Security and Privacy in Smart Grid Advanced Metering Infrastructure Network

PhD Thesis

Diovu, Remigius Chidiebere
216073494

A thesis submitted in fulfilment of the requirement for the degree of

DOCTOR OF PHILOSOPHY IN ENGINEERING (ELECTRICAL)

School of Engineering
University of KwaZulu-Natal,
Durban, South Africa

October, 2018

Security and Privacy in Smart Grid Advanced Metering Infrastructure Network

Diovu, Remigius Chidiebere

Supervisor:Prof. J.T Agee

A thesis submitted in fulfilment of the requirement for the degree of

DOCTOR OF PHILOSOPHY IN ENGINEERING (ELECTRICAL)

School of Engineering

University of KwaZulu-Natal

South Africa

EXAMINER'S COPY

October, 2018

As the candidate's supervisor, I have approved this thesis for submission.

Signed.....Date.....

Name: Prof. J.T Agee

Declaration 1 - Plagiarism

I, **Diovu, Remigius Chidiebere**, declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   (a) Their words have been re-written but the general information attributed to them has been referenced.
   (b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed.....Date.....

Declaration 2 - Publication

The following papers emanating from this research work have either been published or accepted for publication (in press):

UKZN DOHET Accredited Journals and Conferences

1. **R.C Diovu and J.T Agee** "An Overview of Data Aggregation Schemes in a Smart Grid AMI Network," *Journal of Communications, ISSN: 1796-2021*, (Published).
2. R.C Diovu and J.T Agee "ZigBee-Based Smart Grid Advanced Metering Infrastructure: Overview of Security Issues," *Sensor Letters Journal* , (Accepted and in press).
3. R.C Diovu and J.T Agee "Smart Grid Advanced Metering Infrastructure: Overview of Cloud-Based Security Solutions," *International Journal on Communications Antenna and Propagation (IRECAP)*, vol. 8, No. 4, 2018 (Published).
4. R.C Diovu and J.T Agee "Data Aggregation in Smart Grid AMI Network for Secure Transfer of Energy Consumption Data," *International Journal of Engineering Research in Africa (JERA)*, vol. 35, pp. 108-124, March, 2018 (published).
5. R.C Diovu and J.T Agee "Congestion Minimizing Scheme for Enhanced Data Aggregation in ZigBee-Based Smart Grid AMI Network," *International Journal on Communications Antenna and Propagation (IRECAP)*, vol. 8. No. 4, 2018 (Published).
6. R.C Diovu and J.T Agee "A Cloud-Based Openflow Firewall for Mitigation against DDoS Attacks in Smart Grid AMI Networks," *IEEE 2017 Power Africa Conference, Power and Energy Society (PES)*, pp. 28-33, June 2017 (published in IEEE Xplore).
7. R.C Diovu and J.T Agee “Enhancing the Security of a Cloud-Based Smart Grid AMI Network by Leveraging on the Features of Quantum Key Distribution,” *Transactions on Emerging Telecommunications Technologies*, vol. 30, Issue 3, 2019 (Published).
8. R.C Diovu and J.T Agee "Quantitative Analysis of Firewall Security under DDoS Attacks in

Smart Grid AMI Networks," *2017 IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON)*, pp. 697-701, Nov. 2017 (published in IEEE Xplore).

9. R.C Diovu and J.T Agee "Mitigating the Effects of Data Availability Attacks in a Smart Grid AMI Network by using Effective Network Planning Approach," *26th Southern African Universities Power Engineering Conference*, University of Witwatersrand Johannesburg, South Africa, 24-26 January, 2018.

The candidate is the main and corresponding author respectively for all the publications.

Dedication

This thesis is dedicated to the Holy Spirit, my counsellor and teacher per excellence, and Our Lady of Perpetual Help for all the favours and success received through her motherly intercessions during my period of studies.

Acknowledgements

First and foremost, I wish to acknowledge the Almighty God who gave me the opportunity to undertake this study and granted me sound health throughout my period of study. My sincere gratitude also goes to my supervisor, Prof. John T. Agee for his supervision, advice, and directions.

I am truly indebted to my lovely wife, Chinenye for her encouragement, support and prayers. My children, Chiagoziem, Chizaram, Chiduto, Chimela, and Chidiogo are highly acknowledged for being calm in obvious times of depleted parental care and attention. I am not so sure whether I could ever make up for this care and attention undeservedly taken away from you all.

My deepest appreciation and gratitude goes to my late parents for their passion and love for education, to my brothers, sisters, bro-in-law, sisters-in-law for their prayers, support, and care. Your love and encouragement have been and will always be a great source of inspiration in my life. I am particularly thrilled by my mother-in-law's support, care and prayers for me throughout this period of study. Your commitment towards my wife and children in my absence is second to none.

I earnestly appreciate the support and encouragement received from the present director, Prof. David Dorel, and the immediate past director, Prof, Inno Davidson at the Smart Grid Centre of Excellence in HVDC Engineering, UKZN, Westville campus. As a result, all financial support received from the Eskom Power Plant Engineering Institute (EPPIE) through our center is highly acknowledged. To my lovely friends and colleagues at the smart grid center, Grace, Patrick, Leonard, Frank, Emmanuel and others, the good and difficult times we shared together will remain memorable in life.

I cannot forget in a haste the support and prayers from Mr & Mrs. Ferdinand Okechukwu, Mr & Mrs. Augustine Odo, Mrs. Okwubanego Fransisca, Dr. Andrew-Mary Eloka-Eboka, Mr & Mrs Ebhota Williams, Mr & Mrs. Ali Marcel and my lovely friend and brother in the Lord, Mr. Stanley Onwunje. I also wish to express my heartfelt appreciation for all members of the Legion of Mary at St. Joseph Parish Emene, members of the Servants of Light Prayer Group, Catholic Charismatic Renewal of Nigeria, St Joseph Catholic Parish Emene, Enugu; members of Christ the King Catholic Charismatic

Renewal Alumni UNIZIK, Enugu Chapter, the immediate past and the present Diocesan Service Team of the Catholic Charismatic Renewal of Nigeria, Enugu Diocese. Your prayers towards the success of this study are well appreciated.

Abstract

The modernization of the electric power grid via the incorporation of smart systems and devices driven by information and communication technology (ICT), has been conceived as a major feat towards the realization of the smart grid vision. The smart grid advanced metering infrastructure (SG AMI) is an integrated network comprising smart meters, communication networks, and other numerous intelligent electronic devices. The SG AMI provides multi-way communication between energy consumers, grid operators, energy suppliers, and other authorized third-party operators, who may be responsible for billing, metering data management, and other grid control operations. The SG AMI provides new functions, which were either impossible previously or had to be done manually. For instance, energy consumption reading can be taken remotely, hence, making the task of visiting consumer's apartments for the same purpose unnecessary.

Consequently, the deployment of the AMI will greatly improve the reliability of the grid and reduce the costs of power delivery. However, it will expose the grid to cyber threats, vulnerabilities, and attacks that may have severe consequences for the smart grid. On one hand, this could lead to system failures with a serious cascading effect that could result in massive blackout and destruction of grid infrastructures. On the other hand, sensitive customer's personal information can be hijacked by cyber-criminals for some malicious and fraudulent intentions. In addition, a compromise of AMI's communication networks or associated information management systems may allow an attacker access to the control information that can be corrupted and used to threaten the availability of the data in the system. This can consequently lead to a serious violation of the integrity of the system. These concerns about security and privacy are huge and can undermine the smooth functioning of the AMI and the smart grid at large.

The focus of this thesis is to propose solutions to the security and privacy challenges facing the SG AMI which will help in realizing the full potentials and benefits of the AMI. Firstly, cybersecurity problem bothering on data availability attacks against the SG AMI network is investigated. Problems resulting from data availability attacks against the SG AMI will be compounded by an anticipated

increase in the voluminous amount of data traversing the network and beyond. This will bring an additional burden for storage and computations on the network. In dealing with these issues, a cloud-based OpenFlow firewall was proposed in this thesis not only for the mitigation of the effects of data availability attacks but also to provide an increased capacity for the AMI with regards to computations and storage. The proposed solution utilizes the concept of software-defined networking which makes it easy for controls to be enforced in the network. An added advantage of the proposed solution is that infrastructures can be abstracted using virtualization technologies. This can increase the scalability of the network and eliminates the need for the procurement of on-premises hardware in the event of network expansion. The OpenFlow firewall was designed with Riverbed Modeler and performance evaluation of the firewall was carried out using extensive simulations. Simulation results show that the designed firewall can be used to mitigate the effects of a volumetric distributed denial of service (250 Gps) attack against the AMI and provide QoS guarantee for the network in an attack scenario. The effectiveness of firewall security for the AMI was also analyzed in this thesis using a mathematical tool called PRISM model checker. This tool was used to model the AMI network and to carry out the performance of firewall security under different detection probabilities.

Secondly, a cloud-based system model for the SG AMI leveraging on the features of quantum key distribution (QKD) cryptography was proposed. The proposed model incorporates carefully designed protocols based on symmetric cryptosystem only. The previously designed OpenFlow Firewall was also incorporated into the model to provide comprehensive security and privacy solutions for the SG AMI. The proposed model was necessitated by the fact that without quantum safe encryption which can be provided by QKD, data transmitted over a network would be vulnerable to eavesdropping of the data and/or the security keys. In addition, most protocols proposed for the SG AMI are based on PKI which guarantees security depending on the computational complexity of the protocols. However, this study reveals that most of those protocols adjudged to be secure today would be broken easily by quantum algorithms. This is the key motivation for incorporating into the designed model only protocols based on symmetric cryptosystem. It was shown in this study that the QKD is a lightweight key distribution protocol and therefore suitable for the cloud-based SG AMI. Its security strengths against eavesdropping attack were verified through simulations and have been adjudged to be strong.

