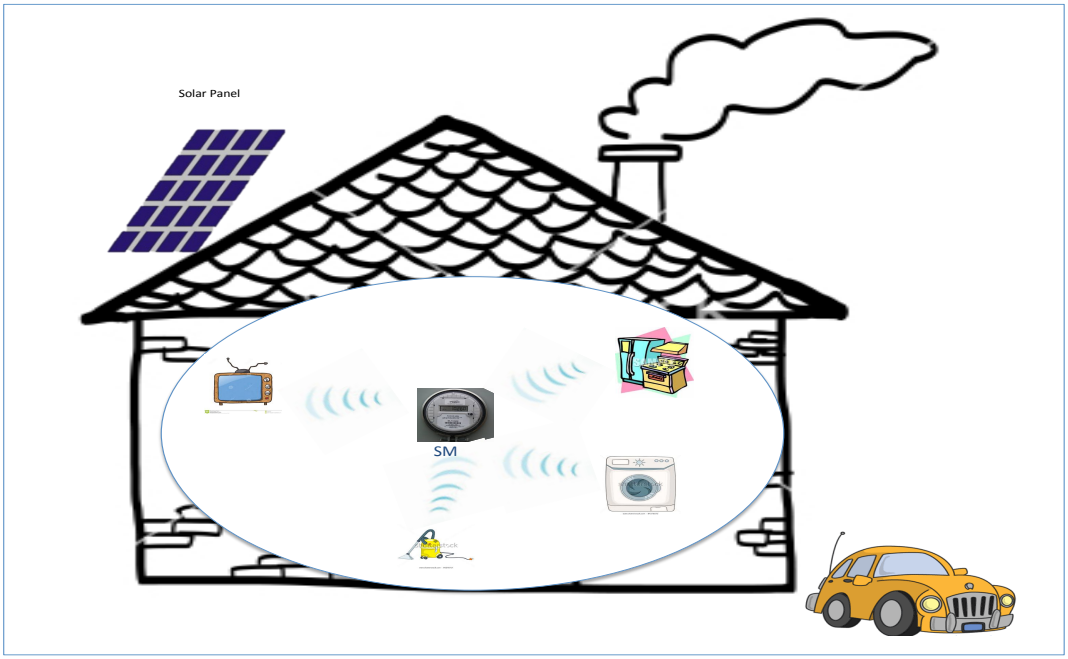

Figure 2.2: An example of HAN.

2.3.2 Neighbourhood Area Networks (NANs)

NAN is responsible for connecting HANs in a specific area to the main CC. It forwards the electricity consumption reports for the region to the service provider. In addition, it sends the electricity payments' value from the utility to all HANs in the area. NAN's applications, such as smart metering and demand response, need higher data rate from 100 kbps to 10 Mbps and larger coverage distance up to 10 km. Therefore, ZigBee mesh networks, WiFi, PLC, and cellular can be suitable for NAN [3].

2.3.3 Vehicle-to-Grid (V2G) Connections

As known, the optimal utilization of generated power and reduction of electricity losses is one of the major objectives of smart grid; this objective requires the presence of storage units that save the extra electricity in case of high power generation and provide the electricity back to the grid in case of high power consumption. Many types of energy storage are used as short-term storage devices, such as fuel cells, flywheels, and EVs (as battery electric vehicles (BEVs) or plug-in hybrid electric vehicle (PHEV) batteries). EVs' batteries consider promising storage media because of the rapid increase in the number of these vehicles in the near future. As well, the batteries are stable storage units; the losses ratio for the stored power in EVs' batteries is low. In addition, the charging and discharging operations for EVs' batteries are much faster than increasing or decreasing the generation level of traditional power plants to satisfy the electricity loads. In other words, the vehicles can work as distributed generation resources; they can quickly supply electricity to the grid if the consumers' demands increased, also they can rapidly store the extra power from the grid if the electrical requirements decreased. As a result, EVs supply certain services to the electricity grid, such as providing peak power, spinning reserves, regulation reserves, and storing the renewable energy. Consequently, V2G networks term is coined to represent the communication between EVs and the power grid. The

communication between the power grid and EVs is bidirectional; when the power transfers from the vehicle's battery to the grid, the connection to manage this operation is called vehicle-to-grid or V2G connection. While if the power transmits from the grid to the battery of the vehicle, the connection is called the grid-to-vehicle or G2V connection. In the thesis, the term V2G connection is used to refer to both connections. V2G connection suffers from some problems related to scheduling the charging/discharging processes; it also experiences particular security and privacy threats, such as the disclosure of EV's owner identity or current location, and DoS attacks.

- V2G concept: Two types of EVs are utilized in V2G connection. The first one is PHEVs, which are mainly powered by battery but have a fossil fuel combustion engine too. This feature increases the vehicle's capacity. The second type is BEVs, which use the battery as the only source of energy. Therefore, BEVs' capacity is less than PHEVs. For simplicity, we will refer to both PHEVs and BEVs as EVs. In addition, specifying EVs' driving pattern is significant for scheduling the charging/discharging operations, because it assists in determining the optimal charging time and location for each EV to gain the optimal power price and reduce the wasted electricity [16].

  EVs are not only a significant solution for the environmental problems, i.e., CO2 emissions, but also have potential economic benefits. EVs can work as a temporary storage for the extra electricity. Subsequently, it can guarantee the reliability of the power grid and reserve the consumed energy. EVs provide four functions to smart grid. The first one is supplying energy during peak power load. The central power plants generate their peak power occasionally, thus some of the produced power is lost due to the low demands. One of the solutions to save the electricity is using several aggregated EVs as storage for this power; these EVs perform peak power shaving during high-generated electricity periods by storing the extra power, and they implement valley-filling process during high consumption periods by restoring the power back to the grid. Accordingly, the load is smoothed. The second function is spinning reserves; EVs can deliver power to the grid at faster rate than the power plants do so that the grid operator uses EVs as a power supply in cases that require fast response. Moreover, EVs offer regulation services to power grid; they can regulate the energy supply/demand rate under the supervision of the grid operator. Finally, EVs act as backup storage for electricity derived from the renewable resources, such as photovoltaic and wind turbines. The main concern about these resources is their intermittent nature. Therefore, usages of EVs overcome the fluctuation of the renewable resources and conserve the produced power load in a certain level [17].

  Although V2G connection can solve the problem of electricity losses and offer a fast supply/store electricity service to power grid, it provides certain problems related to the scheduling of charging/discharging processes. Also, it threaten security and privacy of EVs too. For the charging/discharging operations, the uncoordinated charging/discharging can lead to unbalance in the electrical load and cause many problems. So, the management of V2G network attracts the researchers to provide many techniques to control V2G networks, such as in [18, 19, 20, 21, 22, 23, 24].

- V2G architecture: The architecture of V2G connections, as illustrated in many research works [16, 17, 18, 19, 20, 21, 22, 23, 24, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86], consists of clusters of EVs connected to devices called aggregators. The function of the aggregator is to manage the power exchange operation between the grid and EVs; there are two main types of aggregators; local and central aggregators. The

local aggregators are located in the same region with EVs; their functions are collecting information about EVs, such as their batteries' state of charge, and forwarding the current electricity price to the connected EVs. The local aggregator also receives the EVs' requests to charge or discharge their batteries, and compute the payment amount for each EV. On the other hand, the central aggregators work as a link between power grid and EVs; they receive the grid requests to supply or consume electricity and the current price; they also control the local aggregators to satisfy the requirements of both power grid and vehicles' owners. EVs connect to aggregators during their parking time, which is approximately 95% of the day, so that aggregators and charging stations are suggested to be installed in parking lots; whether these lots are private, such as a residential building or a company parking lot, or public lots, such as the shopping malls parking. Moreover, V2G connection equipment can be installed in gas stations to serve PHEVs. In addition, V2G connection should contain representatives for power grid; their function is to communicate with EVs via central aggregators in order to store the extra electricity in the EVs' batteries during the peak power shaving and restore the power from them during the valley filling. These processes should take into account the real-time electricity price for the optimal benefits of both power grid and vehicles' owners.

Figure 2.3 demonstrates the main architecture of V2G Network. Every cluster of EVs in the same place, e.g., parking lots, is connected to the local aggregator, which receives their charging/discharging requests and control the charging station in the lot. The local aggregator also connects to the central aggregator to report the status of EVs in the lot, e.g., the total free capacity of storage units or the total amount of electricity to supply from EVs. On the other hand, the central aggregator communicates with the grid operator to know the grid requirements; the total demand of electricity that grid wants to sell/purchase and the corresponding price. However, V2G networks face various types of attacks; these attacks can threaten EV owners' privacy and integrity of the exchanged messages. Accordingly, many research studies are proposed to detect the attacks and guarantee the security requirements.

2.3.4 Wide Area Networks (WANs)

WAN, which already exists, is utilized by NANs to forward the electricity reports from their local regions to the main CC in utility company. WAN applications, such as wide-area control, monitoring and protection, requisite higher data rate from 10 Mbps to 1 Gbps long coverage distance up to 100 km. Technologies, such as optical fibre communication, cellular, and WiMAX, are most commonly used between transmission/distribution substations and the utility's CC due to their high capacity, low latency, and wide coverage range [3].

2.4 The Power Control System and State Estimation

Fundamentally, power grid is responsible for generation, transmission, and distribution of electricity to customers. To achieve these functions, the power grid CC should perform certain auxiliary tasks to guarantee the required quality of service and prevent hazards and disasters, such as blackouts. One of the major tasks is monitoring the grid status using local sensors or measurement units. Therefore, CC should assure the accuracy of these measurements by state estimation operation. The traditional state estimators are based on computing the difference

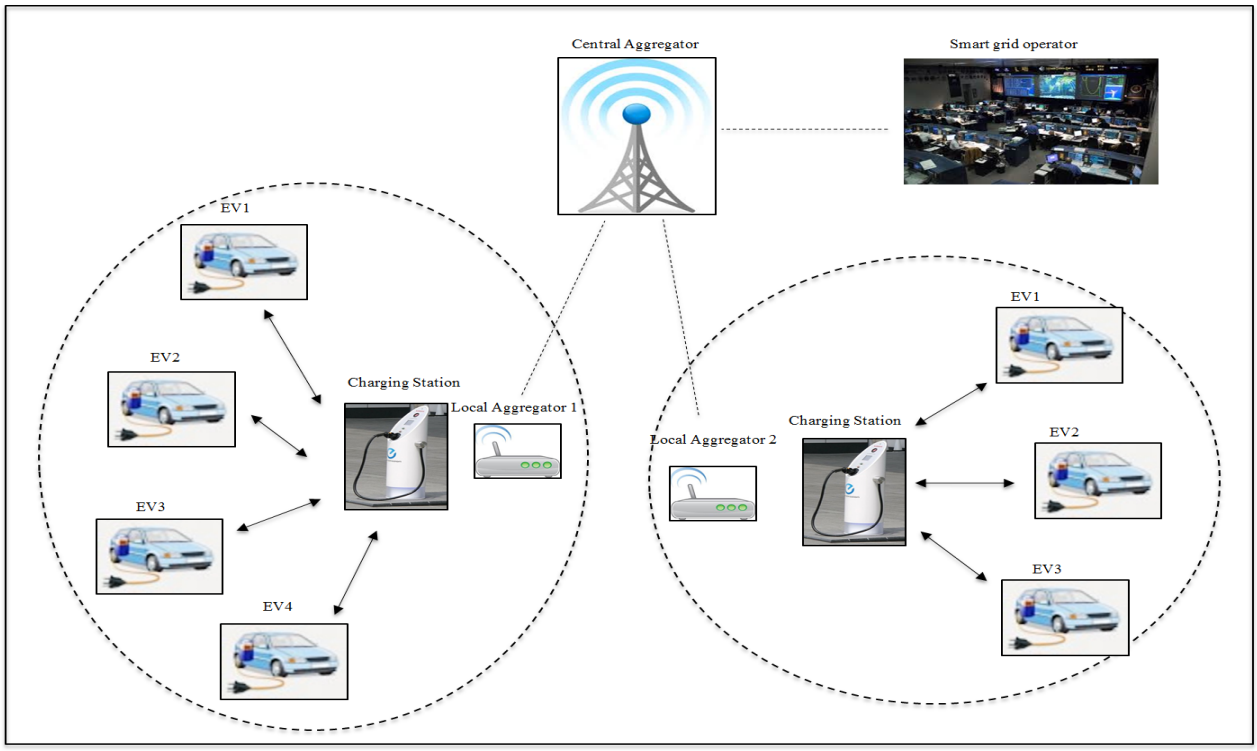

Figure 2.3: V2G network architecture.

between observed and expected measurements and comparing the residual by a specific threshold. However, this technique is not realistic for the novel smart grid, as the power grid exposure to communication networks leads to a new type of attacks that target the infrastructure of the grid by injecting false measurements; these new attacks are called FDI attacks or stealthy attacks. Accordingly, these FDI attacks can mislead the CC to make wrong decisions for the grid and consequently cause catastrophic results; for instance, on August 14, 2003, a large area of United States and Canada experienced an electric power blackout, which affected on about 50 million people and caused economic losses between $4 billion and $10 billion in United States and $2.3 billion in Canada [25]. In this section, we describe power system components and services, and then define power system different models, and traditional state estimation process. In addition, we introduce FDI attack, its types and its impacts on power grid.

2.4.1 Supervisory Control and Data Acquisition (SCADA) System

Supervisory control and data acquisition (SCADA) system, which is the brain of power system, is a type of industrial control system (ICS). ICS is computer-controlled system that physically monitors and controls industrial processes. SCADA systems are distinguished from other ICS systems by being large-scale systems with multiple sites and large distances. SCADA system consists of certain remote units connected to variety of sensors and actuators, and some master stations; it operates by sending signals over communication channels to control these remote equipment. The supervisory operation may be combined with a data acquisition function by sending signals over communication channels to acquire information about the status of the remote devices. SCADA system mainly consists of:

- Remote terminal units (RTUs) that connect to sensors or measurement units that spread in different locations in the grid, convert their signals to digital data, and then send the

data to the supervisory system. As well, they receive the digital commands from the supervisory system and forward them to the sensors.

- Telemetry system, which connects the field devices, i.e., RTUs, with CCs via wired or wireless communication media.
- Data acquisition server that provides some services to the human operator and other parties and allows them to access the field devices' data.
- Human machine interface (HMI), which displays the data in interpretable format for human operator so that he/she monitors and interacts with the grid's status. The operator via HMI can request the data from the data acquisition server.
- Historian, which is a software service that records all time-stamped data and events in a database and utilizes them to graphically show the power trends via HMI.
- Supervisory system, which is the CC, that gathers data about the status of different parties in the grid, also sends control commands to the system via communication infrastructure.

  By these subsystems, SCADA acquires the RTUs' data, and then aggregates and formats them to ease the supervisory system decisions making operation, e.g., to adjust RTUs. Data may also be fed to historian to allow trending and other analytical auditing [26].

SCADA system is used since 1970s in power grid but nowadays more devices that provide more functions are attached to it. The modern grid, smart grid, supports new tasks, such as automatic generation control and optimal power flow analysis; also new types of sensors, e.g., PMUs, are employed for wide-area monitoring and control systems for the grid. In addition, SCADA communication network is heterogeneous; it consists of fibre optics, satellite, and microwave connections. Although the traditional SCADA systems are originally designed to be centralized and closed systems, i.e., original SCADA systems have limited connection with the open networks like Internet, SCADA system in smart grid becomes distributed and connected to different networks. As a result, it is exposed to various cyber security threats. Therefore, SCADA requires protection techniques to provide its services without security risks. Analytical instruments, such as the state estimators, are significant in the modern SCADA system. The measurements units that are spread in the grid collect data about the grid's status and forward it to SCADA system via RTUs. State estimator is an analytical tool in the SCADA's CC that is responsible for checking the accuracy of the received measurements. Consequently, accurate state estimators are significant for the future smart grid to fulfill its tasks. However, the state estimation operation is threatened by cyber and physical security threats, because the exchanged data is often sent without encryption so that malicious attackers exploit that weakness and launch powerful attacks, such as FDI attacks [27, 28, 29, 30].

2.4.2 Power System Model and State Estimation Process

Any power system consists of a collection of power flow meters, transmission lines, and buses. Assume we have a power system with $n+1$ buses and only consider active power flows P_{ij}, active power injections P_i, and bus phase angles θ_i, where $i, j = 1, \ldots, n+1$. Assuming that the resistance in the transmission line connecting buses i and j is small compared to its reactance X_{ij}, then the active power flow from bus i to bus j equals:

$$P_{ij} = \frac{V_i V_j}{X_{ij}} \sin \theta_i - \theta_j$$

At each bus i, positive active power P_i is injected from a power generator. While the negative P_i indicates a power load. If there are no losses, the total power of bus i equals:

$$P_i = \sum_{k \in \mathcal{N}_i} P_{ik}$$

, where $\mathcal{N}_i$ is the set of all buses connected to bus i.

The active power flow measurements z for any system in general are computed by:

$$z = P + e = h(x) + e \in R^m$$

, where $h(x)$ is the power-flow model derived from power flow equations, and $x \in R^{m+1}$ is a vector of $n+1$ unknown bus phase angles, and $e = (e_1 \dots e_m)^T \in \mathcal{N}(0, R)$, is a noise with zero mean Gaussian distribution, where R is the diagonal measurement covariance matrix. Generally, one arbitrary bus phase angle is fixed as reference angle, e.g., $\theta_1 = 0$, and the remaining n angles have to be estimated. The voltage level V_i of each bus is known as well as the reactance of each transmission line. In other words, the power flow measurements z depends on the phase angles θ_i, and the state estimator knows all other parameters. Therefore, this equation presents a common formulation of state estimation problem for the AC power flow model, as $h(x)$ is a nonlinear function. However, $h(x)$ function can be linearized, which means approximating AC system to DC system. Then, the active power flow measurements equation becomes:

$$z = Hx + e$$

, where $x = (x_1, x_2, \dots, x_n)^T \in R^n$ represents the true state vector to be estimated, $z = (z_1, z_2, \dots, z_m)^T \in R^m$ represents the observed sensor measurements, $H \in R^{m \times n}$ is an $m \times n$ Jacobian or DC power flow matrix, and $e = (e_1, e_2, \dots, e_m)^T \in R^m$ is the random measurement errors (noise) [27, 31, 32, 33, 34, 37].

The main tasks of the state estimator are to check if there is a unique estimate for each system state and perverse the system's observability, determine the optimal estimate for each bus based on the real-time analog measurements, i.e., the state estimation process, and lastly detect the measurement errors and identifies the bad data injections and eliminates them if possible; this operation is called bad data detection (BDD) process. To guarantee the system's observability, measurement units are spread optimally along the grid; many optimization research works that address the measurement units placement problem in power grid, such as [35].

According to the state estimation process, Gauss-Newton method [36] is often used to estimate the unknown bus phase angles from power flows measurements z as follows:

$$\hat{x}^{k+1} = \hat{x}^k + (\bar{H}_k^T R^{-1} H_k^{-1} H_k^T R^{-1} (z - h(0, x^k))$$

, where $\hat{x}^k \in \mathbb{R}^n$ is the estimate of the n unknown phase angles, k denotes the iteration number, H_k is the Jacobian evaluated at $\hat{x}^k$ and reference angle equals zero, $H_k = \frac{\partial h}{\partial x}(0, \hat{x}^k) \in$

$\mathbb{R}^{m\times(n+1)}$, $\bar{H}_k \in \mathbb{R}^{m\times n}$ is H_k after removing the corresponding column for the reference angle. Assume that the phase differences $\theta_i - \theta_j$ in the power network are small; then, the phase-angle estimate equals

$$\hat{x} = (\bar{H}^T R^{-1} \bar{H})^{-1} \bar{H}^T R^{-1} z \tag{2.1}$$

, and the active power flows can be estimated by

$$\hat{z} = \bar{H}\hat{x} = \bar{H}(\bar{H}^T R^{-1} \bar{H})^{-1} \bar{H}^T R^{-1} z \tag{2.2}$$

BDD system in CC calculates the measurement residual r, i.e., the difference between observed and estimated measurements, as:

$$r = z - \hat{z} = P + e - \bar{H}\hat{x} = (I - K)z \tag{2.3}$$

, where $z = Hx + e$. Because of measurement errors e, the residual r does not equal zero. Accordingly, r is compared with a specific threshold; if r is larger than it, an alarm is triggered and z_i are declared as bad measurements and removed [7, 9, 10].

- State Estimation and Bad Data Detection

  Three statistical estimation criteria are commonly used in state estimation process: the maximum likelihood, the weighted least square, or the minimum variance criteria. As the measurement error is normally distributed with zero mean, these criteria lead to an identical estimator. The minimum mean squared error (MMSE) estimator is the state estimator currently used in the grid by solving the following matrix:

  $$\hat{x} = (H^T W H)^{-1} H^T W z$$

  , where W is a diagonal matrix whose elements are reciprocals of variances of units errors, $W = R^{-1}$ in equation (2.1);

  $$\begin{bmatrix} \sigma_1^{-2} & \dots & \dots & 0 \\ \vdots & \sigma_2^{-2} & \dots & \vdots \\ 0 & \dots & \dots & \sigma_m^{-2} \end{bmatrix}$$

  , where σ_i^{-2} is the variance of the $i-$th unit $(1 \le i \le m)$.

  The normal measurements usually give an estimate of state variables close to their actual values, while abnormal ones may deviate the estimated state variables away from their true values. Consequently, MMSE estimator can detect if there are errors or noise in measurements by calculating $r = z - H\hat{x}$ and comparing its 2-Norm $\| z - H\hat{x} \|$ with a specific threshold τ; the presence of bad measurements is inferred if $\| z - H\hat{x} \| > \tau$. The state variables are independent and the units errors follow the normal distribution so that $\| z - H\hat{x} \|^2$, denoted as $\mathcal{L}(x)$, follows the chi-distribution $\mathscr{X}^2(v)$ [40], where v is the degree of freedom. However, selection of τ is a main issue; it can be determined through a hypothesis test with a significance level α. In other words, the probability that $\mathcal{L}(x) \le \tau^2$ equals α means that $\mathcal{L}(x) \le \tau^2$ detects the presence of bad measurements with false alarm probability equals α [31, 33, 34, 37].

- Likelihood Ratio Test (LRT)

  Most of traditional state estimators are based on the likelihood ratio test (LRT), which is successful to detect noise or random errors, but fails to detect planned malicious bad data.

The simplest detection process is the task of deciding which of two probability models best matches a set of data. Let X be a random variable and denote the two probability models as $p(x\|H_0)$ and $p(x\|H_1)$, where H_0 and H_1 are the hypothesis 0 and 1 respectively. The detection problem is simply to decide which model is more appropriate. Assume a random variable is distributed on $H_0 : X \sim p_0$ or $H_1 : X \sim p_1$; the process of deciding which distribution best fits an observation of X is called a simple binary hypothesis test, because the two distributions are precisely known without unknown parameters or other uncertainties. The decision is made by partitioning the range of X into two disjoint regions: R_0 and R_1. Let x denote the observed value of X. If $x \in R_i$, then H_i is decided to be the best match. There are four possible outcomes for that test: $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$, where the first argument denotes the decision based on R_0 or R_1 and the second argument is the true data distribution. To optimize the decisions' choice, a non-negative cost c for the decisions is specified; let $c_{i,j}$ be the cost of outcome $(i,j), i,j \in \{0,1\}$. The test aims for decisions with few mistakes so that it is reasonable to assume that $c_{1,0}$ and $c_{0,1}$ are larger than $c_{0,0}$ and $c_{1,1}$. The overall cost for the test is called the Bayes Cost:

$$C = \sum_{i,j=0}^{1} c_{i,j}\pi_j \mathbb{P}(decide\ H_i \| H_j\ is\ true) = \sum_{i,j=0}^{1} c_{i,j}\pi_j \int_{R_i} p_j(x)dx$$

$$= \int_{R_0} (c_{0,0}\pi_0 p_0(x) + c_{0,1}\pi_1 p_1(x))dx + \int_{R_1} (c_{1,0}\pi_0 p_0(x) + c_{1,1}\pi_1 p_1(x))dx$$

, where $\pi_j = \mathbb{P}(H_j\ is\ true), j = 0,1$, is the probability that an observation is generated according to p_j.

$$R_0 \coloneqq \{x : c_{0,0}\pi_0 p_0(x) + c_{0,1}\pi_1 p_1(x) < c_{1,0}\pi_0 p_0(x) + c_{1,1}\pi_1 p_1(x)$$
$$R_1 \coloneqq \{x : c_{0,0}\pi_0 p_0(x) + c_{0,1}\pi_1 p_1(x) > c_{1,0}\pi_0 p_0(x) + c_{1,1}\pi_1 p_1(x)$$

Therefore, the optimal test will be as:

$$\frac{p_1(x)}{p_0(x)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0(c_{1,0} - c_{0,0})}{\pi_1(c_{0,1} - c_{1,1})}$$

The right hand side term is constant that depends on probabilities and costs while the left hand side term is a ratio of probability densities evaluated at x. The value of probability density at the observed x is called likelihood of x under that model. Thus, $\frac{p_1(x)}{p_0(x)}$ is called the likelihood ratio and the LRT test takes the form:

$$\frac{p_1(x)}{p_0(x)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

, where $\gamma > 0$ is the test threshold. Therefore, LRT with an appropriate threshold is optimal for error and noise detection [39].

2.4.3 False Data Injection (FDI) Attacks

FDI attack can be launched by inserting fake measurements to the observed measurements vector. The FDI attacks' objective is to mislead CC to make wrong decisions, such as redistributing electricity according to that wrong injected information, and consequently cause huge

damage; i.e., blackouts. The attacker may also aim to gain a financial profit by redistributing the power loads along the grid to control the electricity price. For example, adversary can buy the electricity by cheap price from a location in the grid, and then change the loads to cause higher power price in another location to sell its electricity. The essential condition for successful FDI attack is to pass the state estimator test and avoid the alarm. In other words, even with the injected false measurements in the observed measurements vector, $r = z - H\hat{x}$ should still be less than τ. To satisfy this condition, the attacker should have some information about the Jacobian matrix of power system H, which is determined by the topology and line impedances of the system. Many research works assume that attackers have full knowledge of H; this information can be extracted by spying on the grid company or monitoring the grid for a period of time.

For undetected attack, the attacker forms the false injected measurements a, also called attack vector, as a function of H, i.e., $a = Hc$, where c is an arbitrary nonzero vector. Then, the observed false measurements vector z_a is represented as $z_a = z + a$, where $z = (z_1, \ldots, z_m)^T$ is the vector of original measurements and $a = (a_1, \ldots, a_m)^T$ is the malicious data added to the original measurements. If $i-$th element a_i is nonzero, then the attacker compromises the $i-$th unit, i.e., replaces its original measurement z_i with a fake measurement $z_i + a_i$. When a malicious data $a = Hc$ is inserted, the vector of the estimated state variables $\hat{x}_{bad}$ is obtained from z_a as:

$$\hat{x}_{bad} = (H^T W H)^{-1} H^T W z_a = (H^T W H)^{-1} H^T W (z + a) = \hat{x} + (H^T W H)^{-1} H^T W a$$

, and the 2-Norm of the measurement residual is

$$\| z - H\hat{x}_{bad} \| = \| z + a - H(\hat{x} + (H^T W H)^{-1} H^T W a) \| = \| z - H\hat{x} + (a - (H^T W H)^{-1} H^T W a) \|$$

$$= \| z - H\hat{x} + (Hc - (H^T W H)^{-1} H^T W H c) \| = \| z - H\hat{x} + (Hc - Hc) \| = \| z - H\hat{x} \| \leq \tau.$$

According to their goals, FDI attacks can be divided into random and target FDI attacks. In random FDI, attackers aim to inject any bad data to cause wrong state estimation in the state variables. While the target FDI attack is injecting an attack vector that causes a specific error into certain state variables [7, 13, 14]. [37] studies two attack scenarios, where attackers are either constrained by compromising specific meters or by their limited resources.

2.5 Smart Grid Security Concerns

Smart grid suffers from security threats that exported from communication networks, such as falsifying the electricity consumption readings, extracting customers' private information, or attacking the availability of the grid's resources. In addition, smart grid is exposed to the novel FDI attacks.The security concerns of smart grid can be briefly categorized into three major groups [1, 4, 8, 9, 10, 11, 12, 13, 14, 15, 16]:

Information privacy is a big concern in smart grid especially for customer-side networks in houses, residential and industrial buildings. Personal information and daily habits for consumers can be revealed to outsiders from the electricity consumption pattern of the unit.

Therefore, any eavesdropper, with moderate data analysis tools, can threaten the privacy of customers. The eavesdropper can extract significant information about householders, such as when they are in/out the house, or the types of electrical appliances they usually use. According to industrial institutes, the observer may get important knowledge about the institutions' production from its consumed power and sell this information to the competitors. The electricity users' privacy requires efficient security schemes to preserve it by guaranteeing the confidentiality of exchanged messages.

Second, data integrity is a significant concern. Adversaries can attempt to alter or fabricate the exchanged messages. There are two obvious examples of integrity attacks in smart grid. First one is when a malicious consumer attempts to forge his/her electricity reading to reduce the electricity bill. Second attack is FDI attack, in which the adversary compromises several measurement sensors and exploits them to inject false information about the grid conditions; this attack misleads CC to wrongly evaluate the grid status and make incorrect decisions. The consequences of this attack extremely affect all parties in the grid.

Network availability is another concern. Malicious adversaries can target network resources by DoS attacks. They attempt to block or corrupt the network resources to delay the transmitted information or even to make it unavailable to authorized parties. For instance, attackers can launch DoS attacks by falsifying a large number of electricity request messages using compromised smart meter, and asking for huge amount of energy. Therefore, the utilized networks in smart grid should be robust to network availability attacks, because the network unavailability could cause severe consequences, such as loss the real-time monitoring of critical power infrastructures, and subsequently lead to huge blackouts.

In summary, merging communication technologies with power grid introduces novel security problems to the grid. In general, the main security risks for smart grid are network resources and information availability, data integrity, and users privacy. In next chapter, we demonstrate the currently proposed solutions for the three main smart grid communication architectures: customer-side networks, V2G connections, and power control and state estimation systems.

Chapter 3

Security and Privacy Concerns in Smart Grid

This chapter presents the existing research works that proposed to guarantee security and privacy requirements for different communication architectures in smart grid. First, we begin with the research studies that are proposed to deal with security concerns in customer-side networks. Mainly, the current studies are utilizing hardware devices, distorting the message's contents, or cryptographic schemes, such as anonymization or homomorphic encryption techniques; the central target is concealing the confidential data and preserving the messages' integrity. Second, we introduce the efforts to secure the vehicles' owner privacy and data integrity during charging and discharging operations in V2G networks. The previous studies are focused on authentication mechanisms, anonymization techniques, or physical layer security methods. Finally, the proposed schemes to detect FDI attacks are described. These solutions are ranging between exploiting alternative estimation tests, distributing the estimators all over the grid, utilizing various optimization techniques, using cryptographic schemes. Figure 3.1 lists the previous proposed solutions and related works categories for each architecture.

3.1 Customer-side Networks Security and Privacy Problems and Related Works

Smart grid technology has to consider many concerns due to network performance and security requirements. The security concerns for smart grid are varied according to the applications. In customer-side networks, the aggregation of electricity consumption and transmission of billing information operations either in individual HAN or among HANs and utility are the main processes to exchange messages; the major security concerns during these operations are the privacy of householders as well as the integrity of aggregated information. Another concern is the restricted-resources devices, i.e., smart meter and smart appliances, in customer-side networks that cannot perform complex crypto-operations. At a result, the suggested schemes to handle these security concerns should not only preserve security and privacy requirements, but also be lightweight due to computation and communication overhead. Several solutions have been proposed to deal with that problem, which can be divided into three different categories:

- *Hardware Equipment:* The first category employs hardware devices that are connected to smart meters to conceal the real electricity consumption [41, 45, 46]. These devices are one

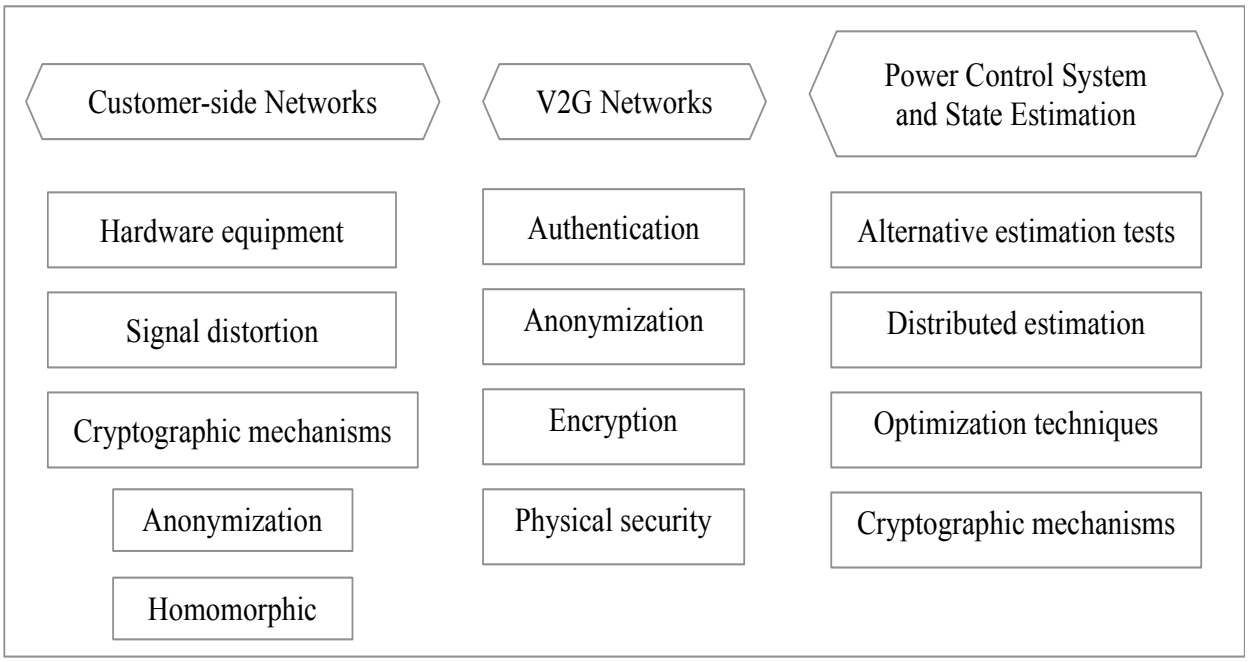

Figure 3.1: The proposed solutions categories in each architecture.

of two classes; the first class is devices that offer more security for smart meters, such as temper-resistance devices, so that no attacker can compromise the meters. Furthermore, the measurement units that spread through the grid to monitor its status can detect any anomaly behaviours in the grid and can be utilized to resist replay attacks [41]. The second class is devices that supply the house by electricity, such as electrical batteries [45, 46]; the battery is connected to smart meter to satisfy some of the electricity demand for the house by self-reliance. Consequently, the requested electricity amount from the grid does not reveal the real consumption. The main target of these procedures is to not depend on cryptographic schemes to alleviate the computation and communication complexity. However, using these expensive devices is not practical, as it is hard to connect a temper-resistance device or a special electric battery to each smart meter of millions of deployed smart meters in addition to the necessary maintenance and replacement operations.

- *Signal Distortion:* Second category distorts the power consumption value before transmission [47, 48, 49]; this process adds noise to the message at smart meter side and removing it at CC. This category saves the power and computation abilities for smart meters. For instance, [49] adds noise to each reading value to further resist the human-factor-aware differential aggregation attack. However, in these approaches, CC requires complex data mining techniques, which consumes a lot of time and resources, to reconstruct the original message from the received one. This method not only impedes the reading aggregation operation but also produces an approximate value for the received message, while CC requires the accurate readings for billing purposes.

- *Cryptographic Mechanisms:* Finally, various existing works have utilized cryptographic schemes to guarantee information security and customers' privacy. [50] provides unified framework to address security and privacy problems in HAN and selects the appropriate defense mechanisms. In [51], the authors present a complete framework to secure HAN; the used authentication scheme is based on password. While, jamming attacks can be countermeasured by hopping between predefined frequencies. Timestamp and sequence

numbers resists replay attacks. Moreover, keeping a log file for some historical operations protects from nonrepudiation threats, and directional reception capabilities are used to limit the interference effect. Asymmetric encryption schemes are utilized to encrypt the electricity consumption message due to their higher security level. [52] suggests public key infrastructure (PKI) scheme and applies attestation protocols, firewalls, and mutual authentication to limit the influence of compromised devices. When a group of users needs to exchange messages so that no outsiders can decrypt these messages, group signature-based schemes are utilized to guarantee the security and privacy [53, 54]. [53] employs the key-policy attribute-based encryption (KP-ABE) scheme for multicasting purposes. KP-ABE [55] utilizes the key based on specific attributes shared between a group of users to decrypt the message. [56] ensures the accountability by considering smart appliances and power generation source in HAN as witnesses to watch the smart meter's behaviour. Each witness uses a log file to record all the information at current time and sends it periodically to service provider.

Homomorphic Public Key Schemes: Several studies exploit the homomorphic features for some public key schemes to aggregate the electricity consumption for a specific region without revealing individual consumption values. [57] proposes a full framework to aggregate the electricity readings for customers in a specific region and guarantees the privacy of customers using additive homomorphic encryption; at the same time, it provides access control scheme based on attribute-based encryption (ABE) scheme. The homomorphic hash function [58] is utilized in mutual authentication between smart meters and the main control server. [59] proposes an efficient privacy-preserving demand response (EPPDR) scheme to aggregate the electricity demand messages for users in local area. In [60], the authors propose a new privacy-preserving scheme for aggregating multi-dimensional metering data using homomorphic encryption scheme (EPPA), which is based on homomorphic Paillier crypto-system scheme. While [61] reviews different security threats in smart grid and suggests intrusion detection system and fast authentication scheme. However, the applied homomorphic schemes provide high computation and communication overhead. For instance, EPPA [60] scheme based on homomorphic Paillier encryption requires time from 100 ms to 220 ms as messages number increases, which considers high load especially with the increase of smart meters' number. The proposed homomorphic schemes are not scalable; their performance degrades, as number of smart meters within the same cluster increases [11].

Authentication Schemes: Other research works employ authentication schemes to confirm the authenticity of smart meters during network establishment; also key management techniques are combined to revoke the compromised keys. [62] proposes a key management theoretical framework that uses extensible authentication protocol (EAP) as an authentication protocol. According to [63], the authors propose a key management scheme that assigns a set of keys for each node in each transmission mode: unicast, multicast, broadcast. While [64] proposes a lightweight scheme to guarantee the authenticity of nodes using Diffie-Hellman scheme and assure the integrity of messages by message authentication code. Merkle tree is utilized for authentication too [65]. [66] proposes an integrated authentication and confidentiality protocol, which first provides a mutual authentication between smart meters and the gateway via the authenticator smart meter. Subsequently, the authenticated smart meters are organized as a chain. Each smart meter encrypts its reading message, checks the received message's validity, combines it with it own message, and forward the result to the next node until reach to the collector node. [67] proposes an authentication scheme based on identity-based cryptography

scheme (IBC). IBC satisfies the security requirements; it also reduces the PKI overhead and resists various attacks. Nevertheless, it requires the smart meter to have a unique identity value, such as IP address or email, to create its own pair of keys. In addition, the authentication operation generally composes of several steps; it consumes the device's computation and communication abilities and increases the overhead; it requires smart meter to have a secret value before the data aggregation phase. This value could be implanted in its memory during the manufacture or sent through a secure channel to use as a seed in authentication key or password generation. The lightweight authentication scheme based on Diffe-Hellman scheme [64], for example, causes an average delay varied between 1 to 10 second, as the smart meters' number increases.

Anonymization Schemes: Other studies employ anonymization schemes to conceal the real identity of smart meter so that any other unauthorized parties cannot link between the smart meter's real identity and its electricity consumption. Anonymity techniques are based on the presence of a third trusted party, which issues two identities to each smart meter: real and pseudorandom identities. Only the third party knows the identity pairs for each smart meter. The third party can also create binding factors to the network nodes to conceal their identities. [68] proposes a privacy-preserving aggregation scheme, which protects the user privacy from internal attackers too. The trusted party first creates a blinding factor for each user, sums all these blinding factors, and then sends the summation to the aggregator. When the users encrypt their electricity consumption by RSA, they combine their blinding factors before forward the value to the aggregator, which aggregates all reading values and adds its blinding factor to cancel the users' blinding factors and obtains the aggregated reading for all users. While [69] presents a privacy-preserving aggregation scheme, which utilizes two crypto-systems: shamir secret sharing (SSS) scheme, and cramer-shoup (CS) crypto-system. SSS divides the secret value into shares distributed among different parties. To recover the secret, specific number of shares should be cooperated. On the other hand, CS uses one public key and two private keys, i.e., weak and strong decryption keys, which decrypt the message in two different nodes. The collaboration of both nodes is required to recover the plaintext; a single node cannot obtain it independently. In credential-based anonymization techniques [70], each smart meter generates enough number of credentials, adds blinding factors to them, and then sends these credentials to CC for signature. Therefore, smart meter joins these credentials to its messages to CC, which checks the validity of the credential before replies by the requested power. At the end of billing period, each user sends the unused credentials to CC, which computes the used credential for each user and calculates the user's bill. The anonymization techniques can guarantee users' privacy. However, these techniques involve performing several processes especially during initialization phase; so they increase the overhead. Also, the third party should be online most of the time.

3.2 V2G Connections Security and Privacy Threats and Related Works

Several research studies are proposed to conserve the demanded security level and alleviate the threats impact on V2G connection; these studies attempt to preserve vehicle's owner privacy, network resources availability, and exchanged message integrity and confidentiality. Many research works have determined the security requirements and concerns for V2G networks, while

other researches offer solutions for their security problems. V2G networks face various types of attacks, which can threaten the vehicle owners' privacy and the integrity of exchanged messages, such as impersonating EVs, illegally modifying their messages, launching DoS attacks, or threatening the owner's privacy. The security requirements for V2G connections, such as authentication for messages or parties, guaranteeing the connection's availability, and anonymizing the EVs owners' identities and preserving their privacy, should be specified so that the safety threats on V2G connections are bounded, and their impacts are limited [71].

The significant concern in V2G networks is securing the service provider, i.e., EV, privacy, also authenticating it to the power grid. Before any charging/discharging or payment operations, the grid should authenticate the vehicle, and the vehicle requires preserving its private information from the grid's operator or any intermediate devices in the connection, such as aggregators. Several solutions are proposed to resist these threats; they are based on various procedures, such as different authentication mechanisms, encryption schemes, physical security methods, and anonymization techniques. For instance, [80] utilizes ECC public key scheme to provide PKI for a multi-domain architecture V2G network, where each domain contains an aggregator and some EVs. While, [85] utilizes combination of several cryptography schemes to propose a new privacy (V2GPriv) scheme to V2G connection. V2GPriv addresses security and privacy requirements of V2G connections by utilizing some techniques, such as double encryption, which conceals the different parties' communication relations from eavesdroppers. In addition, every vehicle's meter has a unique identifier and public-private key pair for signing purposes; these secret parameters are assigned to vehicles during the manufacturing and stored in the meter in a secured module. Symmetric key $128-bit$ AES encryption scheme is used too. The $k-$anonymity service are also utilized so that the pseudonym identity assigned by an aggregator for the gateway is shared with $k-1$ other gateways; there are k gateways have the same pseudonym identity. During the connection establishment, mutual authentication operation is performed; the gateway challenges the vehicle's meter by basic challenge-response protocol to prove that it has not been tampered with. The authors too suggest deploying multiple servers in series; so the aggregator's destination address can be hidden from all except the last server.

- *Authentication Schemes:* Some research studies utilize authentication schemes to guarantee the validity of connected EVs. [72] proposes a unique batch authentication protocol for vehicle-to-grid (UBAPV2G) communications; the authors employ the batch authentication to guarantee privacy and integrity of messages between aggregator and EVs. The main idea is that the aggregator broadcasts a request message to EVs in the range; vehicles respond by messages with their identification information; the responses are encrypted by the aggregator's public key. The aggregator collects the vehicles' responses during a specific time interval, and then verifies the authenticity of vehicles by one verification process using batch authentication. The aggregator then broadcasts one confirmation message to all vehicles. The batch authentication is better that one-by-one authentication method, as it saves time and reduces the computation overhead. In addition, it decreases the communication time, as the number of confirmation messages is reduced to one [73]. However, [74] discovers some security flaws in UBAPV2G scheme that could be launched by adversaries or dishonest aggregators. An attacker can wait until the time interval, in which the aggregator collects the vehicles responses, is about to finish and sends his messages. Consequently, the forged message is stored in the aggregator's buffer, but the aggregator cannot check the message's validity individually, because it will be included in the batch authentication packet. As a result, a malicious user can be authenticated to the grid and subsequently establish more severe attacks. Moreover, a malicious aggrega-

tor can generate multiple forged signatures for innocent vehicles, and they cannot prove that they were not involved in producing these forged signatures. According to [75], it divides the EVs into two modes: home and visiting modes. It utilizes Diffie-Hellman key exchange protocol to perform mutual authentication between EVs and local aggregators. An extra virtual battery is also used to generate the pseudo-identities for EVs, and the trusted central authority (CA) performs the mapping between real and pseudo-identities. The local aggregator exploits the aggregated-proof technique to authenticate a group of EVs to CA so that any EV can enter or leave the group at any time without effecting on this operation, because it is updated periodically. [76] too utilizes mutual authentication to confirm the vehicles' legitimacy by dividing the battery statuses into: charging, fully charged, and discharging. At each state, the first step should be the mutual authentication between EV and aggregator; also, timestamp and identifiers are attached to check the message's validity.