Finally, the problem of privacy in SG AMI is investigated further in this research work. Research has shown that machine learning algorithms applied to end user's consumption data can reveal sensitive information which can be exploited by cyber-criminals to launch various kinds of attacks on electricity consumers. Data aggregation is a popular approach utilized for reducing privacy breaches in the AMI network. This approach relies on the popular assumption that aggregated metering data of

a group of consumers are sufficient for data recipients to perform their duties. Unfortunately, this approach usually relies on cryptographic algorithms which increases the overhead on computations and communication for the network. This often drives the network to a congestive state, thereby resulting in delays which can minimize the task of a cyber-criminal targeting the availability of data in the network. The congestion problem was tackled in this work by designing a robust state-of-the-art communication architecture herein referred to in this thesis as Ring Triangulation Communication Architecture (RTCA). Congestive scenarios in Wireless Fidelity (Wi-Fi) and ZigBee wireless communication technology standards were modeled. The designed architecture was then applied to those networks and its performance was analyzed through extensive simulations. The results of the simulations show that notwithstanding the congestive effects of data aggregation, the designed architecture provides a good QoS guarantee for the two considered networks.

Contents

List of Figures

List of Figures

List of Tables

List of Tables

List of Acronyms

AES	Advanced Encryption Standard
AMI	Advanced Metering Infrastructure
AKC	Asymmetric Key Cryptosystem
CSMA-CA	Carrier Sense Multiple Access with Collision Avoidance
CC	Congestion Control
CMS	Congestion Minimizing Scheme
DA	Data Aggregation
EMS	Energy Management System
FAN	Field Area Network
GOF	Grid Openflow Firewall
HAN	Home Area Network
HMAC	Hashed Message Authentication Code
IBE	Identity Based Encryption
ICT	Information Communication Technology
ITS	Information Theoritic Security
IaaS	Infrastructure as a Service
IED	Intelligent Electronic Device
KMS	Key Management Server
MAP	Markov Decision Process

IBS	Identity Based Signature
NAN	Neighborhood Area Network
NIST	National Institute for Standards and Technology
OTP	One Time Pad
OLAP	Online Analytical Processing
paaS	Platform as a Service
PPDA	Privacy Preserving Data Aggregation
PMC	Probablistic Model Checking
POPI	Protection of Personal Information
RCS	Regional Cloud Server
RTCA	Ring Triangulation Communication Architecture
SKC	Symmetric Key Cryptosystem
SaaS	Software as a Service
SCADA	Supervisory Control and Data Acquisition
SG	Smart Grid
SM	Smart Meter
TCS	Top Cloud Server
TTP	Trusted Third Party
WAMMS	Wide Area Measurement and Monitoring System
QoS	Quality of Service
QKD	Quantum Key Distribution
ZCCF	ZigBee Congestion Control Frame
ZTCC	ZigBee Time Congestion Control

Chapter I

Introduction

Introduction

1 Introduction

1.1 Introduction

The smart grid is considered by many as one of the most complex cyber physical systems to be conceived in history. According to [1], smart grid is the term applied to a class of technology designed to modernize the existing utility grid to intelligently and efficiently respond to available power generation, power transmission, and consumer demand. Fig. 1 shows a framework containing the components of the smart grid which are made up of energy infrastructures, ICT technologies/resources and potential applications [2]. This framework which incorporates technologies like the advanced metering infrastructure (AMI) agrees with the conceptual model for the smart grid developed by the National Institute for Standards and Technology (NIST) [3]. The advanced metering infrastructure is believed to be the most fundamental component of this complex smart grid network. Fig. 2 shows the recommended AMI architecture which was the first initiative for standardization in South Africa, to guide Eskom and Municipalities in their requirements for an AMI deployment [4]. The deployment of technologies like the SG AMI will greatly improve the reliability of the grid and reduce costs of power delivery. With the smart grid advanced metering infrastructure, accurate meter readings can be delivered remotely and timely too. This can reduce to the barest minimum or even eliminate inaccuracies and estimations in meter readings and billings. Utilities are also able to perform remote connections or disconnection in a timely manner. In addition, the AMI provides capabilities for time-of-use metering, pre-paid billing, and tamper detection. These capabilities will definitely increase revenue generation and lower operational costs for utilities and grid operators. Unfortunately, the dependence of the AMI and similar technologies on cyber resources will expose the smart grid to various threats, vulnerabilities, and

cyber-attacks [5–7]. While the attack surface for the AMI can be high, this thesis is more concerned about threats, vulnerabilities, and attacks targeting the security and privacy of metering data, billing information and control commands. These are the core information exchanged in SG AMI and beyond. Violating the security and privacy of these core information can be costly to the consumers and utilities. Solutions that seek to protect this core information with respect to confidentiality, integrity, and availability, is the main focus of this thesis. Confidentiality requires that unauthorized persons must be prevented from obtaining the data. The integrity of the data requires that unauthorized persons must be prevented from modifying the data. On the other hand, availability entails that the data must be made available to authorized persons in a timely manner. In addition, the data must be processed or transmitted in such a way that legitimate network users are not starved of resources. While a compromise on any/all of the three previous objectives could also result in violation of privacy, an additional requirement on privacy entails that data must be used for the purpose for which they were sought for and must be retained for only the period necessary for the fulfillment of the contractual obligation for which the data was sought for in the first instance.

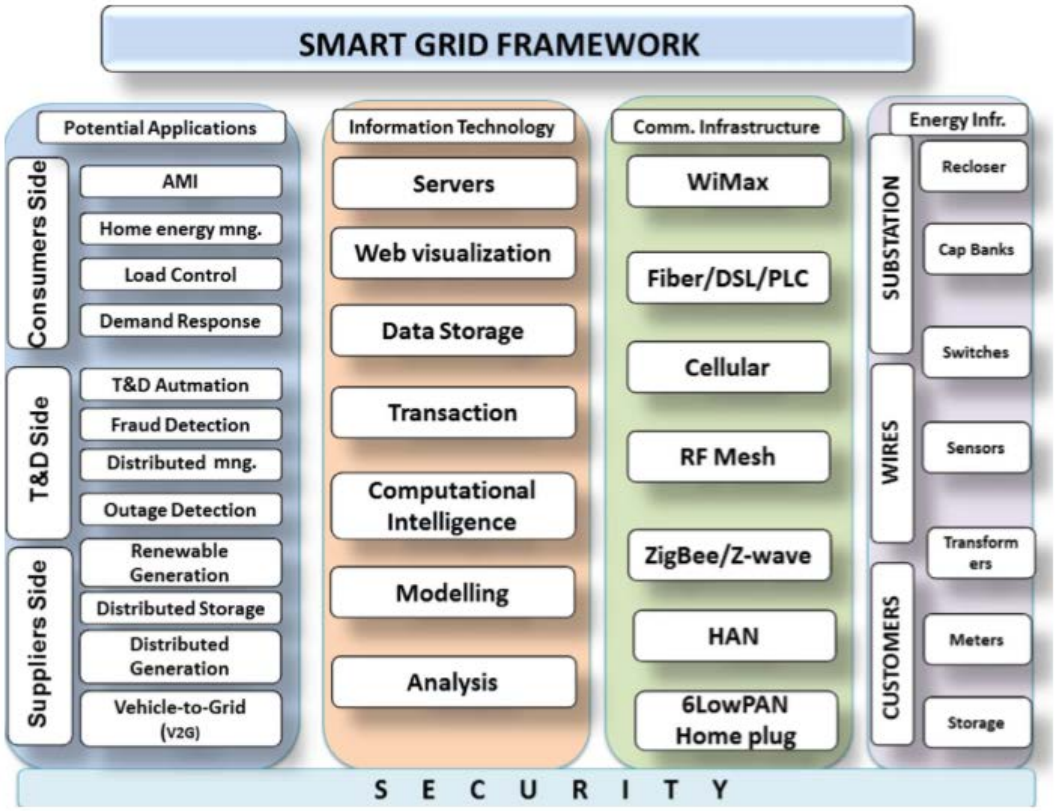

Fig. 1: Smart grid framework with associated technologies and applications [2]

This thesis will therefore help to expand knowledge in the domain of smart grid AMI cyber-security. This was done by proposing security solutions that can mitigate the effects of cyber-attacks targeting

availability, integrity, confidentiality of consumption data/grid control information and the privacy of consumers connected to the network. In the succeeding sections of this chapter, discussion on the motivation for the study, aims and objectives of the study, scope of the study/thesis statement, thesis organization, and contribution of the thesis were presented. Each chapter of this thesis offers an insight into the contents, objectives and the depth of work carried out.

1.2 Motivation for the study

Like other traditional networked systems, vulnerabilities, threats, and cyber-attacks targeting the SG AMI if not detected early, and stopped in a timely manner can be exploited and launched by experienced cyber-criminals to cause the failure of some critical infrastructures of the smart grid. Experience has shown that it is not easy to recover from the effects of such attack. In other words, this can result in a total blackout that can lead to serious economic consequences depending on the cascading effect of such failure. It can also result in a serious network downtime for the entire network that can last for days if the necessary security measures were not put in the right places. In such a situation, the availability of data needed for billing or for urgent control decisions will be severely threatened. This situation will surely hamper the smooth functioning of the system.

Furthermore, the transmission and re-transmission of end user's energy consumption data from consumer's apartment to utility companies, authorized third parties responsible for billing and control stations can bring up routes for vulnerabilities that can violate the privacy of end users. In addition, End user's consumption data stored for load profiling purposes which are needed for energy management system (EMS) operations can be stolen or hacked by cyber criminals for fraudulent purposes. As a result, sensitive end user's data such as names, users address, and other personal information can possibly be obtained by unauthorized persons. Currently, there is a general fear amongst final consumers that this sensitive information can be used against them probably for litigation purposes or for launching different degrees of robbery attacks on them. This might encourage final consumers to resist the mass deployment of the AMI in many countries if these current issues remain unresolved. This situation is further compounded in many countries like South Africa where a bill known as Protection of Personal Information (POPI) has been signed into Law since 2013. This implies that private as well as corporate organizations including Eskom and municipalities in charge of electric power distributions must comply with the bill in their ongoing mass deployment of AMI in the country.

Finally, this thesis is also motivated by the fact that many proposed solutions for the security and

privacy of an SG AMI were done without giving proper consideration for legislation and standardization efforts from regulating bodies in different countries. For instance, many of the proposals found in the literature on data aggregation for the AMI were designed for one entity data recipient. However, many AMI architectures proposed by standard bodies in different countries were designed for multiple data recipients as can be seen [4]. As a result, the general structure of the AMI used for this research study is presented in Fig. 2. The adopted structural model of the AMI comprises three main sections which include: end users, communication and back end systems. End user's devices such as smart meters, in-home-displays (IHDs) and other intelligent electronic devices (IEDs) are usually located at the end users section. These devices connect to the Home Area Network (HAN). The HAN can expand to bigger networks like Neighborhood Area Network (NAN) or Wide Area Networks (WAN) depending on the population of connected end users. The end user domain connects to back end systems through a communication network (based on Wi-Fi, ZigBee, etc) and the appropriate gateway from the end user's side (HAN GW, NAN GW and WAN GW). In other words, the communication side is responsible for transmitting data from end users domain to the back end systems. Finally, the back end systems is responsible for storing and processing of consumption data or control information.

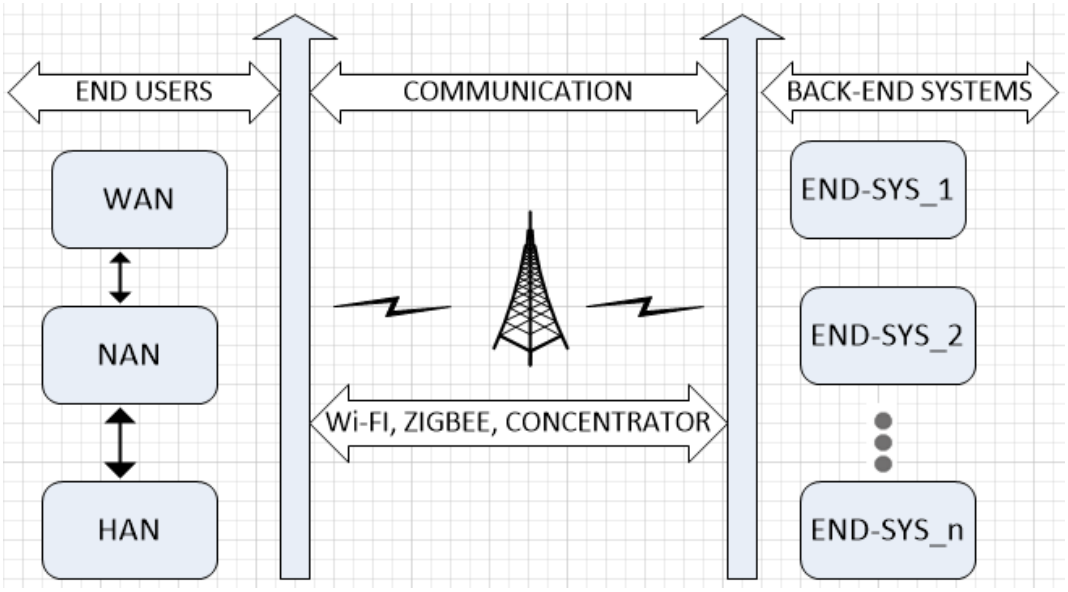

Fig. 2: General structure of AMI used in this study

1.3 Aims and Objectives of Study

The general aim of this research is to propose novel solutions that can make the SG AMI more resilient against cyber-attacks, threats, and privacy violations so that the full benefits and potentials of the AMI

can be fully realized. Specifically, the objectives of this thesis include:

1. To critically review the existing cloud-based solutions for the smart grid advanced metering infrastructure with a focus on cybersecurity solutions.
2. To carry out a systematic review of data aggregation schemes with the intention of discovering how it can be enhanced and applied to an SG AMI based on Wireless Fidelity (Wi-Fi) standard.
3. To analyze the security strengths of ZigBee in order to ascertain the possibility of enhancing it and applying it to the SG AMI.
4. To propose a data aggregation scheme for secure transfer of energy consumption data in a Wi-Fi based SG AMI network.
5. To propose congestion minimizing scheme for enhanced data aggregation in an SG AMI network based on ZigBee wireless standards.
6. To propose a cloud-based firewall for mitigation against data availability attacks in a smart grid advanced metering infrastructure.
7. To propose a cloud-based SG AMI system model leveraging the features of quantum key distribution that can enhance confidentiality, integrity, and availability of metering/control data.
8. To provide a quantitative analysis of firewall security in the face of data availability attacks (like distributed denial of service attacks) in smart grid AMI networks.

1.4 Scope of Study and Thesis Statement

There are indeed many cyber-security challenges in the entire smart grid network starting from the generation system down to the distribution system. Enhancing security of the AMI which is located within the distribution system of the smart grid would enhance the efficiency and reliability of the grid. As a result, the focus of this thesis shall be limited to the distribution system domain with specific attention given to the advanced metering infrastructure network.

Just like the entire smart grid, SG advanced metering infrastructure network has a lot of cyber-threats targeting the network and its ICT driven infrastructure. These threats affect the integrity and confidentiality of data/information traversing the network and can also result in a delay or outright denial of these data when needed. In addition, the privacy of energy consumers can be grossly violated. The effects of these cyber-threats can be mitigated by leveraging on the potentials of cloud computing and the use of properly designed data aggregation schemes.

In this thesis, a state-of-the-art review of data aggregation schemes proposed for the SG AMI was carried out and open issues were highlighted and discussed appropriately. It was discovered from literature review that data aggregation improves security and privacy in the network but can lead to congestion in the network. A robust communication architecture was designed for enhanced data aggregation. This architecture was then applied to IEEE 802.11 and ZigBee (IEEE 802.15.4) wireless standards respectively.

A systematic review on the potentials of using cloud computing for improving the security of the SG AMI network was carried out in this thesis. Research gaps discovered from the review led to the design of openflow firewall for the mitigation of data availability (DDoS) attacks. A quantitative analysis of firewall security was also carried out to understudy the performance of the firewall. Further, a cloud-based system model was also proposed. The proposed system model uses the features of quantum key distribution to enhance security and privacy in the network. The details of the extent of work done in terms of performance analysis of the security solutions proposed in this thesis has been specified in section 1.5 covering the organization of this thesis while taking note of extended comparative analysis of the proposed solutions for possible future research.

1.5 Thesis Organization

The thesis is made up of ten chapters. With the exception of chapter one and chapter ten, other eight chapters contain the published papers which have been earlier declared in this chapter. However, it is to be noted that some of the chapters are slightly different from their published forms. Brief explanations concerning the contents of the thesis chapters have been provided below:

Chapter 1: General Introduction- This chapter presents a general introduction to this thesis. It offers insights on the motivations, objectives, scope, and depth of work carried out in this thesis.

Chapter 2: An Overview of Data Aggregation Schemes in a Smart Grid AMI Network.

In this chapter, a systematic review of the state-of-the-art on data aggregation schemes proposed for the smart grid advanced metering infrastructure is presented. In addition, this chapter contains threats analysis of SG AMI security objects and basic concepts of cryptographic building blocks usually utilized in the design of data aggregation protocols. This chapter also contains highlights and discussion on open research issues on data aggregation for the SG AMI.

Chapter 3: ZigBee-Based SG AMI- Overview of Security Issues

In this chapter, an exhaustive analysis of the security strengths and potential vulnerabilities of ZigBee

was presented. This chapter contains an overview of the state-of-the-art of research on ZigBee based SG AMI with appropriate categorization and classification of research proposals into different domains. Finally, open issues that can be addressed to enhance the security of ZigBee based SG AMI were included and discussed.

Chapter 4: Smart Grid AMI: Overview of Cloud-Based Security Solutions- This chapter presents an overview of cloud-based solutions for the smart grid advanced metering infrastructure network. It contains a review of the potentials of cloud computing for the smart grid. The focus of this chapter is on cloud-based security solutions for the SG AMI. Open research issues on cloud-based security solutions for the SG AMI were highlighted and discussed.

Chapter 5: Data Aggregation in Smart Grid AMI Network for Secure Transfer of Energy Consumption Data.

This chapter presents enhanced data aggregation for the SG AMI for a secure transmission of end user's consumption data. This was achieved by designing a state-of-the-art communication architecture referred to as Ring Triangulation Communication Architecture (RTCA). This architecture was then utilized for improved data aggregation using a Wi-Fi (IEEE 802.11) communication standard as a case study. The discovery made from literature carried out in this thesis shows that data aggregation protocols can be used to enhance security and privacy in SG AMI network. However, these data aggregation protocols impact negatively on the network by increasing the overheads in communications and computations; thereby driving the network to a congestive scenario in the absence of a robust congestion control scheme implemented in the network. Therefore, a comparison of the proposed scheme and transmission control protocol (TCP) congestion control schemes that have been proposed in the literature, and implemented in IEEE 802.11 wireless network standard is included in this chapter.

Chapter 6: Congestion Minimizing Scheme for Enhanced Data Aggregation in ZigBee-Based Smart Grid AMI Network.

In this chapter, the previously designed Ring Triangulation Communication Architecture (RTCA) was utilized for an enhanced data aggregation in a ZigBee based SG AMI. The RTCA was therefore re-configured to suit the protocol specifications of ZigBee (IEEE 802.15.4) at the appropriate layers of the network. The analysis done in this chapter also focused on important quality of service (QoS) that can be affected by a congestive network scenario since data aggregation worsens congestion in communication networks. Two ZigBee congestion control schemes proposed in the literature were compared with the proposed scheme. The two ZigBee congestion control schemes are ZigBee

Congestion Control Frame (ZCCF) and ZigBee Time Congestion Control (ZTCC).

Chapter 7: A Cloud-Based OpenFlow Firewall for Mitigation against DDoS Attacks in an SG AMI Networks.