- *Anonymization Mechanisms:* Other studies attempt to preserve privacy of the vehicle's owners during V2G connection by anonymizing the vehicle so that the outsiders or even the grid's operator and the aggregators cannot extract any distinguishable private information, such as owner identity or vehicle location. Several anonymization mechanisms are proposed; [77] utilizes two aggregators: central and local aggregators. The central aggregator issues a one-day permit for each EV using the partially blind signature scheme. The EV then uses the permit and its pseudorandom identity to communicate with the local aggregator at parking lots. Each EV uses these identity parameters to send periodic reports about its status to local aggregator, also charging/discharging requests. When the grid requires supplying or consuming electricity, it sends a public request. The central aggregators directly make bids, as they know the EVs' capacities, i.e., from the periodic reports. If the bid successes, the central aggregator sends to local aggregator to select some EVs to satisfy the required demand. After the transmission, the local aggregator sends reward messages to the participated EVs. When any EV wants to leave the lot, it sends an aggregated reward request to local aggregator to receive its total reward value for the whole connection period. The EV's privacy is guaranteed by the partially blind signature scheme. However, if any EV cheats, the central aggregator can expose its real identity. A privacy-preserving protocol based on certificate less public key scheme and partially blind signature is proposed in [78]. The utilized schemes reduces the dependence on third party, as in certificate less public key scheme, each entity has a partially private key from trusted party and adds random secret value to it to obtain its private key. [81] suggests a security framework, where each vehicle acts as an independent agent with an electrical chip called trusted platform module that provides the cryptographic parameters to anonymize the vehicle. Also, the authors consider the existence of lateral regional aggregators at the distribution substations to manage electricity-exchange operation. However, [79] presents an adversary model that the attacker can predict (by some probability) the identity of EV from charging locations that the vehicle visits and the distance between them even when EV uses several pseudo identities. Although anonymization techniques and pseudo identities can alleviate the security and privacy problems, they cannot offer a definitive solution. In addition, they increase communication and computation overhead on different parties in the V2G connection.

- *Physical Layer Protection Methods:* Other category of proposed solutions for V2G connection is based on physical layer protection mechanisms. The seamless connectivity of vehicular network could threaten the security of EVs. So [82] utilizes the smart grid's

devices, e.g., smart meter and smart appliances, in home with G.hn technology, which manages the householder to connect with any device using any wire in the house, to secure EV by keeping it connected to the house via any base station whatever the location of it. The authors employ direct-sequence spread spectrum technology and low pass filter to reduce noise and jamming attacks. Although this method uses smart grid as a tool to track stolen vehicles, the technique could used also to secure V2G connections. According to [83], the availability of network resources is the most important requirement that can be threatened by distributed DoS and jamming attacks. So, channel-based key management approach is used to setup a key between transmitter and receiver. The two nodes exchange several beacon signals and apply the same error correction scheme to generate the same key remotely. The authors also suggest employing cognitive radio network because of the dynamic nature of EVs.

- *Other Encryption Schemes:* Several other research works study the impact of price information and propose secure payment management schemes. [84] analyzes the optimal charging policies for PHEV using Markov chain considering different variables, such as mobility and real-time price information. Also, the paper explores the impact of DoS and price manipulation attacks on PHEV charging behaviour. According to [86], it proposes a payment scheme that organizes charging/discharging operation for V2G connection; as well, the proposed scheme preserves the user's privacy, keeps the operation accountability, and traces the stolen cars. First, the vehicle's owner should open an account in the grid's operator. Before any charging/discharging operations, the supplier verifies the balance of the vehicle's account and its capacity, and then performs the operation if it satisfies. If the account balance is below the minimum, the supplier sends to the vehicle to top-up its account. In case of stolen car, the user sends its consent to the judge, which is an independent trusted party, to trace the transactions and reach to the thief. The used cryptographic scheme is based on bilinear pairing and decisional Diffi-Hellman assumption to provide public-private keys pairs to different parties. Also, partially blind signature scheme and zero-proof knowledge are utilized to keep the user privacy during the connection.

In summary, V2G connection security concerns are confined to vehicle's owner information and location privacy, and information integrity and authenticity. Many research works address these problems and recommend different solutions, such as mutual authentication, pseudo-identities and anonymity, and physical layer protection methods. Nevertheless, these provided techniques suffer from certain back-draws, such as high communication and computation burden, and usage of special hardware devices.

3.3 Power Control System and State Estimation Security Problems and Related Works

FDI attacks are based on inserting fabricated measurements in the observed measurements vector to deceive CCs and consequently make wrong decisions. The main target of these attacks is to cause huge damage to power system, such as large blackout, which has negative economic and financial impacts. Moreover, FDI attack can assist attackers to steal power by illegal electricity price manipulation. As a consequence, many researchers attempt to solve the problem and resist these attacks. The related research works can be divided into five categories:

First one utilizes powerful alternative state estimation tests; while second category distributes state estimators all over the grid. according to third category, it approximates the problem to an optimization problem that attempts to place measurement units and PMUs in optimal locations in the grid. Fourth category exploits cryptographic schemes to detect any change in the transmitted measurements. Finally, several research studies addresses load-redistribution (LR) attacks.

- *Alternative Estimation Tests:* The traditional MMSE estimator compares the 2-Norm of measurements residual $\| z - H\hat{x} \|^2$ by a specifc threshold; this estimator can successfully detect any noise or errors that normally distributed with zero mean, but it will not be able to detect the maliciously inserted bad data with non-zero mean. As a result, many researchers study other alternatives that can overcome the limitation of traditional estimators and detect the stealthy FDI attacks, such as generalized likelihood ratio test (GLRT), Rao test, Kalman filter, CUSUM, and others.

  Several studies utilize GLRT-based estimators as an efficient alternative for traditional estimators [32, 87, 88, 89, 90]. GLRT [91, 92] is a general procedure for composite testing problems. It is based on comparing the best model in class H_1 to the best in H_0 where $H_i \colon X \sim p_i(x\|\theta_i), \theta_i \in \Theta_i, i = 0, 1$, and an observation x of X equals $\hat{\lambda}(x) = \frac{\max_{\theta_1 \in \Theta_1} p_1(x\|\theta_1)}{max_{\theta_0 \in \Theta_0} p_0(x\|\theta_0)}$; then GLRT test is $\log \hat{\lambda}(x) \gtrless_{H_0}^{H_1} \gamma$. GLRT is asymptotically optimal, because it offers the fastest decay rate of miss detection probability in case of multiple measurements under the same attack vector a. If the detector has many data samples, the detection performance of GLRT is close to optimal [87]. GLRT detector does not compute explicitly the residue error, but if there is at most one attacked meter, then GLRT-based estimator is using the residue error from MMSE estimator [88]. However, GLRT test is too computationally expensive to implement in practice for fast detection [90].

  Other research works use Rao test [93], because it is a statistical test with a simple null hypothesis to check if the parameter of interest is equal to some value; it does not require an estimate of information under alternative hypothesis. [90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108] present defense mechanism for FDI attacks based on Rao test, which is more efficient than GLRT test, as it does not involve complex computations.

  Kalman filter [94] is an algorithm that uses a series of inaccurate noisy measurements observed over time to produce estimates for unknown variables that tend to be more precise than those based on a single measurement. It operates recursively on streams of noisy data to produce a statistically optimal estimate of the underlying system state. The algorithm estimates the current state variables with their uncertainties, and the outcome is updated using a weighted average so that more weight is given to the estimates with higher certainty. Because of these features, Kalman filter estimator is utilized in many works, such as [95, 96, 97].

  CUSUM test (cumulative sum control chart) [98] is a sequential analysis technique used for change detection. CUSUM determines changes on a parameter of the probability distribution, e.g., the mean, and according to the results, it decides when to take a corrective action. [32] introduces a defense strategy to detect any adversary at CC as quickly as possible utilizing the quickest detection technique, which determines the change based on real-time observations utilizing CUSUM test that detects a change of unknown distribution to known/unknown distributions. The strategy also applies Rao test, because

it is a simple test and does not need to store the observation and re-calculate the unknown parameter every time interval as GLRT test does.

- *Distributed Estimation:* Traditionally, the state estimation process is running in the main CC of SCADA system; the measurement units send their readings to SCADA, which has the central state estimator. However, a penalty of research studies suggests performing distributed estimation [99, 100, 101, 102, 103, 104, 105, 106, 107]; the main objectives of distributed estimators are to increase the state estimation process accuracy and reduce the CC's computation burden. For instance, [104] divides the power system into several small systems and apply the state estimation on each subsystem using lower thresholds; this method improves the accuracy of bad data detection scheme.

- *Optimization Techniques:* Many research works utilize various optimization techniques to detect FDI attacks and perform tasks, such as studying the optimal placement for PMUs and other secured sensors in the grid, or calculating the optimal cost for different protection schemes [27, 33, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122]. For instance, [27] proposes an algorithm to find the least cost for launching FDI attacks and proposes three greedy algorithms to obtain perfect and partial protection against these attacks given limited protection budget. While [33] discusses the optimal number of protected measurements to prevent FDI attacks. According to [122], it discusses the bad data attacks that exploit the change in the electricity price to gain financial benefits; it also introduces a method to protect the real-time electricity pricing based on game theory; the method treats the attacker and the defender as players and each one of them chooses a specific strategy; if the two strategies are identical, the defender can resist the attack. If they share in some parts and differ in other parts, means the defender can protect some measurements, but the attacker can successfully compromise other measurements. When the two strategies are different, the defender cannot resist the attack.

- *Cryptographic-based Techniques:* Certain studies adopt security solutions to detect FDI attacks, such as cryptographic techniques or perturbation approach [123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134]. For instance, [123] assumes that the attacker only possesses a perturbed model for power grid; such model may correspond to a partial model of the true system, or an outdated model. Also, the authors synthesize stealthy deceptions attacks in case of linear and nonlinear estimators and utilize the weighted least square estimator as a bad data detector. [124] introduces an obfuscation method to save the privacy of electricity users during the state estimation operation by adding noise to each observed value so that nobody can discover the estimated value. Every billing cycle, the estimator selects a meter to be the leader; this meter chooses some obfuscation values for each meter and sends each value to the corresponding meter, which in turn adds its obfuscation value to its observed measurement before forwarding it to the estimator. The estimator then computes the total obfuscated observed value and extracts the real estimated value from it. [125] launches unobservable attack against the AC power flow equations using only local information, also proposes a probing defense strategy to detect these attacks by enacting sequences of probes and keeping track of the observed quantities. The defense approach relies on perturbing power system by leveraging distributed flexible AC transmission system devices to change the impedance on the chosen lines to create observable changes in the system that the adversary cannot anticipate. [126] secures the power grid estimation process by exploiting the offline information about the grid's

topology to create all possible attacks scenarios in attack graph template; this graph is converted to a hidden Markov model that uses the previous knowledge about the system to detect the possible future attacks. A novel watermarking technique is introduced in [127] to resist FDI attacks in smart grid. A secret watermark is added to real-time meter readings and the watermarked data is transmitted through the unsecured network. At the utility side, the watermarked data can be correlated with original watermark to detect the presence of injected false data. [131] proposes a filtering scheme for the injected false data in wireless sensor networks using batch verification of several reports. While [132] proposes an efficient and secure data aggregation scheme to address impersonation and FDI attacks in the range of NAN; the scheme uses the hop-by-hop security approach that allows source and data integrity using pairwise key and message authentication code. [133] also presents a data aggregation scheme that protects the wireless sensor networks from FDI attacks. Load altering attacks through Internet by distributed software-intruding agents are studied in [134]. These attacks compromise direct load control command, demand side management price, or cloud computation load distribution algorithms to affect the load at the most crucial locations in the grid to cause circuit overflow or other malfunctions and damage in the power system equipment.

- *Other Techniques and Solutions:* Certain studies address other power system cases, such as using the AC power system model without linearizing it to DC model; [135] introduces the implications of hidden FDI attack at the RTU level on AC state estimation. The non-linearity of the power flow equations provides advantages to system operator; attacking the AC system is more difficult than the DC model, because the attacker needs to know more parameters than DC model to launch a successful attack. [136] shows that if the attacker uses the nonlinear FDI version to attack a linear or nonlinear state estimation system, the attack success and can not be detected. While if the linear FDI attacks target a nonlinear state estimation system, the residual state estimator can easily detect it. According to [137], it implements a more practical FDI attack when the attacker has incomplete information about the power system model.

  Furthermore, certain research works study special sort of FDI attack, which is based on re-distributing the power load to gain financial profits. The electricity market utilizes two pricing techniques; look-ahead forward market and real-time spot market. Look-ahead forward market estimates the power price, i.e., locational marginal price (LMP), for the following day using the price information of the present day; while in the real-time spot market, the market operations will re-calculate the ex-post LMP utilizing the current state estimation values. [138] addresses this pricing technique and studies how the attacker can exploit the virtual bidding and ex-post LMP to gain financial profit. Likewise, [139] studies the load redistribution (LR) Attack through a max-min attacker-defender model. The adversary may have one of two goals; immediate or delayed goal. Immediate LR attacks aim to maximize the operation cost immediately; while, delayed LR attacks aim to maximize the total operation cost after the outage of overloaded lines. Two conditions must be satisfied to establish an LR attack: the magnitude of re-distributed load should be large enough, and the false estimated power flow value should exceed its corresponding transmission capacity limit. [139] also proposes a protection strategy based on securing some measurements to detect any attack. Moreover, [141] studies the damaging effect of LR attacks on power system operation and control, when only load bus power injection and line power flow measurements are attackable. [140] shows how FDI attacks can exploit the information from economic dispatch and real-time state estimation in virtual bidding

operation to gain financial profit during the electricity trading. The attacker can buy the electricity from low price location and sell it in high price location. To achieve that, the adversary compromises some units on transmission lines and then indicates some lines to be congested or non-congested so that prices are changed accordingly; in other words, adversary manipulates the prices to gain the highest profit. [142] presents undetectable ramp-induced data attack and addresses its impacts on state estimation and look-ahead dispatch. This attack manipulates the ramp constraint limits of generators within the generation capacity without being detected, subsequently makes a profit in the real-time power market. While, [143] proposes a distributed control mechanism to protect the power system from anomalous actions that use load-shedding techniques.

In summary, most of the proposed works focus on detecting FDI attacks; they analyze different attack models and propose several protection models and control mechanisms, which are mainly based on optimization techniques. Only few works exploit cryptographic schemes to resist these attacks; the reasons could be the limited computation capabilities field devices, which mainly work as measurement tools and hard to preform complex cryptographic operations. Another reason could be the measurement units' locations; they are located in sparse places and hostile environment that eases the attacker's function to compromise the devices while complicates the defender's task to protect them.

In this chapter, we have presented the related proposed works to security and privacy demands for different communication architectures in smart grid. First, the research studies that are proposed to deal with security concerns in customer-side networks are utilizing hardware devices, distorting the message's contents, or using cryptographic schemes to conceal the confidential data and preserving the messages' integrity. Second, we have explained the proposed techniques to secure the vehicles' owner privacy and data integrity during (dis)charging operations in V2G networks, such as authentication mechanisms, anonymization techniques, or physical layer security methods. Lastly, the FDI attacks detection schemes, such as exploiting alternative estimation tests, distributing the estimators to local areas, utilizing various optimization techniques, using cryptographic schemes, are described.

Chapter 4

Lightweight Security and Privacy Preserving Scheme for Smart Grid Customer-side Networks

Information security and customers' privacy in smart grid are significant concerns. Existing security and privacy preserving schemes consider that the consumption reports for electricity consumption aggregation and billing purposes are sent periodically. These periodic messages increase the computation and communication burden on restricted-capabilities smart meters. In this chapter, we propose a lightweight security and privacy preserving scheme that is based on forecasting the electricity demand for a cluster of houses in the same residential area; it limits the cluster's connection with electricity utility only when the cluster needs to adjust its total demand. The scheme efficiently satisfies the security and privacy requirements in customer-side networks, i.e., communication between customers and power utility. At the same time, it significantly reduce the communication and computation overhead. Moreover, the proposed scheme utilizes NTRU crypto-system to further reduce the computation complexity [144, 145]. The remaining of the chapter is organized as follows. Section 4.1 introduces our system model, security parameters, and design goals. Section 4.2 reviews lattice-based NTRU scheme and its signing NSS scheme. In Section 4.3, we present our proposed scheme. Section 4.4 gives security analysis, while Section 4.5 evaluates the performance of our scheme. Finally, Section 4.6 summarize the work.

4.1 System Model

4.1.1 Network Model

Our system model is shown in Figure 4.1. Specifically, we consider a residential area that consists of a number of BANs $=\{BAN_1, BAN_2, \ldots, BAN_m\}$ connected with the main CC via NAN network, which only forwards the messages between BANs in the region and CC without performing any operations. CC is located in the main center for the utility company, and the communication between CC and NAN is through a secured wired connection.

Each BAN has a server with reasonable memory and processing unit, also it has a gateway to connect CC and involved HANs. BAN also connects to a storage unit, which could be

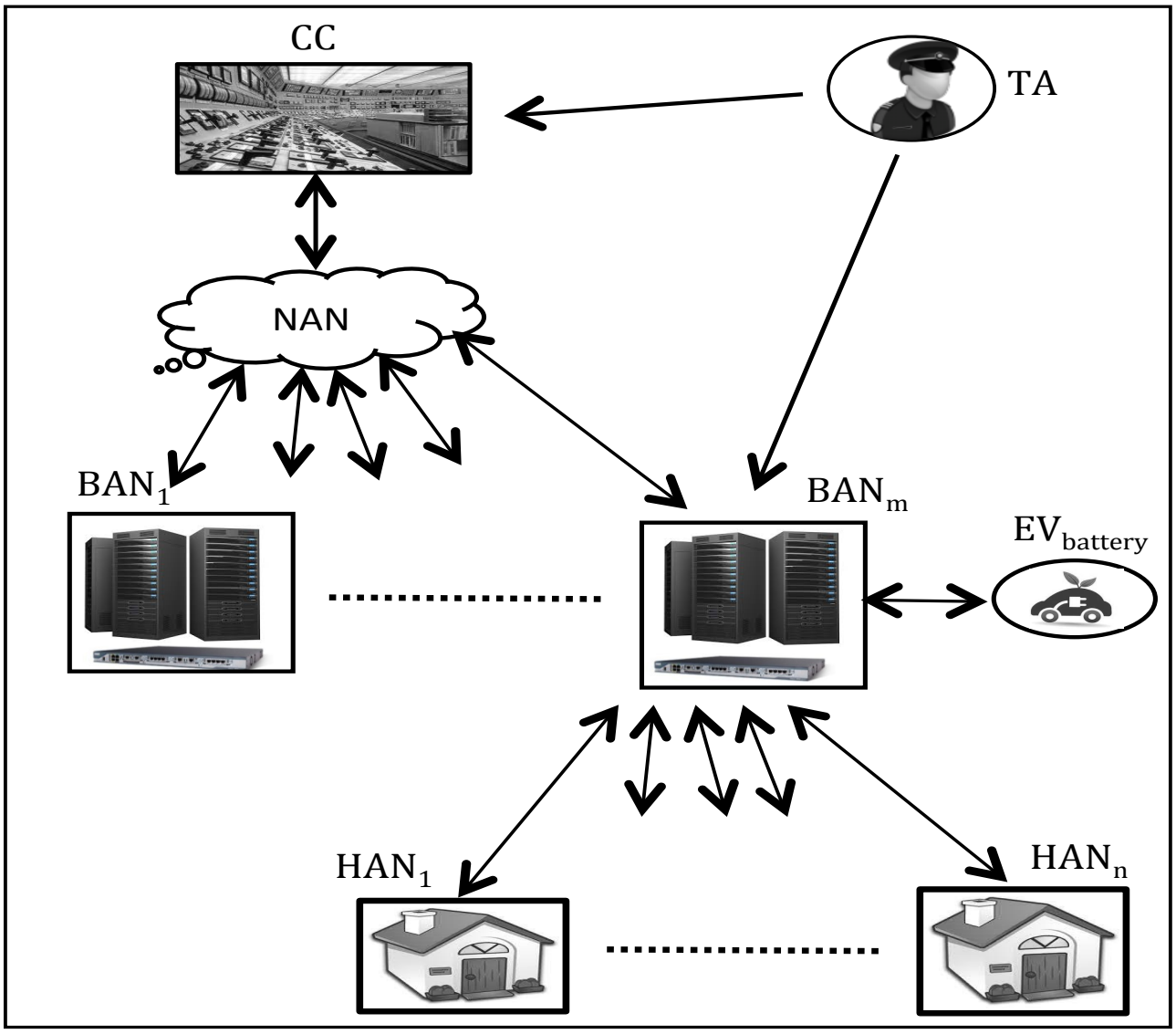

Figure 4.1: System model.

the batteries of EVs owned by some of householders in the area. BAN consists of a cluster of HANs $=\{HAN_1, HAN_2, \ldots, HAN_n\}$; we assume that each BAN has up to 100 HANs to further reduce the overhead on BAN's server. HAN could be a house or a unit in a building; each HAN has a smart meter to estimate its electricity consumption. The communication between BAN gateways and their HANs is through the inexpensive WiFi technology. Both CC and BAN gateways have public keys provided by a trusted authority (TA). Each smart meter has a unique ID issued by TA and stored in a secured place in its memory. CC considers each BAN as one unit and does not know details about the involved HANs in each BAN.

4.1.2 Adversary Model and Security Requirements

We consider that both CC and BANs are honest but curious, i.e., they will not attempt to tamper HANs' data, but they are curious to know the detailed consumption pattern for each user. However, an adversary $\mathscr{A}$ in the region may try to eavesdrop the exchanged messages between different parties or launch some active attacks, such as falsifying the intercepted messages, or

beginning a replay attack. Moreover, $\mathscr{A}$ can launch a DoS attack to make the server unavailable to authorized users. To prevent $\mathscr{A}$'s malicious actions, the following security requirements should be satisfied:

- Customers Privacy: Users' private information are not revealed to outsiders; $\mathscr{A}$ cannot gain any knowledge about individual consumers in the cluster. Also, CC does not need to know the details of individual user's consumption; it considers only total consumption and total bill for each BAN, as it deals with each cluster as a one unit.

- Confidentiality and Messages Integrity: Users' electricity consumption and billing amounts are protected from any adversary. Even if $\mathscr{A}$ eavesdrops any message, he/she cannot extract any information from it. Additionally, integrity should be guaranteed. If $\mathscr{A}$ attempts to resend/modify a message, this malicious action should be detected. In addition, the BAN's database should be secured against any unauthorized access or modification so that any $\mathscr{A}$ cannot intrude or falsify its records.

- Availability: BAN's server should be available to authorized parties all the time, i.e., DoS attacks are prevented.

4.1.3 Design Goals

The objective of the proposed scheme is to preserve the consumers' privacy and information confidentiality in addition to alleviate computation and communication overhead on limited-capabilities smart meters. These objectives can be divided into two folds:

- The proposed scheme should guarantee security requirements for different parties in the network. Customers' privacy should be secured in addition to assure information integrity and confidentiality. Furthermore, the availability of network's resources should be preserved.

- It also should be efficient and lightweight communication and computation overhead so that it is applicable for restricted-capabilities smart meters.

4.2 Preliminaries

Our proposed scheme exploits NTRU crypto-system [146, 147], which is a lattice-based alternative for RSA and ECC public key schemes. In NTRU scheme, encryption and decryption processes are simple polynomial multiplication operations so that NTRU is simple for implementation and very fast compared to other asymmetric schemes. In addition, NTRU scheme is efficiently secured against powerful attacks, such as lattice-basis reduction attacks and attacks through quantum computers.

4.2.1 NTRU Cryptographic Scheme

NTRU utilizes the hardness of shortest vector problem (SVP) [151] and learning with error (LWE) [152] problem. We utilize the revised version of NTRU [148] that exploits the worst-case lattice to define the secret key parameters.

4.2.1.1 NTRU crypto-system

- *Notation*

Let n be a power of 2, $\Phi = xn + 1$, $R = \mathbb{Z}[x]/\Phi, q$ is a prime number that Φ has n linear factors mod q $(q = 1 \, mod \, 2 \, n) : \Phi = \prod_{i \leq n} \Phi_i = \prod_{i \leq n} (x - \Phi_i) mod \, q, R_q = R/qR = \mathbb{Z}[x]_q/\Phi$, and $R_q^{\times}$ is the set of invertible elements of R_q.

- *Key Generation*

Let $n, q \in \mathbb{Z}, p \in R_q^{\times}$, and $\sigma \in \mathbb{R}$. The pair $(sk, pk) \in R \times R_q^{\times}$ is generated by sampling value $\acute{f}$ from the discrete Gaussian distribution $D_{\mathbb{Z}^n,\sigma}$, where $\sigma > Poly(n) \cdot q^{1/2+\epsilon}$ for an arbitrary $\epsilon > 0$, compute the secret key f by:

$$f = p \cdot \acute{f} + 1, \tag{4.1}$$

where $(f \, mod \, q) \in R_q^{\times}$, and $f = 1 \, mod \, p$, and sample secret value g from $D_{\mathbb{Z}^n,\sigma}$ where $(g \, mod \, q) \in R_q^{\times}$. Finally, return secret key $sk = f$ and public key $pk = h$, where:

$$h = pg/f \in R_q^{\times} \tag{4.2}$$

- *Encryption*

To encrypt a message M, sender S generates two random values $s, e \leftarrow \overline{\Upsilon}_{\alpha}$ and computes ciphertext as:

$$C = hs + pe + M \in R_q \tag{4.3}$$

- *Decryption*

Receiver R decrypts C by secret key f as:

$$\acute{C} = f \cdot C \in R_q \tag{4.4}$$

$$M = \acute{C} \, mod \, p \tag{4.5}$$

4.2.1.2 NTRU Signature Scheme (NSS)

NSS [149] is a fast authentication and signature scheme; we utilize the new ring-signature NSS scheme [150].

- *Notations*

Given a prime dimension N, a modulus q, a key size d, and a verification bound parameter NB, there are two polynomials f and g that are invertible modulo q and satisfy that $d+1$ of their coefficients equal 1, d coefficients equal -1, and the remaining equal 0. These parameters are used to compute public key for all users:

$$h = f^{-1} * g(mod\, q). \tag{4.6}$$

Then compute the small polynomials (F, G) satisfying:

$$f * G - g * F = q \tag{4.7}$$

- *Key Generation*

For user i, a random polynomial $r_i \in R_q$ is selected to set:

$$f_i = f * r_i, \quad g_i = g * r_i \tag{4.8}$$

$$F_i = F * r_i^{-1} \tag{4.9}$$

$$G_i = G * r_i^{-1} \tag{4.10}$$

Then the output $Sk_i = (f_i, g_i, F_i, G_i)$

- *Signing Process*

Signer S hashes the message M to create a random vector $(m_1, m_2)(mod\, q)$ and writes m_1, m_2 in:

$$G_i * m_1 - F_i * m_2 = A_i + q * B_i \tag{4.11}$$

$$-g_i * m_1 + f_i * m_2 = a_i + q * b_i \tag{4.12}$$

Then, the signature on M is the polynomial s_i given by

$$s_i = f_i * B_i + F_i * b_i(mod\, q) \tag{4.13}$$

- *Verification*

The verifier V hashes the message M to create the random vector (m_1, m_2), and then computes the value:

$$t_i = s_i * h(mod\, q) \tag{4.14}$$

Afterward, V verifies that if the following condition holds:

$$\| s_i - m_1 \|^2 + \| t_i - m_2 \|^2 \leq NB \tag{4.15}$$

Then, the signature is valid. Two pairs of keys are used in our proposed scheme to provide higher security level; the first pair is used for encryption while the other one is utilized in singing process.

4.3 The Proposed Scheme

Our proposed scheme is divided into two phases. The first phase is initialization phase, which is responsible for establishing the connection among different parties and initializing the electricity supply agreement. The second phase is message exchange phase, which organizes the electricity consumption operation in BAN's region.

4.3.1 Phase 1. Initialization

1) Key generation

TA generates two pairs of keys for CC and BAN gateway as follows.

- Encryption keys

TA computes CC's secret key f_{cc} as:

$$f_{cc} = p \cdot \acute{f}_{cc} + 1$$

, where $(f_{cc} \, mod \, q) \in R_q^{\times}$ and

$$f_{cc} = 1 \, mod \, p$$

and samples g_{cc} from $D_{\mathbb{Z}^n,\sigma}$ so that $g_{cc} \, mod \, q \in R_q^{\times}$.

Then, TA computes

$$h_{cc} = p g_{cc} / f_{cc} \in R_q^{\times}.$$

So, the pair (h_{cc}, f_{cc}) is CC's encryption public and private keys respectively.

For BAN gateway, TA computes its secret key f_{ban} as:

$$f_{ban} = p \cdot \acute{f}_{ban} + 1$$

, where $(f_{ban} \, mod \, q) \in R_q^{\times}$ and

$$f_{ban} = 1 \, mod \, p$$

. Then, samples g_{ban} from $D_{\mathbb{Z}^n,\sigma}$ so that $g_{ban} \, mod \, q \in R_q^{\times}$.

Next, TA computes

$$h_{ban} = p \, g_{ban} / f_{ban} \in R_q^{\times}.$$

Then, the pair (h_{ban}, f_{ban}) is BAN's public and private keys.

- Signing keys

TA chooses polynomials f and g that are invertible modulo q. f and g satisfy that $d+1$ of their coefficients equal 1, d coefficients equal -1, and the remaining equal 0. TA then computes public key for all users:

$$h = f^{-1} * g \,(mod\, q).$$

TA computes small polynomials (F, G), where $f * G - g * F = q$.

To generate signing key for CC, TA selects a random polynomial $r_{cc} \in R_q$ and sets

$$f_{ccs} = f * r_{cc},\ g_{ccs} = g * r_{cc}.$$

Then, TA computes

$$F_{cc} = F * r_{cc}^{-1},\ G_{cc} = G * r_{cc}^{-1}.$$

Therefore, CC's signing key is $Sk_{cc} = (f_{ccs}, g_{ccs}, F_{cc}, G_{cc})$.

To generate signing key for BAN gateway, TA selects a random polynomial $r_{ban} \in R_q$ and sets

$$f_{bans} = f * r_{ban}, g_{bans} = g * r_{ban}.$$

Then, TA computes

$$F_{ban} = F * r_{ban}^{-1},\ G_{ban} = G * r_{ban}^{-1}.$$

Therefore, BAN's signing key is $Sk_{ban} = (f_{bans}, g_{bans}, F_{ban}, G_{ban})$.

- IDs generation

TA assigns a unique secret ID for each smart meter in the BAN's range $ID_1, ID_2, \ldots, ID_n$, where n is the number of HANs within the cluster. TA then sends each ID to the corresponding smart meter. It also forwards all IDs to the BAN gateway, which stores them in the protected BAN's database. Every fixed time period, both BAN gateway and each smart meter apply a specific hash function on the meter's ID to obtain new ID for it. $ID_{new} = h(ID_{old})$

2) Demand forecast

The approximate requirement of electricity for each HAN in the range is computed by a forecasting function, $g()$, from the historical consumption levels of the HAN during a specific time period. For example, $g()$ could be the average electricity consumption every month for the HAN in last 5 years. Applying $g()$, the average electricity consumption value is calculated for all HANs in BAN's cluster so that $HAN_1, HAN_2, \ldots, HAN_n$ have the amounts $x_1, x_2, \ldots, x_n$ respectively, where $x_i = g(HAN_i)$ and n is the number of HANs in BAN region. BAN is responsible for applying $g()$ to obtain the expected electricity share for each HAN. BAN stores the ID for each HAN and the corresponding pair of electricity demand and current price in its database, ID_i, x_i, pc. BAN then aggregates the total demand for all smart meters in the cluster and computes the total required energy amount for BAN for the whole billing period:

$$x = \sum(x_1, x_2, \ldots, x_n) + \epsilon, \tag{4.16}$$

where ϵ is an extra amount of electricity used as backup.

- Backup Value Calculation

After applying $g()$, BAN needs to set a value for ϵ as:

Each BAN gateway connects to a fixed number of HANs m, which runs from 1 to 100 HANs. Also, BAN can predict the expected number of EVs in the area $N_{EV-expected}$. Assume that the available capacity for EV_i to store electricity is C_i, which is known for each EV. So, BAN can compute the total expected available capacity to store extra power as:

$$C_{EV} = \sum_i C_i,$$

where i run from 1 to $N_{EV-expected}$. Then, ϵ is calculated as a ratio of C_{EV}:

$$\epsilon = r * C_{EV},$$

where $0 < r \leq 1$ is a scaling factor. The value of r increases when the number of involved HANs m increases, because more HANs in the cluster requires more backup value for emergency cases. So, as the number of HANs increases, ϵ increases.

During initialization phase, BAN needs to select the optimal number of EVs to work as a storage unit for the cluster. The optimal number of EVs to store ϵ can be computed by:

$$min \quad N_{EV}(m)$$

subject to

$$\begin{aligned} &\epsilon(m) \leq \sum\nolimits_i C_i(m), \quad i \in \{1, \ldots, N_{current}(m)\} \\ &N_{EV}(m) \leq N_{current}(m),\ N_{current}(m) \in \{1, \ldots, N_{Max}(m)\}, \\ &and \qquad m \in \{1, \ldots, 100\} \end{aligned} \tag{4.17}$$

N_{EV} is the optimal number of EVs to store ϵ; $\epsilon(m)$ is the total required electricity backup value for the cluster when the number of HANs equals m, where m can run from 1 to 100 HANs within the same cluster; $C_i(m)$ is the available capacity storage for electricity in EV_i; $i = 1, \ldots, N_{current}$, where $N_{current}$ is the number of EVs that are currently available in the cluster's region from the total number of EVs in the cluster N_{Max}. This optimization model computes the optimal (minimum) number of EVs to store ϵ for BAN's cluster in the case of different number of HANs. ϵ is stored in BAN's storage unit, i.e., number of EVs owned by householders, for emergency. If HANs require more electricity than the assigned share and BAN gateway cannot satisfy the extra share, it supplies the extra electricity from ϵ. However, these cases rarely happen, because the electricity share for each HAN is predefined via accurate forecasting function, and any increase/decrease in the demand is expected to be within a limited range. In certain situations, when one HAN asks for increasing its share, another one may want to decrease the share; so, BAN transfers electricity between them without using ϵ. Generally, if all HANs want to increase their shares, the increase is expected to be within ϵ. But, if the total share's increase is beyond ϵ, a specific procedure is activated to satisfy it.

4) Electricity agreement

BAN considers x as its fixed demand per month. BAN gateway is responsible now for

accomplishing an agreement with CC to supply the connected HANs with their electrical needs per month. CC deals with BAN as a one unit; it has no information about individual HANs in BAN's range.

- *The Agreement Request Message*

- *At BAN gateway*

BAN gateway establishes the connection by sending an agreement request message m_a to CC; BAN signs the electricity amount x, and encrypts it by CC's public key to provide m_a. First, BAN hashes x to create $(x_1, x_2)(mod\, q)$ and writes:

$$G_{ban} * x_1 - F_{ban} * x_2 = A_{ban1} + q * B_{ban1}$$

$$-g_{bans} * x_1 + f_{bans} * x_2 = a_{ban1} + q * b_{ban1}$$

The signature on x is the polynomial s_{ban_1} given by:

$$s_{ban_1} = f_{bans} * B_{ban1} + F_{ban} * b_{ban1}(mod\, q).$$

The result is (x, s_{ban_1}). Then, BAN computes $m_1 = x \| s_{ban_1} \| T_{s1} \| k_1$, where T_{s1} is time stamp and k_1 is a random nonce; they are used to prevent replay attacks. Next, BAN encrypts the message m_1 by CC's public key. BAN sets two random values $s_1, e_1 \leftarrow \overline{\Upsilon}_\alpha$, and uses h_{cc} to obtain the cipher text, $m_a = h_{cc}s_1 + pe_1 + m_1 \in R_q$. Subsequently, BAN gateway sends the agreement request message m_a to CC.

- *At CC*

CC decrypts the received message by its private key. First, CC calculates $\acute{m}_1 = f_{cc}\, m_a \in R_q$, then $m_1 = \acute{m}_1\, mod\, p$. Second, CC verifies BAN's signature s_{ban_1} on the message $m_1 = x \| s_{ban_1} \| T_{s1} || k_1$; CC hashes the message x to create a random vector $(x_1, x_2)(mod\, q)$, computes $t_{ban_1} = s_{ban_1} * h(mod\, q)$, and verifies that $\| s_{ban_1} - x_1 \|^2 + \| t_{ban_1} - x_2 \|^2 \leq NB$. If this condition holds, then the signature is valid. Subsequently, CC checks the validity of timestamp T_{s1} and nonce k_1, and accepts the message if they are acceptable. CC receives many agreement requests from different BANs; it compares the expected total electricity demand for the area with the expected electricity supply and attempts to balance between them. CC should have enough power generation resources to satisfy the electricity requirement for all BANs in the region during the billing period, i.e., one month.

- *The Agreement Response Message*

- *At CC*

If CC accepts the BAN's request, it encrypts the value y, $y = (x, p_e)$, which contains the assigned electricity amount x and the expected price p_e, to obtain the agreement response message m_r. Then, it sends m_r to BAN.

- *At BAN*

BAN receives m_r and decrypts it using its private key and verifies CC's signature. Then, BAN checks the timestamp and nonce validity. Now, BAN guarantees the electricity share x from CC during the whole billing period and knows approximately the expected bill. Figure 4.2 shows the initialization phase.

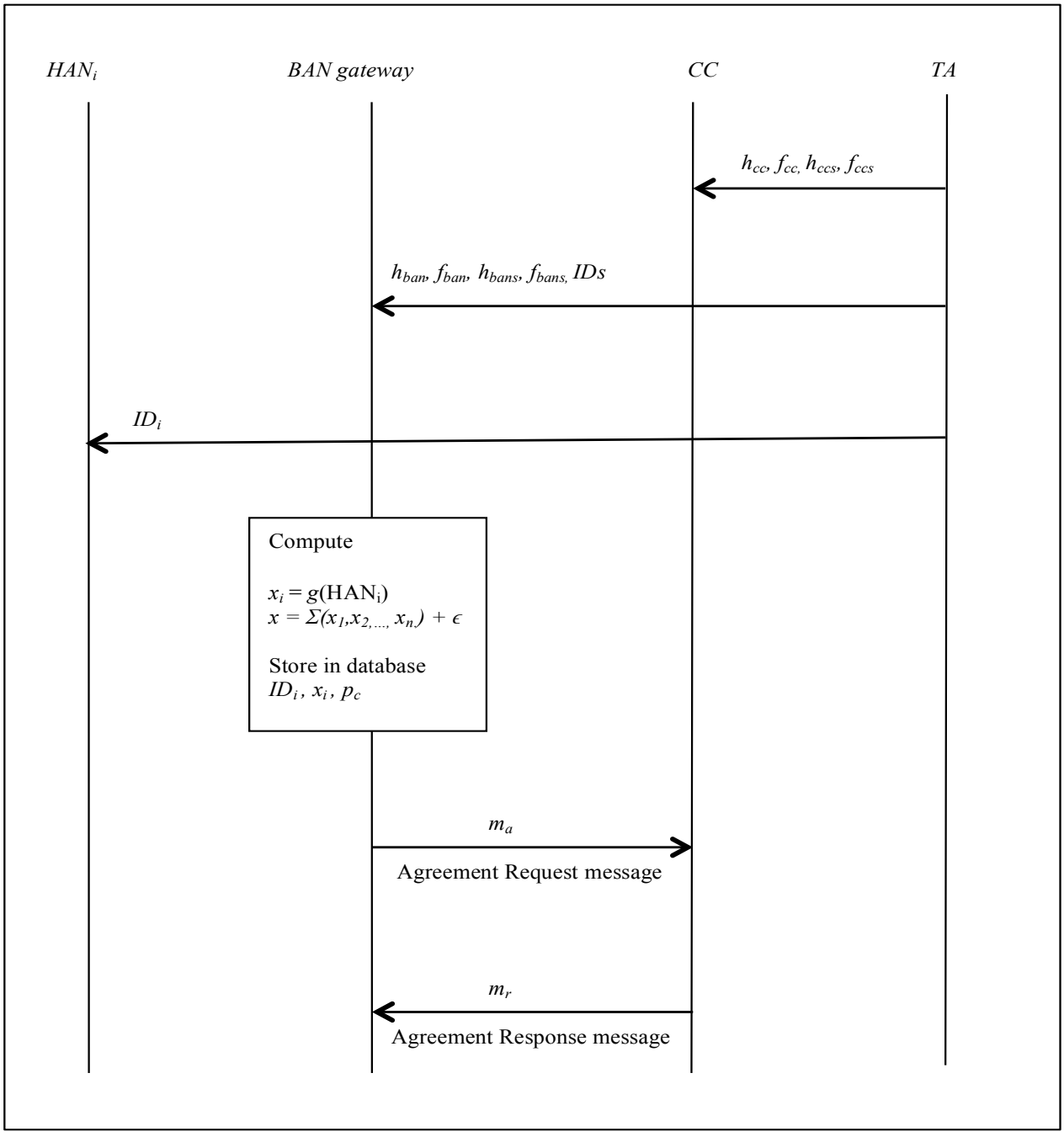

Figure 4.2: Initialization phase.

4.3.2 Phase 2: Exchange Message

At the beginning, BAN gateway supplies each HAN by its electricity share based on the previously calculated amount. BAN computes the current payment b_i for each HAN_i by: $b_i = x_i * p * T_j$, where x_i is the electricity share for HAN_i, p is the current electricity price, and T_j is the time period that the HAN_i consumes its share x_i by the price p. BAN gateway encrypts b_i before storing it in the database, e.g., $b_i = E(b_i)$; only the BAN's operator knows the decryption key.

1) Demand change

- At HAN

If HAN wants to change (increase/decrease) the current share to a new share x_{i-new}, it sends a demand message m_d to BAN gateway. First, a timestamp T_{s3} and nonce k_3 are attached to m_d to prevent replay attacks; $m_3 = x_{i-new} \| Id_i \| T_{s3} \|_{k3}$, where Id_i is the ID for HAN_i. Next, HAN_i encrypts the message m_3 by BAN's public key. HAN_i sets two random values $s_3, e_3 \leftarrow \overline{\Upsilon}_\alpha$; using h_{ban}, it obtains: $m_d = h_{ban} s_3 + p e_3 + m_3 \in R_q$. Subsequently, HAN_i sends demand message m_d to BAN.

- At BAN

BAN gateway decrypts the message m_d. First, BAN calculates: $\acute{m}_3 = f_{ban} \cdot m_d \in R_q$, and then computes $m_3 = \acute{m}_3 \, mod \, p$. Note that the demand message sends only when the power demand for HAN is altered. Thus, the communication overhead is light. BAN supplies HAN with the new share and computes the corresponding new payment for it: $b_{i-new} = x_{i-new} * p * T_j$. BAN encrypts b_{i-new}, and then stores it with pervious payment values in the HAN's record in BAN's database.

2) Price change

When the electricity price changes, BAN receives a price message from CC with the new price p_{new}; this message is broadcasted to all connected BANs. $p_{new} = p_n \| T_{s3} \| k_3$, where p_n is the new price, T_{s3} is a timestamp and k_3 is a random nonce. The price message is sent in plaintext because it is not confidential information; it is only signed by CC's public key. CC hash p_{new} to create $(p_{new1}, p_{new2})(mod \, q)$, and write:

$$G_{cc} * p_{new1} - F_{cc} * p_{new2} = A_{cc2} + q * B_{cc2}$$

$$-g_{ccs} * p_{new1} + f_{ccs} * p_{new2} = a_{cc2} + q * b_{cc2}$$

The signature on p_{new} is the polynomial $s_{cc_2} = f_{ccs} * B_{cc2} + F_{cc} * b_{cc2} (mod \, q)$. The result is the pair (p_{new}, s_{cc_2}).

- At BAN

When BAN receives (p_{new}, s_{cc_2}), it checks the validity of CC's signature s_{cc_2} on $p_{new} = p_n \| T_{s3} \| k_3$: BAN hashes the message p_{new} to create a random vector $(p_{new1}, p_{new2})(mod \, q)$, then computes $t_{cc_2} = s_{cc_2} * h(mod \, q)$, and verify that $\| \, s_{cc_2} - p_{new1} \, \|^2 + \| \, t_{cc_2} - p_{new2} \, \|^2 \leq NB$. If this condition holds, then the signature is valid. Next, BAN checks the validity of T_{s3} and k_3; if they are valid, BAN gateway ensures that the message is legitimate. But, HANs in BAN's range has no connection with CC; they only trust the BAN gateway. So, BAN signs p_{new} by its signing key and then forwards the pair (p_{new}, s_{ban_2}) to the connected HANs.

- At HAN

When HAN receives (p_{new}, s_{ban_2}), it checks BAN's signature s_{ban_2} validity: HAN hashes message p_{new} to create a random vector $(p_{new1}, p_{new2})(mod\, q)$, and compute the value $t_{ban_2} = s_{ban_2} * h(mod\, q)$. Verify that $\parallel s_{ban_2} - p_{new1} \parallel^2 + \parallel t_{ban_2} - p_{new2} \parallel^2 \leq NB$. If this condition holds, then the signature is valid. Next step, HAN checks the validity of T_{s4} and k_4; if they are valid, HAN gateway ensures that the message is legitimate. Notice that the message is sent only when the electricity price changes. If any HAN needs to alter its electricity consumption considering the new price, then it sends a demand message m_d to BAN gateway asking for a new electricity share x_{i-new}. The demand message is sent only when HAN wants to update its electricity share, i.e., the consumed electricity in HAN is increased/decreased. BAN gateway still supplies HAN by the last requested amount of electricity until HAN sends new m_d. Figure 4.3 shows the details of the exchange messages phase.