This chapter presents the design of an OpenFlow firewall for the mitigation of distributed denial of service (DDoS) attacks on an SG AMI. This contains demonstrations of the security strengths (with regards to QoS features) of the designed firewall and how it can be leveraged on for instant detection and mitigation of DDoS attacks in an SG AMI in an attack scenario. The performance analysis of the proposed openflow firewall was presented in two ways with respect to some important QoS. In the first scenario, comparative analysis was done between the proposed grid openflow firewall (GOF) and non-grid openflow firewall. In the second scenario, the GOF was compared with two major classes of firewalls proposed in the literature. The two classes of firewall considered include VLAN based access firewall and IP Gateway based firewall. It is to be noted that the proposed GOF was designed principally for the mitigation of data availability (DDoS) attacks against the SG AMI. In other words, analysis of proposed firewall with respect to other attacks against integrity and confidentiality is beyond the scope of this chapter.

Chapter 8: Enhancing the Security of a Cloud-Based Smart Grid AMI Network by Leveraging on the Features of Quantum Key Distribution.

In this chapter, a cloud-based SG AMI system model leveraging the features of Quantum Key Distribution (QKD) is presented. This key distribution scheme is compatible with the designed OpenFlow firewall and compliments the security features of the firewall. In other words, cyber security objectives such as confidentiality, integrity and availability will be enhanced with the proposed cloud-based SG AMI system model. Performance evaluation of the protocols incorporated into the proposed system model in terms of overheads in communication and computations was carried out in this chapter. The performance evaluation results from the proposed system model were also compared with similar results from proposals in the literature.

Chapter 9: Quantitative Analysis of Firewall Security under DDoS Attacks in Smart Grid AMI Networks.

In this chapter, an SG AMI network was modeled using PRISM and a probabilistic best- and worst-case analysis of the firewall security with regard to DDoS attack success under different firewall detection probabilities was carried out. The results from this quantitative analysis can be useful in determining the extent the DDoS attack can undermine the correctness and performance of a cloud-based firewall. In addition, the study can also be helpful in knowing the extent a firewall can

be improved by applying the knowledge derived from the worst-case performance of the firewall.

Chapter 10: Conclusion and Possible Future work.

This chapter concludes the research undertaken in this thesis with a detailed appraisal of the contributions, significance, and recommendations. In addition, recommendations for possible future research work were highlighted.

1.6 Thesis Contributions

The novel contributions of this research work can be summarized as follows:

- Proposed a cloud-based OpenFlow firewall for mitigation against distributed denial of service (DDoS) attacks in an SG AMI.

  Previous proposals on solutions to the challenges arising from security and privacy violations for the SG AMI have relied heavily on cryptographic algorithms. While these solutions have enhanced the confidentiality and integrity of data/information traversing the SG AMI, they have not contributed significantly in reducing the challenges posed by cyber-attacks targeting the availability of these data/information. Instead, those solutions often times lead to an increase in the overheads on computations and communication for the network, thereby leading to delays that abate cyber-criminals with intention of attacking the availability of these data. The proposed openflow firewall was designed in Riverbed Modeler by leveraging on the features of software defined networking (SDN) which allows for the virtualization of a programmable network using C++ support embedded in Riverbed. The efficacy of this firewall was tested against a 250 Gbps DDoS volumetric attack. The performance of the grid opeflow frirewall was comapared with non-GOF, VLANs access based firewall, and IP Gateway based firewal. Simulation results showed that the firewall outperformed other firewall types and can be used for instant detection and mitigation against DDoS attack. The proposed firewall can also be used to improve the quality of service (QoS) of the network in an attack scenario. This proposed GOF has been published in IEEE Xplore.

- Proposed a cloud-based key distribution scheme for the SG AMI utilizing the features of quantum key distribution (QKD).

  Previous researches on key distribution schemes found in the literature were not designed for a cloud-based SG AMI. In addition, the QKD scheme which has been adjudged by many researchers to be a very strong protocol was originally designed for a two-party system (sender

and receiver) communicating over a channel. The design of a cloud-based SG AMI system model using the security features of quantum key distribution to suit a multi-party system as in the case of a cloud-based network was an important contribution of this chapter. Further, the proposed system model design incorporated lightweight authentication protocols at the upper cloud layers and a data aggregation scheme based on homomorphic encryption and (other symmetric encryption schemes) at the root cloud to preserve the privacy of high-frequency data needed for profiling or for other grid control purposes. Privacy analysis of the proposed model shows that security and privacy can be enhanced in the SG AMI.

Performance evaluation of the Incorporated protocols showed considerably lower overheads in terms of communications and computations when compared with similar protocols proposed in the literature.

Finally, it should be noted that all the incorporated protocols rely on symmetric keys that must be generated by running the BB84 protocol. Unfortunately, literature has revealed that the security of symmetric based protocols can be compromised if symmetric keys are generated with insufficient key material. This is because the key generation process can be affected by eavesdropping attack which can be launched by Eve. The simulation of the BB84 protocol carried out in this chapter revealed an optimum range for the bias ratio which can be utilized by the sending party (Alice) to generate symmetric keys with sufficient key material. This is also an important contribution from this chapter. The proposed system model for SG AMI has been published in a UKZN DOHET accredited Journal.

- Proposed a Ring Triangulation Communication Architecture (RTCA) for enhanced data aggregation in smart grid AMI network. The contribution from RTCA resulted in two publications in two UKZN DOHET accredited Journals.

    1. The problem of congestion that normally results from data aggregation was tackled using Ring Triangulation Communication Architecture (RTCA) which was designed in Riverbed Modeler and then applied to IEEE 802.11n wireless network standard. In designing the RTCA some important components were introduced in Riverbed and modeled with C++ engine in the software. In order to ensure the secure transfer of user consumption data, data minimizing function (DMF) algorithms based on homomorphic encryption were formulated and applied to the designed architecture. The proposed scheme was compared with TCP congestion management schemes proposed in the literature and already implemented in the IEEE 802.11 network standard with respect to some important quality of service (QoS). Results showed that the proposed scheme

delivered a better QoS than the TCP congestion management schemes present in the wireless network standard under consideration. Previous proposals on data aggregation for the SG AMI have not applied their aggregation protocols to network standards recommended for the SG AMI. Such authors usually justify their proposals by analyzing the overheads due to communications and computations resulting from the cryptographic primitives utilized in their protocols.

2. The designed RTCA was also applied to a ZigBee-based smart grid AMI network for enhanced data aggregation which will minimize congestion in the network. Research has shown that there is no efficient mechanism for congestion management in the current ZigBee standard. In order to overcome this shortcoming, the designed RTCA was reconfigured to conform with the ZigBee protocol specification at appropriate network layers. This robust and resilient communication architecture was implemented with a well-constructed ZigBee data and its aggregation algorithms. The quality of service (QoS) results from this proposed scheme outperformed the ZigBee Congestion Control Frame (ZCCF) and ZigBee Time Congestion Control (ZTCC) schemes proposed in [8] in terms of latency and average throughput.

References

[1] M. Greer and M. Rodriguez-Martinez, "Autonomic computing drives innovation of energy smart grids," *Procedia Computer Science*, vol. 12, pp. 314–319, 2012.

[2] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "A survey on smart grid potential applications and communication requirements," *IEEE Transactions on industrial informatics*, vol. 9, no. 1, pp. 28–42, 2013.

[3] N. S. Grid, "Introduction to nistir 7628 guidelines for smart grid cyber security," *Guideline, Sep*, 2010.

[4] H. Groenewald, "Nrs049—advanced metering infrastructure (ami) for residential and commercial customers," *ESKOM, Johannesburg, Presentation*, 2009.

[5] T. Mehra and R. Pateriya, "Cyber security considerations for advanced metering infrastructure in smart grid," *Int. J. Sci. Engg. Res. 4 (8)*, pp. 939–944, 2013.

[6] T. Zhang, X. Lu, X. Ji, and W. Xu, "An identity-based secure communication scheme for advanced metering infrastructure in smart grid," in *Control And Decision Conference (CCDC), 2017 29th Chinese*. IEEE, 2017, pp. 6959–6964.

[7] J. Fei, Y. Ma, X. Huang, Z. Liu, Q. Wang, and Y. Tang, "The research on cyber-attack testbed with hardware-in-loop," in *Energy Internet and Energy System Integration (EI2), 2017 IEEE Conference on*. IEEE, 2017, pp. 1–6.

[8] W. S. Jeong and S. H. Cho, "Congestion control for efficient transmission in zigbee networks," in *Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on*. IEEE, 2009, pp. 1–4.

Chapter II

Overview of Data Aggregation Schemes in a Smart Grid AMI Network

Paper A

Overview of Data Aggregation Schemes in a Smart Grid AMI Network

R.C Diovu and J.T Agee

Published

Journal of Communications (ISSN: 1796-2021)

Abstract

In a fully evolved smart grid advanced metering infrastructure network, grid operators, energy suppliers and other third party operators can be granted access to end users consumption data for purposes such as billing and other grid control operations. This will help to optimize smart grid operations and maximize energy consumer's benefits. However, uncontrolled or unauthorized access to consumption data may put energy consumer's privacy at risk. This is because end user's energy consumption data can be profiled when collected at a high frequency and then subsequently used to make reasonable inferences about the private life of energy consumers. Aggregation of energy consumption data has been a popular research approach towards preserving the privacy of energy consumers. In this paper, we present a survey of different schemes using data aggregation approach for preserving the privacy of energy consumers in a smart grid AMI network. At the end, open research issues are highlighted and discussed.

1 Introduction

The smart grid advanced metering infrastructure (SG AMI) is a technology designed to modernize the traditional electricity network. With the SG AMI, conventional mechanical meters are replaced with smart meters and other intelligent electronic devices (IEDs) which help in the measurement, collection and analysis of energy consumption data. The SG AMI is also designed to provide multi-way communication paths between energy consumers, smart grid operators and other authorized and trusted third party operators [1]. The National Institute for Standards and Technology (NIST) prioritized the AMI as an important functionality in the implementation of the smart grid vision [2]. Experiences from many countries like Italy have shown that the implementation of the AMI can bring some advantages which include reduction in the costs of operations [3]. For instance, periodic readings of user's energy consumption can be taken remotely, thus, eliminating the need for the engagement of persons who perform this task by visiting consumer's apartments. This technology will not only benefit utility companies as consumers will have options of leveraging on AMI's demand response features to switch off their high consuming appliances during peak periods. This will evidently help energy consumers to save a lot of money. Very importantly, the AMI eliminates monotony and creates a competitive market by involving trusted third parties who may be involved in supplying energy to consumers or in-charge of billing and other related operations. A summary of SG AMI's advantages to consumers, energy suppliers and utility has been presented in Fig. A.1.

Unfortunately, the introduction of the AMI into the smart grid vision raises new issues relating to the violation of the privacy of energy consumers. Research has shown that energy consumption data collected at high frequency can be profiled and consequently used to make reasonable inferences or draw accurate conclusions about the private life of energy consumers. Fig.A.2 provides an illustration which shows how smart meter consumption data can be used to predict with near certainty the private behaviours of consumers using machine learning algorithms [4]. From Fig. A.2, it can be clearly seen that consumer's activities such as working hours and vacation periods can be predicted with ease and this can be beneficial to cyber criminals who can utilize such knowledge to further launch various degrees of attack against consumers. In addition, such highly classified private information can be used by insurance companies or/and energy suppliers for price discrimination against energy consumers.

Currently, there are many privacy-related legal frameworks in many countries today which are geared towards the protection of private data. In the United States for example, there is a lack of privacy regulations for smart metering data at the federal level [5]. However, attempts have been made by states like California, Colorado, Ohio and Oklahoma [6] through their respective public utilities commissions to introduce the concept of smart metering data privacy into their privacy-related legal frameworks. Similarly, the protection of personal information (POPI) has been enacted into law in South Africa since 2013 [7]. It is expected that all personal information which includes smart metering data should be regulated in the way and manner they are processed, how long they are to be retained by authorized parties and in ensuring that personal data must be used for the purpose for which they were collected in the first instance.

Unfortunately, there are some drawbacks from these privacy frameworks as there are no guarantees that data owner's consent would be sought and obtained before such data can be released to either public or private organizations in need of them. Another conflicting issue with regards to data owner's consent is the fact that legislations on data protection makes processing legal for the fulfillment of legal or contractual obligations such as billing, energy supplies or even for ensuring network stability. It is therefore safe to state that these privacy frameworks/policies cannot effectively prevent privacy violations expected against energy consumers. In other words, privacy-preserving technologies such as data aggregation techniques which can prevent a lot of privacy violations before they happen are highly desirable. In this, paper, a comprehensive survey of data aggregation schemes proposed for enhancing the privacy of energy consumer's data is carried out. At the end of this survey, open research directions on data aggregation in a smart grid AMI have been highlighted and discussed. The rest of this paper is organized as follows: section 2 contains necessary backgrounds including cryptographic

Fig. A.1: Benefits of advanced metering infrastructure

building blocks utilized in designing data aggregation protocols. In section 3, an extensive review of data aggregation protocols found in the literature is carried out. Section 4 contains open research issues and future challenges while section 5 contains the conclusion of this study.

2 BACKGROUND

2.1 Advanced Metering Infrastructure and Associated Technologies

The advanced metering infrastructure incorporates smart and intelligent electronic devices (IEDs) and communication technologies that help automate metering functions which in the past were accomplished manually and involved intensive operations. The AMI has been integrated with some advanced customer-based technologies which enable power utilities to offer new rate options that produce good incentives to customers, thus, motivating them to reduce their energy consumptions during peak periods. In other words, the AMI and the customer-associated technologies work together to automate functions and improve demand side management. The flow of the relationship between these AMI functions and their associated technologies is presented in Fig. A.3. These

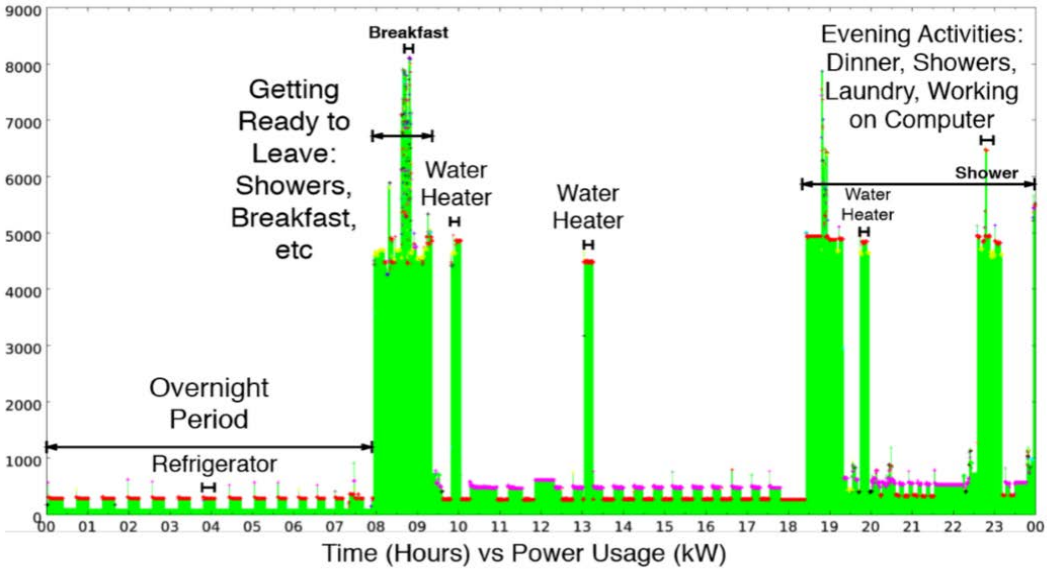

Fig. A.2: An illustration showing how appliance load signatures can be used to predict the private life of users [4]

technologies include but are not limited to the following:

- Information Technologies: With these technologies, customers can be encouraged to better manage their electricity consumption. This can be done by providing them with near real-time data about their electricity consumption and costs through technology platforms like in-home displays (IHDs), web portals, and text/email. Such technologies can provide information in amazing ways which are capable of providing understanding and insights about actions that can save energy and reduce bills.

- Communication Networks: The installation of the AMI necessitates the installation of new communication networks and/or the upgrade of existing ones. These networks ought to be as robust as possible considering the voluminous amount of data that need to be transmitted from the AMI smart meters to the AMI back-end systems and other authorized third parties.

- AMI Back-End Systems: The AMI back-end systems are involved in advanced operations that ensure the smooth functioning and the stability of the smart grid. They are in-charge of functions like billing, meter data management and other grid control operations like power quality monitoring.

- Smart Meters: The smart meters are the most fundamental and core element of the AMI. The smart meters provide functions such as measuring customer electricity consumption at intervals which varies from 5, 15, 30, or 60 minutes. They can also be used to measure voltage levels; and monitoring the on/off status of electric service. Fundamentally, smart meters are

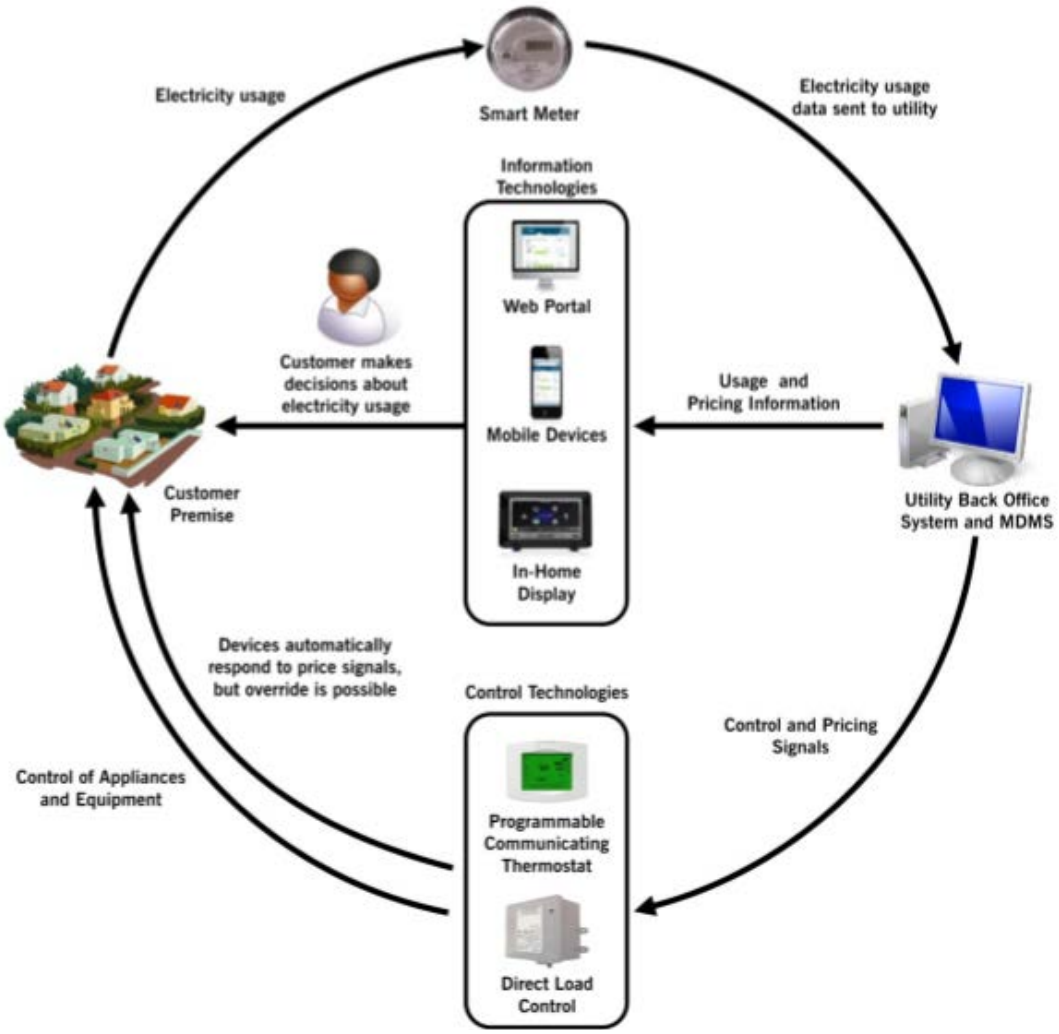

Fig. A.3: Relationship between AMI functions and their associated technologies [8]

designed to communicate these readings to utilities for processing and analysis. Interestingly, important information such as billing charges, energy feedback, and available time-based rates can be communicated back to customers via the smart meters and/or through its associated technologies.