3) Billing Process

BAN gateway computes the payment values for each HAN by multiplying consumed amount by current electricity price and accumulates it with previous values, and saves the result in HAN's record in the database. These records help in tracking the payment amounts for HANs during the billing period to assure accountability. At the end of billing period, BAN computes the total bill for each HAN $B_i (B_i = \sum_l b_l)$, and aggregates the region's total bill $S (S = \sum_i B_i)$. The billing message S is signed by BAN's private key and encrypted using CC's public key: BAN hashes S to create $(S_1, S_2)(mod\, q)$ and writes:

$$G_{ban} * S_1 - F_{ban} * S_2 = A_{ban3} + q * B_{ban3}$$
$$-g_{bans} * S_1 + f_{bans} * S_2 = a_{ban3} + q * b_{ban3}$$

The signature on S is $s_{ban_3} = f_{bans} * B_{ban3} + F_{ban} * b_{ban3}(mod\, q)$. The result is (S, s_{ban_3}). Then, BAN computes $m_5 = S \| s_{ban_3} \| T_{s5} \| k_5$. Next, BAN encrypts m_5. BAN sets two random values $s_5, e_5 \leftarrow \overline{\Upsilon}_\alpha$, and uses h_{cc} to obtain: $m_b = h_{cc} s_5 + p e_5 + m_5 \in R_q$. Subsequently, BAN gateway sends the billing message m_b to CC.

- At CC

CC uses f_{cc} to decrypt m_b. CC then verifies BAN's signature s_{ban_3}, and checks the validity of timestamp and nonce, and accepts the message if they are acceptable. Figure 4.4 shows the billing process.

C) Electricity Share Adjustment Procedure

Electricity share adjustment procedure is a procedure to handle the case when the fixed assigned share for BAN (x) does not fit the current electricity requirements (y). As in algorithm 1, there are four different cases:

- *Case 1:* If x is slightly larger than y, then the extra electricity is stored in EVs' batteries.
- *Case 2:* If x is slightly smaller than y, then remaining demand of electricity is consumed from stored power in EVs.

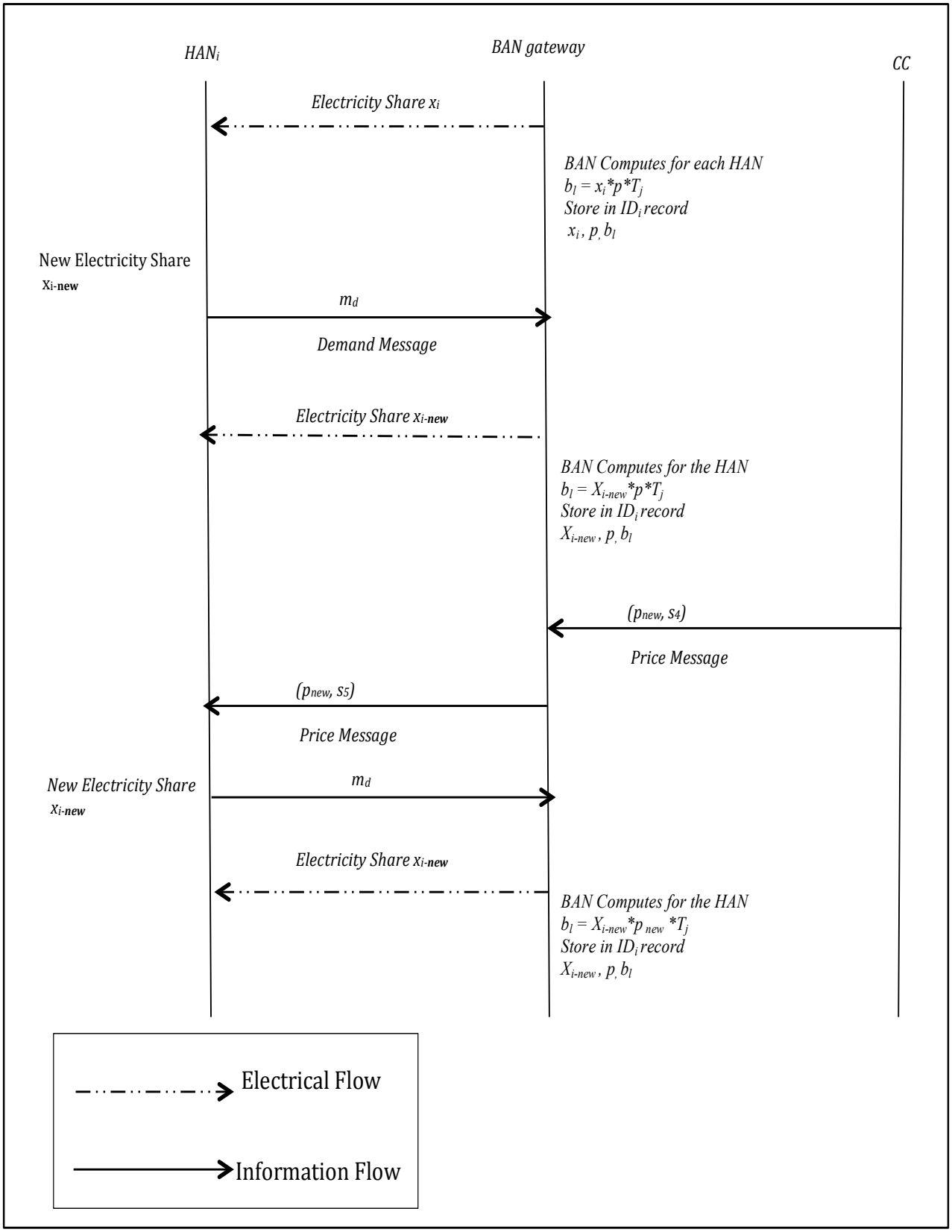

Figure 4.3: Exchange messages phase.

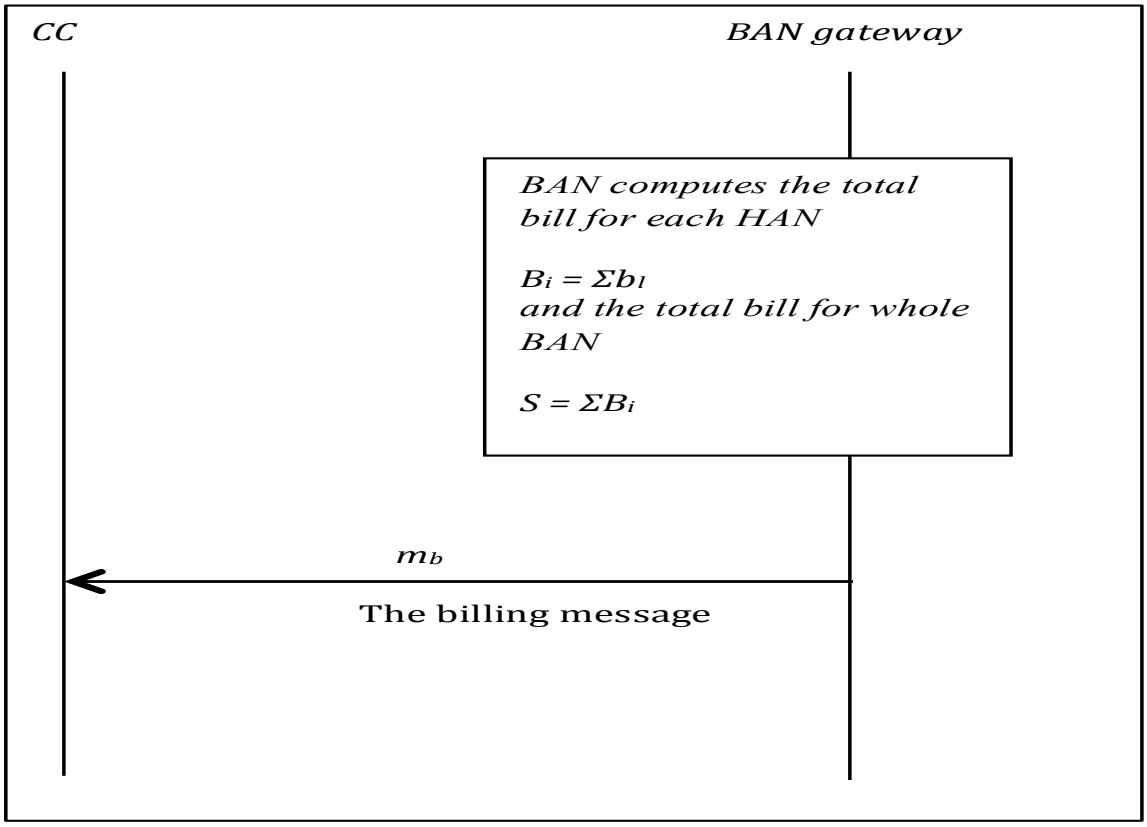

Figure 4.4: Billing phase.

- *Case 3:* If x is much smaller than y and EVs' batteries cannot cover the remaining demand, BAN gateway asks for more electricity from CC.

- *Case 4:* If x is much larger than y and EVs' batteries doesn't have a room for the remaining demand of electricity, BAN gateway sells the extra power to CC.

As the accuracy of employed forecasting function increases, the probability for cases 3 and 4 to occur decreases. However, if case 3 or 4 is repeated, BAN will ask CC to change its fixed value x in the agreement to y.

For further illustration, suppose a network as in Figure 4.5. CC is connected to an BAN via the local NAN, and the BAN includes three HANs. During the initialization phase, the BAN gateway computes the HANs electricity demand from the historical consumed electricity for each HAN. For instance, HAN_1, HAN_2, and HAN_3 consume $100, 200, 300 KWH$ respectively. Mainly, the demand for each of HANs is expected to remain within these levels most of time. The gateway computes the total demand for the month $= [100 + 200 + 300] * 24 * 30 = 432000 KWH$. Accordingly, the BAN gateway asks CC for $435000 KHW$; the remaining amount $\epsilon = 3000 KWH$ will be stored in the EV battery if the BAN urgently needs extra electricity and cannot wait until the CC sends more electricity. If the HANs' demands change, they can exchange the electricity within each other; the HAN with low needs, which has extra power, sends power to the one with large needs. If the exchanged power does not satisfy its need, it consumes the remaining power from the storage units, ie., EVs. So, the BAN cluster rarely needs to ask for more electricity from CC.

Algorithm 1 BAN Electricity Share Adjustment Procedure

BAN Electricity Share Adjustment Procedure
x: The fixed demand for BAN
y: The current actual demand for BAN
z: The EV remaining capacity
$\beta : \beta = \|x - y\|$ The difference between x and y
if $\left(x > y \,\&\, \beta < z\right)$ **then**
    $\beta \rightarrow EV_{battery}$
else if $\left(x < y \,\&\, \beta < z\right)$ **then**
    $\beta \leftarrow EV_{battery}$
else if $\left(x < y \,\&\, \beta > z\right)$ **then**
    $\beta - z \leftarrow CC$
else if $\left(x < y \,\&\, \beta > z\right)$ **then**
    $\beta - z \rightarrow CC$
end if

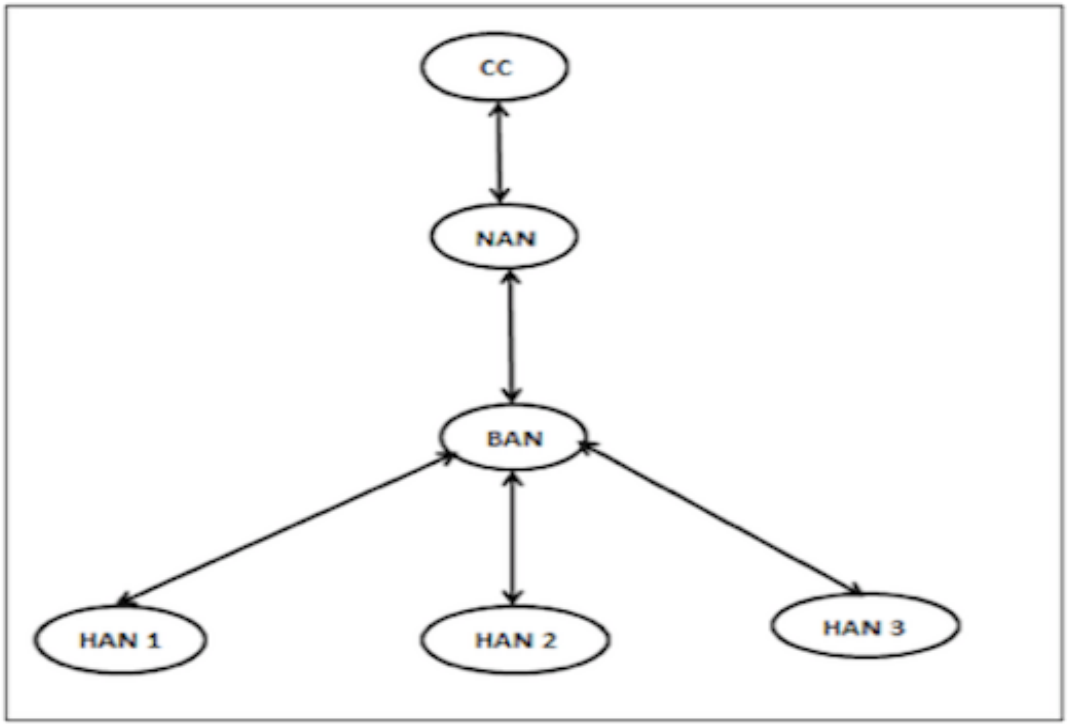

Figure 4.5: An example for a BAN cluster.

4.4 Security Analysis

The proposed scheme is expected to preserve customers' privacy in the cluster. In addition, it guarantees the security requirements: confidentiality, integrity, availability, authenticity, and accountability. This section analyzes the security characteristics of our proposed scheme.

Customers privacy is preserved. Private information that clearly identifies the customers' habits or life style is preserved in the proposed scheme. CC does not know the electricity consumption for each customer in BAN; instead, it receives the bill for the whole BAN S as a one unit. In other words, the individual bills b_is are not exposed to any party outside BAN even the utility. Consequently, any outsider cannot obtain any information about each customer's bill. Moreover, the outsider adversaries cannot know the total bill S, because the billing message is encrypted by CC's public key and CC only can decrypt it. For the adversaries who attempt to compromise the exchanged messages between BAN gateway and different parties, they cannot reveal any information since the exchanged messages between BAN and CC are encrypted using BAN and CC's public keys. Similarly, the messages from HANs to BAN are encrypted by BAN's public key. As a result, no unauthorized party can decrypt these messages

For instance, an adversary $\mathscr{A}$ intercepts the demand request sent from HAN_i to BAN, and tries to detect its share x_i, $\mathscr{A}$ cannot acquire any knowledge about x_i, because the demand message is encrypted by h_{ban}, and only BAN gateway has the decryption key f_{ban}. It is an NP-hard problem to extract f_{ban} from h_{ban}. Even if $\mathscr{A}$ has unlimited computation resources, more powerful than quantum computers, and could compromise f_{ban}, and detect x_i, then he/she also cannot extract any private information about HAN_i, because x_i is HAN_i's share in a relatively large time period and does not reveal the detailed user's consumption pattern. According to BAN's database, it is located in a secured place; no attacker can reach to it. However, if an attacker attempts to compromise the database, he/she needs to discover the applied encryption key to decrypt its contents. Suppose $\mathscr{A}$ succeed to obtain the record for a specific customer from the database, $\mathscr{A}$ knows only the electricity needs for this customer for a long period of time (from one month to 6 hours) and cannot acquire any detailed information about the real-time electricity consumption pattern for that consumer.

Messages' confidentiality is guaranteed. The confidentiality of exchanged messages is guaranteed in the proposed scheme; the agreement messages between CC and BAN are confidential because of using public keys for CC and BANs. Also, the messages from HANs to BAN are encrypted using BAN's public key. If $\mathscr{A}$ tries to impersonate a smart meter to compromise its messages, he/she fails because the smart meter's ID is secured, and $\mathscr{A}$ cannot obtain it. As a result, no impersonation attacks succeed. According to price message, if $\mathscr{A}$ intercepts it, it is not a concern, because he/she can only know the current price of electricity, which is not secret, but cannot modify the message, as it is signed by CC and then BAN's signing keys. Moreover, the message cannot be resend again by $\mathscr{A}$ because of the involved timestamp and nonce values. No man in the middle (MITM) attacks success. If $\mathscr{A}$ only eavesdrops the messages as in passive MITM attack, he/she cannot decrypt the messages. While in active MITM attack, if $\mathscr{A}$ attempts to modify/falsify the message, the attack is detected; as $\mathscr{A}$ cannot mimic BAN's signature or discover the secret IDs for smart meters.

Messages' integrity is assured. The messages from CC/BAN are hashed and signed using its signing key. While the messages from smart meters contain the hashed value of their secret IDs; only BAN can check the validity of smart meter's ID by comparing it with the corresponding ID stored in its database. The integrity of the stored data in the database is assured too; if

any attacker attempts to modify the stored data, BAN's operator detects the attack, because the stored data are encrypted by a secret key known only to the operator.

Authenticity for different parties is guaranteed. Both CC and BAN are authenticated by their public keys. Therefore, the messages encrypted/signed by them are authenticated. As well, the attached unique IDs authenticate smart meters' demand messages.

Availability of resources is confirmed. BAN gateway is always available and no DoS attacks succeed. DoS attack may be launched by a malicious node that sends a huge number of messages to the server until it goes down. In each BAN, there is a specific number of HANs that provides a limited number of messages, and BAN's server capabilities are prepared to deal with that number; this expected number of received messages cannot cause overflow or congestion. Consequently, if BAN notes that the number of messages is larger than expected or an individual HAN sends a huge number of messages, BAN does not response to these requests and isolates the malicious HAN. For instance, consider a BAN's cluster with 80 HANs and the price is expected to change three times a day. Then, BAN gateway assumes that all HANs will send a demand change message for every price change (e.g., three demand messages per HAN per day). Therefore, the maximum number of demand messages that BAN gateway can receive when the price changes is 80, and the total number of demand messages during the day is around 240. Accordingly, if BAN gateway receives larger number of messages in a short period of time, it discards the messages and blocks the malicious HAN

Accountability and tracking historical processes are guaranteed. If any householder wants to validate the bill's value, he/she can check his/her monthly record in BAN's database. HAN's record indicates the amount of consumed electricity and the corresponding price. These records assure the correctness of payment amount for each HAN in the BAN cluster.

The limited number of transmitted messages enhances security and customer's privacy and reduces adversary's chance to acquire any knowledge about the system. In addition, the deployed NTRU crypto-system prevents attackers from extracting any knowledge about private keys from the corresponding public keys or any intercepted message. Suppose an $\mathscr{A}$ with reasonable computation capabilities captures a message exchanged between BAN gateway and CC or between BAN and one of the HANs, such as the billing message m_b sent from BAN to CC. $\mathscr{A}$ cannot extract any information from the message, because $\mathscr{A}$ needs to know CC's private key f_{cc} to decrypt m_b and obtains $m_5 (m_5 = \acute{m}_5 \, mod \, p$, and $\acute{m}_5 = f_{cc} \cdot m_b \in R_q)$. In addition, $\mathscr{A}$ cannot modify the message, as he/she requires BAN's private signing key parameters (f_{bans}, g_{bans}) to forge its signature on the message. $\mathscr{A}$ also cannot resend the message, as it contains a timestamp and nonce number. CC and BANs' secret parameters $(f_{cc}, g_{cc}, f_{ccs}, g_{ccs}, f_{ban}, g_{ban}, f_{bans}, g_{bans})$ are designated exploiting the hardness of SVP and LWE problems. If $\mathscr{A}$ attempts to compromise CC's private key f_{cc}, he/she requires to check all the non-zero vectors in $R \times R_q$ field, and according to SVP, if a lattice L with norm $\mathscr{N}$ and a basis of a vector space V are given, it is an NP-hard problem to find the shortest non-zero vector υ in V, given that $\mathscr{N}(\upsilon) = \lambda(L)$, even by the powerful lattice basis reduction algorithms. Even if $\mathscr{A}$ manages to formulate a number of approximate equations n to determine f_{cc}, the problem will be converted to LWE problem as:

$$\langle f_{cc}, a_1 \rangle \approx_\chi b_1 (mod \, p),$$
$$\langle f_{cc}, a_2 \rangle \approx_\chi b_2 (mod \, p),$$
$$\ldots\ldots\ldots,$$
$$\langle f_{cc}, a_n \rangle \approx_\chi b_n (mod \, p),$$

Then, $\mathscr{A}$ requires $2^{O(n)}$ equations/time using best known algorithm to solve LWE problem to obtain f_{cc} value, which is an NP-hard problem [152, 153]. Consequently, $\mathscr{A}$ cannot compromise the secret key f_{cc} even via a quantum computer. As a result, the data confidentiality and integrity are guaranteed. In addition, the authenticity of different parties is confirmed.

In conclusion, our proposed scheme preserves customers' privacy and fulfills different security requirements for the involved parties in customer-side network.

4.5 Performance Evaluation

This section studies communication and computation complexity for our scheme.

4.5.1 Communication overhead

The number of exchanged messages between different parties (CC, BAN, and HANs) is very small. During initialization phase, CC and each BAN are exchanging two messages to setup the electricity-share agreement, but HANs do not participate in this phase. In second phase, HANs send demand messages only if they require altering their electricity share due to change in HANs' consumption or in electricity price. Note that HAN's electricity share may remain the same during the whole billing period; therefore, no demand messages are sent. Furthermore, two price messages are disseminated only in case of price modification; first one is broadcasted from CC to connected BANs; and second message is forwarded from BAN to its HANs. In addition, BAN should send one billing message to CC indicating the payment amount for whole BAN. However, this message is sent once at the end of billing period, i.e., every month.

Consequently, the total number of messages is changed from three messages (two agreement messages, and one billing message) to $d+3$ messages, where d is the number of demand messages during the month. Therefore, d is a small number as the electricity share for each HAN is not expected to change frequently. Since the electricity share is predefined for each HAN, we assume that HAN updates its share when electricity price changes only. If Time-Of-Use pricing plan [154] is used, there are three different prices during the day known as off-peak, mid-peak and on-peak prices. Therefore, we consider that the maximum number of demand messages is three demand messages per HAN every day. While in the traditional periodic-pattern schemes, each HAN sends its reading message every one hour or 15 minutes.

Figure 4.6 shows communication overhead for our proposed scheme versus a periodic-pattern scheme in terms of different number of connected HANs. As shown, our proposed scheme saves a significant number of messages compared with the traditional periodic schemes. In a cluster of 100 HANs, our scheme requires at maximum 300 demand messages per day, while periodic schemes need 2400 messages (if demand messages are sent every hour).

Figure 4.7 shows the variation in communication burden for our proposed scheme at different number of demand messages. The figure includes six different scenarios: *Case 1:* only portion of HANs send one demand message a day, while the remaining does not send any messages. *Case 2:* all HANs send one demand message per day. *Case 3:* number of HANs send two messages, other group of HANs send one message, and the remaining does not send any. *Case 4:* all HANs send two demand messages per day. *Case 5:* group of HANs send three messages, other number of HANs send two messages, some HANs send one message, and the remaining does not send any messages. *Case 6:* all HANs send three demand messages every day. Axiomatically,

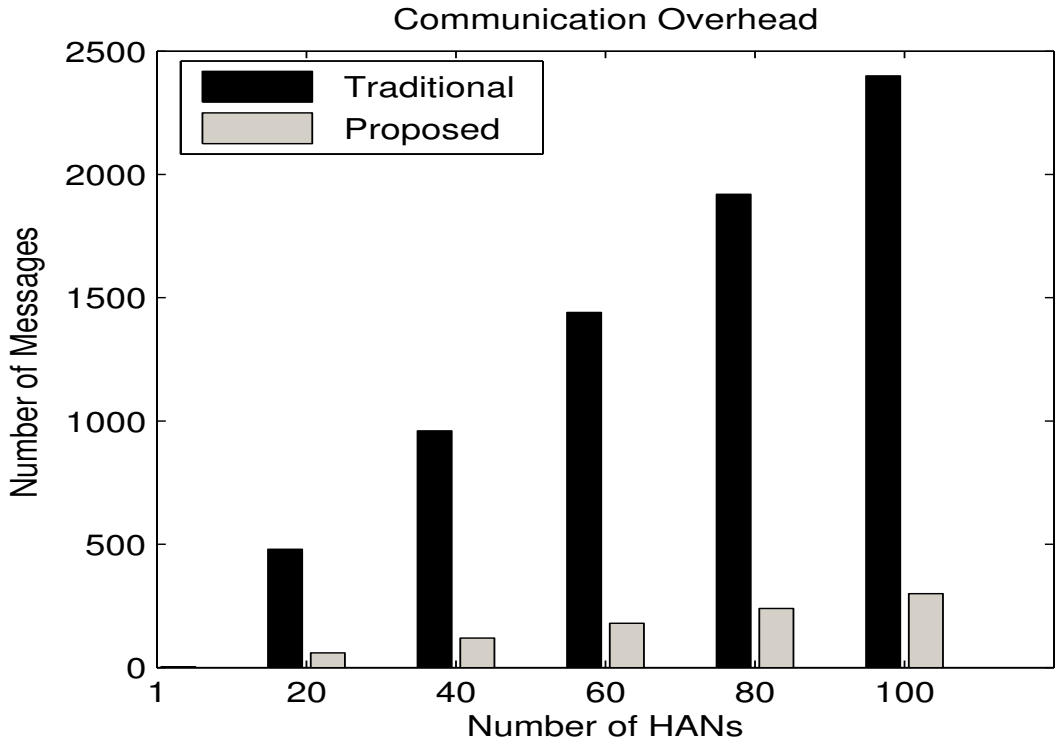

Figure 4.6: Communication overhead
Traditional .vs. Proposed scheme.

the communication complexity grows as number of demand messages increases but within a limited value. The number of messages rises from one message when the cluster has only one HAN to maximally 300 messages when the cluster has 100 HANs. Thus, the maximum number of messages is no more than 300 messages per month, which is a trivial communication overhead for the network.

4.5.2 Computation complexity

Suppose that computation time for encryption, decryption, signing, and verification operations are T_e, T_d, T_s, and T_v respectively. During initialization phase in our scheme, CC and each BAN exchange agreement request and agreement response messages. Thus, both CC and BAN need to perform one encryption, one decryption, one signing and one verification process. Accordingly, the computation time for this phase is $2*(T_e+T_d+T_s+T_v)$ units. In second phase, if HAN's share changes, HAN preforms one encryption operation per each demand message, while BAN decrypts the message. Thus, we have (T_e+T_d) units per message. In case of modified price, two price messages are sent from CC to BANs and from each BAN to its connected HANs. As a consequent, both CC and BAN sign the price message once; as well, BAN and HAN verify the message once. Hence, the computation time for price message is $2 * (T_s + T_v)$ units. So, the total computation operations in this phase equals $m * (T_e + T_d) + (2 * T_s + (m + 1) * T_v)$, where m is the number of HANs in the cluster. During payment process, only billing message is sent from BAN to CC; it requires one encryption, one decryption, one sign and one verification process. The computation time for this message is $(T_e + T_d + T_s + T_v)$ units. While, the performed operations on BAN's database, i.e., computing electricity shares and bills, are trivial computation loads and can be neglected.

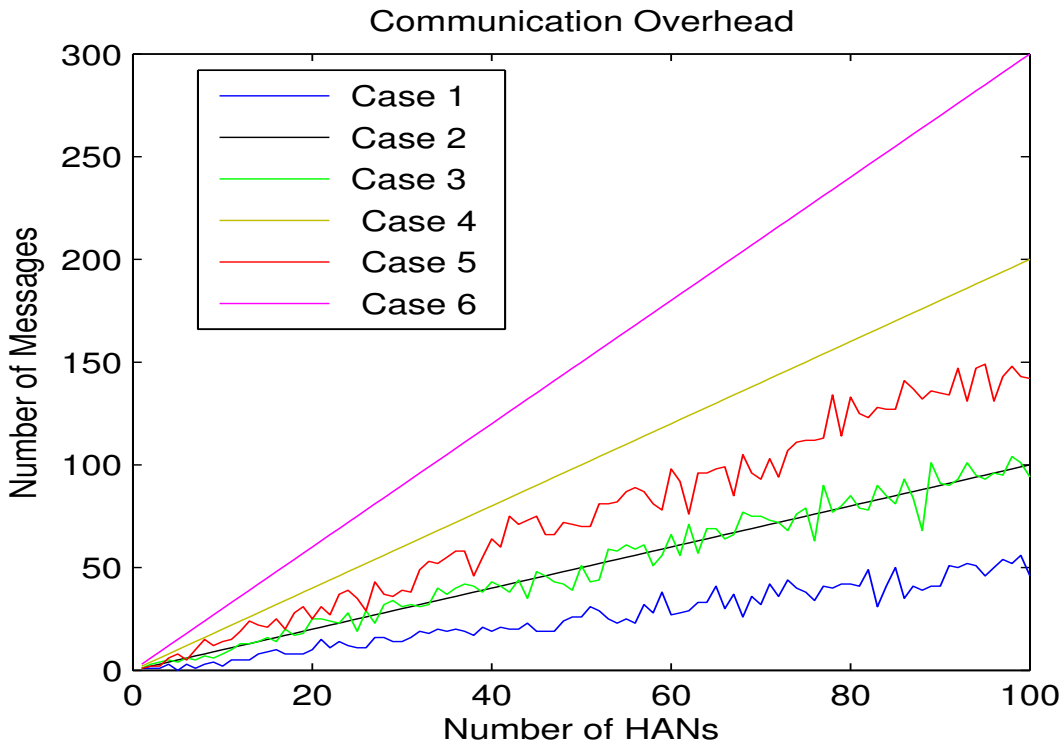

Figure 4.7: Communication overhead for proposed scheme different cases.

We exploit the moderate security mode for public key NTRU crypto-system with private key = 530 bits, public key = 1169 bits, and plaintext size = 187 bits. For NSS signing parameters, the used private key = 502 bits, public key = 1757 bits, and signature size = 1757 bits. We expect three demand messages per HAN every day in our scenario; each demand message requires two operations: one encryption and one decryption. Also, three price messages are applied; each one requires two signing and two verification operations. BAN also requires three signing and three verification operations every day for the first send of the price message from CC to BAN. Moreover, the second price message from BAN to each HAN requires three signing and three verification operations per day. So, the operations for one HAN per day are three encryption, three decryption, three signing and three verification operations.

On the other hand, each HAN sends its reading message every hour in the periodic-pattern schemes. Each reading message necessitates four processes: one encryption, one decryption, one signing and one verification operation. In addition, CC should reply by control message; if we assume these control messages are sent as price changes. Thus, there are three control messages per day for each HAN. As result, each HAN needs 27 encryption, 27 decryption, 27 signing and 27 verification operations per day. TABLE 4.1 demonstrates the total computation overhead for our proposed scheme versus a periodic-pattern scheme for each HAN per month.

As the number of HANs in the cluster increases, the computation overhead increases. Consequently, the total computation operations for the whole cluster in our proposed scheme per month equal $[2*(T_e+T_d+T_s+T_v)]+30*([3*m*(T_e+T_d)]+[6*(T_s)+3*(m+1)*(T_v)])+[(T_e+T_d+T_s+T_v)]=[(90*m+3)*(T_e+T_d)]+183*T_s+[(90*m+93)*T_v]$ operations, where m is the number of HANs in the BAN cluster. While, the periodic-pattern scheme computes $810*n*(T_e+T_d+T_s+T_v)$ operations per month, where n is the number of connected HANs

Table 4.1: Total computation overhead per HAN per month
Traditional .vs. Proposed scheme

	Computation Overhead
Traditional	$810 * T_E + 810 * T_D + 810 * T_S + 810 * T_V$
Proposed	$90 * T_E + 90 * T_D + 90 * T_S + 90 * T_V$

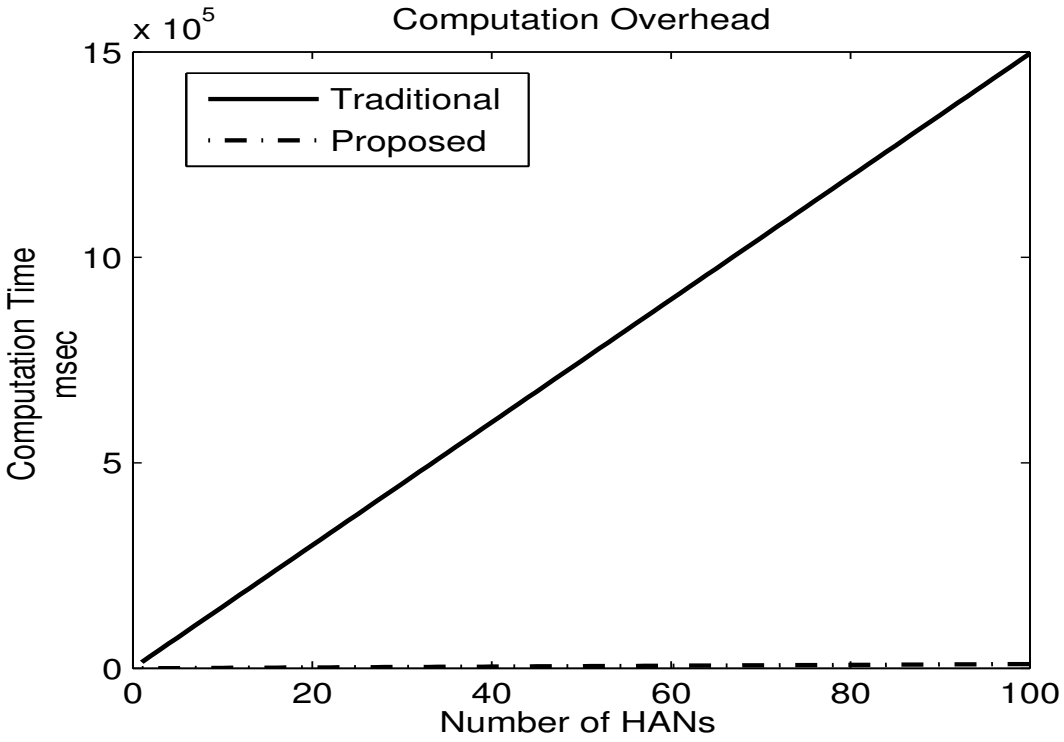

Figure 4.8: Computation overhead
Traditional .vs. Proposed scheme.

(we assume that m and n have the same value. However, n value could be much greater that m).

Figure 4.8 shows the total computation time of our proposed scheme implementing NTRU as encryption scheme and NSS as signing scheme versus a traditional periodic scheme implementing RSA for encryption and signing operations. As indicated in the figure, when the number of HANs increases, the increase in computation time in the proposed scheme is significantly less than the one in traditional scheme. In the proposed scheme, the computation time increases from 161.82 to 7557.12 msec per month. While the traditional scheme computation time increases from 14960 to 1496070 msec per month. In other words, our proposed scheme notably reduces the overall computation time and saves the limited resources of the network.

We evaluate the performance of our proposed scheme versus the traditional periodic scheme when both schemes exploit NTRU crypto-system as encryption scheme and NSS as signing scheme. Figure 4.9 compares the worst case of our proposed scheme, when all HANs in the cluster send their maximum number of demand messages, with the moderate case of the traditional scheme, when every HAN sends periodic demand message every hour. It can be seen that there is significant difference in computation overhead between the two schemes; our pro-

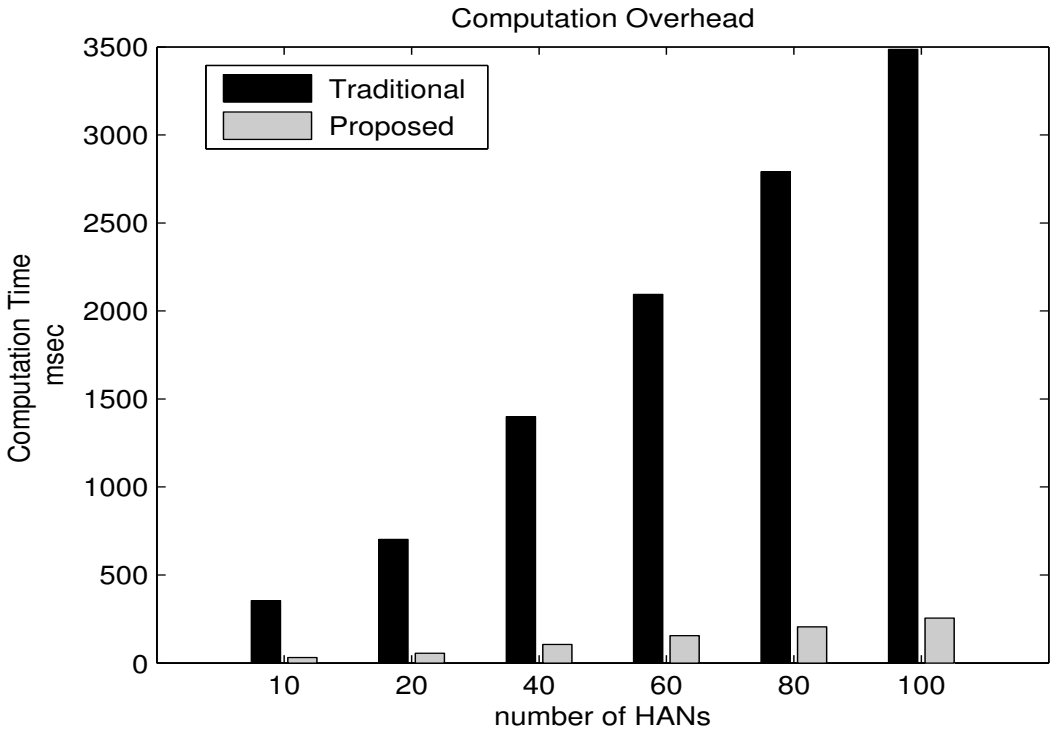

Figure 4.9: Computation overhead
Traditional .vs. Proposed scheme.

posed scheme consumes much less computation time than the traditional one, especially as the number of HANs increases. In the proposed scheme, the computation delay increases from 8.7 to 255.21 ms per day as the number of HANs increases from one to one hundred. While the traditional scheme computation time increases from 410 to 3486.2 ms per day. Consequently, our proposed scheme remarkably decreases the overall computation time.

If the number of demand messages varies from zero (when the HAN does not want to change its share during the day) to the maximum number of messages (e.g. three messages a day), the total computation operations for the whole cluster in our proposed scheme per month equals $[2*(T_e+T_d+T_s+T_v)]+30*([d*m*(T_e+T_d)]+[6*(T_s)+3*(m+1)*(T_v)])+[(T_e+T_d+T_s+T_v)] = [(30*d*m+3)*(T_e+T_d)]+183*T_s+[(90*m+93)*T_v]$ operations, where m is the number of HANs in the BAN cluster, and d is the number of demand messages, $d \in \{0,1,2,3\}$.

Figure 4.10 shows the impact of demand messages' number on the computation time in our proposed scheme (Figure 4.10 includes the same six cases as in Figure 4.7). Although computation complexity rises as number of demand messages increases, this increase is not a heavy computation overhead on the network's resources. For instance, computation time per day increases from 7.14 ms when the cluster has only one HAN to 225.21 ms as maximum when the cluster has 100 HANs. In conclusion, our proposed scheme not only guarantees security and privacy requirements for customer-side network, but also ensures low communication and computation overhead.

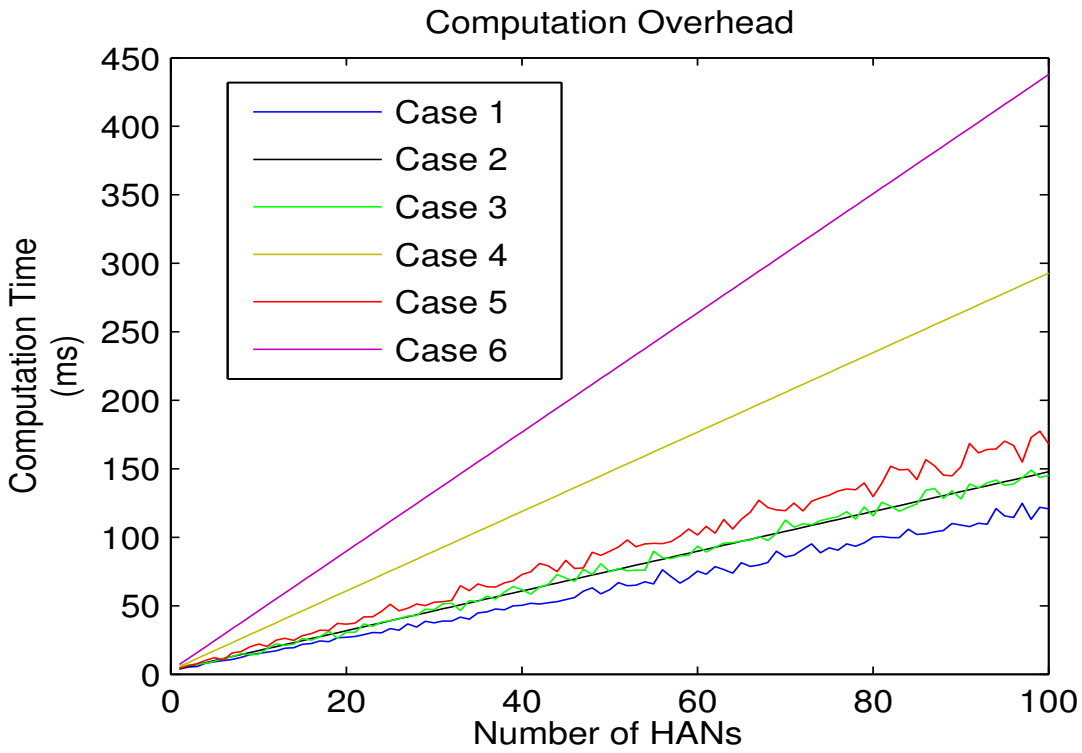

Figure 4.10: Computation overhead for proposed scheme different cases.

4.6 Summary

Consumers privacy and information confidentiality are major concerns for customer-side networks in smart grid. In contrary to the existing solutions, we have proposed a lightweight security and privacy preserving scheme based on predicting the expected electricity demand for a cluster of HANs. The proposed scheme guarantees the electricity customers' privacy in addition to assure the confidentiality and integrity of the exchanged electricity consumption messages. It also restricts the connection with the electricity provider only when the total cluster's demand needs to be adjusted. Security analysis and simulation results demonstrate that the proposed scheme satisfies security and privacy requirements for householders, at the same time, guarantees light communication and computation burden.

Chapter 5

A Lightweight Lattice-based Homomorphic Privacy-Preserving Data Aggregation Scheme for Smart Grid

As previously mentioned, consumers privacy and their consumption information confidentiality and integrity are the main security concerns for smart grid connection with the residential electricity consumers. Only few research works attempt to conserve consumers' privacy during appliances' readings collection, i.e., before smart meter sends house's consumption to the local CC. [155] proposes a secure in-network data aggregation utilizing an orthogonal chip code and circuit shifting operation to guarantee the confidentiality and anonymity of APs' information. However, this scheme requires sharing the chip code with APs and performing mutual authentication. In addition, SM can reconstruct the original reading for each AP from the mixed data. In our work [156, 157], we propose a lightweight lattice-based homomorphic privacy-preserving electricity consumption aggregation scheme, in which smart household appliances aggregate their readings without involving the related smart meter. Although smart meters or the intermediate base station cannot decrypt this aggregated consumption, they can validate the message's authenticity. The proposed scheme also investigates the impact of different types of smart appliances on the HAN overhead. The total communication and computation load for the proposed scheme is trivial and tolerable by different parties in the connection, i.e., smart appliances, smart meters, and the base station. In addition, the deployed crypto-system, which depends on simple arithmetic operations, can further reduce the computation duty for smart appliances. Simulation results and security analysis show that our proposed scheme guarantees consumers privacy and messages authenticity and integrity with lightweight communication and computation complexity. The remaining of the chapter is organized as follows: Section 5.1 describes the system model and security requirements. Sections 5.2 reviews the lattice-based homomorphic encryption scheme. Sections 5.3 presents our proposed scheme. The security analysis and performance evaluation are illustrated in Sections 5.4 and 5.5 respectively. Finally, Section 5.6 concludes the work.

5.1 System Model

5.1.1 Network Model

Consider a residential area that consists of the main CC for service provider, i.e., utility company, that is connected to a number of base stations (BSs) located in different areas $BSs = \{BS_1, BS_2, \ldots, BS_h\}$. Each BS is responsible for a cluster of HANs in its local region $HANs = \{HAN_1, HAN_2, \ldots, HAN_m\}$. HAN could be a townhouse or a unit in a building; each HAN has a smart meter (SM) that connects to the house's appliances (APs); $APs = \{AP_1, AP_2, \ldots, AP_n\}$. APs also can communicate to each other directly without involving SM. The communication inside HAN is through inexpensive short coverage distance technology, such as Bluetooth or ZigBee. While, the connection between HANs and the corresponding BS is through inexpensive WiFi technology. CC, BSs, and SMs have public keys provided by an independent TA. Each AP has a unique ID that issued to it by TA and stored in a secured memory. Figure 5.1 shows the system model.

5.1.2 Adversary Model and Security Requirements

We consider that CC, BSs, and SMs are honest but curious. However, an adversary $\mathscr{A}$ can eavesdrop the exchanged messages between different parties, i.e., the messages among APs and between them and SM, also the forwarded messages from SMs to BS, to extract consumers' personal information. $\mathscr{A}$ may establish some active attacks, e.g., falsify the captured messages or begin a replay attack; also, $\mathscr{A}$ may compromise SMs. Thus, we should thwart $\mathscr{A}$'s malicious actions by guaranteeing:

- *Consumers' Privacy:* assure that any attacker could not gain any knowledge about HAN's consumption. In addition, CC should not know the detailed consumption pattern for each customer in the region.

- *Authenticity and Data Integrity:* guarantee the confidentiality and authenticity of customers' consumption; even if $\mathscr{A}$ already intercepts a message, he/she cannot extract any knowledge. Likewise, we should ensure messages' integrity; suppose $\mathscr{A}$ attempts to resend/modify a message, we should detect these malicious actions.

5.1.3 Design Goals

The main objective of the proposed scheme is fulfilling the security requirements for the network; it should guarantee consumers' privacy. It also should prevent any illegal access/modification of messages/devices in addition to be efficient and lightweight in terms of communication and computation overhead.

5.2 Preliminaries

Lattice-based Homomorphic Encryption Scheme

Our scheme exploits the lightweight lattice-based homomorphic encryption scheme [158], which utilizes the vector space structure to encrypt messages as noisy lattices. So, it guarantees

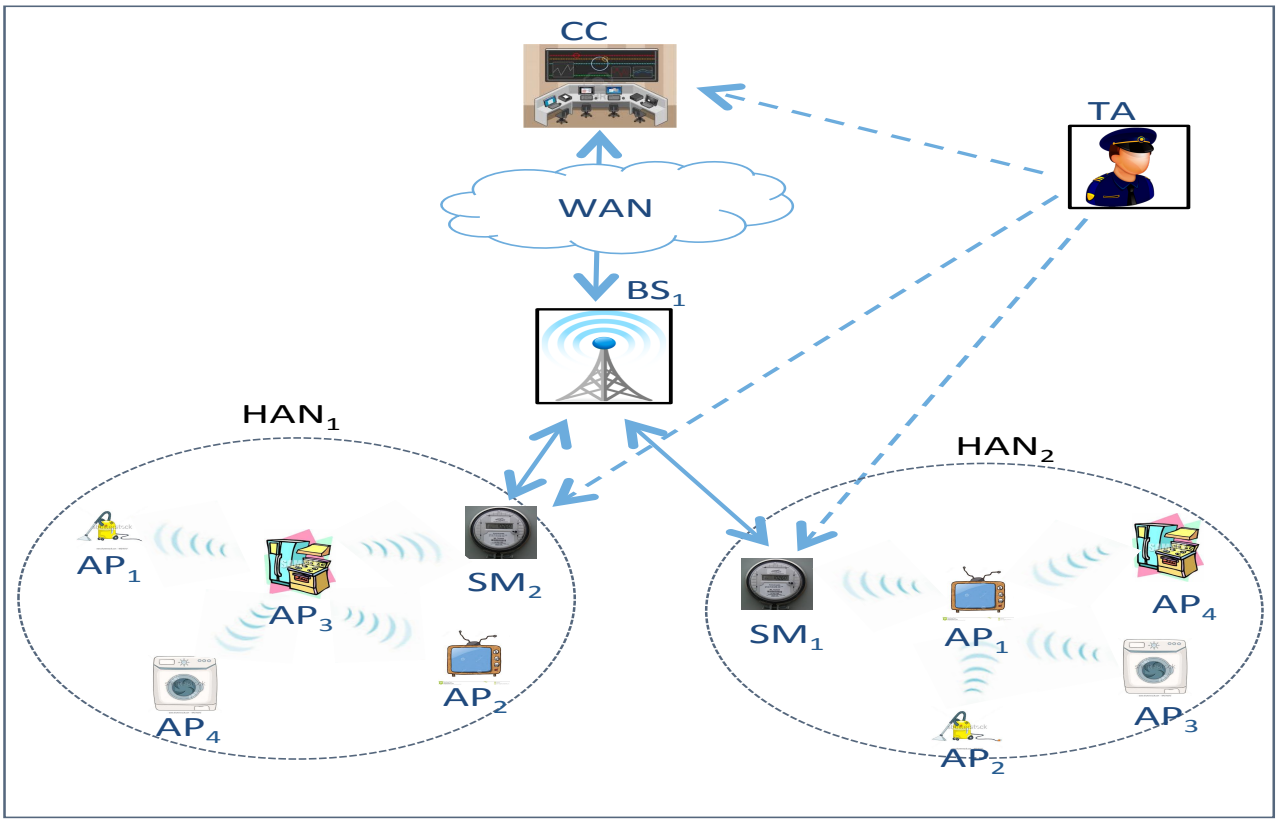

Figure 5.1: System model.

messages' security with low computation complexity, as it mainly performs simple addition and multiplication operations in vector space.

Key Generation:

The scheme defines five global integer parameters:

N is the number of coordinates of plaintext vectors,

r is the characteristic of the ring over which they are constructed,

l is the maximum number of homomorphic operations that can be done,

n is the number of softly disturbed matrices in the public key,

and ε_{max} is an upper bound for the coordinates of random vectors used to insert noise.

Let $l_0 = n \times N \times \varepsilon_{max} + (N-1) \times r$, $q = 2 \times l_0 \times (2l+1)$, and $p = q \times r + \varepsilon$ is a prime number, $\varepsilon < l_0$. Then, generate two random $N \times N$ matrices over $GF(p)$: $\boldsymbol{A}$ and $\boldsymbol{B}$, where $\boldsymbol{A}$ is invertible and $\boldsymbol{M} = [\boldsymbol{A} \mid \boldsymbol{B}]$. Also, generate a random scrambling matrix $\triangle$, which is an $\boldsymbol{N \times N}$ diagonal invertible matrix over $\boldsymbol{GF(p)}$. Compute $\ddot{\boldsymbol{M}}_{\boldsymbol{i}} = [\boldsymbol{A}_{\boldsymbol{i}} \mid \boldsymbol{B}_{\boldsymbol{i}}]$ by multiplying $\boldsymbol{M}$ to the left of a random invertible matrix $\boldsymbol{P}_{\boldsymbol{i}}$. Subsequently, generate a soft noise matrix $\boldsymbol{D}_{\boldsymbol{i}}$, a random $\boldsymbol{N \times N}$ matrix over $\boldsymbol{\{-1, 1\}}$, for each $\boldsymbol{i \in \{1, 2, \ldots, n\}}$. Next, compute softly distributed matrix $\dot{\boldsymbol{M}}_{\boldsymbol{i}} = [\boldsymbol{A}_{\boldsymbol{i}} \mid \boldsymbol{B}_{\boldsymbol{i}} + \boldsymbol{D}_{\boldsymbol{i}} \triangle]$. Similarly, compute a hard noise matrix $\boldsymbol{D_0}$ by generating a soft noise matrix then replacing the diagonal values by $\boldsymbol{q}$. Then, compute the hardly distributed matrix $\dot{\boldsymbol{M}}_{\boldsymbol{0}} = [\boldsymbol{A_0} \mid \boldsymbol{B_0} + \boldsymbol{D_0} \triangle]$. Next, choose a permutation operation $\mathscr{P}\boldsymbol{(.)}$, and compute $\boldsymbol{M}_{\boldsymbol{i}} = \mathscr{P}(\dot{\boldsymbol{M}}_{\boldsymbol{i}})$,$\boldsymbol{i \in \{1, 2, \ldots, n\}}$.

Finally, the $\boldsymbol{n+1}$ matrices $\boldsymbol{\{M_0, M_1, \ldots, M_n\}}$ are the public key. While the private key consists of the permutation $\mathscr{P}\boldsymbol{(.)}$, the hidden matrix $\boldsymbol{M}$, and the scrambling matrix $\triangle$.

Encryption:

First, the plaintext message is constructed as a message vector $\boldsymbol{m}$ in $\boldsymbol{Z_r^N}$. Then, $\boldsymbol{m}$ multiplies by the hard noise matrix $\boldsymbol{M_0}$. The result is disturbed by adding $\boldsymbol{mM_0}$ to the summation of $\boldsymbol{n}$ soft noise vectors $\sum_i \boldsymbol{r_i} * \boldsymbol{M_i}$, where $\boldsymbol{r_i}$ are $\boldsymbol{n}$ random vectors with coordinates smaller than $\boldsymbol{\varepsilon_{max}}$. Then, the ciphertext is

$$\boldsymbol{c} = \boldsymbol{mM_0} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{r_i} * \boldsymbol{M_i} \tag{5.1}$$

Decryption:

The decryption operation is based on filtering the added noise. First, the permutation is reversed as

$$\dot{\boldsymbol{c}} = \mathscr{P}^{-\boldsymbol{1}}(\boldsymbol{c}) \tag{5.2}$$

, where $\boldsymbol{c} \in \boldsymbol{GF(p)^{2N}}$ is the ciphertext. Then, the receiver computes the scrambled noise

$$\boldsymbol{e} = \dot{\boldsymbol{c_D}} - \dot{\boldsymbol{c_U}}\boldsymbol{A^{-1}B} \tag{5.3}$$

, where $\dot{\boldsymbol{c_D}}, \dot{\boldsymbol{c_U}}$ are the undisturbed and disturbed halves of $\dot{\boldsymbol{c}}$. Then the unscrambled noise is

$$\dot{\boldsymbol{e}} = \boldsymbol{e}\triangle^{-}\boldsymbol{1}. \tag{5.4}$$

For each $\dot{\boldsymbol{e_j}}$ in $\dot{\boldsymbol{e}} = [\dot{\boldsymbol{e_1}} \dots \dot{\boldsymbol{e_N}}]$, get

$$\ddot{\boldsymbol{e_j}} = \dot{\boldsymbol{e_j}} - \boldsymbol{\mu} \tag{5.5}$$

, where

$$\boldsymbol{\mu} = \begin{cases} \dot{\boldsymbol{e_j}} \ \boldsymbol{mod}\ \boldsymbol{q} & \dot{\boldsymbol{e_j}} \ \boldsymbol{mod}\ \boldsymbol{q} < \frac{\boldsymbol{q}}{\boldsymbol{2}} \\ (\dot{\boldsymbol{e_j}} \ \boldsymbol{mod}\ \boldsymbol{q}) - \boldsymbol{q} & \text{otherwise}. \end{cases}$$

$$\boldsymbol{m_j} = \ddot{\boldsymbol{e_j}}\ \boldsymbol{q^{-1}} \tag{5.6}$$

, where $\boldsymbol{i} \in \boldsymbol{\{1, 2, \dots, N\}}$. Lastly, return the original plaintext:

$$\boldsymbol{m} = (\boldsymbol{m_1}, \dots, \boldsymbol{m_N}) \tag{5.7}$$

This crypto-system resists lattice-based and chosen plaintext attacks so that it assures data security. The system is suitable for APs with limited capabilities because of its low computation complexity.

5.3 The Proposed Scheme

Our proposed scheme has two phases: initialization phase, which sets-up the secure connection between APs and CC via SMs and BSs. While reading aggregation phase organizes the aggregation operation of electricity consumption readings.

5.3.1 Initialization Phase:

TA assigns a pair of public private keys for CC, each BS and each connected SM.

- For CC, its public key parameters are $\{\boldsymbol{M_{cc0}}, \boldsymbol{M_{cc1}}, \ldots, \boldsymbol{M_{ccn}}\}$, where $\boldsymbol{M_{cc0}}$ is the hard noise matrix, and $\{\boldsymbol{M_{cc1}}, \ldots, \boldsymbol{M_{ccn}}\}$ are the $\boldsymbol{n}$ soft noise matrices. APs use this key to encrypt their readings. The CC's private key parameters are $\mathscr{P}_{\boldsymbol{cc}}(.), \boldsymbol{M_{cc}}, \triangle_{\boldsymbol{cc}}$.

- For each BS, the public key is $\{\boldsymbol{M_{bs0}}, \boldsymbol{M_{bs1}}, \ldots, \boldsymbol{M_{bsa}}\}$, where $\boldsymbol{M_{bs0}}$ is the hard noise matrix, and $\{\boldsymbol{M_{bs1}}, \ldots, \boldsymbol{M_{bsa}}\}$ are the $\boldsymbol{a}$ soft noise matrices. Its private key parameters are $\mathscr{P}_{\boldsymbol{bs}}(.), \boldsymbol{M_{bs}}, \triangle_{\boldsymbol{bs}}$.

- According to each SM, its public key is $\{\boldsymbol{M_{sm0}}, \boldsymbol{M_{sm1}}, \ldots, \boldsymbol{M_{smi}}\}$, and the private key is $\mathscr{P}_{\boldsymbol{sm}}(.), \boldsymbol{M_{sm}}, \triangle_{\boldsymbol{sm}}$.

- Each AP has a unique ID issued by TA, $\{\boldsymbol{AP_1}, \ldots, \boldsymbol{AP_m}\}$, where $\boldsymbol{m}$ is the total number of APs in HAN. APs are arranged in a fixed order according to their IDs. The aggregator, $\boldsymbol{AP_s}$, for each aggregation round is known to SM and APs. This order is fixed and securely sent to all APs in the HAN so that each one automatically knows its turn to be the aggregator. For instance, if there are five APs in the house, $\boldsymbol{AP_1}$, $\boldsymbol{AP_2}$, $\boldsymbol{AP_3}$, $\boldsymbol{AP_4}$, and $\boldsymbol{AP_5}$, then SM arranges them so that $\boldsymbol{AP_1}$ is the aggregator for first round, $\boldsymbol{AP_2}$ is the aggregator for second round,..., $\boldsymbol{AP_5}$ is the aggregator for fifth round, and then $\boldsymbol{AP_1}$ is the aggregator for sixth round, and so on.

- Each AP stores its encrypted ID $\boldsymbol{ID_{j-enc}}$ in a secured place.

$$\boldsymbol{ID_{j-enc}} = \boldsymbol{ID_j} * \boldsymbol{M_{sm0}} + \sum_{\boldsymbol{i}} \boldsymbol{r_i} * \boldsymbol{M_{smi}} \tag{5.8}$$

The ID is encrypted by SM's public key, because AP needs to prove its identity to SM when it becomes the aggregator during the aggregation phase.

Moreover, TA assigns certain extra number of IDs for each HAN in the area, e.g., TA sets 20 IDs per HAN so that HAN can have 20 APs at maximum. Then, if the customer needs to add or remove APs, the following procedure is applied:

- If a new extra AP is added to the network, the customer selects one ID for the new AP from the assigned range of IDs for that house (i.e, the IDs are assigned by TA).

- If an old AP is replaced by a new same AP, such as an old AC is replaced by a new AC, then the new AP uses the same ID as the old one.

- If an AP is removed from the network, no procedure is required. As other APs in HAN still can aggregate their readings utilizing the homoromphic feature of the applied crypto-system.

5.3.2 Reading Aggregation Phase:

1) Inside Home Area Networks (HANs)

- At the beginning of each readings' aggregation round, each $\boldsymbol{AP_j}$ in HAN encrypts its reading vector $\boldsymbol{m_j} = (\boldsymbol{m_1}, \ldots, \boldsymbol{m_w})$ using CC's public key.

$$\boldsymbol{c_j} = \boldsymbol{m_j} * \boldsymbol{M_{cc0}} + \sum_{\boldsymbol{i}} \boldsymbol{r_i} * \boldsymbol{M_{cci}} \tag{5.9}$$

Then, it sends $\boldsymbol{c_j}$ to $\boldsymbol{AP_s}$, i.e., the aggregator for the current round.

$$\boldsymbol{AP_j} \xrightarrow{\boldsymbol{c_j}} \boldsymbol{AP_s}$$

- $\boldsymbol{AP_s}$ computes the total reading $\boldsymbol{c}$ by aggregating the received readings employing the homomorphic addition feature.

$$\boldsymbol{c} = \sum_{\boldsymbol{j}} \boldsymbol{c_j} \tag{5.10}$$

- $\boldsymbol{AP_s}$ attaches its encrypted ID ($\boldsymbol{ID_{s-enc}}$) to the aggregated message and then forwards the message to SM.

$$\boldsymbol{AP_s} \xrightarrow{\boldsymbol{c,\, ID_{s-enc}}} \boldsymbol{SM}$$

- After checking the validity of $\boldsymbol{AP_s}$'s ID, SM attaches timestamp $\boldsymbol{T_v}$ and nonce $\boldsymbol{f}$ vectors, and then signs the received message using $\boldsymbol{\mathscr{P}_{sm}(.), M_{sm}, \triangle_{sm}}$:

$$\boldsymbol{x} = \boldsymbol{c} \| \boldsymbol{T_v} \| \boldsymbol{f} \tag{5.11}$$

$$\dot{\boldsymbol{x}} = \boldsymbol{\mathscr{P}_{sm}^{-1}(x)} \tag{5.12}$$

$$\boldsymbol{e} = \dot{\boldsymbol{x_D}} - \dot{\boldsymbol{x_U}} \boldsymbol{A_{sm}^{-1} B_{sm}} \tag{5.13}$$

$$\dot{\boldsymbol{e}} = \boldsymbol{e} \triangle_{\boldsymbol{sm}}^{\boldsymbol{-1}} = [\dot{\boldsymbol{e_1}}, \ldots, \dot{\boldsymbol{e_N}}] \tag{5.14}$$

For each $\dot{\boldsymbol{e_j}}$, $\boldsymbol{i \in \{1, 2, \ldots, N\}}$, SM computes

$$\ddot{\boldsymbol{e_j}} = \dot{\boldsymbol{e_j}} - \boldsymbol{\mu} \tag{5.15}$$

, where $\boldsymbol{\mu} = \begin{cases} \dot{\boldsymbol{e_j}} \ \boldsymbol{mod\ q} & \dot{\boldsymbol{e_j}} \ \boldsymbol{mod\ q} < \frac{\boldsymbol{q}}{\boldsymbol{2}} \\ (\dot{\boldsymbol{e_j}} \ \boldsymbol{mod\ q}) - \boldsymbol{q} & \text{else} \end{cases}$

$$\boldsymbol{y_j} = \ddot{\boldsymbol{e_j}}\, \boldsymbol{q^{-1}}, \boldsymbol{i \in \{1, 2, \ldots, N\}} \tag{5.16}$$

$$\boldsymbol{Y} = (\boldsymbol{y_1}, \ldots, \boldsymbol{y_N}) \tag{5.17}$$

SM then forwards $\boldsymbol{Y}$ to the local BS.

$$\boldsymbol{SM} \xrightarrow{\boldsymbol{Y}} \boldsymbol{BS}$$

2) At Local Base Station (BS)

- BS first verifies each SM's signature and obtains its message $\boldsymbol{(x = c \mid T_v \mid f)}$:

$$\boldsymbol{x} = \boldsymbol{Y} * \boldsymbol{M_{sm0}} + \sum_{\boldsymbol{i}} \boldsymbol{r_i} * \boldsymbol{M_{smi}} \tag{5.18}$$

Also, it checks the validity of timestamp $\boldsymbol{T_v}$ and nonce $\boldsymbol{f}$.

- Then, BS concatenates the received aggregated consumptions from different SMs in the area

$$\boldsymbol{C} = \sum_{\boldsymbol{k}} \boldsymbol{c_k} \tag{5.19}$$

, where $\boldsymbol{k}$ is the number of connected HANs to local BS.

- Next, BS signs the total consumption of the area by its private key $\boldsymbol{\mathscr{P}_{bs}(.), M_{bs}, \triangle_{bs}}$:

$$g = C \| T_u \| q \tag{5.20}$$

$$\dot{g} = \mathscr{P}_{bs}^{-1}(g) \tag{5.21}$$

$$w = \dot{g_D} - \dot{g_U} A_{bs}^{-1} B_{bs} \tag{5.22}$$

$$\dot{w} = w \triangle_{bs}^{-1} = [\dot{w_1}, \ldots, \dot{w_N}] \tag{5.23}$$

For each $\dot{w_j}$, $i \in \{1, 2, \ldots, N\}$, BS computes

$$\ddot{w_j} = \dot{w_j} - \mu \tag{5.24}$$

, where $\mu = \begin{cases} \dot{w_j} \bmod q & \dot{w_j} \bmod q < \frac{q}{2} \\ (\dot{w_j} \bmod q) - q & \text{else} \end{cases}$

$$d_j = \ddot{w_j}\, q^{-1}, i \in \{1, 2, \ldots, N\} \tag{5.25}$$

$$D = (d_1, \ldots, d_N) \tag{5.26}$$

BS then forwards the aggregated D to CC as an electricity reading message for one unit, i.e., CC deals with BS and its connected cluster of HANs as one HAN.

- CC verifies BS's signature on D and then decrypts C using its private key parameters:

$$\dot{c} = \mathscr{P}_{cc}^{-1}(C) \tag{5.27}$$

$$s = \dot{c_D} - \dot{c_U} A_{cc}^{-1} B_{cc} \tag{5.28}$$

$$\dot{s} = s \triangle_{cc}^{-1} = [\dot{s_1} \ldots \dot{s_N}] \tag{5.29}$$

For each $\dot{s_k}$, $k \in \{1, 2, \ldots, N\}$, CC computes

$$\ddot{s_k} = \dot{sk} - \mu_0 \tag{5.30}$$

, where $\mu_0 = \begin{cases} \dot{s_k} \bmod q & \dot{s_k} \bmod q < \frac{q}{2} \\ (\dot{s_k} \bmod q) - q & \text{else} \end{cases}$

$$m_k = \ddot{s_k}\, q^{-1}, k \in \{1, 2, \ldots, N\} \tag{5.31}$$

$$m = (m_1, \ldots, m_N) \tag{5.32}$$

CC now obtains the total aggregated consumption for BS's area in plaintext m.

3) Control Messages

- If any AP from group 3 or 4 needs to send a request to CC, e.g., to change its load, or duration of usage, it can directly send its request to CC via SM and BS and does not wait for the new aggregation round. These messages are called control messages.

For instance, if AP_j wants to send a request R, it first adds a timestamp T_d and random nonce L to R, concatenates its ID (ID_{j-enc}), $n_j = R \| ID_{j-enc} \| T_d \| L$, and then encrypts n_j by CC's public key:

$$z_j = n_j * M_{cc0} + \sum_i r_i * M_{cci} \tag{5.33}$$

- AP_j then sends control message z_j to SM, which signs the message and forwards it to CC

via BS, i.e., BS verifies SM's signature and signs z_j again before forwards it to CC.

5.4 Security Analysis

The main objective of the proposed scheme is to preserve the privacy of HAN's customers and does not expose daily habits and lifestyle of houses' owners from their electricity consumptions. In addition, our scheme aims to satisfy basic security requirements, such as confidentiality and messages' integrity.

Privacy: The electricity consumption for HAN can reveal the daily behaviours of householders so that preserving their privacy is a major concern. The proposed scheme guarantees that no party even SM or BS knows the individual reading for each AP. $\boldsymbol{AP_s}$ cannot analyze the daily life pattern for householders too, as the received individual readings are encrypted; in addition, the aggregator is different for each reading round. Although we assume that APs' secret IDs are protected, $\mathscr{A}$ will not gain a lot of information if he/she manages to compromise one of the APs, i.e., he/she can only know the reading for that AP and cannot analyze the private life for householders by that data only. Furthermore, if the compromised AP by chance is the aggregator $\boldsymbol{AP_s}$, $\mathscr{A}$ cannot extract any data, as the messages sent to $\boldsymbol{AP_s}$ are encrypted and only CC has the decryption key. However, we assume that $\mathscr{A}$ cannot compromise APs and cannot obtain their secure IDs, because if $\mathscr{A}$ can physically compromise APs, e.g., $\mathscr{A}$ can enter the house, he/she can easily obtain the readings for APs by him/herself and does not need to snoop/modify and analyze the messages. The same happens when $\mathscr{A}$ attempts to compromise SM. Since SM is just a relay node, it only forwards the received encrypted messages. According to BS, it receives the total house's consumption, but this received message is encrypted so that BS cannot extract any knowledge about electricity consumption for householders. While, CC receives the total aggregated consumption for the whole area so that it cannot extract any private data about real-time consumption pattern of a specific house/customer.

Authenticity and Messages' Confidentiality and Integrity: Only authorized parties have access to messages' contents. APs' readings cannot be revealed to anybody even $\boldsymbol{AP_s}$, which receives only the encrypted version of these readings. According to the total aggregated consumption for HAN, neither $\boldsymbol{AP_s}$ nor related SM can decrypt it, only CC can. As well, BS cannot extract any knowledge from the total consumption of the whole area, as it processes the messages in their encrypted versions. Moreover, messages integrity and confidentiality are also guaranteed. The proposed scheme guarantees the data confidentiality, as none of the participated parties can know the reading of each AP; only the authorized party, CC, can decrypt the total aggregated consumption for the whole area. Furthermore, the message is protected against different attacks. If $\mathscr{A}$ successes to compromise SM, he/she cannot interpret the message's contents, as SM does not own the decryption key. The same happens if $\mathscr{A}$ is powerful enough to compromise the BS (BS is more protected and can resist attacks more than SM). Therefore, any eavesdropping attack does not succeed. In addition, the proposed scheme assures that no illegal party can modify/access the exchanged messages. $\mathscr{A}$ cannot forge the transmitted messages from HANs to the connected BS, as $\mathscr{A}$ does not know SM's private key to falsify its signature. In addition, $\mathscr{A}$ does not have access to BS's private key so that he/she cannot mimic its signature. Replay attacks do not succeed too, because of the attached timestamps and nonce values.

Security of utilized crypto-system is guaranteed by the hardness of hidden lattice problem (HLP) [159] that is based on disorganizing the lattice in a way so that no party can extract the

original lattice from its disturbed version, i.e., HLP disturbs lattice $\boldsymbol{l}$ by a specified technique to be $\bar{\boldsymbol{l}}$ so that $\mathscr{A}$ cannot extract $\boldsymbol{l}$ from $\bar{\boldsymbol{l}}$. Lattice-based homomorphic encryption scheme exploits HLP to select the private key parameters: the hidden lattice $\boldsymbol{M}$, the scrambling matrix $\triangle$ and the permutation $\mathscr{P}(.)$. Accordingly, to break the system, $\mathscr{A}$ should obtain the secret permutation and then retrieve the disturbed columns' indexes to solve the corresponding HLP. To guarantee the security and robustness of the system against attacks specially chosen plaintext and lattice-based attacks, the main four parameters of the crypto-system, i.e., $\boldsymbol{l}$, $\boldsymbol{r}$, $\boldsymbol{N}$, and $\boldsymbol{p}$, must be chosen carefully so that the cost of non-disturbed columns searching operation is unbearable for powerful attackers. If $\boldsymbol{N \geq 50}$, this operation requires $\binom{2N}{N} \geq \binom{100}{50} \approx \mathbf{2^{100}}$. In addition, choosing a high-dimension lattice, e.g., **600**, can further increase the hardness of HLP problem. Following these constraints during parameters' selection can enhance the proposed scheme's resistance to attacks [159].

5.5 Performance Evaluation

This section analyzes the performance of the proposed scheme in terms of communication and computation overhead.

5.5.1 Communication Overhead

The number of exchanged messages between different parties every readings' aggregation round is too small. The number of messages that should be sent by the limited-capability devices SM and APs is trivial. During each readings' aggregation round, each AP sends only one reading message. As well, $\boldsymbol{AP_s}$ just sends the aggregated message. According to SM, it only forwards the total aggregated readings message for the house to BS, which in turn forwards the total aggregated readings message for the whole area to CC. According to control messages, APs from group 3 and 4 only need to send their requests to CC. These messages are sent directly to SM, which forwards them to CC via BS. Generally, the house could contain two or three of these APs, e.g., a house includes one EV, one AC, and one clothes washer. Assume that each HAN has three of these APs for maximum; that means three control messages. These messages are sent occasionally; assume that each AP needs to send one or two control messages per day so that the maximum number of control messages per day for one HAN is six messages. In summary, the total communication burden for BS, SM, and for each AP is one message per reading round, which considers insignificant load.

Figure 5.2 shows the communication overhead inside HAN for every reading round. It can be seen that the communication delay is increased from **2** messages in two-appliances case to **20** messages in twenty-appliances case. Although the communication overhead is expected to increase as the number of APs increases, its growth is limited and affordable by the restricted-resources devices in the house. Figure 5.3 shows the total communication delay for the area per day, after adding the control messages overhead, as the number of APs in the house and HANs in the area increases. Although the APs' number increases, each AP, SM, and BS still have to send one message only per round, which means a fixed number of messages are sent from each HAN in the area per day. According to number of HANs, the total communication overhead for the area is increased as the number of connected HANs increases. However, the communication load is affordable by different parties in the network; in other words, the proposed scheme provides light communication overhead.

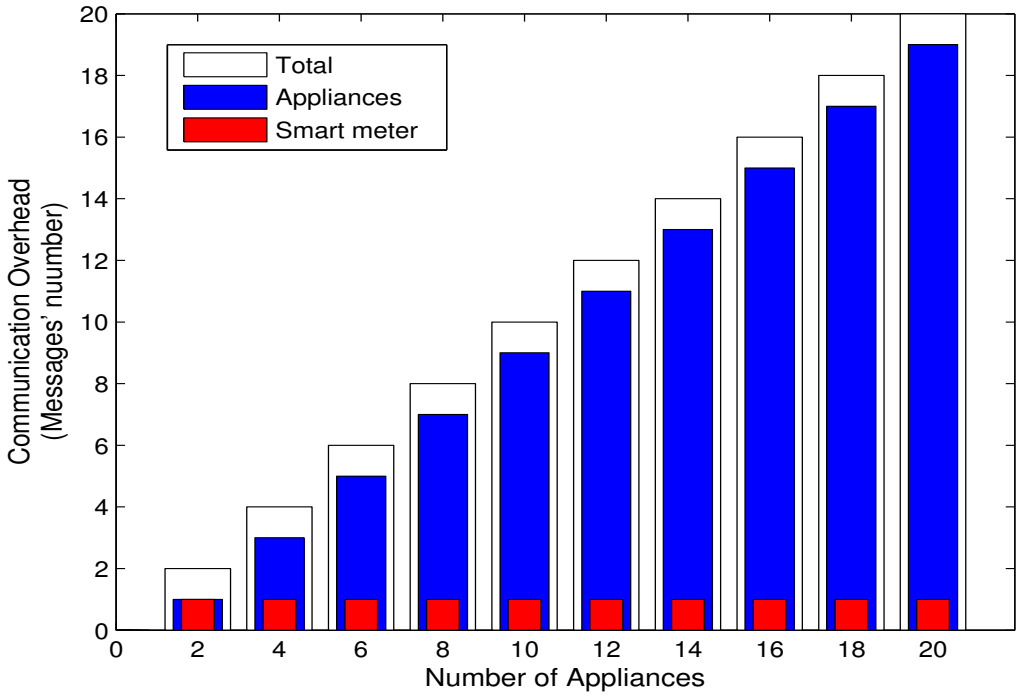

Figure 5.2: Communication overhead per reading round.

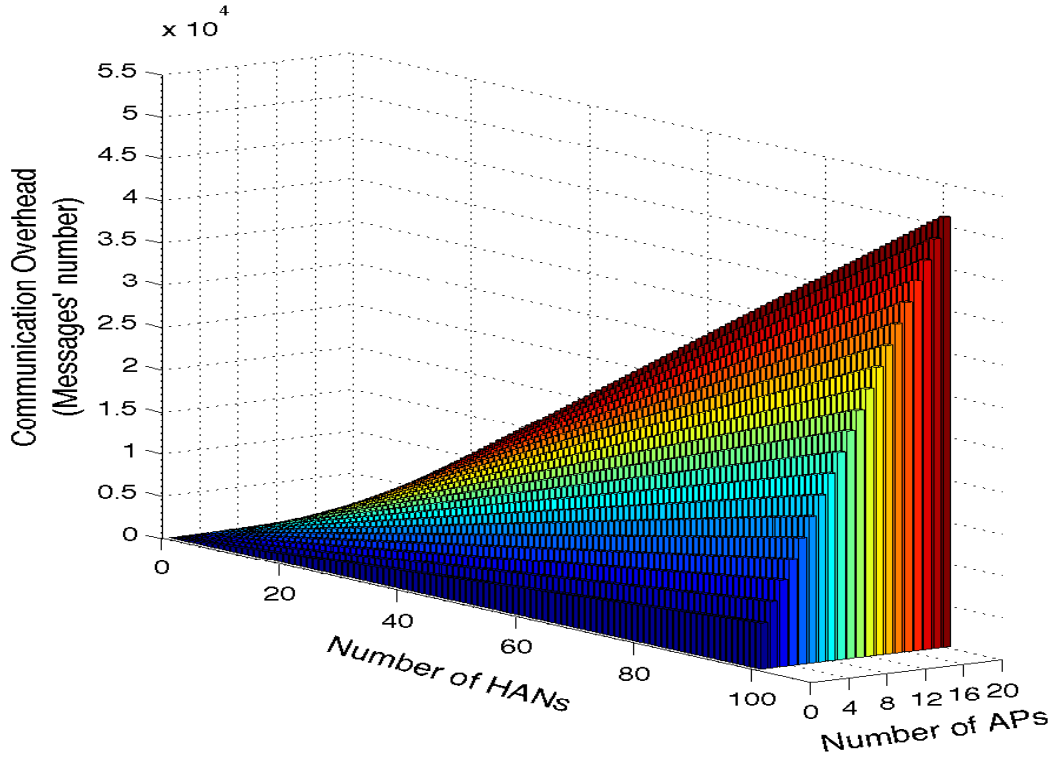

Figure 5.3: Communication overhead per day.

5.5.2 Computation Overhead

Each AP in the HAN has to perform one encryption process per round for its reading except $\boldsymbol{AP_s}$, which also has to perform simple neglected summation operation. The total number of encryption operation per HAN is $\boldsymbol{n}$, where $\boldsymbol{n}$ is the number of APs in the HAN. These encryption processes do not require high computation capabilities, because the deployed cryptographic scheme is lightweight scheme especially for encryption; it only consists of simple addition and multiplication operations. Next, SM signs the received aggregated message from $\boldsymbol{AP_s}$ before forwarding it to CC via BS. That means one signing process for each HAN per round. Then, BS aggregates the received readings from the connected cluster, signs the result, and then sends it to CC. So, if the total number of HANs in BS's area is $\boldsymbol{m}$, then the total computation overhead for the area equals $[\boldsymbol{m} * (\boldsymbol{n} * \boldsymbol{T_e} + \boldsymbol{T_s} + \boldsymbol{T_v})] + \boldsymbol{T_s} + \boldsymbol{T_v} + \boldsymbol{T_d}$, where $\boldsymbol{T_e}$ is the computation time for one encryption process, $\boldsymbol{T_d}$ is the time for one decryption process, $\boldsymbol{T_s}$ is the time for one signing process, and $\boldsymbol{T_v}$ is the time for one verification process.

In addition, If HANs have group 3 or 4 APs, then control messages have to be encrypted by these APs and sent to CC via SM and BS. Assume that each house has up to three group 3 or 4 APs, then six control messages are sent from HAN per day that need six encryption operation maximally. SM, i.e., also BS, needs to sign these control messages. However, these messages will not impact on the total computation duty for different parties per day, because they are few and sent occasionally. Table 5.1 presents the number of operations per reading round and per day for each AP and SM, where $\boldsymbol{h}$ is the number of rounds per day.

Table 5.1: The number of operations for smart devices.

Number of Operations	Per Round	Per Day
$\boldsymbol{APs(Group-1\&2)}$	$\boldsymbol{1 * T_e}$	$\boldsymbol{h * T_e}$
$\boldsymbol{APs(Group-3\&4)}$	$\boldsymbol{1 * T_e}$	$\boldsymbol{(h+2) * T_e}$
$\boldsymbol{SM}$	$\boldsymbol{1 * T_s}$	$\boldsymbol{(h+6) * T_s}$

Assume that the hidden lattice dimension is **600** to resist the lattice-based attack, $\boldsymbol{n = 9, r = 2, p \approx 260, \varepsilon_{max} = 1024, l_0 \approx 219,}$ and $\boldsymbol{q = 2^{21}.2^{38}}$. The size of the public parameters is $\boldsymbol{2N^2(n+1)Log_2(p)}$. The speed of one encryption/verification operation is the cost of $\boldsymbol{(n+1).N}$ addition processes of vectors of lengths $\boldsymbol{2N}$ and $\boldsymbol{Log_2(p)}$ bits plus the cost of multiplying two random vectors with length $\boldsymbol{Log_2(\varepsilon_{max})}$ bits. While the decryption/signing speed is the cost of two $\boldsymbol{N \times N}$ matrix multiplication in $\boldsymbol{GF(p)}$, which equals $\boldsymbol{2N^2.Log_2(p)^2}$. Using a MATLAB simulator on a 3.20 GHz-processor with 6.00GB RAM, we study the computation delay for our proposed scheme.

Figure 5.4 presents the computation load for each AP and for SM every reading round. It can be noticed that the computation overhead for each AP is the same and does not affected by the number of APs in the house. Moreover, the load of SM does not change, which is expected, as it requires to perform one signing process regardless the number of messages included. Figure 5.5 points out the total computation load for the whole cluster per day as numbers of APs and HANs increase. As indicated, the computation overhead increases by the increase of APs and HANs' numbers, but still within a bounded limit; the total computation delay for a cluster of **100** HANs that each one of them has **20** APs is around **90** second per day.

Practically speaking, our proposed scheme is feasible for the restricted-computation capabilities APs. Considering the Smart Grid Smart City data set provided by the Australian

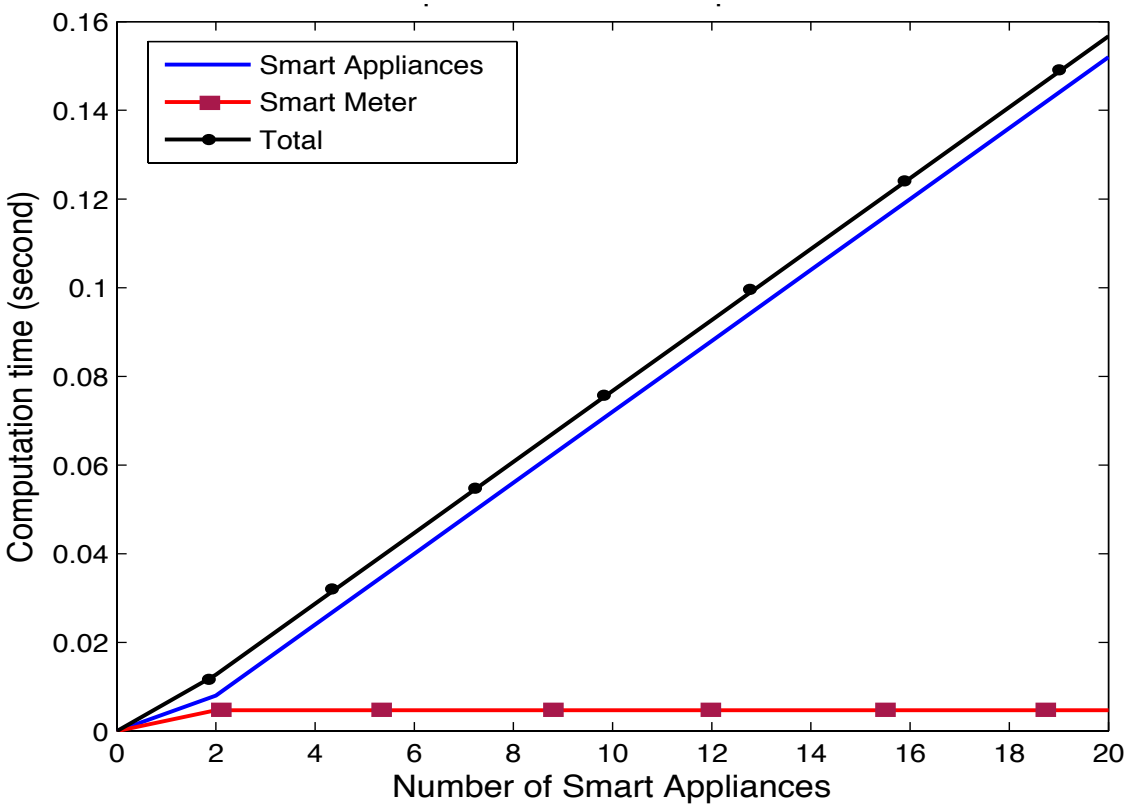

Figure 5.4: Computation delay per reading round.

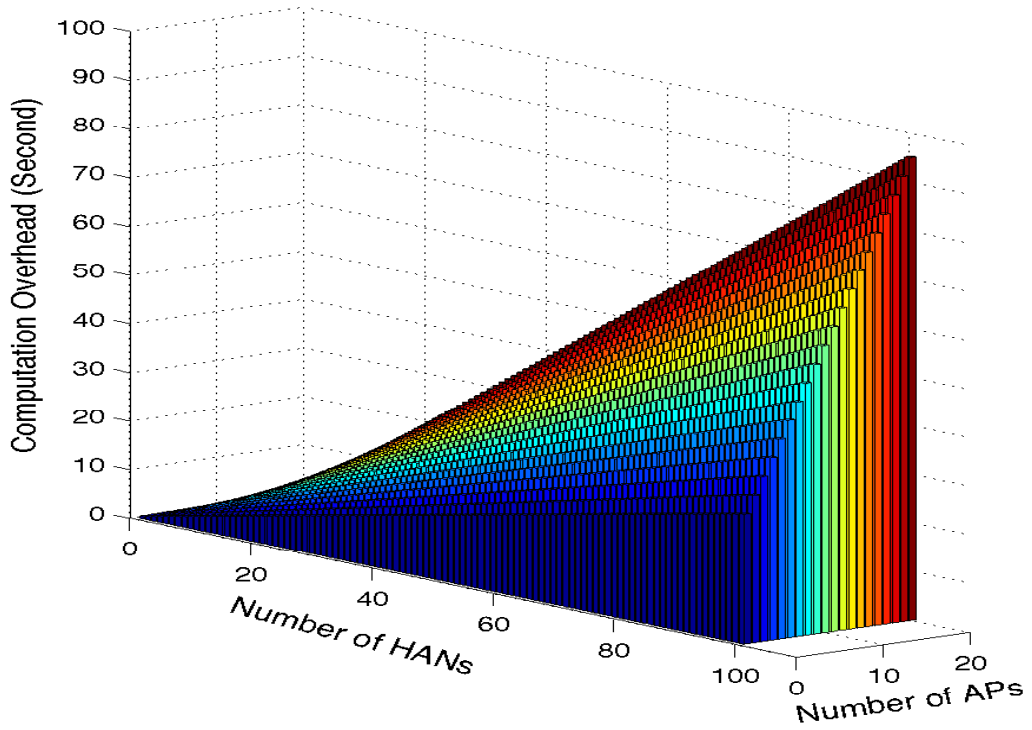

Figure 5.5: Computation delay per day.

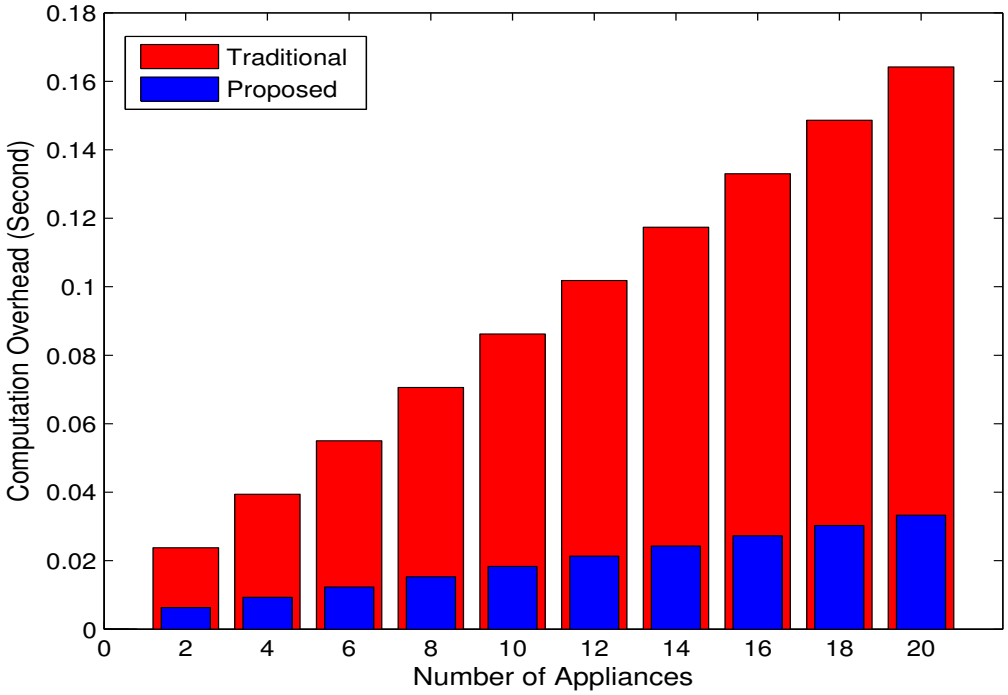

Figure 5.6: Computation delay per reading round.

Department of Industry, Innovation and Science [160], for instance, which includes a data set of different household APs' readings at different times per one HAN in Sydney, e.g., at November 16 - 2013, almost at 3:12 pm, the current APs' readings for customer **11178213** are measured as: television **388.68**, computer **124.799**, stove **44.474**, AC **1.711**, and light **11.474** kWh, it can be seen that the APs readings' range is limited by the maximum value **731625** kWh. So, the readings' values do not require a significant storage memory. In addition, encrypting the APs' readings does not provide high computation burden on the AP, as the encryption operation consists of generating a fixed series of random numbers and then computing trivial summation and multiplication operations. Consequently, simple cheap processing device embedded on the APs, such as Raspberry PI [161], is enough to implement that light encryption process.

The existing privacy schemes cannot be applied on APs, since they have to perform complex crypto-operations, such as exponentiation and pairing, which require high computation capabilities not owned by APs. Although the current crypto-systems are not applicable in these restricted-resources APs, we compare the performance of the proposed scheme with a traditional homomorphic Paillier-based scheme, as several privacy-preserving schemes in the literature utilize Paillier crypto-system because of its additive homomorphic feature, also it is robust against privacy attacks.

In Figure 5.6, we compare the total computation time per HAN for the proposed scheme versus the traditional one per round, as the number of APs at house increases. Clearly, the computation delay for the proposed scheme is much less than the traditional one, especially at the high number of APs in the house. The computation delay goes from 6.3 to 33.3 ms for the proposed scheme while the traditional scheme's delay increases from 23.8 to 164.2 ms, when the number of APs increases from **2** to **20** APs.