2.2 The SG AMI and its Security Challenges

The smart grid advanced metering infrastructure (SG AMI) was conceived in the smart grid vision to provide near-real time monitoring of energy consumption/usage. The SG AMI has been designed to provide dynamic pricing support in its demand response scheme especially for the residential customers. The SG AMI is interconnected with other smart grid domains with an open, distributed and highly networked infrastructural technologies. Thus, the AMI consists of several communication hardware/software, severally connected intelligent electronic devices (IEDs) and data management

system. In order to keep the smart grid AMI very secure, genuine but collaborative efforts are required from all the stakeholders which include [9]-customers, grid/utility operators, billing companies, energy providers, other third party operators who may be in-charge of some control operations and representatives of government. The smart metering data, control commands from smart meters, billing information and energy consumer's private information are the key security objects that need to be protected against security breaches and/or privacy violations [10]. The threat analysis of these key objects have been provided briefly.

2.2.1 Threats Analysis of SG AMI Security Objects

Notwithstanding the unquantifiable benefits of the SG AMI to energy consumers, energy suppliers and grid operators, the SG AMI provides opportunities for cyber criminals and malicious insiders to attack the smart grid or violate the privacy of energy consumers for intentions which can either be fun or for ill-gotten gains. The threats to the security objects include but not limited to the following:

2.2.2 Denial of Service/Distributed Denial of Service Attack

Denial of service or distributed denial of service attack is an attack against the availability of smart metering data or control signals. The smooth functioning of the SG AMI depends on the availability of these data [11]. Denial of service/distributed denial of service attacks are executed mainly by sending fake requests which can be voluminous to servers or other network nodes such that the affected nodes would be driven to a point of exhaustion where little or no resources are left to process requests from legitimate nodes connected to the network [12]. Denial of service attacks can result in an improper scheduling of data delivery between meters and data concentrators. This situation can lead to buffer overflow and data loss at the concentrator's side. Moreover, this can cause delay in data delivery or even data loss at the AMI back-end due to limited link bandwidth [13]. On the other hand, control data are needed in a timely manner so that urgent decision can be taken by grid operators to ensure general stability of the system.

2.2.3 Eavesdropping Attack

Eavesdropping is an attack where a cyber-criminal listens or gathers data intended for the smart grid AMI. In this kind of attack, the cyber-criminal, or eavesdropper illegally monitors and taps into the transmission signal between the data source and the smart grid AMI back-end. This can be done

between the time the data are encoded and the time it is decoded. That may not be simple for a cyber-attacker who is not experienced but some cyber-criminals could have access to the common decoding algorithms, and by trying so many options, they can succeed in determining how to read the data [14].

2.2.4 Injecting False Information (Replay Attack)

In this type of attack, the cyber-criminal can send packets to inject false information in the network. This can range from false metering data, false prices, fake emergency event. The major motivation for an attacker with this malicious intension may be to evade the payment for consumer's electrical energy or to pay a drastically reduced bills [15]. An attacker may also decide to record sequence of smart metering data and then replays this sequence later. This kind of attack targets the integrity of the measured data and also prevents data freshness [12]. Similarly, the attacker can transmit false data to the control center. Fundamentally, most control systems are designed to question or ignore data whose mean square difference from the normal or expected value is too high [16]. An attacker with this basic knowledge, can analyze data for a period of time, figure out an acceptable range of values, and inject data maliciously which will be accepted by the control system [17].

2.2.5 Energy Consumer's Privacy Violations

Collecting energy consumer's metering data very frequently may reveal sensitive information about the consumer, such as energy consumption patterns [18], the type of appliances used at the consumer's premises [19], or information that can help to know when the premises are occupied or not. This kind of information may be very useful to a number of external entities. For example, criminal elements may leverage on information revealed from metering data to identify and target temporary unoccupied premises. This would increase their possibility of burgling such apartments undetected. Organizations like insurance companies may be fascinated by such data. For instance, an insurance company may use a consumer's energy data to infer the consumer's ruinous life styles which can be utilized by the insurance company to come to a decision on parameters of the consumer's life insurance policy.

2.2.6 Access Rights Violations

Energy consumer's fine-grained energy consumption data may be very useful for authorized third parties or other internal SG entities too. For example, if an entity in charge of energy supply has access to the consumption data of other suppliers' customers (in addition to its own customers), the entity (supplier) would have a competitive advantage over its competitors in the electricity markets. As

such, their competitors can be easily short-changed in the open, liberalized and ambitious electricity markets. Such competitive advantage could be helpful to the supplier in making decisions that can increase its market share. Therefore, such authorized entities might be tempted to to have undue access to the metering data of customers connected to their competitors.

2.3 Basic Cryptographic Concepts

In this section, the basic cryptographic concepts or building blocks that have been used in different data aggregation schemes have been reviewed. These cryptographic building blocks include but are not limited to the following: symmetric key cryptosystems, hashed message authentication codes (HMACs) and asymmetric key cryptosystems like digital signature schemes, Rivest–Shamir–Adleman (RSA) cryptosystem, El-Gamal Cryptosystem and Elliptic curve key cryptography.

2.3.1 Symmetric Key Cryptosystem

In symmetric cryptographic system, the same key is used by a sending party and the receiving party. With this cryptosystem, the confidentiality of the transmitted message can be guaranteed using the encryption and decryption process of the cryptosystem [20]. Conventionally, the sending party sends a message, m which is encrypted with a unique key, k in an encryption process where m and k are inputs to the encryption algorithm, E. The encryption algorithm generates an encrypted message, e, known as the ciphertext. This process can be represented as: $e = E(k, m)$. This ciphertext is then sent to the receiving party through a communication channel. In other to recover the original message that was encrypted, the receiving party uses the ciphertext and the unique key as inputs to a decryption algorithm, D. This process can be represented as: $m = D(k, e)$. A typical illustration of the symmetric key cryptosystem that can be utilized for providing message confidentiality has been illustrated in Fig. A.4 while the Advanced Encryption Standard is well known example of a symmetric encryption algorithm [21].

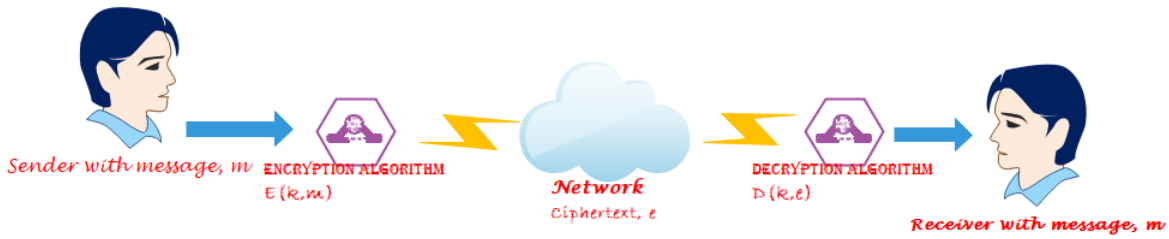

Fig. A.4: Symmetric key cryptosystem

2.3.2 Hashed Message Authentication Code (HMAC)

A hashed message authentication code (HMAC) is obtained by merging a hash function with a shared symmetric key. HMAC can be used to guarantee the authenticity of the transmitted message. On the other hand, a hash function, H(.), is that function which receives as input data with an unpredictable length and produces as output a fixed-sized string of hash value, h [20]. This process can be represented as: $h = H(m)$, where m is the message to be transmitted. Typical examples of well-known standard hash functions include SHA-1, SHA-224, SHA-256, SHA-384 and SHA-512 specified in the Secure Hash Standard (SHS) [22]. A HMAC otherwise known as a keyed-hash function because it is key dependent hash function. In other words, a symmetric key is always used in the generation and verification of the hash value of the message to be transmitted [20]. The HMAC can be used to provide authenticity for the transmitted message because the generated hash value can only be verified by somebody in possession of the identical symmetric key, k. This guarantees that the message originated from the claimed sender. In addition, the integrity of the transmitted message can also be guaranteed as the message receiver can verify whether or not the transmitted message has been modified in the process on transmission. The above guarantees can only be sustained if the hash function has the following properties:

1. Given a message, m, the hash value of the message $h = H(m)$ can be computed with ease but given the hash value, h, it would be extremely hard to compute the message, m, such that $H(m) = h$.

2. Given a message, m, it would be very difficult to find another message, m', such that $H(m) = H(m')$.

As shown in Fig. A.5, a hash function is used with a confidential key known only by the sender and receiver. The sending party applies this hash function in combination with the secret key to produce a hash value $h = H(k, m)$ which is then transmitted along with the message through a communication network (channel). With the help of the secret key the receiver can calculate the hash value and then compares this with the received hash value. This makes it easier to confirm that the the received message is authentic and has not been modified.

2.3.3 Asymmetric Key Cryptosystem

While symmetric key cryptosystem makes use of one unique key for the encryption and decryption algorithms, the asymmetric key cryptosystem uses two separate keys to perform cryptographic

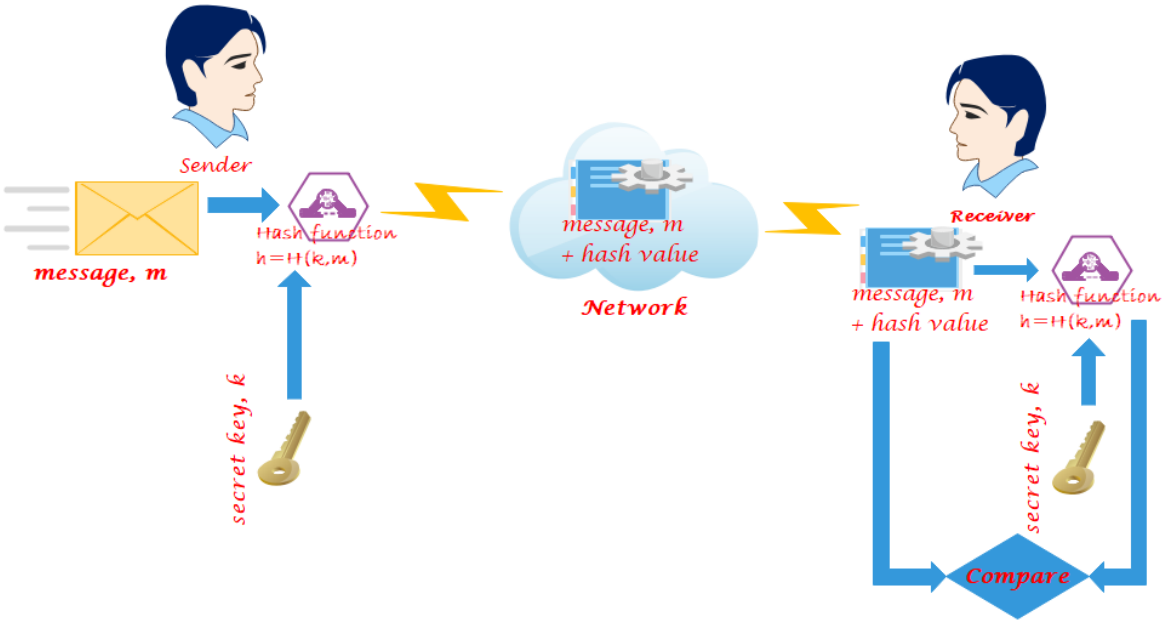

Fig. A.5: An Illustration of Hashed Message Authentication Code (HMAC)

operations. These two keys are conventionally referred to as public and private keys respectively. This type of cryptosystem can be used to guarantee the confidentiality and authenticity of the transmitted message [20]. Typical examples of well-known asymmetric encryption algorithms include the Rivest-Shamir-Adleman (RSA) [23], El-Gamal [24], Cramer-Shoup [25] and Paillier [26] algorithms. In this type of cryptosystem, the sending party uses both the message to be transmitted and the public key, Kp as inputs to the encryption algorithm which in turn generates a ciphertext, e. This process can be represented as: $e = E(K_p, m)$ where K_p represents the public key. Once ciphetext (encrypted message) is generated, it can then be transmitted through a communication network. In other to recover the original message sent by the sender, the receiver uses his/her private key and the received ciphertext as inputs to a decryption algorithm. This decryption process can be represented as: $m = D(K_s, e)$, where K_s represents the private key. A simple illustration of asymmetric key cryptosystem has been presented in Fig. A.6.

2.3.3.1 Paillier Cryptosystem: The paillier cryptosystem invented by a French researcher Pascal Paillier in 1999 is an example of asymmetric cryptosystem that is worth reviewing in this paper primarily because it is efficient and secure. Pailllier cryptosystem has received good attention in its application to various privacy-preserving technologies because of its nice homomorphic property [27]. Unlike most cryptosystems already reviewed in this paper, paillier cryptosystem comprises of three algorithms which include a key generation algorithm, encryption and decryption algorithms. A brief explanation of the steps taken for the construction of the three algorithms have

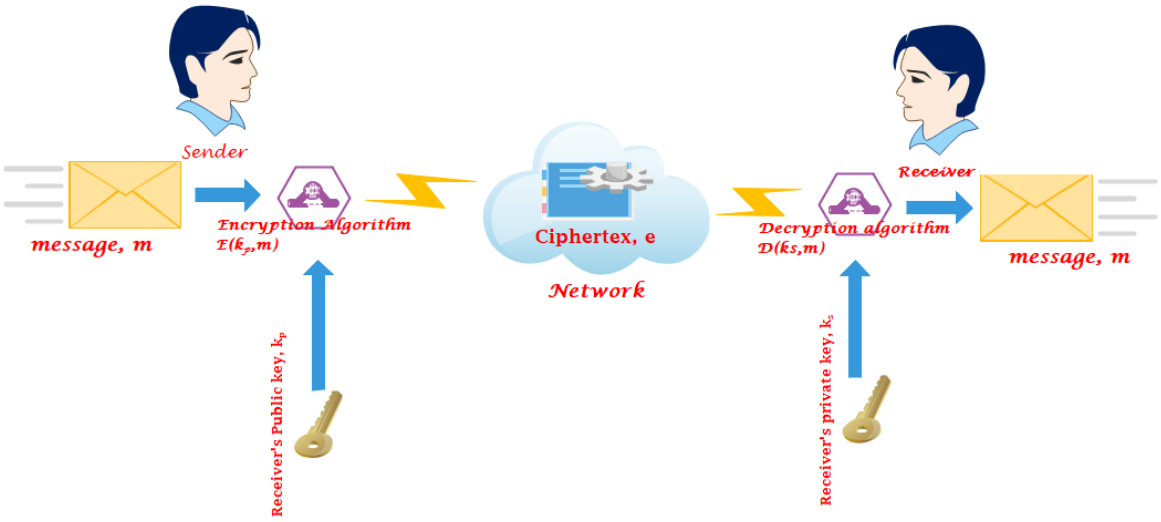

Fig. A.6: An Illustration of Asymmetric Key Cryptosystem)

been provided below [28, 29]:

1. Key Generation Algorithm: The steps for key generation include:
   (a) Two large prime numbers p and q are chosen randomly and independently of each other.
   (b) The modulus, n, of the two prime numbers, their product, i.e $n = p.q$ and λ, which is called camichael's function are computed.
   (c) select a random variable, g, such that $g \in Z*_n^2$ and $g's$ order is a non-zero multiple of n since $g = 1 + n$ works and can be calculated easily.
   (d) It is to be ensured that n divides the order of g by checking that the following modular multiplicative inverse exists:$u = (L(g^\lambda \mod n^2)^{-1} \mod n$ where the function L is a langrage function, $L(u) = (u-1)/n$.
   (e) following (d) above, the public key would be (u, g) while the private key would be λ for $u = 1 \mod n$.
2. Encryption Algorithm: In the case of encryption algorithm, the following steps can be taken:
   (a) Given a message (plaintext), m, such that $m \in Z_n^*$ and $m < n$.
   (b) Select a random number such that $r \in Z_n^*$.
   (c) Calculate the ciphertext, e, such that $e = g^m.r^n mod\ n^2$.
   (d) The expression for the encryption algorithm can be represented as: $e = E(K_p, m, r)$.
3. Decryption Algorithm: Finally, the steps for executing the decryption algorithm include:

(a) (i) Given a ciphertext, e, such that $e \in Z_n^2*$ and $e < n^2$.

(b) Recover the original plaintext, m, such that $m = L(e\lambda \mod n^2)/L(g\lambda \mod n^2) \mod n$.

(c) The expression for the decryption algorithm can be represented as: $m = D(K_s, e)$.

In summary, the Paillier cryptosystem has two important properties which include: additive homomorphism and a capability of random number recovery. The mathematical elements of these two important properties are provided below [30]:

1. Additive Homomorphism: This property stipulates that the multiplication of the encrypted messages would result in the sum of the original messages (plaintexts). Mathematically, it can be stated that:

$$e(m_1).e(m_2).e(m_3).e(m_4) = (g^{m1}.r_1^n).(g^{m2}.r_2^n).(g^{m3}.r_3^n).(g^{m4}.r_4^n) \mod n^2 \quad \text{(A.1)}$$

$$= g^{m1+m2+m3+m4}.(r_1.r_2.r_3.r_4)^n \mod n^2$$

$$= e(m_1 + m_2 + m_3 + m_4)$$

It can be noted that this additive property of the Paillier cryptosystem can be utilized for aggregation in a progressive manner. This can be stated mathematically as follows:

$$e(m_1).e(m_2) = (g^{m1}.r_1^n).(g^{m2}.r_2^n) \mod n^2 \quad \text{(A.2)}$$

$$= g^{m1+m2}.(r_1.r_2)^n \mod n^2$$

$$= e(m_1 + m_2)$$

$$e(m_3).e(m_4) = (g^{m3}.r_3^n).(g^{m4}.r_4^n) \mod n^2 \quad \text{(A.3)}$$

$$= g^{m3+m4}.(r_3.r_4)^n \mod n^2$$

$$= e(m_3 + m_4)$$

$$e(m_1 + m_2).e(m_3 + m_4) = (g^{m_1+m_2}.(r_1.r_2)^n.(g^{m_3+m_4}.(r_3.r_4)^n \mod n^2 \quad \text{(A.4)}$$

$$= g^{m_1+m_2+m_3+m_4}.(r_1.r_2.r_3.r_4)^n \mod n^2$$

$$= e(m_1 + m_2 + m_3 + m_4)$$

2. The second property of Paillier cryptosystem shows that raising an encrypted message to the power of a second message would result in the multiplication of the plaintext messages [28]:

$$E(m_1, K_p)^m 2 = g^{(m_1.m2)}.(r_1^m 2)^n. \mod n^2 \tag{A.5}$$

$$= e(K_p, m_1, m_2, \mod n)$$

Summarily, the steps for the encryption process is as follows:

- The sender obtains the receiver's public key (n, e).
- The plaintext is expressed as a positive integer, m.
- The ciphertext is calculated such that ciphertext $c = m^e \mod n$.
- The ciphertext, c, is then transmitted to the receiver.