While Figure 5.7 shows the total computation delay for the whole BS's connected area per day and presents the effect of the increase in APs' number also the increase in the number of HANs in the BS's cluster. It is shown that our scheme takes less computation time compared to the Paillier-based schemes, especially as numbers of APs and HANs increase. Although the computation overhead increases in our scheme, the resulted computation delay is limited and affordable by APs. The computation delay per day for a cluster of **100** HANs with **20** APs each is around **90** second for proposed scheme versus **450** second for traditional schemes. In conclusion, the proposed scheme guarantees privacy and security requirements for the residential consumers with low computation and communication overhead even for limited-computation capabilities APs.

5.6 Conclusion

In this work, we have proposed a lightweight lattice-based homomorphic security and privacy-preserving scheme that secures the electricity consumption aggregation operation for HANs in residential areas. The proposed scheme depends on house's APs to aggregate their consumption among themselves without involving the connected SM utilizing the lightweight lattice-based homomorphic crypto-system to secure their readings. However, SMs and the intermediate BS can validate the messages' authenticity without decrypting them. The security analysis and simulation results show that the proposed scheme guarantees consumers' privacy and messages' confidentiality and integrity, at the same time, ascertains lightweight communication and computation overhead. So, our proposed scheme is suitable for limited-computation resources APs.

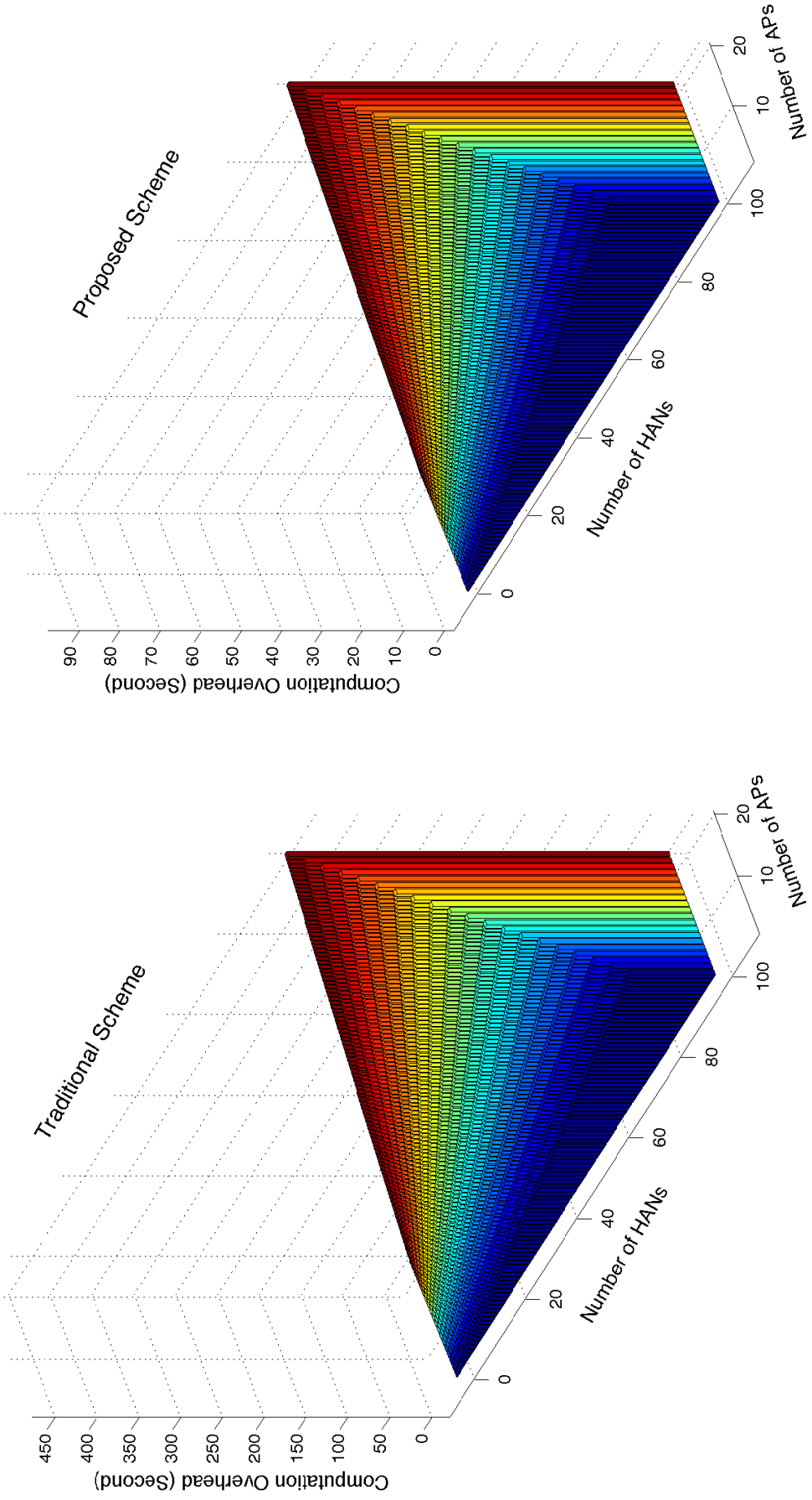

Figure 5.7: Computation delay per day.

Chapter 6

Lightweight Authentication and Privacy-Preserving Scheme for V2G Connections

V2G connection allows two-way electricity transmission between EVs and power grid for achieving many known benefits. However, V2G connections suffer from certain security threats, such as EV's privacy and authenticating it to the grid. In our work [165, 166], we propose a lightweight secure and privacy-preserving V2G connection scheme, in which the power grid assures the confidentiality and integrity of exchanged information during (dis)charging electricity sessions and overcomes EVs' authentication problem. The proposed scheme guarantees the financial profits of the grid and prevents EVs from acting maliciously. Meanwhile, EVs preserve their private information by generating their own pseudonym identities. In addition, the scheme keeps the accountability for the electricity-exchange trade. Furthermore, the proposed scheme provides these security requirements by lightweight overhead; as it diminishes the number of exchanged messages during (dis)charging sessions. Simulation results demonstrate that the proposed scheme significantly reduces the total communication and computation load for V2G connection especially for EVs. The remaining of the chapter is organized as follows. Section 6.1 introduces our system model, security parameters, and design goals. Section 6.2 reviews BlueJay ultra-lightweight hybrid crypto-system. In Section 6.3, we present our proposed scheme. Section 6.4 gives security analysis, while Section 6.5 evaluates the performance of our scheme. Finally, Section 6.6 summarizes the work.

6.1 System Model

6.1.1 Network Model

Our V2G network consists of the CC that belongs to the utility company. CC is connected to several local aggregators (LAs), which are owned by independent private service provider companies. They are located at power lines or buses in the neighbour region. Each LA connects to fleets of EVs, which are located in different parking lots in the area; LAs = $\{\boldsymbol{LA_1}, \boldsymbol{LA_2}, \ldots, \boldsymbol{LA_m}\}$, where $\boldsymbol{m}$ is the number of power lines. CC communicates with LAs through wired connection, e.g., Internet. In each parking lot, there is a cluster of EVs; EVs = $\{\boldsymbol{EV_1}, \boldsymbol{EV_2}, \ldots, \boldsymbol{EV_n}\}$, where $\boldsymbol{n}$ is the number of EVs in the cluster, so $\boldsymbol{n}$ could be

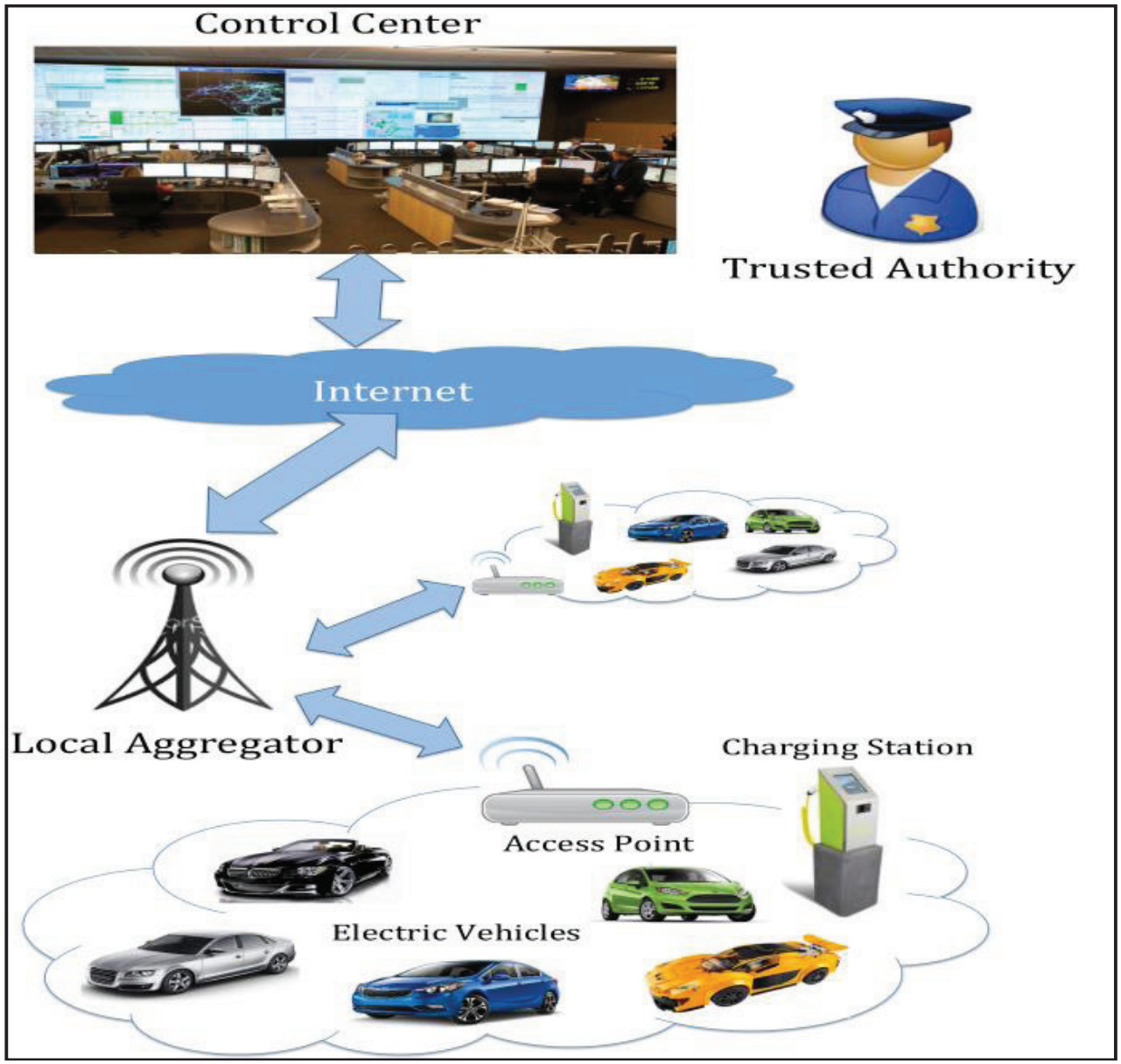

Figure 6.1: System model.

different for each lot. The EVs' cluster is connected to LA via an access point (AP), which works as a relay node to forward the exchanged messages between LA and EVs through WiFi connection. LA connects to several CSs wirelessly via APs too. CSs may locate in the parking lots or specific charging places; CSs = $\{CS_1, CS_2, \ldots, CS_j\}$, where j is the number of CSs controlled by LA. Keying parameters are provided to different parties by independent TA. V2G system model is shown in Figure 6.1.

6.1.2 Adversary Model and Security Requirements

CC and LAs are honest but curious. They will not try to act maliciously toward EVs, but they may attempt to extract EVs' private information from exchanged messages. According to EVs, they are non-trusted parties. Some EVs may act selfishly to gain a benefit or prevent other EVs from obtaining advantages; also, malicious EVs may try to impersonate innocent EVs. In addition, malicious adversaries threaten V2G connection; adversary $\mathscr{A}$ can eavesdrop the

exchanged messages between LA and EVs. Moreover, $\mathscr{A}$ may establish some active attacks; such as attempting to fabricate the captured messages, or begin a replay attack. Moreover, $\mathscr{A}$ may try to impersonate an honest EV to seize its connection with LA. To thwart these malicious actions, our proposed scheme will fulfill the following security requirements:

- Authentication: secure LA's messages against any unauthorized action, i.e., prevent any illegal parties from accessing or modifying LA's messages.

- EV's Privacy: assure that the private information of EV is not revealed to other parties and guarantee that nobody can link between EV's location and its owner's identity. Neither attackers nor CC and LAs can gain any distinguishable knowledge about a particular EV. So, they cannot link EV's battery status or location with the owner's identity.

- Confidentiality and Messages Integrity: guarantee that EVs' (dis)charging and service payment messages are confidential; they are only accessible by legitimate parties, e.g., LA and related EV. The messages' integrity should also be guaranteed. Even if $\mathscr{A}$ already intercepts the message, he/she has no access to the decryption key. Additionally, $\mathscr{A}$ cannot resend a message or modify its contents.

- Accountability: previous electricity trade sessions should be traceable; CC guarantees the accuracy of former processes. No malicious LAs or EVs can forge previous bills to increase their profits.

6.1.3 Design Goals

The main objectives of our proposed scheme can be divided into two folds:

- It should guarantee the security requirements for V2G connection. EVs owners' and location privacy should be preserved, and information confidentiality and integrity should be assured. Likewise, authentication of different parties should be guaranteed. Finally, accountability of the electricity trade operation should also assured.

- The proposed scheme also should be efficient and lightweight due to communication and computation overhead.

6.2 Preliminaries

We utilize BlueJay ultra-lightweight hybrid crypto-system [167], which is a lightweight fast crypto-system especially in its encryption process. BlueJay combines PASSERINE public key crypto-system [168] and Hummingbird-2 lightweight symmetric scheme [169].

6.2.1 PASSERINE crypto-system

PASSERINE scheme is a lightweight version of Rabin public-key scheme [170] that has two advantages: the message space is fully utilized, and the encryption operation provides lightweight computation load.

6.2.1.1 Key generation

Generate two large random distinct primes $\boldsymbol{p}$ and $\boldsymbol{q}$, with roughly the same size, as private key parameters. Then, compute the public key $\boldsymbol{n = pq}$.

6.2.1.2 Encryption

Generate a set of co-prime numbers $\boldsymbol{b_1, b_2, \ldots, b_k}$ denoted as a base and their product is $\boldsymbol{B = \prod_{i=1}^{k} b_i}$. Encrypt the plaintext $\boldsymbol{m}$ as:

$$\boldsymbol{C_0 \equiv m^2 + Y\ n (mod\ b_0),}$$

$$\boldsymbol{C_1 \equiv m^2 + Y\ n (mod\ b_1),}$$

$$\boldsymbol{\ldots\ldots\ldots,}$$

$$\boldsymbol{C_m \equiv m^2 + Y\ n (mod\ b_m).}$$

6.2.1.3 Decryption

Reconstruct the message $\boldsymbol{C}$ as:

$$\boldsymbol{C = \left(\sum_{i=1}^{k} C_i . \frac{B}{b_i} . \left(\frac{B}{b_i} \right)^{-1}_{b_i} \right) mod\ B.}$$

Compute $\boldsymbol{m_p = (c^{((p+1)/4)} mod\ p).q.q_p^{-1}}$, and $\boldsymbol{m_q = (c^{((q+1)/4)} mod\ q).p.p_q^{-1}}$. Select the right root from $\boldsymbol{m = \{m_p + m_q, m_p - m_q, -m_p + m_q, -m_p - m_q\}\ (mod\ n)}$ as the plaintext message.

6.2.1.4 Signing

We utilize Rabin signature algorithm [170], where $\boldsymbol{H}$ is a collision-resistant hash function $\boldsymbol{H : \{0,1\}^* \rightarrow \{0,1\}^k}$, The public key is $\boldsymbol{n}$, and the private key is the pair $\boldsymbol{(p, q)}$. To sign a message $\boldsymbol{m}$, pick random padding $\boldsymbol{U}$ and calculate $\boldsymbol{H(mU)}$ so that $\boldsymbol{H(mU)}$ is a square modulo $\boldsymbol{n}$, $\boldsymbol{x^2 = H(mU)\ mod\ n}$. The signature on $\boldsymbol{m}$ is the pair $\boldsymbol{(U, x)}$.

6.2.1.5 Verification

Given $\boldsymbol{m}$ and $\boldsymbol{(U, x)}$, calculate $\boldsymbol{x^2}$ and $\boldsymbol{H(mU)}$. Then verify if $\boldsymbol{x^2 = H(mU)\ mod\ n}$.

BlueJay is a combination of Hummingbird-2 lightweight authenticated encryption scheme and PASSERINE crypto-system optimized for a **1024**-bit public modulus $\boldsymbol{n}$ and **32**-bit register size.

The proposed scheme also utilizes AKARI-2 [171], a lightweight pseudorandom number generator, to generate pseudonym IDs (PIDs) and symmetric keys. AKARI-2 is simple and easy to implement; it is appropriate for limited-computation devices, e.g., radio frequency identification, and sensors.

6.3 The Proposed Scheme

Our proposed scheme has three phases: the first one is the initialization phase, which setups the system components and generates the keys and security parameters; the second phase is the operation phase, which is responsible for electricity-trade operation; and the final phase is the billing phase to pay the electricity expenses.

6.3.1 Initialization Phase

The initialization phase establishes the network and defines the required security parameters.

First, TA issues two pairs of public-private keys to both CC and LAs as follows:

- For CC, TA generates CC's private key pair $\{p_{cc}, q_{cc}\}$ and corresponding public key $n_{cc}(n_{cc} = p_{cc}\, q_{cc})$.

$$TA \xrightarrow{n_{cc},\, p_{cc},\, q_{cc}} CC.$$

- For each LA, TA generates LA's private key pair $\{p_{la}, q_{la}\}$ and public key $n_{la}(n_{la} = p_{la}\, q_{la})$.

$$TA \xrightarrow{n_{la},\, p_{la},\, q_{la}} LA.$$

- Also, TA assigns a secret ID $\{\mathcal{L}_{cs}\}$ and secret session key $\{k_{cs}\}$ for each CS; the station includes its secret ID in its messages to LA to proof its identity. Thus, this ID should be stored in a secured memory module.

$$TA \xrightarrow{\mathcal{L}_{cs_j}, k_{cs_j}} CS_j.$$

- Each EV generates a PID utilizing AKARI-2, $EV_i = AKARI(x0, x1)$, where $x0, x1$ are initial seeds. Each EV changes its ID from time to time for more anonymity, e.g., a different PID for each connection session.

- In addition, EVs generate symmetric keys to secure the connection sessions with LA. $k_i = AKARI(h0, h1)$, where $h0, h1$ are initial seeds. This key is changed frequently so that the probability of compromising it is diminished. Then, the messages' confidentiality is enhanced.

6.3.2 Operation Phase

The operation phase is responsible for the electricity trade operation; it organizes the tasks between different participants during the power purchase. First, we need to define a type of messages called *request messages*, which are utilized to manage the electricity exchange process. The request message is a request to buy/sell electrical power; it consists of the required amount of electricity and the current price. The request message could be sent from EVs to LA to (dis)charge their batteries or from CC to LA to supply electricity to EVs or consume electricity from them. There are four request messages: the first one is *the CC supply request message*,

which is a request from CC to sell a share of its electricity. The second type is *the CC consume request message*, which is a request from CC to purchase an amount of electricity from EVs. The third message is *the EV charge request message*, which is a request from an EV to charge its battery, i.e., the EV wants to buy a portion of power from the grid. Last request message is *the EV discharge request message*, which is a request from an EV to discharge its battery; that means the EV wants to sell some of its stored energy to the grid. It can be inferred that there are four cases of electricity transfer operations:

6.3.2.1 Case 1. The CC Supply Request

In this case, CC wants to sell a portion of grid's electricity to EVs. Consequently, CC should follow a specific procedure:

- CC first sends *the CC supply request message* to the connected LAs; the message $\boldsymbol{M_s}$ should indicate the amount of sold electricity $\boldsymbol{A}$ and the price that the grid asked for $\boldsymbol{x}$ (the price is in form $\boldsymbol{x}$ cents/KWH), $\boldsymbol{M_s = (A, x)}$.

- CC only signs the message by its private key to authenticate it, as the contents of that message are not confidential. First, a timestamp and a nonce are involved to the message to prevent replay attacks, $\boldsymbol{m_s = (M_s \| T_{s1} \| L_1)}$. Then, CC picks a random padding $\boldsymbol{U_1}$ and calculates $\boldsymbol{H(m_s U_1)}$ so that $\boldsymbol{H(m_s U_1)}$ is a square modulo $\boldsymbol{n_{cc}}$, $\boldsymbol{x_1^2 = H(m_s U_1)\, mod\, n_{cc}}$. The signature on $\boldsymbol{m_s}$ is the pair $\boldsymbol{(U_1, x_1)}$. Consequently, CC sends the signed message $\boldsymbol{(m_s, U_1, x_1)}$ to the connected LAs.

- When LA receives *the CC supply request message*, it checks the validity of CC's signature by recalculating $\boldsymbol{(x_1, U_1)}$ and checks if $\boldsymbol{x_1^2 = H(m_s U_1)\ mod\ n_{cc}}$ is hold. Then, LA verifies the timestamp $\boldsymbol{T_{s1}}$ and nonce $\boldsymbol{L_1}$ values. Subsequently, LA signs $\boldsymbol{M_s}$ by its private key to authenticate it to its connected EVs. The message also involves a timestamp and a nonce, $\boldsymbol{m_{s1} = (M_s \,|T_{s2}\, |L_2)}$. Using a random padding $\boldsymbol{U_2}$, LA computes $\boldsymbol{x_2^2 = H(m_{s1} U_2) mod n_{la}}$. Then, it forwards $\boldsymbol{(m_{s1}, U_2, x_2)}$ to the existing EVs via APs.

- If an EV wants to participate in that session and charge electricity from the grid, it verifies the signature of LA by calculating $\boldsymbol{(U_2, x_2)}$ and checking if the $\boldsymbol{x_2^2 = H(m_{s1} U_2) mod n_{la}}$ is hold and then verifying the validity of timestamp $\boldsymbol{T_{s2}}$ and nonce $\boldsymbol{L_2}$. Then, it checks the message's contents. While if it does not want to charge its battery, it simply discards/ignores the message.

- If one or more EVs are satisfied by the current price and want to purchase the offered electricity from CC, they response to LA by *the EV charge request message*, which contains the desired amount of electricity to buy. When EV responses to LA, it means an implicit acceptance for the offered price; only EVs that are interested to buy by the offered price will reply while the remaining EVs do not reply. EVs encrypt their charge requests by LA's public key. In addition to the required amount of power, the message includes the EV's PID and a secret session key. If the EV, with current PID $\boldsymbol{EV_i}$, wants to purchase $\boldsymbol{W}$ amount of power by the price $\boldsymbol{X}$, then it generates a secret key $\boldsymbol{k_i}$. Next, EV combines the required amount $\boldsymbol{W}$, its current PID $\boldsymbol{EV_i}$, and the secret key $\boldsymbol{k_i}$ and includes timestamp $\boldsymbol{T_{s3}}$ and nonce $\boldsymbol{L_3}$ values, $\boldsymbol{M_c = W\ |EV_i\ |k_i\ |T_{s3}\ |L_3}$. Subsequently, it encrypts the result $\boldsymbol{M_c}$ using $\boldsymbol{n_{la}}$:

EV generates a set of co-prime numbers $\boldsymbol{b_1, b_2, \ldots, b_j}$, and a random value $\boldsymbol{Y_1}$ and then encrypts $\boldsymbol{M_c}$ to obtain:

$$\boldsymbol{m_{c1} \equiv M_c^2 + Y_1\, n_{la} (mod\, b_0),}$$

$$m_{c2} \equiv M_c^2 + Y_1\ n_{la} (mod\ b_1),$$
$$\ldots\ldots\ldots,$$
$$m_{cj} \equiv M_c^2 + Y_1\ n_{la} (mod\ b_j).$$
$$m_c = \{m_{c1}, m_{c2}, \ldots, m_{cj}\}.$$

Afterward, it sends m_c to LA via the connected AP.

- If a particular EV is chosen for the current charging session, LA uses its suggested secret key to encrypt the messages to EV. LA sends the acceptance messages to the selected EVs encrypted by the previously shared session keys; each message includes the charge/discharge order (i.e., charge order in this case), the location of target CS, and the payment approach. For instance, if LA chooses the EV with current PID EV_i, then LA encrypts the acceptance message with its session key k_i; $M_a = E_{k_i}(charge, L_{cs}, cash/token)$, where L_{cs} is the location of target CS, and $cash/token$ determines the payment way. The used symmetric encryption scheme is the lightweight Hummingbird-2 crypto-system that further reduces the computation overhead.

- LA selects a number of EVs that satisfies the offered power from the grid. EV's selection depends on the number of responses and the location of the replied EVs. If the amount of electricity from responses is more than the offered amount, LA selects certain number of EVs that satisfies the supply and nearer to the CSs.

- The payment approach determines the way that EVs should follow to pay the electricity expenses. EVs pay by either cash or coupons (tokens). The coupons are anonymous authenticated tokens that the grid gives to the EV as a payment for a previous discharging session. The token only shows its financial value without any information about its holder; for instance, EV_1 discharges **300** KW of electricity to the grid via CS_3, and the grid's operator pays to EV_1 the price, which equals **100** dollar, by a **100**-dollar token. The token shows only its value, the **100** dollar, but no data about the process details or involved parties is included.

- EV first pays the required price to LA via a payment message M_p, $M_p = E_{k_i}(cash/token)$. Then, LA sends a confirmation message to the assigned CS to charge EV's battery. The confirmation message M_f contains the assigned power to that EV and its current PID. The message also involves CS's ID $\mathcal{L}_{cs}$ and timestamp and nonce values $m_f = (M_f\ |\mathcal{L}_{cs}\ |T_{s4}\ |L_4)$, then m_f is encrypted by the pre-shared symmetric key between CS and LA k_{cs}, $MF = E_{k_{cs}}(m_f)$. After CS checks the message's validity, it charges the EV by the approved amount of electricity. In other words, EV pays the declared price to LA first before receiving any electricity.

At the end of the current supply round, LA only stores the total amount of supplied electricity A and the corresponding total profit for this round X; it keeps this information in its database to calculate the total bill for the whole month later. At the same time, CC saves the same data in its record; this step conserves the accountability and guarantees the correctness of LA's total bill. As LA calculates the total bill at the end of the month and sends it to CC, which checks its correctness by comparing it with the corresponding value in its record. Figure 6.2 shows the CC Supply Request Case.

6.3.2.2 Case 2. The CC Consume Request

When CC wants to purchase a portion of electricity from the existing EVs, it should follow a specific procedure:

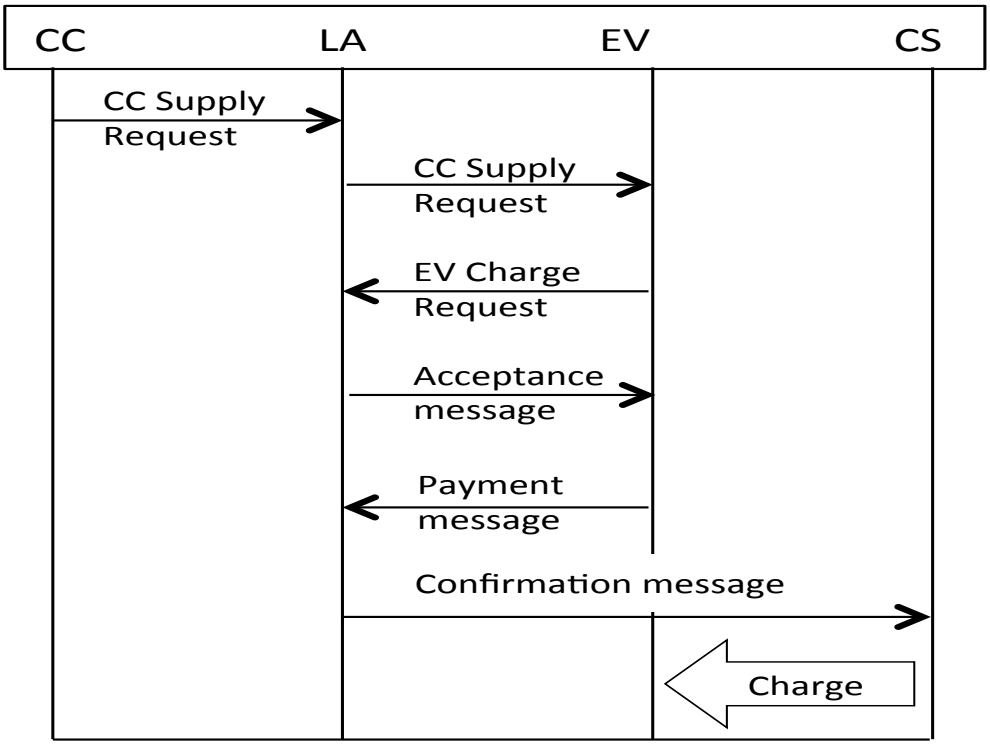

Figure 6.2: The CC supply request case.

- CC first sends *the CC consume request message* to the connected LAs; the message $\boldsymbol{M_n}$ should indicate the amount of required electricity $\boldsymbol{C}$ and the price that the grid offers $\boldsymbol{y}$ cents/KWH, $\boldsymbol{M_n = (C, y)}$. CC only signs the message to authenticate it. The message also involves a timestamp and a nonce to prevent replay attacks, $\boldsymbol{m_n = M_n \,|T_{s5}\,|L_5}$. CC picks a random padding $\boldsymbol{U_4}$ and calculates $\boldsymbol{H(m_n U_4)}$ so that $\boldsymbol{H(m_n U_4)}$ is a square modulo $\boldsymbol{n_{cc}}$, $\boldsymbol{x_4^2 = H(m_n U_4) mod n_{cc}}$. The signature on $\boldsymbol{m_n}$ is the pair $\boldsymbol{(U_4, x_4)}$. Therefore, CC sends the signed message $\boldsymbol{(m_n, U_4, x_4)}$ to LAs.

- When LA receives *the CC consume request message*, it checks the validity of the signature by recalculating the values $\boldsymbol{(U_4, x_4)}$, checks if the $\boldsymbol{x_4^2 = H(m_n U_4) mod n_{cc}}$ is hold, and then verifies timestamp $\boldsymbol{T_{s5}}$ and nonce $\boldsymbol{k_5}$ values. Next, LA signs the message to authenticate it to EVs. The message also involves a timestamp $\boldsymbol{T_{s6}}$ and nonce $\boldsymbol{k_6}$ to prevent replay attacks, $\boldsymbol{m_{n1} = (M_n \,|T_{s6}\,|K_6)}$, and then LA signs the message as: using the random padding $\boldsymbol{U_5}$, LA computes $\boldsymbol{x_5^2 = H(m_{n1} U_5) mod n_{la}}$. Then, it forwards the signed message $\boldsymbol{(m_{n1}, U_5, x_5)}$ to the existing EVs.

- If one or more EVs want to sell their stored electricity to CC, they first verify the signature of LA by computing $\boldsymbol{(U_5, x_5)}$, checking if the $\boldsymbol{x_5^2 = H(m_{n1} U_5) mod n_{la}}$ is hold, and then verifying the timestamp $\boldsymbol{T_{s6}}$ and nonce $\boldsymbol{k_6}$ values. While the other uninterested EVs ignore the request message from the first place.

- The interested EVs then response to LA by *the EV discharge request message*, which contains the desired amount of power to sell. EVs encrypt their discharge requests by LA's public key. In addition to the amount of sold power, the messages include the EVs' PID and the secret session keys. If the EV, which currently has the PID $\boldsymbol{EV_r}$, wants to sell $\boldsymbol{D}$ amount of power to the grid by the price $\boldsymbol{y}$, then it generates a secret key $\boldsymbol{k_r}$. Next, the EV combines the required amount $\boldsymbol{D}$, its current PID $\boldsymbol{EV_r}$, and the secret key $\boldsymbol{k_r}$, includes the timestamp

T_{s7} and nonce K_7 values, $M_d = D \,|EV_r\,|k_r\,|T_{s7}\,|K_7$, and then encrypts the result M_d using n_{la}:

EV generates a set of co-prime numbers $a_1, a_2, \ldots, a_e$ and a random value Y_2 and then encrypts M_d by LA's public key as follows:

$$m_{d1} \equiv M_d^2 + Y_2\, n_{la} (mod\, a_0),$$

$$m_{d2} \equiv M_d^2 + Y_2\, n_{la} (mod\, a_1),$$

$$\ldots\ldots\ldots,$$

$$m_{de} \equiv M_d^2 + Y_2\, n_{la} (mod\, a_e).$$

$$m_d = \{m_{d1}, m_{d2}, \ldots, m_{de}\}.$$

Afterward, EV sends the encrypted message m_d to LA.

- If a particular EV is chosen for the current discharging session, LA uses its suggested secret key to encrypt the messages to the selected EV. For instance, if LA chooses the EV with current PID EV_r, then LA encrypts the acceptance message with the session key k_r; the LA's acceptance message to EV contains the discharge order, the CS's location L_{cs}, and the payment order, $M_{a1} = E_{k_r}(discharge, L_{cs}, cash/token)$.

- EV discharges the agreed amount of electricity to the assigned station first, and then CS sends a confirmation message M_{f1} to LA. CS also involves its ID $\mathcal{L}_{cs}$ and timestamp and nonce values, $m_{f1} = (M_{f1}\,|\mathcal{L}_{cs}\,|T_{s8}\,|K_8)$, before encrypts the message using the pre-shared symmetric key k_{cs1}, i.e., $MF_1 = E_{k_{cs1}}(m_{f1})$. Then, LA pays the price of sold power to EV at the moment via a payment message M_{p1} (e.g., $M_{p1} = E_{k_r}(token)$ is a token payment message to EV_r). In other words, EV first transfers the contracted electricity to the grid, and then LA directly pays the obligated price to it.

At the end of the current consume round, LA, CC as well, only stores the total amount of sold electricity C and the corresponding total price for the round Y. Figure 6.3 shows the CC Consume Request Case.

6.3.2.3 Case 3. The EV Charge Request

In this case, EV asks for electricity from the grid, i.e., wants to purchase a portion of electricity from the grid. So, EV follows a specific procedure:

- EV first sends *the EV charge request message* to LA; the message should indicate the amount of required electricity S and the price z that EV can afford. EV encrypts the message by LA's public key; it attaches its current PID and a secret session key to the request. The message also involves a timestamp and a nonce to prevent replay attacks. If the EV, which currently has the PID EV_q, wants to sell E amount of power to the grid by the price w, then it generates a secret key k_q. Next, EV combines the required amount E, the suggested price w, its current PID EV_q, and the secret key k_q, includes timestamp T_{s9} and nonce K_9 values, $M_{cr} = E\,|w\,|EV_q\,|k_q\,|T_{s9}\,|K_9$, and then encrypts the result M_{cr} using the connected LA's public key n_{la}:

EV generates a set of co-prime numbers $d_1, d_2, \ldots, d_v$ and a random value Y_3 and then encrypts the M_{cr} by LA's public key to obtain a series of cipher texts:

$$m_{cr1} \equiv M_{cr}^2 + Y_3\, n_{la} (mod\, b_0),$$

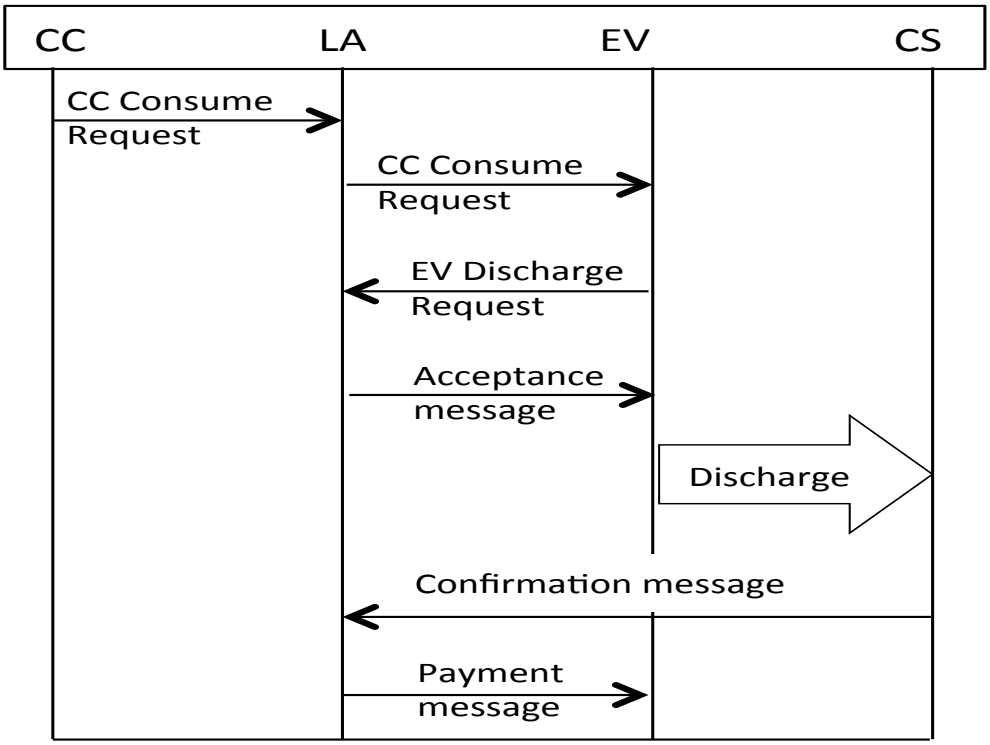

Figure 6.3: The CC consume request case.

$$m_{cr2} \equiv M_{cr}^2 + Y_3\, n_{la} (mod\, b_1),$$
$$\ldots\ldots\ldots,$$
$$m_{crv} \equiv M_{cr}^2 + Y_3\, n_{la} (mod\, b_j).$$
$$m_{cr} = \{m_{cr1}, m_{cr2}, \ldots, m_{crv}\}.$$

Afterward, EV sends m_{cr} to LA.

- When LA receives *the EV charge request message*, it decrypts the message and verifies the attached timestamp and nonce values. Then, LA aggregates the total amount of requested electricity from all EVs in the connected clusters. $Q = \sum_{EV_{charge}} S_{EV_{charge}}$, where EV_{charge} is the total number of EVs that ask to charge their batteries in LA's connected clusters. LA then sends a request message by the total demanded power to CC. The request also includes the different suggested prices by EVs. The message should involve timestamp T_{s10} and nonce K_{10} to prevent replay attacks; $m_t = Q\ |T_{s10}\ |K_{10}$. Then, it is signed by LA's private key and encrypted by CC's public key. LA picks a random padding U_6 and calculates $H(m_n U_6)$ so that $H(m_n U_6)$ is a square modulo n_{la}, $x_6^2 = H(m_n U_6) mod n_{la}$. The signature on m_t is the pair (U_6, x_6). Therefore, the signed message $M_T = (m_t, U_6, x_6)$ is encrypted by n_{cc}.

LA generates a set of co-prime numbers $c_1, c_2, \ldots, c_h$ and a random value Y_4 and then encrypts the M_T by CC's public key to obtain a series of cipher texts:

$$m_{T1} \equiv M_T^2 + Y_4\, n_{cc} (mod\, c_0),$$
$$m_{T2} \equiv M_T^2 + Y_4\, n_{cc} (mod\, c_1),$$
$$\ldots\ldots\ldots,$$

$$m_{Th} \equiv M_T^2 + Y_4\, n_{cc} (mod\, c_h).$$
$$m_T = \{m_{T1}, m_{T2}, \ldots, m_{Th}\}.$$

Afterward, LA sends the message m_T to CC.

- CC decrypts the message and verifies the signature of LA and then checks validity of timestamp and nonce, then it checks if its resources can cover the required amount of electricity. It also verifies the offered prices from different EVs and set the final price for its sold power. CC then sends a confirmation message to each LA; this message contains the final price and the location of CSs for each region. This message is signed by CC's private key and then encrypted using LA's public key.

- LA sends an order message to each EV; the order message contains the location of target CS and the corresponding price. LA encrypts the order message to EV by the previously shared session key; also, the message includes EV's PID. For example, if LA chooses the EV with current PID EV_q, then LA encrypts the order message with its shared secret key k_q; $M_o = E_{k_q}(charge, L_{cs}, cash/token)$, where L_{cs} is the location of target CS.

- EV first pays the required price to LA; it sends the price in a payment message, e.g., EV_q pays to LA by $M_{p2} = E_{k_q}(cash/token)$. Next, LA sends a confirmation message to CS. The confirmation message M_{f2} contains the assigned power to that EV and its current PID EV_q. The message involves a timestamp and nonce values $m_{f2} = (M_{f2} \,|\mathcal{L}_{cs}\, |T_{s11}\, |K_{11})$. LA then encrypts m_{f2} by the pre-shared secret key between CS and LA k_{cs}, $MF_1 = E_{k_{cs}}(m_{f2})$. Consequently, CS charges the agreed amount of electricity to EV after checking the message's validity. In other words, EV pays to LA before charging its battery.

At the end of the current charging round, LA only stores the total amount of sold electricity Q and the corresponding total profits H for this round. LA keeps this information to calculate the total bill for the month. CC does the same calculations and saves the results in its record too. Figure 6.4 shows the EV Charge Request Case.

6.3.2.4 Case 4. The EV Discharge Request

In this case, EV wants to discharge its battery to the grid, i.e., sell a portion of its electricity to the grid. Then, EV follows a specific procedure:

- EV first sends *the EV discharge request message* to LA; the message should indicate the amount of sold electricity Wt and the price that EV offers l. EV encrypts the message by n_{la}; it attaches to the request its current PID and the suggested secret session key.

- LA aggregates the total offered electricity amount from the interesting EVs and sends a request message by the total amount to CC. The message is signed by LA's private key and encrypted by CC's public key. CC computes the current needs of the grid and estimates the suitable corresponding price. It then sends confirmation messages to LAs; the confirmation message contains the electricity amount, the price, and the location of CSs in LA's region.

- Next, LA sends order messages to EVs that are interested to discharge their batteries to the grid; each order message contains the location of target CS and the price for discharging power. LA encrypts the order message by the previously shared session key with that EV.

- EV first discharges the agreed amount of electricity to the required CS, which sends a confirmation message to LA. Then, LA pays to EV the price of sold power at the moment via

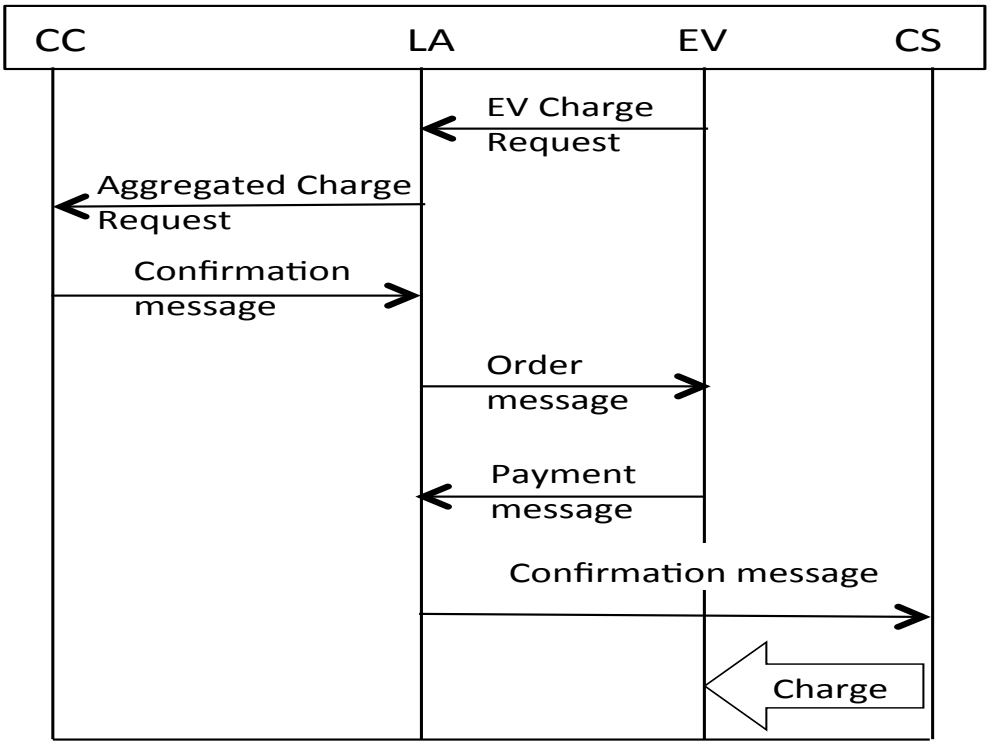

Figure 6.4: The EV charge request case

a payment message. In other words, EV discharges the assigned electricity first, and then LA directly pays the corresponding price to it.

At the end of the current discharge round, LA (also CC) only stores the total amount of sold electricity $\boldsymbol{Qc}$ and the corresponding total price $\boldsymbol{L}$. Figure 6.5 shows the EV Discharge Request Case.

6.3.3 Billing phase

First, both of LA and CC compute the total amount of exchanged power and the corresponding price or profit for different cases for the whole month. For *the CC Supply Request Case*, the total sold electricity is $\boldsymbol{A_{all}} = \sum_{\boldsymbol{b_s}} \boldsymbol{A_{b_s}}$ and the corresponding profit $\boldsymbol{X_{all}} = \sum_{\boldsymbol{b_s}} \boldsymbol{X_{b_s}}$, where $\boldsymbol{b_s}$ is the total number of the CC Supply Request sessions per month. According to *the CC Consume Request case*, the total purchased electricity $\boldsymbol{C_{all}} = \sum_{\boldsymbol{b_c}} \boldsymbol{C_{b_c}}$ and its total price $\boldsymbol{Y_{all}} = \sum_{\boldsymbol{b_c}} \boldsymbol{Y_{b_c}}$, where $\boldsymbol{b_c}$ is the total number of the CC Consume Request rounds for the month. In *the EV Charge Request case*, the total amount of charged electricity equals $\boldsymbol{Q_{all}} = \sum_{\boldsymbol{b_h}} \boldsymbol{Q_{b_h}}$ and the corresponding profit $\boldsymbol{H_{all}} = \sum_{\boldsymbol{b_h}} \boldsymbol{H_{b_h}}$, where $\boldsymbol{b_h}$ is the total number of the EV Charge Request sessions per month. For *the EV Discharge Request case*, the total amount of discharged electricity equals $\boldsymbol{Qc_{all}} = \sum_{\boldsymbol{b_d}} \boldsymbol{Qc_{b_d}}$ and the corresponding profit $\boldsymbol{L_{all}} = \sum_{\boldsymbol{b_d}} \boldsymbol{L_{b_d}}$, where $\boldsymbol{b_d}$ is the total number of the EV Discharge Request sessions per month. Next, they compute the total sold electricity $\boldsymbol{E_S} = \boldsymbol{A_{all}} + \boldsymbol{Q_{all}}$ and the corresponding profit $\boldsymbol{P_F} = \boldsymbol{X_{all}} + \boldsymbol{H_{all}}$, and the total purchased electricity $\boldsymbol{E_U} = \boldsymbol{C_{all}} + \boldsymbol{Qc_{all}}$ and its total price $\boldsymbol{P_R} = \boldsymbol{y_{all}} + \boldsymbol{L_{all}}$. Then, LA signs the billing information $\boldsymbol{m_B} = (\boldsymbol{E_S}, \boldsymbol{P_F}, \boldsymbol{E_U}, \boldsymbol{P_R})$ by its private key, i.e., after adding timestamp and random nonce values, and then encrypts it by CC's public key. Next, LA sends the resulted bill message $\boldsymbol{B}$ to CC.

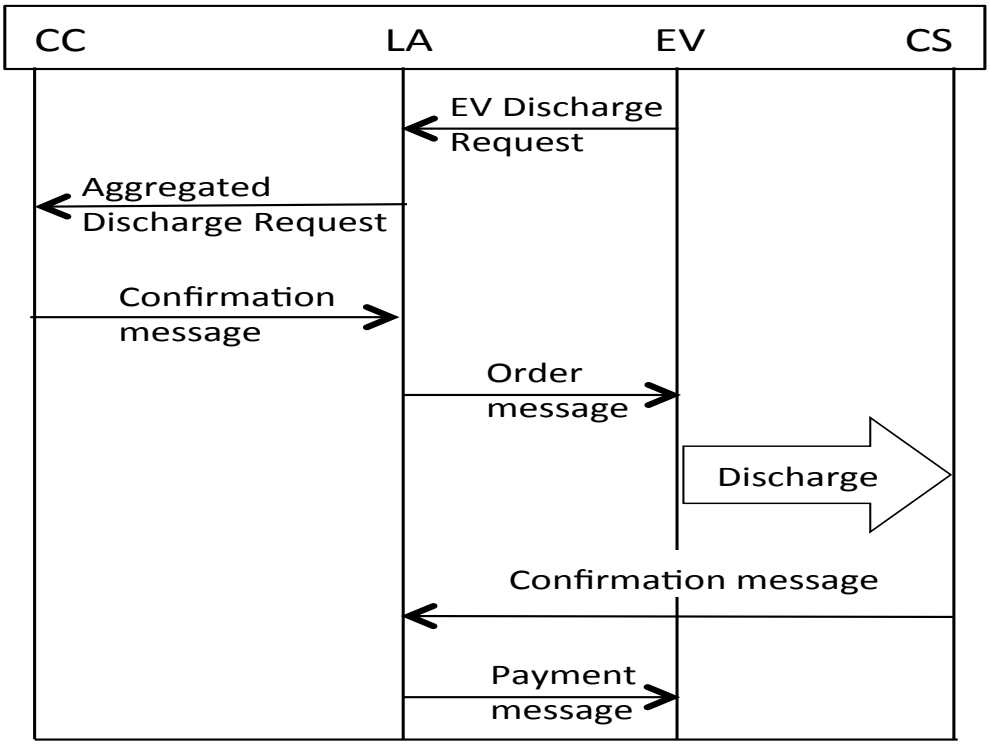

Figure 6.5: The EV discharge request case

$$\boldsymbol{LA} \xrightarrow{B} \boldsymbol{CC}.$$

After decrypting the message and checking LA's signature, CC compares LA message with its computed information. Subsequently, CC computes the net price and pays to LA via a payment message $\boldsymbol{M_P}$, which is signed by $(\boldsymbol{p_{cc}}, \boldsymbol{q_{cc}})$ and encrypted by $\boldsymbol{n_{la}}$.

$$\boldsymbol{CC} \xrightarrow{M_P} \boldsymbol{LA}.$$

It is obvious that EVs deal only with the related LA and have no connection with CC. Moreover, LA communicates with EVs via their PIDs, which are changing frequently. EV may use different PID for each communication session with LA. Furthermore, EV should take action first in all cases, because it is the untrusted party. When CC needs to supply power or EV wants to discharge its battery, EV transmits the electricity to grid first, and then receives the payment from LA. Accordingly, EV should pay the electricity price first before charging its battery. In addition, LA does not even need to save the various PIDs for different connected EVs; it only saves the total amount of traded power and the corresponding total price/profit for each round, e.g., if LA connects with **50** EVs in a CC supply request case, it stores the total power sold to the **50** EVs and the corresponding total profit. As a result, the privacy of EVs is preserved; EV's real ID, exact location and personal information are protected. At the same time, LAs and CC assure their profits and overcomes the trust problem, i.e., avoid malicious EVs. The scheme also ensures the traceability of electricity trade operation.

6.4 Security Analysis

The proposed scheme attempts to guarantee several security requirements for V2G connections simultaneously: EVs' privacy (i.e., owner identity and vehicle location), confidentiality and messages integrity, and overcomes EVs' authenticity problem and achieves the power trade accountability.

EVs' Privacy. The proposed scheme guarantees EV's privacy by preserving EV's real identity from being exposed to any party, i.e., adversaries, aggregators, or even grid's operator. Our scheme allows each EV to generate its own PIDs; these PIDs can be changed per session or just from time to time. This frequent change assists in preserving EV's private information so that no party can link between the EV and a specific location or ID. For instance, $\boldsymbol{EV_i}$ needs to charge its battery by the value $\boldsymbol{x}$ at time $\boldsymbol{t_1}$. So, $\boldsymbol{EV_i}$ selects a pseudo identity $\boldsymbol{PID_1}$ and new secret session key $\boldsymbol{k_1}$ for this session. At time $\boldsymbol{t_2}$, $\boldsymbol{EV_i}$ wants to sell a certain electricity amount $\boldsymbol{y}$ to the grid; then, it chooses another pseudo identity $\boldsymbol{PID_2}$ and secret session key $\boldsymbol{k_2}$ to that new session. Therefore, no party can link between these two sessions and $\boldsymbol{EV_i}$, as $\boldsymbol{EV_i}$ has finished all the related processes including payment during the session. In the charging session, it pays for $\boldsymbol{x}$ amount, and then charges its battery. While in the discharging session, $\boldsymbol{EV_i}$ injects $\boldsymbol{y}$ to the grid before receiving the payment. No need to store information about different PIDs and session keys for $\boldsymbol{EV_i}$ in LA or grid's records, as it is not useful for LA or CC after the sessions end. As a result, neither LA nor CC can link between $\boldsymbol{EV_i}$, i.e., its real identity or pseudo ones, and these sessions.

In addition, if an adversary $\mathscr{A}$ could extract certain messages from the two sessions, $\mathscr{A}$ cannot discover that the involved EV in the two sessions is the same vehicle. In other words, $\mathscr{A}$ cannot link between the EV and its activities. Furthermore, grid's operator and associated LAs do not need to know the real identities of EVs; they guarantee grid's profit without tracing EVs' real identities. So, they will not ever ask for EV's real identity. Then, if an EV receives any request to expose its real identity, it knows that the requester is an adversary and blocks that malicious request. Suppose $\mathscr{A}$ compromises the legal $\boldsymbol{LA_i}$ and contacts with its EVs. $\mathscr{A}$ may try to obtain information about EVs' real identities to seize certain financial gain. In that case, EVs know that $\boldsymbol{LA_i}$ is compromised. So, they ignore its request and report to the grid's operator about that LA. The operator verifies LA's status after receiving a certain number of bad reports about it.

Authenticity. The signatures of CC and LA authenticate their messages. For example, the CC supply request message is authenticated by CC's signature; no malicious party can forge it. On the other hand, grid's operator guarantees that the involved EVs operate honestly, because they have to take action first by paying the electricity expenses before charging their batteries or discharging to the grid before receiving the payment, consequently the operator assures the financial profits of the grid. In addition, EV's current PID and shared key authenticate its messages during the running session. The proposed scheme forces EVs to follow a specific procedure to prevent them from acting maliciously, i.e., EV has no chance to misbehave, because it has to finish its part first, and then receive the corresponding action from the grid. Even if $\mathscr{A}$ impersonates an EV, he/she cannot seize its benefits, because $\mathscr{A}$ has to follow the procedure by accomplishing the EV's part in the electricity exchange operation first. Consequently, $\mathscr{A}$ is forced to behave honest to receive the electricity or the price. In summary, the scheme overcome EVs' authentication problem and guarantees the grid's financial profit.

Confidentiality and Messages Integrity. The proposed scheme assures confidentiality and integrity of the exchanged messages using a combination of public and symmetric key schemes.

In the CC Supply Request case, for instance, before CC sends the CC supply request message $\boldsymbol{M_s}$, CC authenticates it by its signature. No party can modify its contents, but they can access it. In addition, no replay attacks can success, as the message contains timestamp $\boldsymbol{T_{si}}$ and nonce $\boldsymbol{L_i}$. Similarly, when LA forwards that message to EVs. When interested EVs response to LA by charge request messages $\boldsymbol{M_c}$. They encrypt it by $\boldsymbol{n_{la}}$. So, confidentiality and integrity of their messages are guaranteed. For the acceptance message $\boldsymbol{M_a}$, LA encrypts it by the secret key suggested by EV. So, this message is secured against confidentiality and integrity attacks. For instance, if LA chooses EV with $\boldsymbol{PID_i}$, and encrypts $\boldsymbol{M_a}$ by its session key $\boldsymbol{k_i}$, then any $\mathscr{A}$ cannot extract information, such as the target CS's identity, from the message. As well, the confirmation message $\boldsymbol{M_f}$ is encrypted by CS secret key so that its confidentiality and integrity is assured too. In addition, the used secret keys are frequently changed so that the probability to compromise them is limited.

The proposed scheme utilizes the efficient BlueJay ultra-lightweight crypto-system, which is a combination of lightweight public key PASSERINE scheme and Hummingbird-2 lightweight symmetric encryption scheme.

The lightweight public key PASSERINE scheme that used to encrypt the exchanged messages during the establishment of the secured sessions between the LA and EVs, such as EV (dis)charging request messages, is an enhanced lightweight version of Rabin cryptosystem. It has been practically proven that Rabin-based crypto-systems hardness problem is equivalent to the integer factorization problem.

Theorem 1. *Let* $\boldsymbol{N = pq}$*, where* $\boldsymbol{p \equiv q \equiv 3 (mod\ 4)}$ *are primes, and define* $\boldsymbol{S_{N,l} = \{1 \leq x < N : gcd(x, N) = 1, 2^l \mid (x + 1)\}}$*, where* $\boldsymbol{x \in \mathbb{Z}_N^* - S_{N,l}}$*, then the probability that there exists* $\boldsymbol{y \in S_{N,l}}$ *with* $\boldsymbol{x \neq y}$ *but* $\boldsymbol{x^2 \equiv y^2 (mod N)}$ *equals* $\boldsymbol{1/2^{l-1}}$*. Then breaking the one-wanness security property of Rabin-based crypto-systems (i.e., factoring* $\boldsymbol{N}$ *into* $\boldsymbol{p}$ *and* $\boldsymbol{q}$*) is in polynomial time if the redundancy bits in the ciphertext is* $\boldsymbol{l = O(log(log(N)))}$*.*

Proof. Let $\boldsymbol{\mathcal{O}}$ be an oracle that takes the public key $\boldsymbol{N}$ and a ciphertext message $\boldsymbol{c}$ (with $\boldsymbol{l}$ redundant bits) as input and returns either the corresponding plaintext $\boldsymbol{m}$ or an invalid value $\boldsymbol{Null}$.

Choose a random $\boldsymbol{x \in \mathbb{Z}_N^*}$ such that neither $\boldsymbol{x}$ nor $\boldsymbol{N - x}$ satisfy the redundancy scheme (i.e., the $\boldsymbol{l}$ least significant bits are not all $\boldsymbol{1}$) and set the ciphertext $\boldsymbol{c = x^2 (mod\ N)}$ as input for $\boldsymbol{\mathcal{O}}$. According to theorem's assumption, if the probability that one of the two unknown square roots of $\boldsymbol{c\ mod\ N}$ has the correct $\boldsymbol{l}$ least significant bits equals $\boldsymbol{1/2^{l-1}}$, then $\boldsymbol{\mathcal{O}}$ can output the plaintext $\boldsymbol{m}$ from $\boldsymbol{c}$, where $\boldsymbol{\acute{x} = 2^l m + (2^l - 1)}$, then we have $\boldsymbol{(\acute{x})^2 \equiv x^2 (mod\ N)}$ and $\boldsymbol{\acute{x} \not\equiv \pm x (mod\ N)}$. Hence $\boldsymbol{gcd(\acute{x} - x, N)}$ will split $\boldsymbol{N}$.

So, if $\boldsymbol{l = O(log(log(N)))}$, then it is required approximately $\boldsymbol{2^{l-1}}$ trials to factor $\boldsymbol{N}$ in polynomial time. As a result, factoring $\boldsymbol{N}$ is NP-hard problem when $\boldsymbol{l \neq O(log(log(N)))}$. □

While the lightweight symmetric encryption scheme Hummingbird-2, with an innovative hybrid structure of block cipher and stream cipher, is utilized to encrypt the exchanged messages during the (dis)charging sessions, such as acceptance or payment messages. Hummingbird-2 with its four optimal 4-bit S-Boxes belongs to a group of lightweight symmetric schemes that are proven to be resistible to differential, linear and algebraic attacks [172]. Moreover, the used secret keys to encrypt the exchanged messages during (dis)charging sessions are frequently changed, i.e., a new key for each session. For instance, $\boldsymbol{EV_t}$ needs to charge its battery at time $\boldsymbol{t_1}$. So, it selects a pseudo identity $\boldsymbol{PID_w}$ and new secret session key $\boldsymbol{k_w}$ for this session. At time $\boldsymbol{t_2}$, $\boldsymbol{EV_t}$ wants to sell a certain electricity to the grid so that it chooses another pseudo

identity $\boldsymbol{PID_v}$ and new secret session key $\boldsymbol{k_v}$ to that new session. The key is used to encrypt only two messages during the session; consequently, the probability of compromising the session key is diminished.

Accountability. Some previous works attempt to achieve the accountability and traceability. They require saving the detailed information about previous sessions for each EV, and most of time, they need to reveal EV's real identity by the end of the connection. While our proposed scheme guarantees the operation's accountability without revealing any private information about EVs. EVs' real identities are concealed and never revealed to any party neither during connection sessions nor during the tracing operation. Also, grid's operator does not need to maintain all the detailed information about numerous previous power trade sessions for tracing purposes. Therefore, the proposed scheme saves the storage capacity of both LA and CC. They just need to keep the total amount of purchased/sold electricity and the corresponding price/profit for each session.

6.5 Performance Evaluation

In this section, we evaluate the performance of the proposed scheme in terms of communication overhead and computation complexity.

6.5.1 Communication Complexity

The proposed scheme assures security and privacy demands for different parties in V2G connection with low communication overhead. During the initialization phase, TA sends only three messages to the parties: CC, LAs, and CSs. This process considers a trivial communication load for the high-capabilities TA. In addition, EVs, i.e., the restricted-capabilities parties, do not participate in this phase. Then, the communication overhead for that phase can be neglected.

According to the operation phase, the number of exchanged messages per session does not exceed six messages if only one EV is participated. Only one or two messages are sent by EV, which remains a minor communication duty for it. In the CC supply request case, CC sends one message to the connected LAs, and they forward this message to the involved EVs. Next, LA and the interested EV exchange three messages. Typically, EV receives three messages from LA and replies by one message. Also, LA receives one message related to this EV from the assigned CS. This process is repeated for all participated EVs in the session. Suppose the number of participated EVs per session is $\boldsymbol{q}$, then the total number of exchanged messages in that case equals $[\boldsymbol{2 + (4 * q)}]$ messages. Other three cases have also the same total overhead. In summary, the total communication overhead for each case is $[\boldsymbol{2 + (4 * q)}]$ messages most of them handled by CC and LAs. While, the maximum communication overhead for each EV is constant and equals sending one or two messages per session. According to LA, its overhead is linearly increased, as the number of involved EVs increased. Thus, the total communication load increases linearly with the increase in the number of selected EVs. However, the increase in the total communication load is bounded by the maximum number of EVs in the parking lot, which can be roughly determined. Figure 6.6 shows the communication load for each EV and the total overhead for each (dis)charging session. As shown, the communication overhead for EV is constant and very low. Although the total communication burden, most of it is handled by LA, is linearly increasing with the increase in EVs' number, it is still lightweight and bearable by the connection because of the limited number of participated EVs per session.

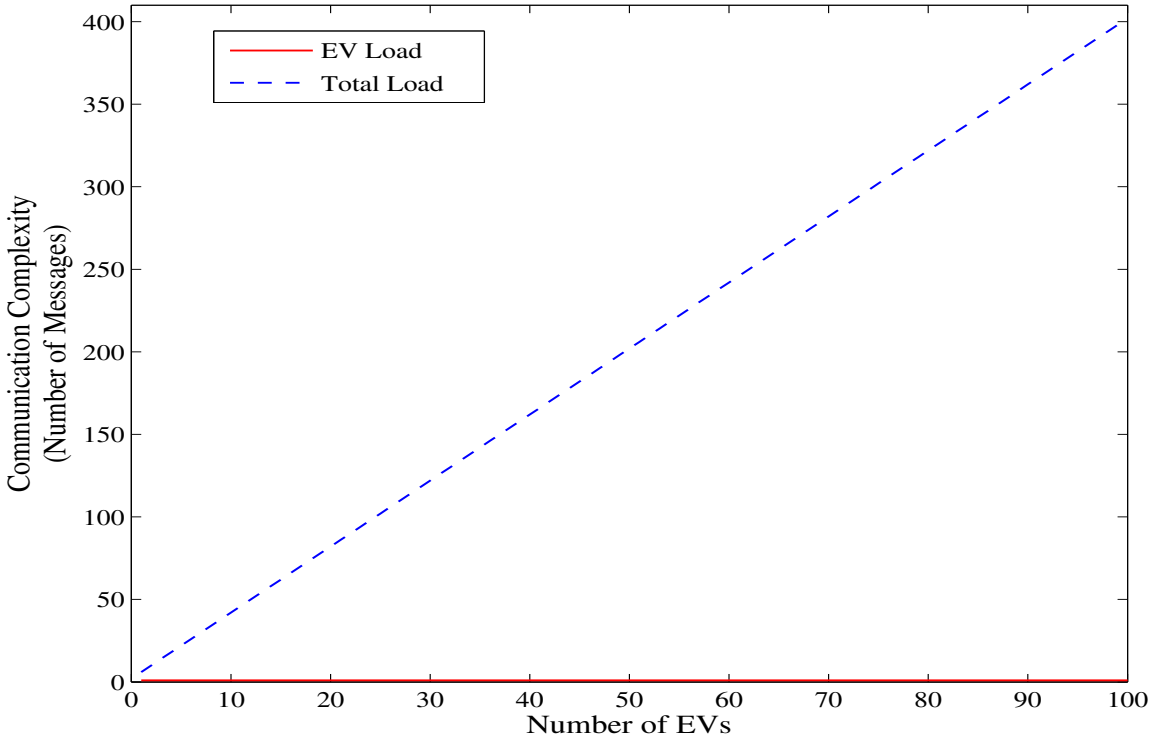

Figure 6.6: Communication complexity per session.

The total communication complexity for (dis)charging sessions per month equals $\boldsymbol{N = ([b_s + b_c + b_h + b_d] * [2 + (4 * q)])}$. In billing phase, CC and LA exchange two messages. Then, the total communication overhead per month equals $\boldsymbol{([b_s+b_c+b_h+b_d]*[2+(4*q)])+2}$ messages. Figure 6.7 illustrates the impact of EVs' number, also the effect of different number of electricity sessions on the total communication overhead per month. The overhead is increased by the increase of participated EVs' number. Also, the load is increased, if large number of sessions is performed during the month. However, the increase in communication complexity is limited and tolerable by different parties. As a result, our proposed scheme requires lightweight communication duty and suits the limited-capabilities EVs.

In addition to guarantee the security requirements for V2G connection, such as EVs' owners and location privacy, the electricity trade information confidentiality and integrity, and the involved parties authentication, our proposed scheme offers more lightweight communication and computation overhead. We compare our proposed scheme with the scheme [77], which is an efficient authentication scheme in the literature that also preserves EVs' real identities. Before the electricity (dis)charging operations begin, the central authority (**CA**), which is a third trusted party, issues a one-day permit for each EV so that EV can authenticate itself to LA by this permit, while it uses a PID to preserve its privacy. EV and LA then exchange several messages to authenticate EV and create the required symmetric key for that session. Each EV is obligated to send periodic reports about its status to the connected LA during the session, i.e., we assume that each session lasts for one hour and EV sends its periodic report every minute. At the end of the (dis)charging session, LA sends a signed reward message to EV, which uses it to claim its profit from **CA**. On the other hand, our scheme does not require the different parties to have pre-shared secret parameters, also it does not need the presence of the

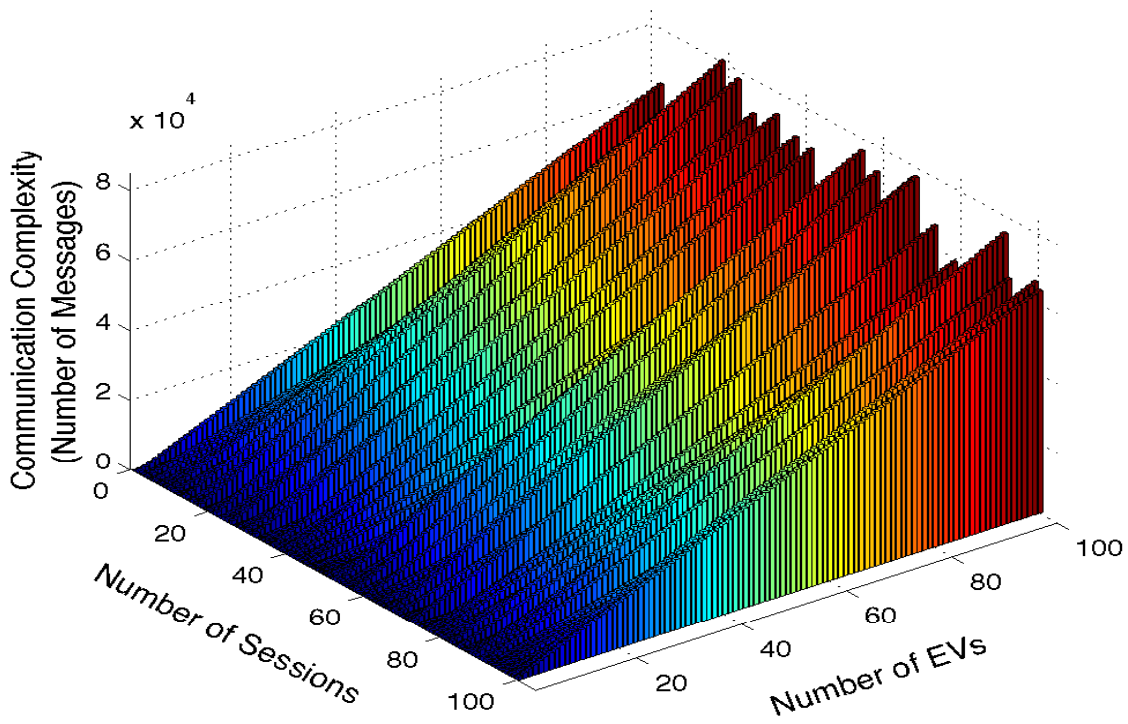

Figure 6.7: Total communication complexity.

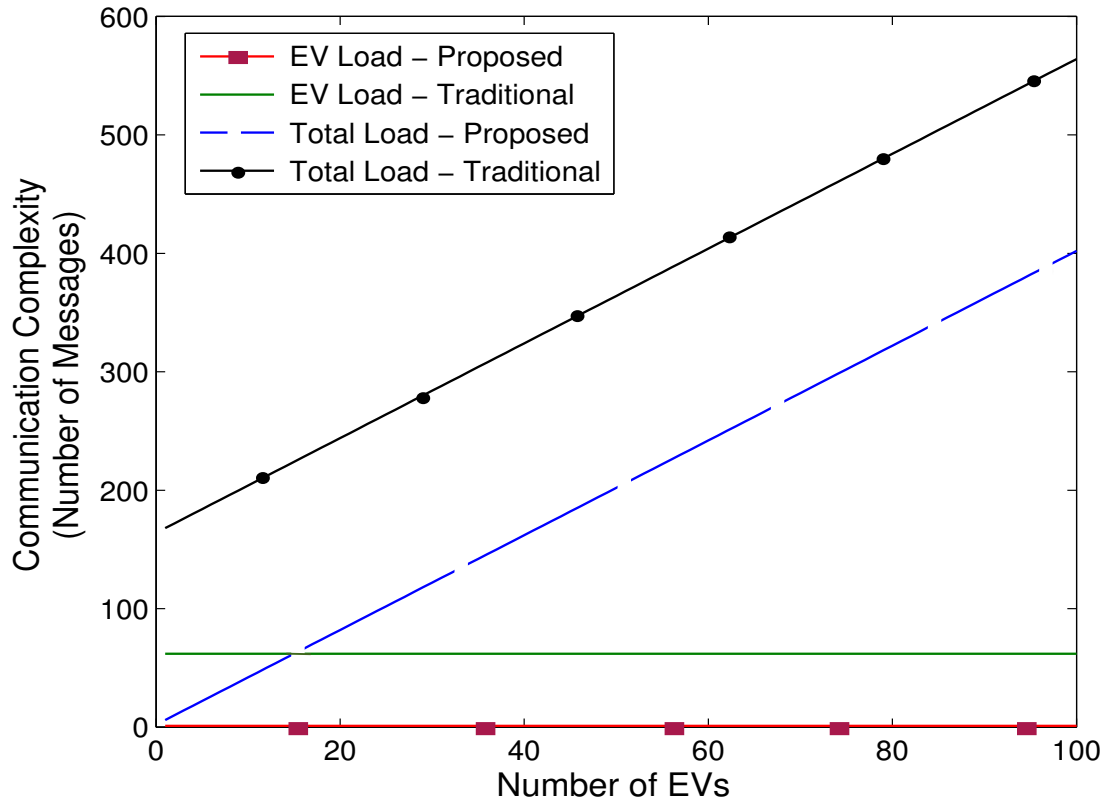

Figure 6.8: Communication complexity per session
Proposed .vs. Traditional Scheme.

third party during the operation phase. In addition, our scheme requires to exchange only six messages per session. Figure 6.8 shows the comparison between one session for the CC supply request case in our proposed scheme versus one charging session for [77]. The total number of exchanged messages during the charging session for [77] equals $(\mathbf{164 + 4 * q})$ messages, while the total exchanged messages' number in the CC supply request cases equals $(\mathbf{2 + 4 * q})$ messages. It can be seen that our proposed scheme tends to have less communication load than scheme [77] especially for the load on EVs. As a result, the proposed scheme saves the communication overhead for EVs and LAs; specifically when the number of connected EVs increased. In conclusion, our proposed scheme requires lightweight communication duty and suits the limited-capabilities EVs.

6.5.2 Computation complexity

EVs require lightweight crypto-systems because of their limited-computation capabilities. The proposed scheme guarantees that feature by reducing the number of exchanged messages, i.e., limited number of messages to be encrypted, in addition to use a combination of lightweight public and symmetric key schemes. Assume that $\boldsymbol{T_s, T_v, T_{Ep}, T_{Dp}, T_{Es}}$, and $\boldsymbol{T_{Ds}}$ are signing, verification, public key encryption, public key decryption, symmetric key encryption, and symmetric key decryption computation times in ms, respectively. During the initialization phase, TA provides the secret parameters to CC and LAs, while EVs are not involved in that operation. They only require to generate PIDs using the tiny-overhead AKRI-2 scheme. Thus, the computation overhead for that phase can be neglected.

According to the operation phase, CC, LAs, and EVs should perform few crypto operations: certain processes are public key while the remaining are symmetric key operations. Only three messages per session are encrypted using the lightweight public key Passerine crypto-system while the remaining messages are ciphered by the symmetric key Hummingbird-2 system. EVs are only burdened by trivial computation load. In the CC supply request case, CC sends the CC supply request message to the connected LAs; this message is signed using its private key. LA first verifies CC's signature and signs the message before forwarding it to the connected EVs. Next, the interested EV encrypts a charge request message by $\boldsymbol{n_{la}}$ and replies to LA. So, LA has to perform a public key decryption process for that message. After that, the shared symmetric key is used to encrypt the remaining messages, i.e., three messages, for that session. Typically, EV receives three messages from LA and replies by one message. Consequently, the total computation time for the CC supply request case per EV is $\boldsymbol{T_{supply} = [(2 * T_s) + (2 * T_v) + (1 * T_{Ep}) + (1 * T_{Dp}) + (3 * T_{Es}) + (3 * T_{Ds})]}$ ms. The EV's share from that load is $\boldsymbol{T_{EV_{supply}} = [(1 * T_v) + (1 * T_{Ep}) + (1 * T_{Es}) + (1 * T_{Ds})]}$ ms. While, the total computation time for the CC supply request case for $\boldsymbol{q}$ participated EVs is $\boldsymbol{T_{supply_{all}} = [(2{*}T_s){+}[(q{+}1){*}T_v]{+}(q{*}T_{Ep}){+}(q{*}T_{Dp}){+}(3{*}q{*}T_{Es}){+}(3{*}q{*}T_{Ds})]}$ ms. The CC consume request case has the same total computation overhead per EV as the supply case $\boldsymbol{T_{consume} = T_{supply} = [(2{*}T_s){+}(2{*}T_v){+}(1{*}T_{Ep}){+}(1{*}T_{Dp}){+}(3{*}T_{Es}){+}(3{*}T_{Ds})]}$ ms. While, the EV's share in that load is different $\boldsymbol{T_{EV_{consume}} = [(1{*}T_v){+}(1{*}T_{Ep}){+}(2{*}T_{Ds})]}$ ms. However, the total computation time for the CC consume request case for $\boldsymbol{q}$ participated EVs is the same as $\boldsymbol{T_{supply_{all}}}$; $\boldsymbol{T_{consume_{all}} = [(2 * T_s) + [(q + 1) * T_v] + (q * T_{Ep}) + (q * T_{Dp}) + (3 * q * T_{Es}) + (3 * q * T_{Ds})]}$ ms.

According to the EV charge request case, EV first sends the EV charge request message to LA; this message is encrypted by LA's public key. LA decrypts it, and then sends a request message by the total request from all EVs that are interested in purchasing electricity from CC.

The message is signed by LA's private key and encrypted by CC's public key. CC decrypts the message, verifies the signature of LA, and then sends a confirmation to LA; the confirmation message is signed by CC's signing key and encrypted using LA's public key. LA sends an order message to each EV, which is encrypted by the previously shared session key. EV then pays the assigned price to LA by a payment message, which is encrypted by the given symmetric key too. LA then sends a confirmation message to the assigned CS; that message is encrypted by the pre-shared key between LA and CS. In this case, the total number of crypto operations per EV is two signing and two verification processes in addition to three encryption and three decryption public key processes and three encryption and three decryption symmetric key processes. $\boldsymbol{T_{charge} = [(2 * T_s) + (2 * T_v) + (3 * T_{Ep}) + (3 * T_{Dp}) + (3 * T_{Es}) + (3 * T_{Ds})]}$ ms, where $\boldsymbol{T_{charge}}$ is the total computation time for the EV charge request case per EV. The computation overhead for each charged EV $\boldsymbol{T_{EV_{charge}}}$ is one public key encryption, one symmetric key encryption, and two symmetric key decryption processes. $\boldsymbol{T_{EV_{charge}} = [(1 * T_{Ep}) + (1 * T_{Es}) + (1 * T_{Ds})]}$ ms. While, the total computation time for the EV charge request case for $\boldsymbol{q}$ EVs equals $\boldsymbol{T_{charge_{all}} = [(2 * T_s) + (2 * T_v) + [(q + 2) * T_{Ep}] + [(q + 2) * T_{Dp}] + (3 * q * T_{Es}) + (3 * q * T_{Ds})]}$ ms. Finally, in the EV discharge request case, the involved parties follow almost the same procedure as in the EV charge request case but for discharging EVs. So, the total number of operations per EV $\boldsymbol{T_{discharge}}$ during that case is the same as the EV charge request case $\boldsymbol{T_{discharge} = T_{charge} = [(2 * T_s) + (2 * T_v) + (3 * T_{Ep}) + (3 * T_{Dp}) + (3 * T_{Es}) + (3 * T_{Ds})]}$ ms. The computation overhead of discharging EV $\boldsymbol{T_{EV_{discharge}}}$ is a little different from $\boldsymbol{T_{EV_{charge}}}$; it includes one public key encryption, and two symmetric key decryption processes. $\boldsymbol{T_{EV_{discharge}} = [(1 * T_{Ep}) + (2 * T_{Ds})]}$ ms. While, the total computation time for the EV discharge request case for $\boldsymbol{q}$ EVs is the same as $\boldsymbol{T_{charge_{all}}}$; $\boldsymbol{T_{discharge_{all}} = [(2*T_s)+(2*T_v)+[(q+2)*T_{Ep}]+[(q+2)*T_{Dp}]+(3*q*T_{Es})+(3*q*T_{Ds})]}$ ms.

The total computation time for the operation phase per month equals $\boldsymbol{T_{op} = [(b_s * T_{supply_{all}}) + (b_c * T_{consume_{all}}) + (b_h * T_{charge_{all}}) + (b_d * T_{discharge_{all}})]}$. In billing phase, LA sends a billing message to CC, and CC replies by a payment message. Then, the computation load for billing phase is $\boldsymbol{T_{bill} = 2 * [T_s + T_v + T_{Ep} + T_{Dp}]}$. In summary, the total computation overhead per month equals $\boldsymbol{T = T_{op} + T_{bill} = ([b_s + b_c + b_h + b_d] * [(2 * T_s) + [(q + 1) * T_v] + (q * T_{Ep}) + (q * T_{Dp}) + (3 * q * T_{Es}) + (3 * q * T_{Ds})]) + (2 * [T_s + T_v + T_{Ep} + T_{Dp}])}$ ms. Most of the computation load is performed by LAs and CC, and this load is insignificant for their capabilities. While EVs' computation operations are mainly tiny symmetric crypto-operations.

The performance of our proposed scheme has been analyzed and evaluated deploying a hardware implementation of BlueJay crypto-system with a $\boldsymbol{1024 - bit}$ public modulus $\boldsymbol{n}$ and $\boldsymbol{32 - bit}$ register size. It runs on a Cortex M0 platform, i.e., simple and fast, cheap, low power, and smallest ARM processor, which is embedded on different parties in the V2G connection. Figure 6.9 shows the computation time per EV versus the total computation time per (dis)charging session. Clearly, each EV performs a small fixed number of crypto operations; while, the remaining computation load are handled by LA. Although the computation load for LA is linearly increased by the increase in EVs' number, its load is bounded and manageable. While Figure 6.10 shows the impact of different number of EVs and sessions on the total computation overhead per month. The total overhead is increased by the increase in EVs' number as well as the increase in the charging sessions' number. However, the increase in computation load is limited and tolerable by different parties. The proposed scheme saves the computation time and preserves the processing abilities for the participated parties in the electricity trade operation especially EVs.

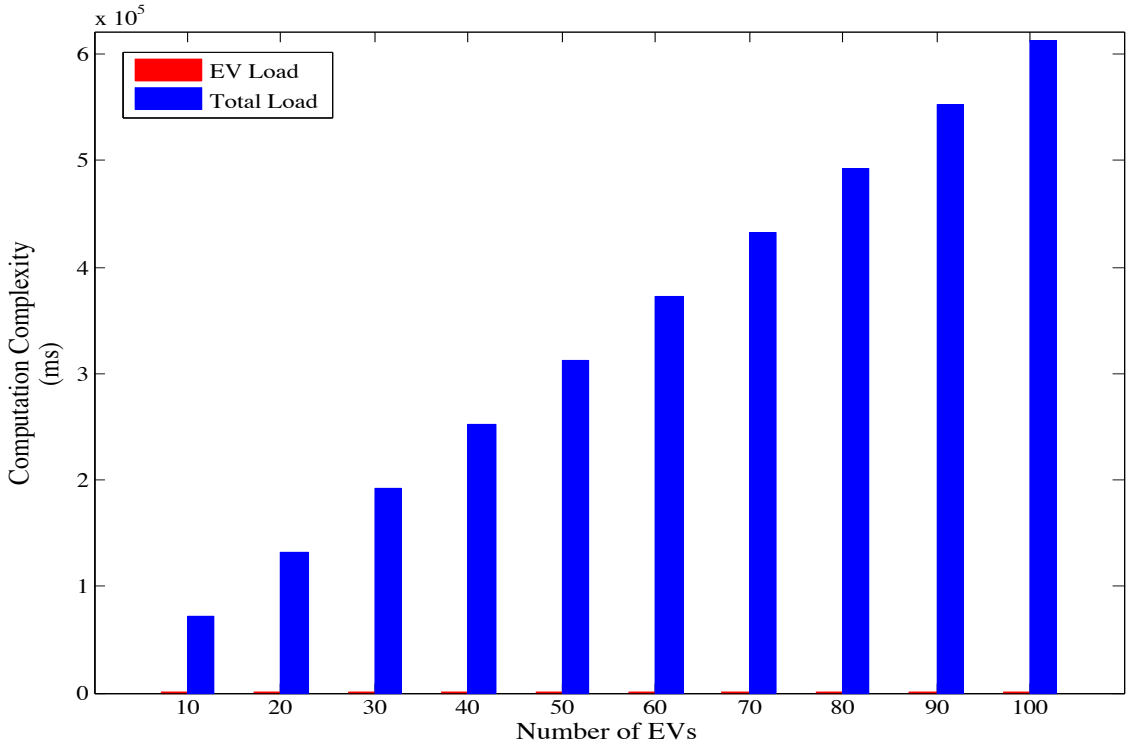

Figure 6.9: Computation complexity per session.

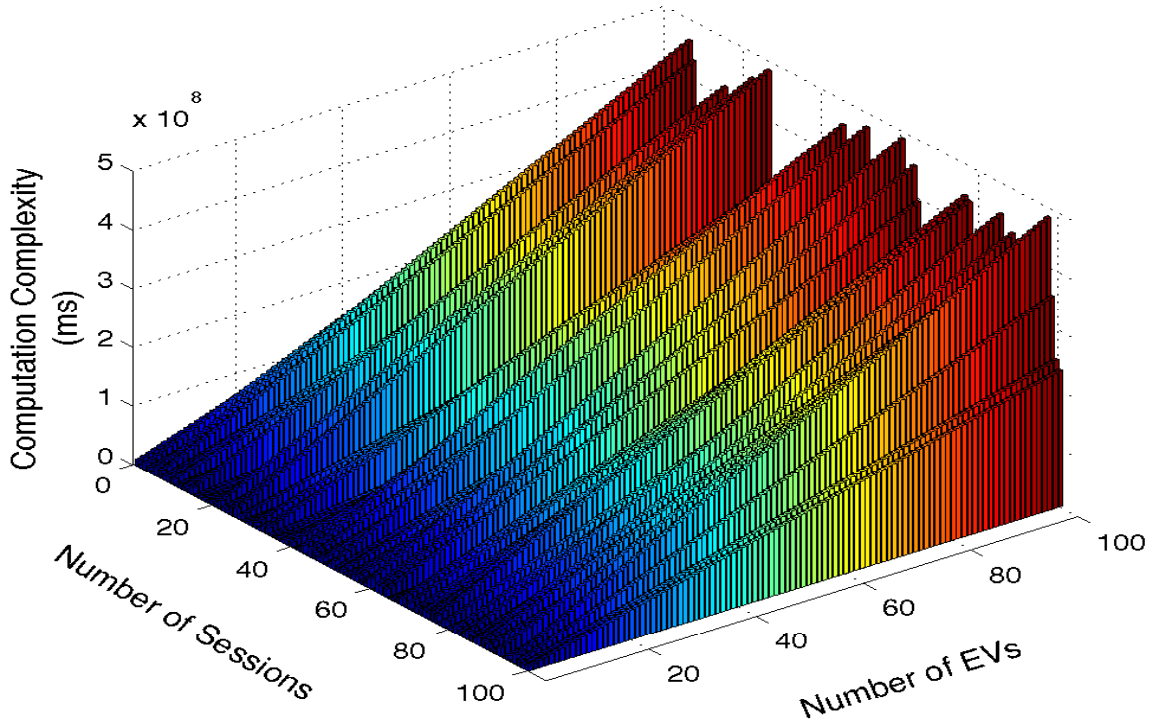

Figure 6.10: Total computation complexity.

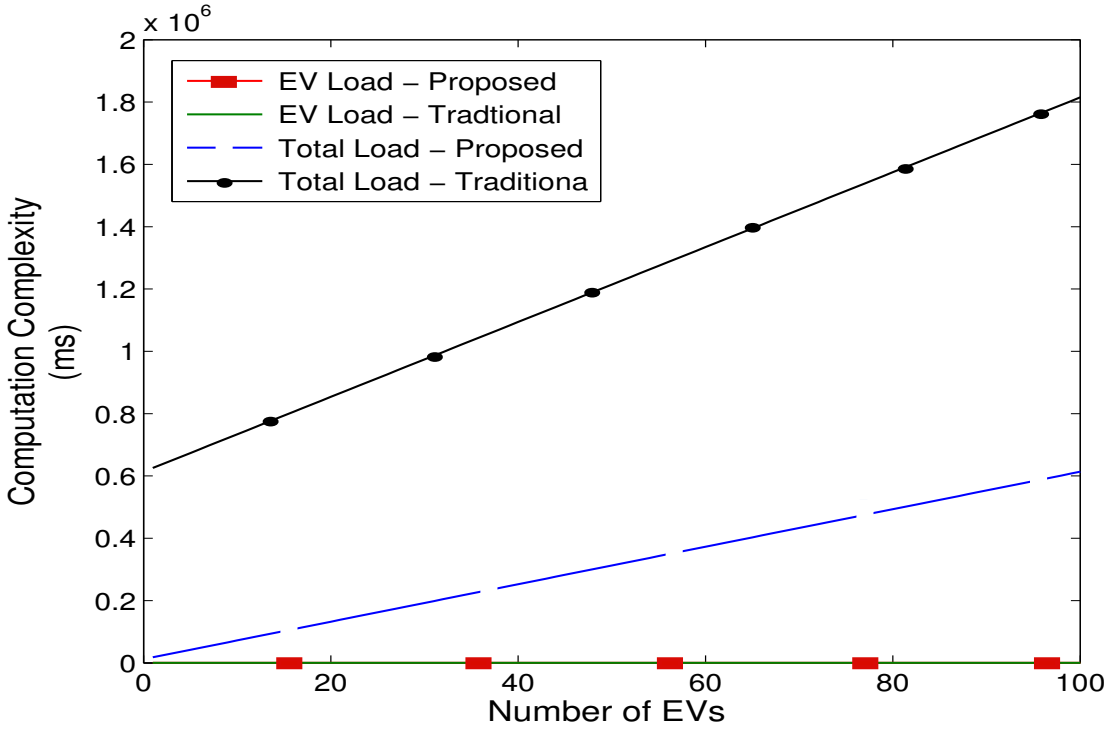

Figure 6.11: Computation complexity per session Proposed .vs. Traditional Scheme.

In Figure 6.11, we compare between the proposed scheme and the scheme in [77]. As shown, there is a big gap between the total computation delay in two schemes for the benefit of our scheme. The total delay increases from **18023.03** to **613560.64** msec for the proposed scheme versus **625558.87** to **1815078.95** msec in [77] for the whole session. However, there is no huge difference in the computation delay per EV in both schemes; **15.18** msec for the proposed scheme versus **20.52** msec for the traditional one; still, the computation load for EVs is tiny and tolerable by them. Then, our scheme saves the total computation time and preserves the processing abilities for the participated parties in electricity trade operation.

In summary, the proposed scheme not only guarantees the security requirements for all involved parties, i.e., CC, LAs, and EVs, but also provides light computation and communication overhead.

6.6 Summary

In this work, we have proposed a lightweight security and privacy-preserving scheme for V2G connection. The proposed scheme can guarantee several security requirements of V2G connections simultaneously. It preserves EV's owners and location privacy, diminishes the impact of malicious EVs, and overcomes EV's frequent authentication concern; it also assures confidentiality and integrity of the exchanged electricity trade messages. Moreover, the scheme keeps accountability and electricity-exchange operations traceability. Simulation results demonstrate that the proposed scheme reduces the overall communication and computation overhead for V2G connection, as it decreases the number of exchanged messages between different parties;

especially the messages sent by EVs. In addition, using a combination of symmetric key and lightweight public key schemes is further reducing the computation complexity. Thus, the proposed scheme is lightweight in terms of communication and computation complexities especially for EVs.

Chapter 7

Efficient Prevention Technique for False Data Injection Attack in Smart Grid

Power grid status is monitored by many measurement units spread all over the grid; these units periodically send their measurements to the grid's CC, which utilizes them to make the right decisions for the grid. FDI attack, which is one of the severe attacks that threatens the smart grid's efficiency and reliability, inserts fake measurements among the correct ones to mislead CC to make wrong decisions and consequently impact on the grid's performance. Several works are proposed to only detect FDI attack utilizing various estimation tests and optimization techniques. In our work [173], we propose an FDI attack prevention technique that protects the integrity and availability of the measurements at measurement units and during their transmission to the CC, even with the existence of compromised units. The proposed scheme alleviates the negative impacts of FDI attack on grid's performance. Security analysis and performance evaluation show that our scheme guarantees the integrity and availability of the measurements with lightweight overhead, especially on the restricted-capabilities measurement units. The remaining of the chapter is organized as follows. Section 7.1 introduces our system model, security parameters, and design goals. Section 7.2 reviews McEliece cryptosystem. In Section 7.3, we present our proposed scheme. Section 7.4 gives security analysis, while Section 7.5 evaluates the performance of our scheme. Section 7.6 presents a case study that applying the proposed scheme on IEEE -14 bus system. Finally, Section 7.7 concludes the work.

7.1 System Model

7.1.1 Network Model

Our model divides the power grid's distribution and consumption subsystems into small areas. Each area consists of a local substation (LS) that is responsible for collecting the readings of the measurement units in the area and guarantees their accuracy/integrity. $\boldsymbol{LS} = \{\boldsymbol{LS_1}, \boldsymbol{LS_2}, \ldots, \boldsymbol{LS_m}\}$, where $\boldsymbol{m}$ is the number of the local substations in the grid. Each LS connects to a large cluster of the measurement units that spreads all over the area to monitor its status. LS connects to these units via a wired connection, such as fiber optics, or wireless connection, e.g., WiFi. Most probably, the expensive wired connections are utilized in the

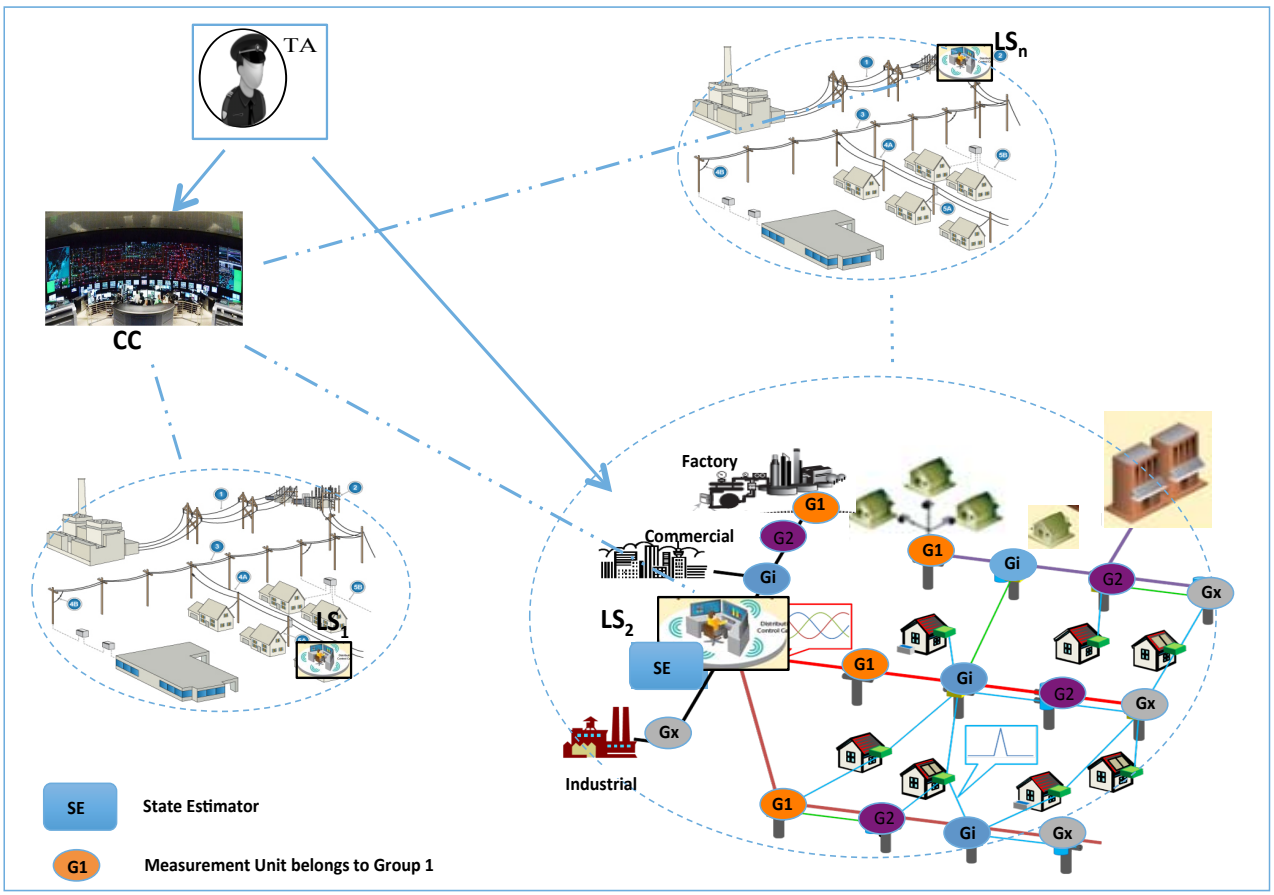

Figure 7.1: System model.

distribution subsystems, which has big industrial institutions and large factories. While, the cheap wireless connections are used in the consumption areas that have diverse consumers. $\boldsymbol{MU = \{MU_1, MU_2, \ldots, MU_n\}}$, where $\boldsymbol{n}$ is the total number of measurement units in the area.

In each cluster, LS divides the connected units into groups; $\boldsymbol{G = \{G_1, G_2, \ldots, G_h\}}$, where $\boldsymbol{h}$ is the number of the groups connected to the LS. Each group contains a specific number of measurement units that are responsible for calculating a specific measurement value; in other words, all the units in the group monitor the same value; $\boldsymbol{G_i = \{MU_1, MU_2, \ldots, MU_j\}}$, where $\boldsymbol{j}$ is the number of the measurement units in the group $\boldsymbol{G_i}$, $\boldsymbol{j}$ is different for each group. For example, $\boldsymbol{G_1}$ may include 50 units $\boldsymbol{\{MU_1, MU_2, \ldots, MU_{50}\}}$ to monitor the same measurement, while $\boldsymbol{G_2}$ may have 60 units $\boldsymbol{\{MU_1, MU_2, \ldots, MU_{60}\}}$, ..., and *etc.* However, the units that are belonged to the same group are not conjugated, they are spread over the whole area. Only the related LS knows the number and the locations of units in each group. The number of measurement units in each group is different according to several parameters, such as the importance and nature of measurement, the location of units, and the surrounding environment status. Therefore, LS is the only party that can link between the measurement units and its task. Each LS has also an efficient SE that filters out the error or noise from the measurements. On the other hand, each LS sends the resulted measurements to the main CC of the grid via a secured wired connection. Finally, the model has a TA to provide the keying parameters for different parties in the connection. Figure 7.1 shows the network model.

7.1.2 Adversary Model

In FDI attack, adversaries threaten the measurements' integrity and availability by compromising the measurement units or intercepting the transmitted measurements and then negatively impact on the grid's stability and efficiency. The CC and LSs in power grid belong to the utility company and located in protected places. They are trusted parties; they will not attempt to falsify the received measurements.

The measurement units are non-trusted parties because of their location in hostile environment. The attacker $\mathscr{A}$ can compromise and impersonate the measurement units to inject false measurements, also can intercept and block the units' measurements. In addition, $\mathscr{A}$ can begin a replay attack, or attempt to intercept and forge the transmitted messages. However, $\mathscr{A}$ has limited resources; he/she cannot compromise all measurement units in the area; only a limited number of them.

7.1.3 Security Requirements and Design Goals

The proposed scheme aims to prevent the FDI attack from happening in the first place by guaranteeing the measurements' integrity and availability at units and during their transmission to CC. Accordingly, our scheme should give the first priority to the measurements' integrity; it needs to guarantee the accuracy of the measurements.

In addition, the scheme should assure the measurements' availability and prevents delaying or blocking them. Furthermore, the proposed scheme needs to be lightweight in terms of communication and computation overhead because of the presence of the limited-capability measurement units.

7.2 Preliminaries

7.2.1 McEliece cryptosystem

McEliece cryptosystem [174, 175] is an asymmetric encryption algorithm that is based on the hardness of decoding a general linear code, i.e., an NP-hard problem. The McEliece cryptosystem has fast encryption and decryption operations, also it is immune to many attacks so that it is a candidate for the post-quantum cryptography. However, the key sizes of the original McEliece cryptosystem are too large. So, we utilize in our work an improved version of McEliece cryptosystem with smaller keys [177, 178].

7.2.1.1 Notions

A binary linear $[\boldsymbol{n}, \boldsymbol{k}]$ error-correcting code $\boldsymbol{C}$ of length $\boldsymbol{n}$ is a subspace of $\mathbb{F}_2^n$ of dimension $\boldsymbol{k}$ and co-dimension $\boldsymbol{r} = \boldsymbol{n} - \boldsymbol{k}$. Code $\boldsymbol{C}$ can be defined either by a generator matrix or by a parity-check matrix. The generator matrix $\boldsymbol{G} \in \mathbb{F}_2^{k \times n}$ defines $\boldsymbol{C} = \{\boldsymbol{mG} \in \mathbb{F}_2^n \mid \boldsymbol{m} \in \mathbb{F}_2^k\}$ and the parity-check matrix $\boldsymbol{H} \in \mathbb{F}_2^{r \times n}$ defines $\boldsymbol{C} = \{\boldsymbol{c} \in \mathbb{F}_2^n \mid \boldsymbol{cH_T} = \boldsymbol{0}^r\}$. The syndrome $\boldsymbol{s} \in \mathbb{F}_2^r$ of a vector $\boldsymbol{x} \in \mathbb{F}_2^n$ is defined as $\boldsymbol{s} = \boldsymbol{Hx^T}$. It follows that if $\boldsymbol{x} \in \boldsymbol{C}$, then $\boldsymbol{s} = \boldsymbol{0}^r$, and otherwise $\boldsymbol{s} \neq \boldsymbol{0}^r$.

If there exists some integer n_0 such that every cyclic shift of a codeword $c \in C$ by n_0 positions results another codeword $\acute{c} \in C$, then code C is called quasi-cyclic. If $n = n_0 p$ for some integer p, then both the generator and the parity-check matrix are composed of $p \times p$ circulant blocks. It suffices to store one row (usually the first) of each circulant block to completely describe the matrices.

(n, r, w)-MDPC code is a binary linear $[n, k]$ code defined by a parity-check matrix with constant row weight w, and (n, r, w)-QC-MDPC code is an (n, r, w)-MDPC code that is quasi-cyclic with $n = n_0 r$.

7.2.1.2 Key generation

To generate an (n, r, w)-QC-MDPC code with $n = n_0 r$, select the first rows $h_0, \ldots, h_{n_{0-1}} \in \mathbb{F}_2^r$ of the n_0 parity-check matrix blocks $H_0, \ldots, H_{n_{0-1}} \in \mathbb{F}_2^{r \times r}$ with weight $\sum_{i=0}^{n_{0-1}} wt(h_i) \leq w$ uniformly at random. The parity-check matrix blocks $H_0, \ldots, H_{n_{0-1}}$ are then generated by $r - 1$ quasi-cyclic shifts of $h_0, \ldots, h_{n_{0-1}}$. The parity-check matrix H is formed by concatenating $H_0, \ldots, H_{n_{0-1}}$.

Generator matrix $G = [I_k \mid Q]$ is computed from H in row reduced echelon form by concatenating the identity matrix I_k with

$$Q = \begin{pmatrix} (H_{n_{0-1}}^{-1}.H_0)^T \\ (H_{n_{0-1}}^{-1}.H_1)^T \\ \cdots \\ (H_{n_{0-1}}^{-1}.H_{n_{0-2}})^T \end{pmatrix}$$

Since both matrices are quasi-cyclic, it suffices to store their first rows instead of the full matrices.

The public key is the generator matrix G and the secret key is the parity-check matrix H.

7.2.1.3 Encryption

To encrypt a message $m \in \mathbb{F}_2^k$, generate an error vector $e \in \mathbb{F}_2^n$ with at most t set bits uniformly at random and compute $x = mG \oplus e$.

7.2.1.4 Decryption

To decrypt a ciphertext $x \in \mathbb{F}_2^n$, compute $mG \leftarrow \Psi_H(x)$. Since G is of systematic form, extract m from the first k positions of mG.

7.3 The Proposed Scheme

Our proposed scheme attempts to prevent FDI attack by guaranteeing the measurements' integrity and availability at measurement units (during measuring and sensing operation) and during measurements' transmission to LS and then to CC. So, the scheme can be divided into the following two phases:

7.3.1 Initialization phase

The initialization phase is responsible for establishing the network and defining the required security parameters.

TA provides the keying parameters for each party in the connection. TA issues a public private key pair for CC. The public key is $\boldsymbol{G_{cc}}$ while the private key is $\boldsymbol{H_{cc}}$. For each LS, TA provides $\boldsymbol{G_{ls}}$ as a public key and $\boldsymbol{H_{ls}}$ as a private key.

Each measurement unit in the LS's cluster has a unique ID issued by the TA, e.g., unit **1** in the first group in the LS's area has the ID $\boldsymbol{MU_1}$. The units' secret IDs are signed by LS so that each unit can use its ID to prove its identity during measurements' transmission. These IDs are also sent securely to the LS, which stored them in its powerful memory.

LS is connected to all measurement units in its local area; it divides the connected units into groups; each group is responsible for measuring a specific measurement, i.e., all units in the group compute the same value. LS is the only party that knows ID, location, function, and the group that each unit belongs to.

7.3.2 Operation phase

The operation phase is responsible for guaranteeing the integrity and availability of the measurements. So, it performs two functions: The first one is *operation for transmission* to guarantee the measurements' integrity and availability during the transmission from units to CC. The second function is *operation for compromised measurement units* to protect the accuracy and correctness of the received measurements if a number of units in the area is compromised.

7.3.2.1 Operation for transmission

The first function guarantees that the measurements will not be tampered with during their transmission to CC by utilizing the efficient improved McEliece public key crypto-system. The measurements are encrypted so that no party can modify them; only the connected LS decrypts them.

Every reading round, each unit measures the required value $\boldsymbol{MS_i}$, attaches its ID $\boldsymbol{MU_i}$, adds timestamp $\boldsymbol{T_{s1}}$ and nonce value $\boldsymbol{q_1}$, $\boldsymbol{ms_i} = \boldsymbol{MS_i} \| \boldsymbol{MU_i} \| \boldsymbol{T_{s1}} \| \boldsymbol{q_1}$ and then encrypts the result by LS's public key; $\boldsymbol{X_i} = \boldsymbol{ms_i G_{ls}} \oplus \boldsymbol{e_1}$.

Next, it sends the encrypted message to LS via unsecured network connection.

LS first decrypts the received message $\boldsymbol{X_i}$ by computing $\boldsymbol{ms_i G_{ls}} \leftarrow \boldsymbol{\Psi_{H_{ls}}(x)}$ using a $\boldsymbol{t}$-error correcting (QC-)MDPC decoder $\boldsymbol{H_{ls}}$. Since $\boldsymbol{G_{ls}}$ is on systematic form, LS extracts $\boldsymbol{ms_i}$ from the first $\boldsymbol{k}$ positions of $\boldsymbol{ms_i G_{ls}}$. It then checks the validity of the attached timestamp and random nonce.

7.3.2.2 Operation for compromised units

When the LS receives all the measurements from all units in the associated group; $\boldsymbol{M_{G_x}} = \boldsymbol{\{ms_1, ms_2, \ldots, ms_j\}}$, where $\boldsymbol{j}$ is the number of the measurement units in group $\boldsymbol{G_x}$, it first compares between the received measurements to check if some values are deviated so

far from the majority of measurements. Then, LS discards these deviated measurements $\{ms_1, ms_2, \ldots, ms_c\}$, and marks the related units $\{MU_1, MU_2, \ldots, MU_c\}$ as suspicious nodes.

LS adds the suspicious units to a specific list called the black list. It also maintains another list for the units with delayed values $\{MU_1, MU_2, \ldots, MU_d\}$, which is called the gray list. It contains the units that do not send their measurements to LS at the assigned time.

LS keeps an activity record for each unit in the local area that contains the previous measurements' transmission for the unit. These records assist to easily detect and replace the malicious nodes. These suspicious units are kept in the black/gray list until they exceed a predefined threshold t_p of sending malicious/delayed messages, then the LS physically replaces them by new honest units. These two lists are important to help guaranteeing measurements' integrity and availability.

The accepted number of deviated measurements that considers minority is different from one group to another; it depends on the importance of the measured data, the channel conditions and the environment around the sensors. It can be calculated as $h = \frac{j}{2} - i$, where j is the total number of measurement units in the group, and i runs from $\frac{j}{2}$ down to 1. The value of i is different for each group. If the number of deviated measurements is more than the accepted threshold, then LS discards the whole measurements of the group. While if the number of deviated measurements is a minority, the remaining measurements, the majority, are considered the correct measurements. Consequently, LS computes the average value of them to obtain the final (more accurate) measurement value R as follows:

$$R_{avg} = \frac{\sum_{k=1}^{l} ms_k}{l}$$

$$\sigma = \sqrt{\frac{\sum_{k=1}^{l} (ms_k - R_{avg})^2}{l}}$$

$$\Delta R_{avg} = \frac{\sigma}{\sqrt{l}}$$

$$R = R_{avg} \pm \Delta R_{avg}$$

, where $l \leq j$ is the number of the accurate measurements in group G_j, σ is the uncertainty in a single measurement value, ΔR_{avg} is the uncertainty in the mean R_{avg}, and R is the accurate average value of the collected measurements.

For example, suppose LS_w's connected area is divided into 1000 groups; each group is responsible for measuring one measurement value: group G_1's units are monitoring the value x_1, group G_2's units are monitoring the value x_2,..., group G_{1000}'s units are monitoring the value x_{1000}. Suppose that group G_i in the area has 50 measurement units $\{MU_1, MU_2, \ldots, MU_{50}\}$, then these units are spread over the area in a specific way known only by LS_w. When LS_w receives a measurement from a unit in a specific group, it saves the value in the allocated memory for that group in its database. After LS_w receives all the 50 measurements, it first check if one measurement or a few number of measurements are deviated far from the majority of the received 50 measurements, say 6 measurements are deviated. LS_w then discards them and adds the corresponding sensors $\{MU_1, MU_2, \ldots, MU_6\}$ to the black list, i.e., the units

that exceed the threshold for sending malicious readings will be physically replaced by honest units. $\boldsymbol{LS_w}$ then computes the average value of the remaining **44** measurements (the majority) to obtain the accurate result;

$$\boldsymbol{R_{i-avg}} = \frac{\sum_{k=1}^{44} \boldsymbol{ms_k}}{\mathbf{44}}$$

$$\boldsymbol{\sigma_i} = \sqrt{\frac{\sum_{k=1}^{44} (\boldsymbol{ms_k} - \boldsymbol{R_{i-avg}})^2}{\mathbf{44}}}$$

$$\boldsymbol{\Delta R_{i-avg}} = \frac{\boldsymbol{\sigma}}{\sqrt{\mathbf{44}}}$$

$$\boldsymbol{R_i} = \boldsymbol{R_{i-avg}} \pm \boldsymbol{\Delta R_{i-avg}}$$

, where the number of accurate measurements in group $\boldsymbol{G_i}$ is **44** and $\boldsymbol{R_i}$ is the average value of these measurements.

7.3.2.3 Operation for state estimation

At the end of the reading round, LS has a set of accurate measurements that represents the status of the area, i.e., the accurate observed measurements vector for that part of the grid $\mathscr{S}$. To guarantee that FDI attack does not exist and filter out all error and noise, LS also applies an efficient estimation test based on cosine similarity matching on the set of estimated measurements. Cosine similarity [176] is a measure of similarity between two vectors using the inner product space that measures the cosine of the angle between them, i.e., cosine similarity metric computes how similar two vectors of data are; if both vectors are identical, the cosine similarity value equals one. The provided cosine similarity based SE detects the existing of any deviation between the measured data vector $\overrightarrow{\mathscr{S}}$ from the units and the expected data vector $\overrightarrow{\hat{\mathscr{S}}}$ that is derived from the previous measurements.

$$\boldsymbol{\Lambda} = \cos\boldsymbol{\theta} = \frac{\overrightarrow{\mathscr{S}} \cdot \overrightarrow{\hat{\mathscr{S}}}}{\| \overrightarrow{\mathscr{S}} \| \| \overrightarrow{\hat{\mathscr{S}}} \|}$$

, where $\overrightarrow{\mathscr{S}} \cdot \overrightarrow{\hat{\mathscr{S}}}$ represents the inner product of the vectors $\overrightarrow{\mathscr{S}}$ and $\overrightarrow{\hat{\mathscr{S}}}$, and $\| \overrightarrow{\mathscr{S}} \| \| \overrightarrow{\hat{\mathscr{S}}} \|$ denotes the product of their Euclidean lengths.

When LS runs the cosine similarity test $\boldsymbol{\Lambda}$, the result could have one of three values:

$$\boldsymbol{\Lambda} = \begin{cases} \mathbf{1} & \text{the two vectors are identical} \\ \boldsymbol{> 0 \; and \; < 1} & \text{the two vectors have some differences} \\ \mathbf{0} & \text{the two vectors are totally different.} \end{cases}$$

Case 1: when $\boldsymbol{\Lambda}$ equals **1**; the collected measurements from the units in the grid are exactly the same as expected measurements. So, LS accepts this set of measurements as a correct

representation for the grid's current status and marks the set as a fully-guaranteed accuracy measurements.

Case 2: when $\mathbf{\Lambda}$ equals $\mathbf{0}$; the collected measurements from the units in the grid are totally different than the expected measurements. So, LS rejects this set of measurements and marks it as false measurements. LS physically check the corresponding measurement units; it also reports to CC about the situation.

Case 3: when $\mathbf{\Lambda}$ value is between $\mathbf{0}$ and $\mathbf{1}$; LS determines the accuracy level of the collected measurements according to the exact value of $\mathbf{\Lambda}$. For instance, if $\mathbf{\Lambda}$ value is close to $\mathbf{1}$, such as $\mathbf{0.9}$, then LS can accept that set of measurements as partially-guaranteed accuracy measurements. While if $\mathbf{\Lambda}$ value approaching $\mathbf{0}$, e.g., $\mathbf{0.2}$, then LS rejects this set of measurements and checks the measurement units to detect the place for that error. However, because of the applied security scheme, the probability of case 2 and 3 occurrence is diminished.

Finally, LS contains the resulted measurements for the area in a message $\boldsymbol{M_T}$. Afterward, LS attaches a timestamp $\boldsymbol{T_{s2}}$ and random nonce $\boldsymbol{q_2}$ to the message; $\boldsymbol{m_T = M_T \| T_{s2} \| q_2}$ and then encrypts it by CC's public key; $\boldsymbol{T_{ls} = m_T G_{cc} \oplus e_2}$. The received information from the connected LSs assists CC to make the right decisions for each area and for the whole grid.

7.4 Security Analysis

The main security concerns for that type of connection are the integrity and availability of measurements at measurement units and during their transmission to CC.

Integrity: The measurements' integrity is the major concern, as the injected false measurements can mislead CC to make wrong decisions and consequently impact severely on the grid's stability and reliability.

- *Compromised measurement units' case.* Certain number of measurement units send the same measurement to LS. Only LS knows how these units are spread and where are their exact locations, i.e., this helps to guarantee the accuracy of the measurements values. When the LS receives two different sets of values for the same measurement, it accepts the majority. However, if the number of measurement units in the two sets is close, CC rejects the two values and physically checks the units to detect which set is compromised. The probability that the attacker compromises a large number of measurement units from the same group is low, as they are distributed all over the local area, i.e., it is not necessary that the units close to each other in location measure the same value. LS only knows the measurement units included in each group and their exact locations.

- *Intercepted measurements case.* Adversaries cannot falsify the measurements during their transmission to CC, as the values are encrypted by the powerful McEliece cryptosystem. If adversary $\mathscr{A}$ manages to intercept the measurement $\boldsymbol{X_i}$, he/she can not modify its value, because $\mathscr{A}$ does not have the decryption key $\boldsymbol{H_{ls}}$ and consequently cannot extract the plaintext measurement $\boldsymbol{ms_i}$ from $\boldsymbol{X_i}$. According to the final measurements matrix $\boldsymbol{M_T}$, it is sent from each LS to CC encrypted by CC's public key so that no party except CC can decrypt the message or has any information about $\boldsymbol{M_T}$. Moreover, the message contains a timestamp and random nonce to prevent the replay attack. So, $\mathscr{A}$ cannot interpret/modify the message's content or begin a replay attack.

Availability: To resist FDI attack and guarantee the grid's efficiency and reliability, the measurements should be available to LS, i.e., and consequently to CC, whenever they asked to.

The proposed scheme allows redundancy in the measurements, as several measurement units measure the same value and send their versions of data to the LS. So, if some units in the same group are not available, the remaining still send their measurements to the LS. If $\mathscr{A}$ compromises a certain number of units, LS still can guarantee the correctness of each measurement value by receiving redundant values for the same measurement from other units.

Moreover, LS reduces the probability of attack by eliminating the suspicious units. If the LS does not receive the expected periodic reports from certain units, it declares them as suspicious nodes, i.e., adds them to the gray list, and sends a report to CC about the problem. On the other hand, LS's computation resources are ready to manage a large number of received measurements, as it expects to receive periodic reports from all measurement units in the connected area. Then, if the LS receives a huge number of reports, more than the expected, from one measurement unit or a group of units, then LS blocks them and declares them as malicious units; it also reports to CC to check them.

Confidentiality: Although confidentiality is not a high-priority security concern for that type of connection, it is still an important requirement for the grid's stability. Our proposed scheme guarantees the confidentiality of measurements during their transmission to LS and then to CC. No party can extract the contents of the transmitted measurements because of the powerful McEliece cryptosystem that is utilized to secure them. If adversary $\mathscr{A}$ manages to intercept measurement $\boldsymbol{X_i}$, $\mathscr{A}$ cannot obtain any information about the plaintext measurement $\boldsymbol{ms_i}$, because $\mathscr{A}$ does not have the decryption key $\boldsymbol{H_{ls}}$ and consequently cannot decrypt $\boldsymbol{X_i}$.

Even if $\mathscr{A}$ manages to compromise some units and captures/falsifies their messages or prevents them from sending to LS, LS still can have the same measurement from other measurement units in the group so that compromising measurement units or blocking messages does not have a significant impact on the final measurement value.

7.5 Performance Evaluation

In this section, we evaluate the performance of the proposed scheme in terms of communication overhead and computation complexity. Also, we study the impact of the included SE on the performance of the proposed scheme.

7.5.1 Communication Complexity

To guarantee the accuracy of the gird's decisions, information about the grid's current status need to be sent periodically from measurement units to CC. So, the communication duty for each measurement unit is sending one message every reading period, which is an affordable burden for the limited-capability units. Although our proposed scheme follows the standard technique, i.e., sending periodic reports, it saves more communication load for units, as it guarantees the integrity and availability of the measurements, which consequently reduces the probability of measurements' retransmission due to error or malicious attacks. On the other hand, in traditional grid, the measurements are sent in plaintext so that they are more prone to be compromised by adversaries, and most probably the measurements require to be retransmitted several times. In addition, the adversary can know which units are the responsible for estimating a specific measurement and consequently compromise all of them. In that case, CC cannot distinguish fabricated measurements from the right ones and may need to replace all the units.

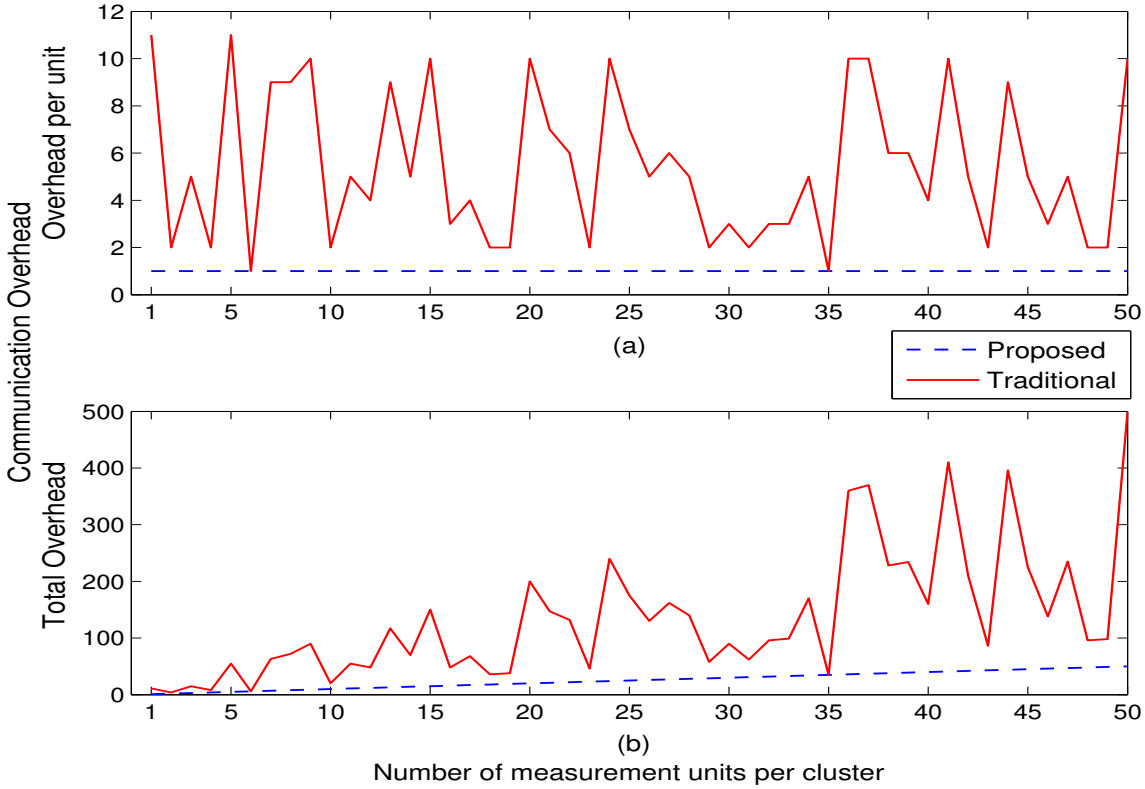

Figure 7.2: Communication overhead.

Figure 7.2 shows the communication overhead for our proposed scheme versus the traditional case. Figure 7.2(a) shows the number of retransmission for each measurement unit in the proposed scheme versus the traditional case per round; the unit in the proposed scheme case needs to send the measurement for one time only so that the communication overhead per unit is fixed and trivial overhead. While, in traditional scheme, the unit has to keep resending the message until a correct accurate version of the measurement reach to the LS. The number of retransmitted messages is varied according to the probability of error, noise, and attacks' occurrence in the communication channel. According to Figure 7.2(b), it demonstrates the total communication overhead for the cluster in the local area. As shown, the total communication load for the proposed scheme is linearly increased as the number of the connected units is increased; however, the overhead is limited by the maximum number of the units per cluster. In the traditional scheme, the communication overhead is higher than the proposed scheme load and increased but in random pattern because of the unforeseen number of retransmission for the measurements. In summary, the proposed scheme significantly reduces the probability of measurements' retransmission and consequently the communication overhead for each unit. In contrary, in the traditional case, the measurements are exposed to outsiders; any party can intercept/modify their values, which leads to unexpected number of measurements' retransmission and consequently increases the communication overhead.

7.5.2 Computation complexity

The limited-computation capabilities of the measurement units prevents them from performing complex cryptographic operations. So, our proposed scheme utilizes a lightweight version of McEliece cryptosystem [177, 178] in addition to implement low-cost hardware platform on the

measurement units to guarantee the security requirements without increasing the computation overhead on the units. Another feature for that lightweight version of McEliece cryptosystem is that its encryption operation requires less computation time than decryption process so that it saves more computation overhead on the measurement units, as they have to perform the encryption process only. For LS, it has more powerful processing unit so that the decryption process on LS does not provide much computation burden too. While, in traditional connection, the measurements are sent in plaintext and the units do not have to perform any cryptographic operations, then the computation overhead on units is tiny. However, they are exposed to various attacks that can block the measurements or insert bad data and force units to remeasure and retransmit the values. Although our proposed scheme increases the computation overhead a little bit, it guarantees the integrity and correctness of the measurements that decreases the possibility of receiving bad data and consequently reduces the probability of measurements' retransmission. In addition, the computation overhead of the proposed scheme is trivial and does not consider a big load on the limited-capability units.

Figure 7.3 shows the computation overhead for our proposed scheme versus traditional unsecured case. The computation overhead per round for every measurement unit is shown in Figure 7.3(a); the overhead for the proposed scheme is fixed tiny delay, i.e., a little bit more than **3** msec, while the computation complexity for the traditional unsecured connection is fluctuated between **1** and **5** msec according to the number of measurements' re-estimation and retransmission, which is effected by the quality of the communication channel and the presence of adversaries. Mostly the computation burden on the units in the unsecured case is less than the overhead in the proposed scheme case. Figure 7.3(b) presents the total computation overhead for the cluster; it can be seen that the total load for the proposed scheme case is linearly increased by the increase in the units' number in the cluster, whereas the unsecured case total overhead is oscillated according to the number of messages' retransmission, i.e., the overhead is varied because of the unpredictable number of messages' retransmission. As expected, the computation overhead for unsecured case is less than the proposed scheme's overhead, but the computation burden for the proposed scheme is still small and tolerable by the measurement units; that overhead considers a good price for more reliable and efficient connection and consequently more accurate and integral measurements.

7.5.3 State Estimator Performance Evaluation

The proposed scheme utilizes a cosine similarity-based SE to further improve the process of preventing FDI attacks. As indicated in the proposed scheme, the cosine similarity-based SEs are distributed at different LSs of the grid; they are used to determine the acceptance level of the received measurements. According to the result of the SE test, LS decides to accept the received set of measurements or reject and re-measure it again. In other words, that test works as final filter for the bad injected data.

In this section, we study the performance of the grid in its traditional case, the grid with cosine similarity-based SEs only, versus the proposed scheme case. The grid in its traditional case has no protection from FDI attacks at all, i.e., does not apply any attacks' protection or detection techniques. The measurement units are not protected; also, the measurements are transmitted in plaintext. So any adversary can compromise any number of units or intercept and falsify any number of transmitted measurements. These attacks have significant negative impact on the CC's decision and consequently on the grid's stability and efficiency.

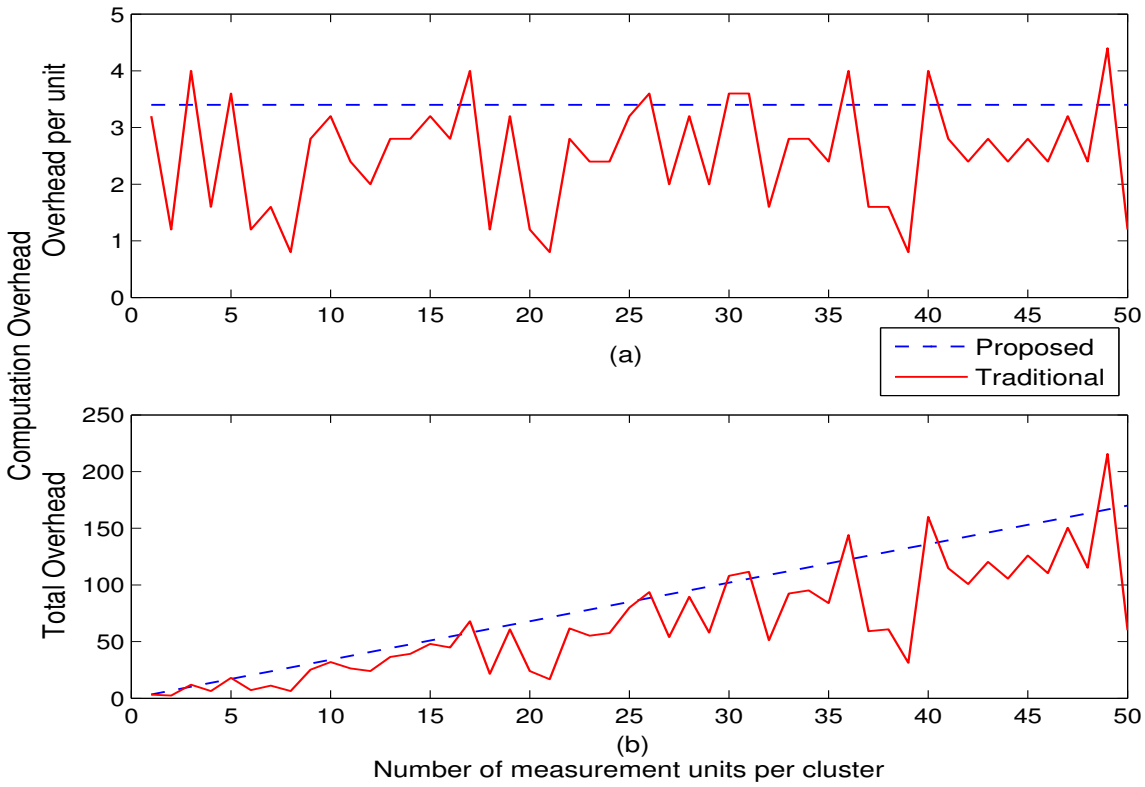

Figure 7.3: Computation overhead.

While the grid with applied cosine similarity-based SE only is in better condition than traditional case. Generally, the efficiency of detection process depends on the applied SE; the cosine similarity-based SE is more efficient than other utilized SE. Thus, the grid in that case is more protected than the traditional case or the regular applied SE [179]. However, the FDI attacks detection is not the most efficient solution for the grid especially for the communication and computation load, as if a FDI attack is detected, the CC should ask for new accurate measurements and that means more communication and computation operations, i.e., measurement units should re-estimate and re-send the value again. In addition, we cannot guarantee the accuracy for the new measurements, i.e., false alarm, as the corresponding units may be compromised.

Thus, our proposed scheme solves these concerns by preventing the FDI attacks from happening and guarantee the accuracy of the received measurements matrix utilizing several techniques. The proposed scheme protects the measurement units at their locations, also guards the measurements during their transmission to LS. LS then applies accuracy technique and cosine similarity-based SE to guarantee that is the most accurate set of measurements for the area before forwarding an encrypted version of that set to CC. At the same time, the proposed scheme conserves the communication and computation capabilities for different parties in the grid especially measurement units.

We assume that the number of FDI attacks that targets the grid is random but increased as number of measurement units in the grid increased. Figure 7.4 demonstrates the number of successful FDI attacks in the three cases. For traditional case, it has been seen that all the FDI attacks are successful, i.e., no filter at all. As shown, the curve is generally increased but fluctuated; the number of successful FDI attacks are changed, i.e., increased or decreased,

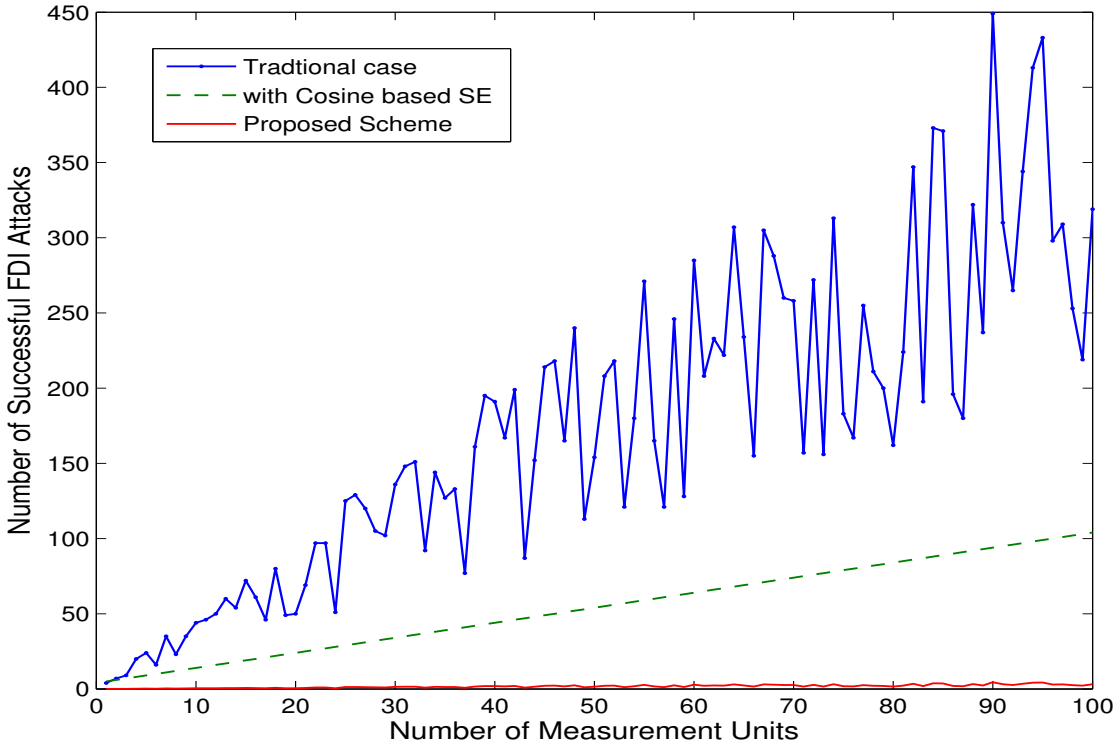

Figure 7.4: Probability of successful FDI attacks.

according to the existence of adversaries and the number of compromised units in the grid. According to the cosine similarity-based SE case, the utilized SE can detect a portion of FDI attacks. However, the number of successful FDI attacks is increased as number of units increased, because the probability of attack is increased, i.e, the number of compromised units increased. Moreover, when SE detects false measurements, it rejects these measurements and asks for re-measurement and re-transmission, which increases the total overhead. For the proposed scheme, the successful FDI attacks number is diminished, as shown, even if the number of attacks is increased. The reason is the different filtering levels that the measurements have to go through before being accepted by LS as an accurate/correct measurement.

7.6 Case Study

The proposed scheme divides the power grid's distribution and consumption subsystems into small areas. We represent each area as an IEEE-14 bus system. The local area consists of one LS, $\boldsymbol{LS_v}$, which collects the readings of the measurement units in the area and guarantees their accuracy/integrity. The total number of measurements that can guarantee the observability for IEEE-14 bus network equals **54** measurements [27]. However, $\boldsymbol{LS_v}$ connects to larger cluster of units that spreads all over the area, i.e., **54** groups of measurement units, as we utilize the measurements' redundancy to guarantee the integrity and availably of the measurements.

The cluster is divided into **54** groups; $\boldsymbol{G = \{G_1, G_2, \ldots, G_{54}\}}$, where the number of the groups connected to the LS is **54**. Each group contains a specific number of measurement units that calculates the same measurement value; $\boldsymbol{G_i = \{MU_1, MU_2, \ldots, MU_j\}}$, where $\boldsymbol{j}$ is

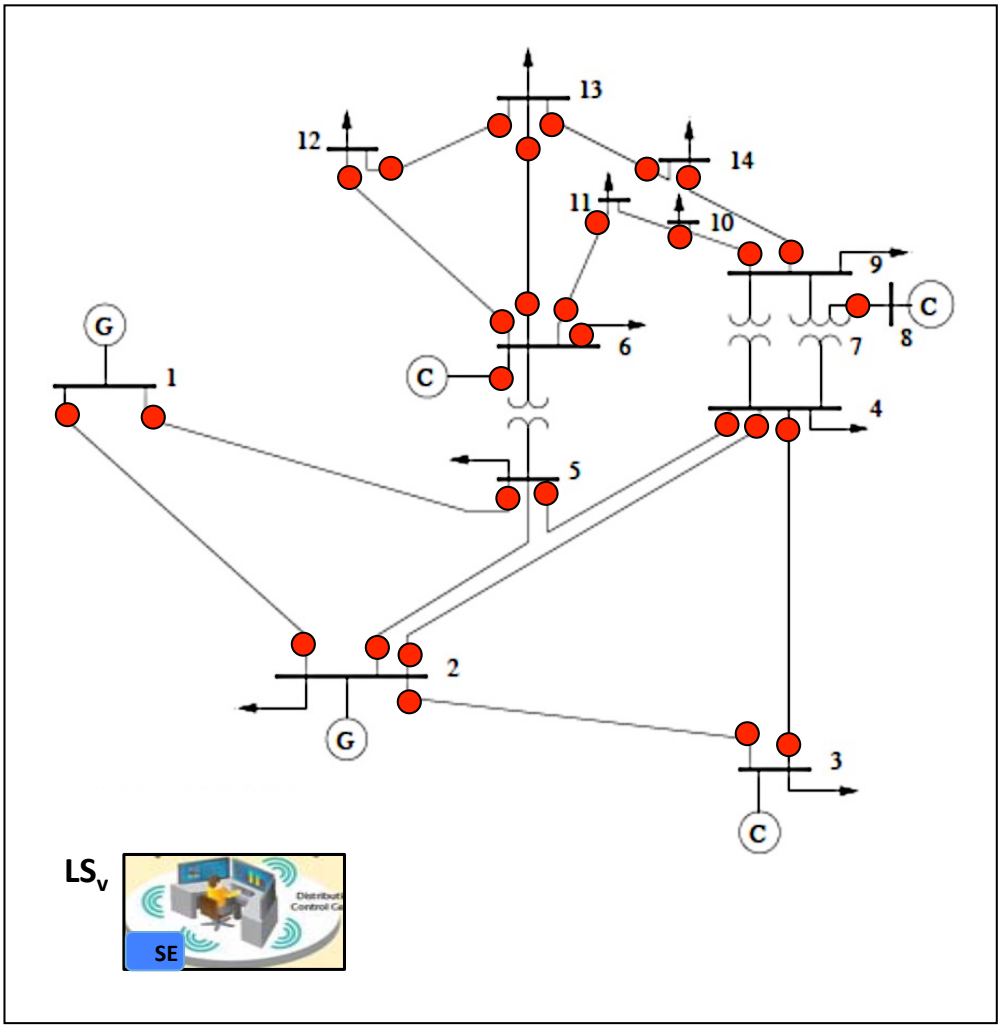

Figure 7.5: Case study.

the number of the measurement units in the group $\boldsymbol{G_i}$. The number of measurement units in each group $\boldsymbol{j}$ is different according to several parameters, such as the importance and nature of measurement, the location of units, and the surrounded environment (several related work propose techniques for the optimal redundancy for the measurements, such as [180]). Only LS knows the number and the locations of units in each group. Therefore, LS is the only party that can link between the measurement units and its task. Figure 7.5 demonstrates the architecture of one area in the grid, where (G) refers to generators, (C) for synchronous compensators, and ● for measurement units.

Each measurement unit $\boldsymbol{MU_y}$, e.g., from $\boldsymbol{G_w}$ that has $\boldsymbol{u}$ measurement units and is responsible for measuring the values of power flow parameters for line $\mathbf{1-2}$, estimates the current active power in its location, $\boldsymbol{P_{1-2(y)}}$, attaches its ID $\boldsymbol{MU_y}$, adds timestamp $\boldsymbol{T_{s5}}$ and nonce value $\boldsymbol{q_5}$, $\boldsymbol{ms_y = MS_y \| MU_y \| T_{s5} \| q_5}$ and then encrypts the result by LS's public key; $\boldsymbol{X_y = ms_y G_{ls_v} \oplus e_5}$. Then, $\boldsymbol{MUy}$ sends its encrypted periodic report $\boldsymbol{X_y}$ to $\boldsymbol{LS_v}$ via unsecured connection. $\boldsymbol{LS_v}$ checks the validity of the message before including the measurement to the pool of measurements from group $\boldsymbol{G_w}$. Then, $\boldsymbol{LS_v}$ filters out the deviated measurements and uses the remaining measurements, the majority, to get the final measurement for the current active power:

$$\boldsymbol{P_{1-2(avg)} = \frac{\sum_{y=1}^{ls} P_{1-2(y)}}{ls}}$$

$$\boldsymbol{\sigma = \sqrt{\frac{\sum_{y=1}^{ls} (P_{1-2(y)} - P_{1-2(avg)})^2}{ls}}}$$

$$\boldsymbol{\Delta P_{1-2(avg)} = \frac{\sigma}{\sqrt{ls}}}$$

$$\boldsymbol{P_{1-2} = P_{1-2(avg)} \pm \Delta P_{1-2(avg)}}$$

, where $\boldsymbol{ls \leq u}$ is the number of the accurate measurements in group $\boldsymbol{G_w}$, $\boldsymbol{\sigma}$ is the uncertainty in a single measurement value, $\boldsymbol{\Delta P_{1-2(avg)}}$ is the uncertainty in the mean $\boldsymbol{P_{1-2(avg)}}$, and $\boldsymbol{P_{1-2}}$ is the accurate average value of the collected measurements.

Next, $\boldsymbol{LS_v}$ forms the accurate observed measurements vector for that part of the grid $\mathscr{S}$ and runs the cosine similarity test to compare between the estimated matrix $\mathscr{S}$ and the expected one $\hat{\mathscr{S}}$ as follows:

$$\boldsymbol{\Lambda_v = \cos\theta_v = \frac{\vec{\mathscr{S}} \cdot \vec{\hat{\mathscr{S}}}}{\| \vec{\mathscr{S}} \| \| \vec{\hat{\mathscr{S}}} \|}}$$

, where $\vec{\mathscr{S}} \cdot \vec{\hat{\mathscr{S}}}$ represents the inner product of the vectors $\vec{\mathscr{S}}$ and $\vec{\hat{\mathscr{S}}}$, and $\| \vec{\mathscr{S}} \| \| \vec{\hat{\mathscr{S}}} \|$ denotes the product of their Euclidean lengths.

If $\boldsymbol{\Lambda_v = 1}$, $\boldsymbol{LS_v}$ forwards that fully-guaranteed accuracy measurements to CC. But, if $\boldsymbol{\Lambda_v}$ equals $\mathbf{0}$, $\boldsymbol{LS_v}$ rejects this set of measurements and check the corresponding units, i.e., this case almost cannot happen. While if $\boldsymbol{\Lambda_v}$ is between $\mathbf{0}$ and $\mathbf{1}$, $\boldsymbol{LS_v}$ determines the accuracy level of the collected measurements according to the exact value of $\boldsymbol{\Lambda_v}$. If $\boldsymbol{\Lambda_v}$ value is greater than $\mathbf{0.5}$, then $\boldsymbol{LS_v}$ can accept that set of measurements as partially-guaranteed accuracy measurements and forwards them to CC. And if $\boldsymbol{\Lambda_v}$ value is less than or equal $\mathbf{0.5}$, then $\boldsymbol{LS_v}$ rejects this set of measurements and check the measurement units to detect the malicious units.

Then, $\boldsymbol{LS_v}$ includes the matrix of all the accurate measurements for the area in a message $\boldsymbol{M_{Tv}}$. $\boldsymbol{LS_v}$ first attaches a timestamp $\boldsymbol{T_{sv}}$ and random nonce $\boldsymbol{q_v}$ to the message; $\boldsymbol{m_{Tv} = M_{Tv} \| T_{sv} \| q_v}$ and then encrypts it by CC's public key; $\boldsymbol{T_{ls_v} = m_{Tv} G_{cc} \oplus e_v}$. Then, $\boldsymbol{LS_v}$ forwards $\boldsymbol{T_{ls_v}}$ to CC.

7.7 Summary

FDI attack injects fake measurements in the grid that mislead the CC to make wrong decisions for the grid and consequently threaten smart grid's efficiency and reliability. Previous research works focus on detecting the attack mainly utilizing estimation detectors and optimization techniques. This work has presented a prevention technique to alleviate the impact of FDI attacks on the CC's decisions; the proposed scheme is based on guaranteeing the integrity and availability of the measurements at measurement units and during their transmission to CC even with the existence of compromised units; it also utilizes an efficient SE as another filtering

level for FDI attacks. In conclusion, our scheme guarantees the accuracy of the measurements received by CC, with lightweight overhead especially on the limited-computation capability measurement units.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

This PhD research works focus on analyzing security concerns for the three principal smart grid's communication architectures: customer-side networks, V2G connections, and power control and state estimation systems. We have proposed lightweight security and privacy-preserving schemes that guarantee the main security objectives for these architectures as follows:

- First, we have studied security and privacy threats for smart grid's customer-side networks, i.e., HANs, BANs/IANs, and NANs and proposed two different solutions to guarantee security and privacy requirements for these networks; at the same time, our proposed approaches are lightweight schemes so that they are appropriate for limited-capabilities devices in the network.

  The first proposed scheme is a lightweight lattice-based security and privacy preserving scheme. Our scheme is based on forecasting the future electricity demand for a cluster of customers in the same residential area; it limits the whole cluster's connection with electricity utility only when the cluster needs to adjust its total electricity share. The proposed scheme guarantees security and privacy demands, i.e., customers privacy, data integrity, and network resources and information availability, for customer-side networks. It is also a lightweight and efficient in terms of communication and computation complexities so that it is suitable for limited-capabilities devices, i.e., smart meters.

  The second solution is a lightweight lattice-based homomorphic privacy-preserving data aggregation scheme for the residential electricity consumers in smart grid. The proposed scheme (unlike the related works) allows the household smart appliances to aggregate their consumption among themselves without involving the related smart meter utilizing a lightweight lattice-based homomorphic crypto-system. Smart meter does not know the reading for each individual smart appliance; it receives the total encrypted aggregated consumption for all smart appliances in the HAN. So, smart meters, also the related base station, work as relay nodes and just forward HAN's total consumption to CC. However, smart meter, i.e., base station as well, has the ability to check the authenticity of messages' senders without revealing their contents. Consequently, the proposed scheme guarantees the security and privacy demands, i.e., customers' privacy, data confidentiality

and integrity, for the connection. It is also lightweight and efficient in terms of communication and computation complexities so that it is suitable for limited-capability devices, i.e., smart meters and smart household appliances in that case.

- In V2G connections, we have analyzed and determined their essential security concerns, and proposed a lightweight secure authentication and privacy-preserving V2G connection scheme. The scheme allows EVs to generate their own pseudonym identities and do not expose their private information to any party even the grid's operator; so, the EVs' privacy is preserved. In addition, the scheme forces EVs to follow a specific procedure, i.e., finishing their part in the connection first, to prevent them from acting maliciously so that the proposed scheme keeps the grid's financial profits. The scheme also assures the confidentiality and integrity of exchanged messages during (dis)charging sessions in addition to keep the accountability of the electricity-trade operation. Furthermore, the scheme achieves these security requirements with lightweight computation and communication overhead.

- According to power control and state estimation system, we have analyzed and studied the main problem in this communication environment: FDI attacks. Moreover, we have proposed FDI attack prevention scheme that protects integrity and availability of the measurements at measurement units and during their transmission to CC with lightweight communication and computation overhead; even if some units in the area are compromised. In addition, the scheme utilizes an accurate cosine similarity-based SE to filter out any error or noise on the measurements and to prevents the attacker from inserting bad data to the system, if the adversary manages to exceed the applied protection mechanism. The scheme provides that protection from then FDI attacks with lightweight communication and computation load. As a result, the proposed scheme bounds the damaging impacts of FDI attack on the power grid, simultaneously saves its resources with low cost.

In other words, we have satisfied the main security objectives of the smart grid's communication architectures. As our proposed schemes guarantee the privacy of different parties during exchanging messages; for instance, in customer-side networks, the personal habits for electricity consumers are concealed from all other parties. According to V2G connections, the service provider, EVs, preserves its private information so that on one can link between EV owner's identity or location to extract some knowledge about the owner's behaviour or life style. In addition, the schemes assure the different security requirements: confidentiality, integrity, availability, and authenticity, for each device in the connection. For example, authenticating EVs in V2G connections is satisfied, while in power control and state estimation, the proposed solution focuses more on guaranteeing the transmitted messages' integrity. Lastly, the proposed works provide these security needs for each party in smart grid exploiting lightweight security crypto-systems; this objective is significant because of the existence of restricted-capabilities devices in the different network architectures of smart grid. In customer-side connections, smart meters and smart appliances are limited-computation capabilities devices. For V2G connections, EVs are not prepared to perform complex crypto-operations. According to power control and state estimation system, the measurement units that monitor the grid's status are restricted-capabilities sensors.

8.2 Further Research Topics

There are certain open security problems in each communication network from the three main smart grid architectures that could be analyzed and studied in the future. In customer-side networks, the first proposed scheme predicts the expected electricity demand for a cluster of HANs, i.e., one BAN. The connection with electricity provider is done throughout the BAN's server only when the total cluster's demand needs to be adjusted. So, the impact of malicious servers on the performance of the customer-side networks is worth further studying. For the second proposed solution, the household smart appliances aggregate their consumption among themselves without involving the connected smart meter or the corresponding base station of the area. However, the effect of using EVs, i.e., the residential costumers' vehicles, as storage for the extra electricity on the performance of the proposed scheme needs more investigation.

According to V2G connection, the proposed scheme preserves the security requirements for the EV's connection with the grid during charging/discharging sessions. However, the optimal selection for the charged/discharged EVs during the electricity-trade process needs to be explored. In power control and state estimation system, we have proposed FDI attack prevention scheme that prevent the integrity and availability attacks from affecting on the accuracy of the final values of measurements. In the future, studying the optimal placement technique for phasor measurement units in the power grid for complete observability requires more research.

Bibliography

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid – the new and improved power grid: A survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944 – 980, 2012.

[2] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart grid," *Computer Networks*, vol. 55, no. 15, pp. 3604–3629, 2011.

[3] M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Communication network requirements for major smart grid applications in HAN, NAN and WAN," *Elsevier Editorial System for Computer Networks*, August 2013.

[4] Z. Fan, P. Kulkarni, C. E. S. Gormus, G. Kalogridis, M. Sooriyabandara, Z. Zhu, S. Lambotharan, and W. Chin, "Smart grid communications: Overview of research challenges, solutions, and standardization activities," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 21 – 38, First Quarter 2013.

[5] C. Gentile, D. Griffith, and M. Souryal, "Wireless Network Deployment in the Smart Grid: Design and Evaluation Issues," *IEEE Network*, pp. 48 – 53, November/December 2012.

[6] Y. Kim, J. Lee, G. Atkinson, H. Kim, and M. Thottan, "SeDAX: A Scalable, Resilient, and Secure Platform for Smart Grid Communications," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1119 – 1136, July 2012.

[7] *Smart Grid Consumer Benefits, IEEE Smart Grid*, [Online]. Available: http://smartgrid.ieee.org/questions-and-answers/964-smart-grid-consumer-benefits. [Accessed 4 September 2013]

[8] *Smart Grid Economic and Environmental Benefits, A Review and Synthesis of Research on Smart Grid Benefits and Costs*, [Online]. Available: http://smartgridcc.org/wp-content/uploads/2013/10/SGCC-Econ-and-Environ-Benefits-Full-Report.pdf. [Accessed 8 October 2013]

[9] V. Aravinthan, V. Namboodiri, S. Sunku, and W. Jewell, "Wireless AMI Application and Security for Controlled Home Area Networks," *in Proc. IEEE Power and Energy Society General Meeting*, USA, July 2011.

[10] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A Survey on Cyber Security for Smart Grid Communications," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 998 – 1010, Fourth Quarter 2012.

[11] H. Khurana, M. Hadley, N. Lu, and D. Frincke, "Smart-Grid Security Issues," *IEEE Computer and Reliability Societies*, pp. 81 – 85, January/February 2010.

[12] J. Liu, Y. Xiao, S. Li, W. Liang, and C. Chen, "Cyber Security and Privacy Issues in Smart Grids," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 981 – 997, Fourth Quarter 2012.

[13] G. Ericsson, "Cyber Security and Power System Communication — Essential Parts of a Smart Grid Infrastructure," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1501 – 1507, July 2010.

[14] Z. Erkin, J. Troncoso-Pastoriza, R. Lagendijk, and F. Perez-Gonzalez, "Privacy-preserving data aggregation in smart metering systems," *IEEE Signal Processing Magazine*, pp. 75 – 86, March 2013.

[15] E. Bou-Harb, C. Fachkha, M. Pourzandi, M. Debbabi, and C. Assi, "Communication Security for Smart Grid Distribution Networks," *IEEE Communications Magazine*, pp. 42 – 49, January 2013.

[16] P. Chen, S. Cheng, and K. Chen, "Smart attacks in smart grid communication networks," *IEEE Communications Magazine*, pp. 24 – 29, August 2012.

[17] J. Soares, T. Sousa, H. Morais, Z. Vale, and P. Faria, "An Optimal Scheduling Problem in Distribution Networks Considering V2G," *in Proc. IEEE CIASG*, France, April 2011.

[18] T. yiyun, L. Can, C. Lin, and L. Lin, "IMPACTS OF ELECTRIC VEHICLES ON POWER GRID," *in Proc. SUPERGEN*, China, September 2012.

[19] M. Galus, R. Waraich, F. Noembrini, K. Steurs, G. Georges, K. Boulouchos, K. Axhausen, and G. Andersson, "Integrating Power Systems, Transport Systems and Vehicle Technology for Electric Mobility Impact Assessment and Efficient Control," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 934 – 949, June 2012.

[20] Y. Fu, L. Wu, and Z. Li, "A Hierarchically Coordinated Operation Framework for Optimally Integrating PHEVs into Power Grids," *in Proc. IEEE VPPC*, USA, September 2011.

[21] S. Acha, T. Green, and N. Shah, "Techno-economical Tradeoffs from Embedded Technologies with Storage Capabilities on Electric and Gas Distribution Networks," *in Proc. IEEE Power and Energy Society General Meeting*, USA, July 2010.

[22] K. Mets, T. Verschueren, F. Turck, and C. Develder, "Exploiting V2G to optimize residential energy consumption with electrical vehicle (dis)charging," *in Proc. IEEE SGMS (IEEE SmartGridComm)*, Belgium, October 2011.

[23] J. Soares, H. Morais, T. Sousa, Z. Vale, and P. Faria, "Day-Ahead Resource Scheduling Including Demand Response for Electric Vehicles," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 596 – 605, March 2013.

[24] T. Soares, H. Morais, Z. Vale, P. Faria, and J. Sousa,"Intelligent Energy Resource Management Considering Vehicle-to-Grid: A Simulated Annealing Approach," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 535 – 542, March 2012.

[25] *U.S.-Canada Power System Outage Task Force. Final report on the August 14, 2003 blackout in the United States and Canada.*, [Online]. Available: https://reports.energy.gov/B-F-Web-Part1.pdf [Accessed 8 April 2013]

[26] Available: http://en.wikipedia.org/wiki/SCADA. [Accessed 8 May 2013]

[27] G. Dan, and H. Sandberg,"Stealth Attacks and Protection Schemes for State Estimators in Power Systems," *in Proc. IEEE SmartGridComm*, USA, October 2010.

[28] Y. Zhang, and J. Chen,"Wide-area SCADA system with distributed security framework," *Journal of Communications and Networks*, vol. 14, no. 6, pp. 597 – 605, December 2012.

[29] D. Wei, Y. Lu, M. Jafari, P. Skareand, and K. Rohde,"Protecting Smart Grid Automation Systems Against Cyberattacks," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 782 – 795, December 2011.

[30] M. Kim,"A Survey on Guaranteeing Availability in Smart Grid Communications," *in Proc. ICACT*, Korea, February 2012.

[31] A. Teixeira, G. Dan, H. Sandberg, and K. Johansson,"A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator," *in Proc. IFAC World Congress*, Italy, August 2011.

[32] Y. Huang, H. Li, A. Campbell, and Z. Han,"Defending False Data Injection Attack On Smart Grid Network Using Adaptive CUSUM Test," *in Proc. 45th Annual Conf. Info. Sciences and Sys*, USA, March 2011.

[33] R. Bobba, K. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. Overbye,"Detecting false data injection attacks on DC state estimation," *in Proc. CPSWEEK*, Sweden, April 2010.

[34] S. Cui, Z. Han, S. Kar, T. Kim, H. Poor, and A. Tajer,"Coordinated Data-Injection Attack and Detection in Smart Grid," *IEEE Signal Processing Magazine, special issue on Signal Processing for Smart Grid*, vol. 29, no. 5, pp. 106 – 115, September 2012.

[35] B. Gou, and A. Abur,"An Improved Measurement Placement Algorithm for Network Observability," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 819 – 824, November 2001.

[36] S. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, U.K, 2004.

[37] Y. Liu, P. Ning, and M. Reiter,"False Data Injection Attacks against State Estimation in Electric Power Grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, Article. 13, May 2011.

[38] M. Rahman, E. Al-Shaer, and M. Rahman,"A Formal Model for Verifying Stealthy Attacks on State Estimation in Power Grids," *in Proc. IEEE SmartGridComm*, Canada, October 2013.

[39] H. Trees, *Detection, Estimation, and Modulation Theory*. Wiley. Press, USA, 1971.

[40] Available: http://mathworld.wolfram.com/ChiDistribution.html [Accessed 8 May 2013]

[41] Y. Mo, T. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli,"Cyber-Physical Security of a Smart Grid Infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195 – 209, January 2012.

[42] S. Sridhar, A. Hahn, and M. Govindarasu,"Cyber-Physical System Security for the Electric Power Grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210 – 224, January 2012.

[43] R. Rivest, A. Shamir, L. Adleman,"A Method for Obtaining Digital Signatures and Public-Key crypto-systems," *Communications of the ACM*, vol. 21, no. 2, pp. 120 – 126, 1978.

[44] P. Paillier,"Public-Key crypto-systems Based on Composite Degree Residuosity Classes," *in Proc. EUROCRYPT*, Czech Republic, May 1999.

[45] O. Tan, D. Gunduz, and H. Poor,"Increasing Smart Meter Privacy Through Energy Harvesting and Storage Devices," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1331 – 1341, July 2013.

[46] Z. Chen, and L. Wu,"Residential Appliance DR Energy Management With Electric Privacy Protection by Online Stochastic Optimization," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1861 – 1869, December 2013.

[47] X. He, X. Zhang, and C. Kuo,"A Distortion-Based Approach to Privacy-Preserving Metering in Smart Grids," *IEEE Access: practical innovations: open solutions*, vol. 1, pp. 67 – 78, May 2013.

[48] L. Sankar, S. Rajagopalan, S. Mohajer, and H. Poor,"Smart Meter Privacy A Theoretical Framework," *IEEE Transactions on Smart Grid*, vol. 4, No. 2, pp. 837 – 846, June 2013.

[49] W. Jia, H. Zhu and C. Xiao, "Human-factor-aware privacy-preserving aggregation in smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 598 – 607, June 2014.

[50] G. Kalogridis, M. Sooriyabandara, Z. Fan, and M. Mustafa, "Toward Unified Security and Privacy Protection for Smart Meter Networks," *IEEE Systems Journal*, vol. 8, no. 2, pp. 641 – 654, June 2014.

[51] V. Namboodiri, V. Aravinthan, S. Mohapatra, B. Karimi, and W. Jewell, "Toward a Secure Wireless-Based Home Area Network for Metering in Smart Grids," *IEEE Systems Journal*, vol. 8, no. 2, pp. 509 – 520, August 2013.

[52] A. Metke, and R. Ekl, "Security Technology for Smart Grid Networks," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 99 – 107, June 2010.

[53] Z. Fadlullah, N. Kato, R. Lu, X. Shen, and Y. Nozaki, "Toward secure targeted broadcast in smart grid," *IEEE Communications Magazine*, pp. 150 – 156, May 2012.

[54] D. He, C. Chen, J. Bu, S. Chan, Y. Zhang, and M. Guizani, "Secure service provision in smart grid communications," *IEEE Communications Magazine*, pp. 53 – 61, August 2012.

[55] V. Goyal, A. Sahai, O. Pandey, and B. Waters, "Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data," *in Proc. ACM CCS*, USA, October/November 2006.

[56] J. Liu, Y. Xiao, and J. Gao, "Achieving accountability in smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 493 – 508, June 2014.

[57] S. Ruj and A. Nayak, "A decentralized security framework for data aggregation and access control in smart grids," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 196 – 205, March 2013.

[58] Y. Kim and J. Heo, "Device authentication protocol for smart grid systems using homomorphic hash," *Journal of Communications and Networks*, vol. 14, no. 6, pp. 606 – 613, December 2012.

[59] H. Li, X. Lin and X. Shen, "EPPDR: An efficient privacy-preserving demand response scheme with adaptive key evolution in smart grid," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 8, pp. 2053 – 2064, August 2014.

[60] R. Lu, X. Liang and X. Shen, "EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1621 – 1631, September 2012.

[61] X. Li, X. Liang, X. Shen and H. Zhu, "Securing smart grid cyber attacks, countermeasures, and challenges," *IEEE Communications Magazine*, pp. 38 – 45, August 2012.

[62] S. Das,Y. Ohba, M. Kanda, D. Famolari, and S. Das, "A Key Management Framework for AMI Networks in Smart Grid," *IEEE Communications Magazine*, pp. 30 – 37, August 2012.

[63] N. Liu, J. Chen, L. Zhu, J. Zhang, and Y. He, "A Key Management Scheme for Secure Communications of Advanced Metering Infrastructure in Smart Grid," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp. 4746 – 4756, October 2013.

[64] M. Fouda, Z. Fadlullah, N. Kato, R. Lu and X. Shen, "A lightweight message authentication scheme for smart grid communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 675 – 685, December 2011.

[65] H. Li, R. Lu, L. Zhou, B. Yang, and X. Shen, "An efficient merkle tree-based authentication scheme for smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 655 – 663, June 2014

[66] Y. Yan, R. Hu, S. Das, H. Sharif and Y. Qian, "An Efficient Security Protocol for Advanced Metering Infrastructure in Smart Grid," *IEEE Network*, pp. 64 – 71, July/August 2013

[67] H. Nicanfar, P. Jokar and V. Leung, "Efficient authentication and key management mechanisms for smart grid communications," *IEEE Systems Journal*, vol. 8, no. 2, pp. 629 – 640, June 2014.

[68] C. Fan and Y. Lai, "Privacy enhanced data aggregation scheme against internal attackers in smart grid," *IEEE Transactions on Industrial informatics*, vol. 10, no. 1, pp. 666 – 675, February 2014.

[69] C. Rottondi and C. Krauß, "Distributed privacy-preserving aggregation of metering data in smart grids," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1342 – 1354, July 2013.

[70] T. Chim, S. Yiu and V. Li, "Privacy-preserving advance power reservation," *IEEE Communications Magazine*, pp. 18 – 23, August 2012.

[71] M. Mustafa, N. Zhang, G. Kalogridis, and Z. Fan, "Smart Electric Vehicle Charging: Security Analysis," *in Proc. IEEE PES ISGT*, USA, February 2013.

[72] H. Guo, Y. Wu, H. Chen, and M. Ma, "A Batch Authentication Protocol for V2G Communications," *in Proc. NTMS*, France, February 2011.

[73] H. Guo, Y. Wu, H. Chen, F. Bao, and M. Ma, "UBAPV2G: A Unique Batch Authentication Protocol for Vehicle-to-Grid Communications," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 707 – 714, December 2011.

[74] H. Tseng, "On the Security of a Unique Batch Authentication Protocol for Vehicle-to-Grid Communications," *in Proc. ITST*, Taiwan, November 2012.

[75] H. Liu, H. Ning, Y. Zhang, and L. Yang, "Aggregated-Proofs Based Privacy-Preserving Authentication for V2G Networks in the Smart Grid," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1722 – 1733, December 2012.

[76] H. Liu, H. Ning and M. Guizani, "Battery status-aware authentication scheme for V2G networks in smart grid," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 99 – 110, March 2013.

[77] Z. Yang, S. Yu and C. Liu, "P2: Privacy-preserving communication and precise reward architecture for V2G networks in smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 697 – 706, December 2011.

[78] H. Tseng, "A Secure and Privacy-Preserving Communication Protocol for V2G Networks," *in Proc. IEEE WCNC*, France, April 2012.

[79] M. Stegelmann and D. Kesdogan, "Design and evaluation of a privacy-preserving architecture for vehicle-to-grid interaction," *Lecture Notes in Computer Science, Public Key Infrastructures, Services and Applications, Springer Berlin Heidelberg*, vol. 7163, pp. 75 – 90, 2012.

[80] B. Vaidya, D. Makrakis, and H. Mouftah, "Security Mechanism for Multi-domain Vehicle-to- Grid Infrastructure," *in Proc. IEEE GLOBECOM*, USA, December 2011.

[81] D. Ghosh and S. Wicker, "A privacy-aware design for the vehicle-to-grid framework," *in Proc. HICSS*, USA, January 2013.

[82] T. Holden and J. Yazdani, "Hybrid security for hybrid vehicles exploring smart grid technology, power-line and wireless communication," *in Proc. IEEE PES ISGT Europe*, UK, December 2011.

[83] H. Su, M. Qiu, and H. Wang, "Secure wireless communication system for smart grid with rechargeable electric vehicle," *IEEE Communications Magazine*, vol. 50, no. 8, pp. 62 – 68, August 2012.

[84] Y. Li, R. Wang and Z. Han, "Resilient PHEV charging policies under price information attacks," *in Proc. IEEE SmartGridComm*, Taiwan, November 2012.

[85] M. Stegelmann and D. Kesdogan, "V2GPriv: Vehicle-to-grid privacy in the smart grid," *in Proc. CSS*, Australia, December 2012.

[86] M. Au, J. Liu and J. Zhou, "A new payment system for enhancing location privacy of electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 3 – 18, January 2014.

[87] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures," *in Proc. IEEE SmartGridComm*, USA, October 2010.

[88] O. Kosut, L. Jia, R. Thomas, and L. Tong, "On malicious data attacks on power system state estimation," *in Proc. UPEC*, UK, August 2010.

[89] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious Data Attacks on the Smart Grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645 – 658, December 2011.

[90] Y. Huang, M. Esmalifalak, H. Nguyen, R. Zheng, Z. Han, H. Li, and L. Song, "Bad Data Injection in Smart Grid: Attack and Defense Mechanisms," *IEEE Communication Magazine*, pp. 27 – 33, January 2013.

[91] M. Rabbat, M. Coates, and R. Nowak, "Multiple-Source Internet tomography," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 12, pp. 2221 – 2234, 2006.

[92] S. Wilks, "The Large-Sample distribution of the likelihood ratio for testing composite hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. 60 – 62, March 1938.

[93] A. Bera, and Y. Bilias, "Rao's score, Neyman's $\boldsymbol{C(\alpha)}$ and Silvey's LM tests: an essay on historical developments and some new results," *Journal of Statistical Planning and Inference*, vol. 97, no. 1, pp. 9 – 44, August 2001.

[94] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35 – 45, 1960.

[95] S. Zheng, T. Jiang, and J. Baras, "Robust state estimation under false data injection in distributed sensor networks," *in Proc. IEEE GLOBECOM*, USA, December 2010.

[96] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "False data injection attacks against state estimation in wireless sensor networks," *in Proc. IEEE Conference on Decision and Control*, USA, December 2010.

[97] D. Choi and L. Xie,"Ramp-Induced Data Attacks on Look-Ahead Dispatch in Real-Time Power Markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1235 – 1243, September 2013.

[98] E. Page,"Continuous Inspection Scheme," *Biometrika*, vol. 41, no. 1/2, pp. 100 – 115, June 1954.

[99] M. Ozay, I. Esnaola, F. Vural, S. Kulkarni, and H. Poor,"Sparse Attack Construction and State Estimation in the Smart Grid: Centralized and Distributed Models," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1306 – 1318, July 2013.

[100] A. Tajer, S. Kar, H. Poor, and S. Cui,"Distributed joint cyber attack detection and state recovery in smart grids," *in Proc. IEEE SmartGridComm*, Belgium, October 2011.

[101] F. Pasqualetti, R. Carli, and F. Bullo,"A Distributed Method for State Estimation and False Data Detection in Power Networks," *in Proc. IEEE SmartGridComm*, Belgium, October 2011.

[102] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao,"On false data injection attacks against distributed energy routing in smart grid," *in. Proc. of IEEE/ACM ICCPS*, China, April 2012.

[103] V. Kekatos and G. Giannakis,"Distributed Robust Power System State Estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617 – 1626, May 2013.

[104] Y. Gu, T. Liu, D. Wang, X. Guan, and Z. Xu,"Bad Data Detection Method for Smart Grids based on Distributed State Estimation," *in Proc. IEEE ICC*, Hungary, June 2013.

[105] M. Ozay, I. Esnaola, F. Vural, S. Kulkarni, and H. Poor,"Distributed models for sparse attack construction and state vector estimation in the smart grid," *in Proc. IEEE SmartGridComm*, Taiwan, November 2012.

[106] M. Talebi, C. Li, and Z. Qu,"Enhanced protection against false data injection by dynamically changing information structure of micro-grids," *in Proc. Sensor Array and Multichannel Signal Processing Workshop*, USA, June 2012.

[107] I. Matei, J. Baras, and V. Srinivasan,"Trust-based Multi-Agent Filtering for Increased Smart Grid Security," *in Proc. Mediterranean Conference on Control and Automation*, Spain, July 2012.

[108] D.!Levitin, and G. Russell,"Rao's spacing test," *Encyclopedia of Statistical Sciences*, vol. 3, pp. 87 – 89, 1999.

[109] T. Kim and H. Poor,"Strategic Protection Against Data Injection Attacks on Power Grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 326 – 333, June 2011.

[110] A. Giani, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla,"Smart grid data integrity attacks: Characterizations and countermeasures," *in Proc. IEEE SmartGridComm*, Belgium, October 2011.

[111] M. Esmalifalak, H. Nguyen, R. Zheng, and Z. Han,"Stealth False Data Injection using Independent Component Analysis in Smart Grid," *in Proc. IEEE SmartGridComm*, Belgium, October 2011.

[112] Q. Yang, J. Yang, W. Yu, N. Zhang, and W. Zhao,"On a Hierarchical False Data Injection Attack on Power System State Estimation," *in Proc. IEEE GLOBECOM*, USA, December 2011.

[113] S. Bi, and Y. Zhang,"Defending Mechanisms Against False-data Injection Attacks in the Power System State Estimation," *in Proc. IEEE International Workshop on Smart Grid Communications and Networks*, 2011.

[114] L. Liu, M. Esmalifalak, and Z. Han,"Detection of False Data Injection in Power Grid Exploiting Low Rank and Sparsity," *in Proc. IEEE ICC*, Hungary, June 2013.

[115] J. Kim and L. Tong,"On Phasor Measurement Unit Placement against State and Topology Attacks," *in Proc. IEEE SmartGridComm*, Canada, 2013.

[116] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao,"On False Data Injection Attacks against Power System State Estimation: Modelling and Countermeasures," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 717 – 729, March 2014

[117] H. Sedghi, and E. Jonckheere,"Statistical Structure Learning of Smart Grid for Detection of False Data Injection," *in Proc. IEEE Power and Energy Society General Meeting (PES)*, Canada, July 2013.

[118] Z. Qin, Q. Li, and M. Chuah,"Unidentifiable Attacks in Electric Power Systems," *in Proc. IEEE/ACM Third International Conference on Cyber-Physical Systems*, China, April 2012.

[119] Y. Law, T. Alpcan, and M. Palaniswami,"Security Games for Voltage Control in Smart Grid," *in Proc. Fiftieth Annual Allerton Conference*, USA October 2012.

[120] M. Ozay, I. Esnaola, F. Vural, S. Kulkarni, and H. Poor,"Smarter Security in the Smart Grid," *in Proc. IEEE SmartGridComm*, Taiwan, November 2012.

[121] H. Sandberg, A. Teixeira, and K. Johansson,"On Security Indices for State Estimators in Power Networks," *in Proc. CPSWEEK*, Sweden, April 2010.

[122] M. Esmalifalak, G. Shi, Z. Han, and L. Song,"Bad Data Injection Attack and Defense in Electricity Market Using Game Theory Study," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 160 – 169, March 2013.

[123] A. Teixeira, S. Amin, H. Sandberg, K. Johansson, and S. Sastry,"Cyber-security analysis of state estimators in electric power systems," *in Proc. IEEE Conference on Decision and Control*, USA, December 2010.

[124] Y. Kim, E. Ngai, and M. Srivastava,"Cooperative State Estimation for Preserving Privacy of User Behaviours in Smart Grid," *in Proc. IEEE SmartGridComm*, Belgium, October 2011.

[125] K. Davis, K. Morrow, R. Bobba, and E. Heine,"Power Flow Cyber Attacks and Perturbation-Based Defence," *in Proc. IEEE SmartGridComm*, Taiwan, November 2012.

[126] S. Zonouz, K. Rogers, R. Berthier, R. Bobba, W. Sanders, and T. Overbye,"SCPSE: Security-Oriented Cyber-Physical State Estimation for Power Grid Critical Infrastructures," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1790 – 1799, December 2012.

[127] S. Bi, and Y. Zhang,"Defending Mechanisms Against False-data Injection Attacks in the Power System State Estimation," *in Proc. IEEE GLOBECOM*, USA, December 2011.

[128] X. Yang, J. Lin, W. Yu, P. Moulema, X. Fu, and W. Zhao,"A Novel En-route Filtering Scheme against False Data Injection Attacks in Cyber-Physical Networked Systems," *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 4 – 18, January 2015.

[129] X. Yang, J. Lin, W. Yu, P. Moulema, X. Fu, and W. Zhao,"A Novel En-route Filtering Scheme against False Data Injection Attacks in Cyber-Physical Networked Systems," *in Proc. IEEE International Conference on Distributed Computing Systems*, China, June 2012.

[130] L. Yang, and F. Li,"Detecting False Data Injection in Smart Grid In-Network Aggregation," *in Proc. IEEE SmartGridComm*, Canada, October 2013.

[131] A. Euodial, M. Joyce Beryl Princess,"EFBV: En-Route Filtering Based Batch Verification Scheme for False Data Injection Attack in Wireless Sensor Networks," *in. Proc. ICGHPC*, India, March 2013.

[132] M. Abdullah, I. Welch, and W. Seah,"Efficient and Secure Data Aggregation for Smart Metering Networks," *in. Proc. IEEE ISSNIP*, Australia, April 2013.

[133] A. Boonsongsrikul, K. Lhee, and M. Hong,"Securing data aggregation against false data injection in wireless sensor networks," *in. Proc. ICACT*, South Korea, February 2010.

[134] A. Mohsenian-Rad, and A. Leon-Garcia,"Distributed Internet-Based Load Altering Attacks Against Smart Power Grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 667 – 674, December 2011.

[135] G. Hug, and J. Giampapa,"Vulnerability Assessment of AC State Estimation With Respect to False Data Injection Cyber-Attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362 – 1370, September 2012.

[136] M. Rahman, and H. Mohsenian-Rad,"False Data Injection Attacks Against Nonlinear State Estimation in Smart Power Grids," *in. Proc. IEEE Power and Energy Society General Meeting (PES)*, Canada, July 2013.

[137] M. Rahman, and H. Mohsenian-Rad,"False Data Injection Attacks with Incomplete Information Against Smart Power Grids," *in. Proc. IEEE GLOBECOM*, USA, December 2012.

[138] L. Xie, Y. Mo, and B. Sinopoli,"False Data Injection Attacks in Electricity Markets," *in. Proc. IEEE SmartGridComm*, USA, October 2010.

[139] Y. Yuan, Z. Li, and K. Ren,"Modelling Load Redistribution Attacks in Power Systems," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 382 – 390, June 2011.

[140] L. Xie, Y. Mo, and B. Sinopoli,"Integrity Data Attacks in Power Market Operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659 – 666, December 2011.

[141] Y. Yuan, Z. Li, and K. Ren,"Quantitative Analysis of Load Redistribution Attacks in Power Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731 – 1738, September 2012.

[142] D. Choi and L. Xie,"Ramp-Induced Data Attacks on Look-Ahead Dispatch in Real-Time Power Markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1235 – 1243, September 2013.

[143] K. Ross, K. Hopkinson, and M. Pachter,"Using a Distributed Agent-Based Communication Enabled Special Protection System to Enhance Smart Grid Security," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 1216 – 1224, June 2013.

[144] A. Abdallah, and X. Shen, "A Lightweight Lattice-based Security and Privacy-Preserving Scheme for Smart Grid," *in Proc. IEEE GLOBECOM*, USA, December 2014.

[145] A. Abdallah, and X. Shen, "Lightweight Security and Privacy Preserving Scheme for Smart Grid Customer-side Networks," *IEEE Transaction on Smart Grid*, to appear.

[146] J. Hoffstein, D. Lieman, J. Pipher, and J. Silverman,"NTRU: A PUBLIC KEY cryptosystem," *in Proc. EUROCRYPT*, Austria, May 2001.

[147] A. Nitaj, "Cryptanalysis of NTRU with two public keys," *International Journal of Network Security*, vol. 16, no. 2, pp. 112 – 117, March 2014.

[148] D. Stehlé and R. Steinfeld, "Making NTRU as secure as worst-case problems over ideal lattices," in *EUROCRYPT*, Estonia, May 2011.

[149] J. Hoffstein, N. Howgrave-Graham, J. Pipher, J. Silverman, and W. Whyte,"Ntrusign: Digital signatures using the NTRU lattice," in *CT-RSA*, USA, April2003.

[150] C. Wang and H. Wang, "A new ring signature scheme from NTRU lattice," in *ICCIS*, China, August 2012.

[151] D. Micciancio, "Shortest vector problem," *Encyclopedia of Cryptography and Security*, pp. 1196 – 1197, August 2011.

[152] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *Journal of the ACM*, vol. 56, no. 6, September 2009.

[153] D. Micciancio, and O. Regev, "Lattice-based Cryptography," *Post Quantum Cryptography*, pp. 147 – 191, February 2009.

[154] *Ontario Hydro*, [Online]. Available: http://www.ontario-hydro.com/index.php?page=current_rates. [Accessed 28 July 2015]

[155] Y. Yan, Y. Qian, and H. Sharif, "A secure data aggregation and dispatch scheme for home area networks in smart grid," *in Proc. IEEE GLOBECOM*, USA, December 2011.

[156] A. Abdallah, and X. Shen, "Lightweight Lattice-based Homomorphic Privacy-Preserving Aggregation Scheme for Home Area Networks," *in Proc. WCSP*, China, October 2014.

[157] A. Abdallah, and X. Shen, "A Lightweight Lattice-based Homomorphic Privacy-Preserving Data Aggregation Scheme for Smart Grid," *IEEE Transaction on Smart Grid*, to appear.

[158] C. Melchor G. Castagnos, and P. Gaborit, "Lattice-based homomorphic encryption of vector spaces," *in Proc. IEEE ISIT*, Canada, July 2008.

[159] C. Melchor, and P. Gaborit, "A lattice-based computationally-efficient private information retrieval protocol," *in Proc. WEWoRC*, Germany, July 2007.

[160] *Smart Grid Smart City*, [Online]. Available: https://data.gov.au/dataset/smart-grid-smart-city-customer-trial-data/resource/63d2b1cd-f453-4440-8bb7-ed083326f5ae'inner_span=True. [Accessed 20 January 2016]

[161] *Rasberry PI*, [Online]. Available: https://www.raspberrypi.org [Accessed 20 January 2016]

[162] M. Li, S. Salinas and P. Li, "LocaWard: A Security and Privacy Aware Location-Based Rewarding System," *IEEE Transaction On Parallel And Distributed Systems*, Vol. 25, No. 2, February 2014.

[163] W. Luo and U. Hengartner, "Veriplace: a privacy-aware location proof architecture," *In Proc. ACM SIGSPATIAL GIS*, USA, November 2010.

[164] C. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati, "An obfuscation-based approach for protecting location privacy," *IEEE Transactions on Dependable and Secure Computing*, Vol. 8, No. 1, pp. 13 – 27, January/February 2011.

[165] A. Abdallah, and X. Shen, "Lightweight Security and Privacy-Preserving V2G Connection Scheme in Smart Grid," *in Proc. IEEE GLOBECOM*, USA, December 2015.

[166] A. Abdallah, and X. Shen, "Lightweight Authentication and Privacy-Preserving Scheme for V2G Connections," *IEEE Transaction on Vehicular Technology*, to appear.

[167] M. Saarinen, "The BlueJay ultra-lightweight hybrid crypto-system," *in Proc. IEEE Symposium on Security and Privacy Workshops*, USA, May 2012.

[168] M. Saarinen, "The PASSERINE public key encryption and authentication mechanism," *Lecture Notes in Computer Science, Springer*, vol. 7127, pp. 283 – 288, 2011.

[169] D. Engels, M. Saarinen and E. Smith, "The Hummingbird-2 lightweight authenticated encryption algorithm," *Lecture Notes in Computer Science, Springer*, vol. 7055, pp. 19 – 31, 2011.

[170] A. Menezes and S. Wanstone, *Handbook of Applied Cryptography.* CRC Press, 1996.

[171] H. Martin, E. Millan and J. Castro, "AKARI-x: a pseudorandom number generator for secure lightweight systems," *In Proc. IEEE IOLTS*, Greece, July 2011.

[172] B. Collard and F. Standaert, "A statistical saturation attack against the block cipher PRESENT," *In Proc. CT-RSA*, USA, April 2009. *Lecture Notes in Computer Science, Springer*, vol. 5473, pp. 195 – 210, 2009.

[173] A. Abdallah, and X. Shen, "Efficient Prevention Technique for False Data Injection Attack in Smart Grid," *in Proc. IEEE ICC'*, Malaysia, May 2016.

[174] R. McEliece,"A Public-Key Cryptosystem Based On Algebraic Coding Theory," *DSN Progress Report. 42-44*, pp. 114 – 116, January/February 1978.

[175] A. Menezes and S. Wanstone, *Handbook of Applied Cryptography.* CRC Press, 1996.

[176] G. Adomavicius, and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the- art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734 – 749, June 2005.

[177] I. Maurich, T. Guneysu, "Lightweight code-based cryptography: QC-MDPC McEliece encryption on reconfigurable devices," *in Proc. IEEE DATE*, Germany, March 2014.

[178] I. Maurich, T. Oder, and T. Guneysu, "Implementing QC-MDPC McEliece Encryption," *ACM Transaction of Embedded Computing Systems*, vol. 14, no. 3, article. 44, pp. 1 – 27, April 2015.

[179] D. Rawat, and C. Bajracharya, "Detection of False Data Injection Attacks in Smart Grid Communication Systems," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1652 – 1656, October 2015.

[180] J. London, L. Alberto, and N. Bretas, "Network observability: identification of the measurements redundancy level," *in Proc. PowerCon2000*, Australia, December 2000.