Decryption Process: Finally, the steps for the decryption process is as follows [31]:

1. The receiver uses the private key, (n, e) to compute $m = c^d \mod n$.
2. The plaintext is then recovered from m.

2.3.3.2 The El-Gamal Public Key Cryptosystem: In 1985, Taher Elgamal proposed El-Gamal public key cryptosystem which is used over finite fields and its security is based on the discrete logarithm problem (DLP) [32]. The cryptosystem provides additional layer of security by making it possible to asymmetrically encrypt keys that have been previously used for symmetric encryption. The reason for the use of the discrete logarithm problem is because of the difficulty in finding discrete logarithm while the inverse operation of exponentiation can be calculated with ease. Bryce Allen defined discrete logarithm to be the inverse of modular exponentiation [33]. This definition states that given a modular exponentiation $y = g^x$ in Z_p* and the base g, the discrete logarithm $\log_g y$ is x. This is a discrete logarithm in the cyclic group $\langle g \rangle$ which may or may not be all of Z_p*. When $|g| = n$ is large and has at least one large prime factor, discrete log problems in $\langle g \rangle$ are considered intractable. An El-Gamal cryptosystem can be defined by a tuple p, g, x, y with p being a large prime number which also describes which group Z_p* is used, g, being an element of order n in Z_p*. In addition, x is a random integer with $1 \leq x \leq n - 1$, and $y = g^x$.

2.3.3.2.1 Encryption and Decryption of El-Gamal Cryptosystem: Before encrypting a plaintext using El-Gamal cryptosystem, it must be converted to an integer between 1 and $p - 1$ where integer between $(1, p - 1) \in Z_p*$. If the message is the key for a symmetric cipher, then it may already be

a number, if on the other hand the message is larger than $p-1$, then it can be broken into blocks. If g is primitive, then, $n = p-1$. However, n can be chosen to be much smaller than $p-1$, g is not a primitive, then only n of the $p-1$ members Z_p* will be in $\langle g \rangle$. Since a discrete log would be required in order to recover the original message, raising g to the power of the message would not be workable. In summary, the public and private keys for the El-Gamal cryptosystem would be (p, g, y) and x respectively. With an integer $K \in [1, n-1]$ and a public key, the encryption and decryption functions would be given by $E_k(m) = (g^k, my^k)$ and $D(u, v) = u^{-x}.v$ respectively where all operations are performed in $\mod p$ in Z_p*. The decryption function can be used to retrieve the original message as follows: $u^{-x}g^{-kx} = y^{-k}$ such that $D(E_k(m) = u^{-x}.v = y^{-k}.my^k = m$.

2.3.3.3 Digital Signature Schemes: Digital signature scheme is another variant of public-key cryptosystem. The notion of digital signature was first conceived by Whitfield Diffie and Martin Hellman in 1976 [34]. The authors formulated the properties which a digital signature scheme has to satisfy in order to be able to substitute for a hand written signature. By definition, a digital signature is a mathematical strategy for demonstrating the authenticity of digital messages. Digital signature schemes can also be used to realize other objectives of cyber security such as integrity and non-repudiation. In this context, the receiver of the message would have every reason to believe that the message actually originated from the sender and that the message was not altered in transit. In addition, it would be difficult for the sender to deny that the message was sent by him. Formally, a digital scheme consists of three different algorithms which include [20]:

1. Key generation algorithm (generates a public key, K_p and a corresponding private key, K_s.
2. Signature generation algorithm.
3. Signature verification algorithm.

In order to sign a message digitally, the following steps can be taken:

1. A hash function is applied to the message by the sender to generate a hash value, $h = H(m)$.
2. The sender uses this hash value together with his/her private key to generate a digital signature of the message, σ.
3. The sender then appends this signature to the original message and then sends both to the receiving party.
4. The receiver verifies this sender's signature by using the sent signature and the sender's public key to obtain the hash value, h. The receiver compares this computed hash value with the

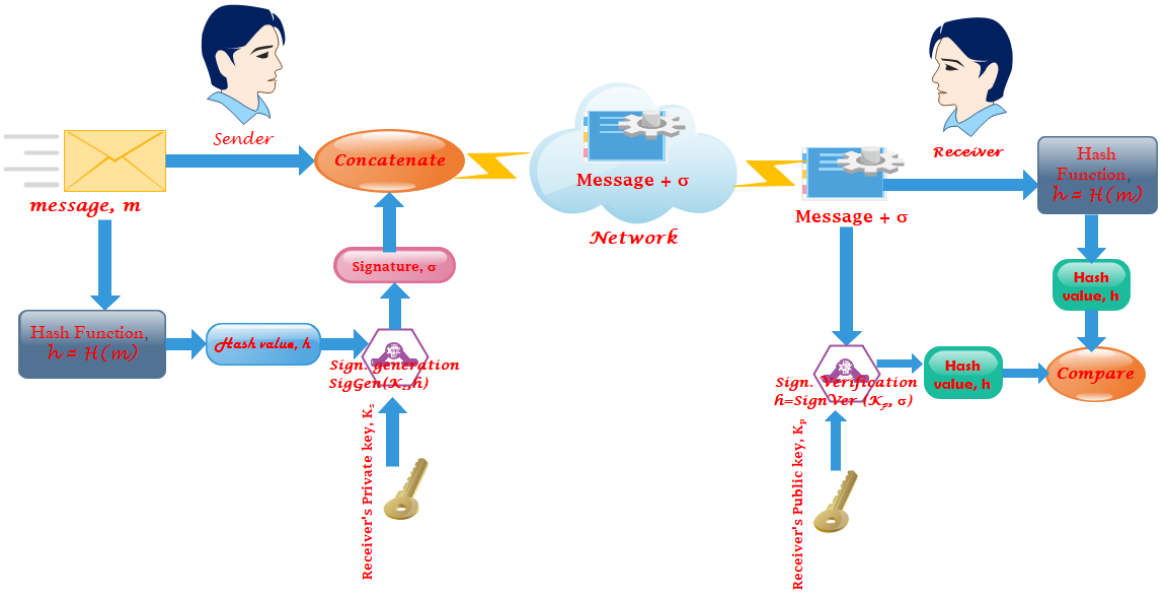

Fig. A.7: Illustration of Digital Signal Generation and Verification)

received value. If the two values are equal, then the signature is adjudged to be valid, else, it is regarded as an invalid signature.

Typical examples of digital signature scheme include: digital signature algorithm [35], Boneh-Lynn-Shacham (BLS) short signature scheme [36], Lamport/Merkle-Winternitz one-time signature schemes [37] and aggregate signature scheme [38]. A general illustration of digital signature generation and verification process is presented in Fig. A.7.

2.3.3.4 Elliptic Curve Cryptography: Elliptic curve cryptography (ECC) is another variant of public-key cryptosystem based on the algebraic structure of elliptic curves over finite fields. Elliptic curve cryptography (ECC) [39, 40] is increasingly used in practice for the design of public-key based cryptographic protocols. ECC requires smaller keys compared to non-ECC cryptography (based on plain Galois fields) to provide equivalent security [41]. The practical benefits of using elliptic curves cannot be neglected after many years of their introduction. They offer smaller key sizes [42] and have more efficient implementations [43] at the same security level as other widely deployed schemes like RSA cryptosystem. A brief explanation of important concepts like elliptic curve definition, elliptic curve key generation process and elliptic curve encryption process has been provided in [44].

2.3.3.4.1 Elliptic Curve Definition: An elliptic curve E over F_p is defined by an equation of the form:

$$y^2 = x^3 + ax + b \tag{A.6}$$

where p is a prime number and F_p denotes the field of integers modulo $p, a, b \in F_p$ satisfy, $4a^3 + 27b^2 \neq 0 \ \ (mod \ \ p)$. Equation A.6 is called Weierstrass normal form for elliptic curves while Fig. A.8 represents two instances of elliptic curve in which $a = -1, b = 0$ and $a = -1, b = -1$ respectively [45]. A pair of (x, y) is a point on the curve if (x, y) satisfies equation A.6 provided $x, y \in F_p$. It is to be noted that the point at infinity (also known as the ideal point), usually denoted by ∞, is also said to be on the curve. As such, equation. A.6 can be refined to produce equation A.7 and given below:

$$E_p(a, b) = (x, y) \in R^2 | y^2 = x^2 + ax + b, 4a^3 + 27b^2 \neq 0 \, \Psi(\infty) \tag{A.7}$$

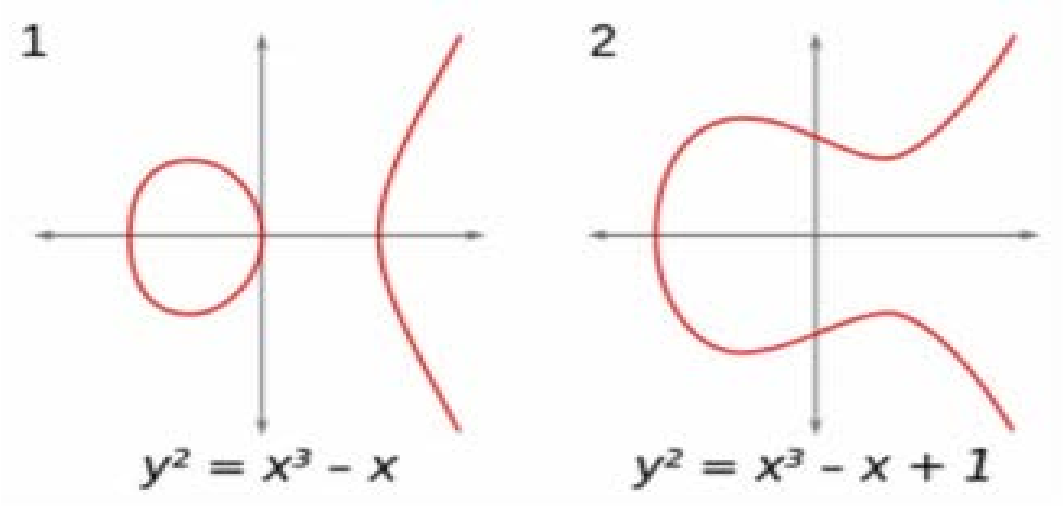

Fig. A.8: Graph showing two Instances of Elliptic Curves [45]

2.3.3.4.2 Elliptic Curve Key Generation: The definition given below provides the necessary basics for the generation of the key for the elliptic curve:

Definition 1: Let E be an elliptic curve defined over a finite field F_p. Let p be a point in $E(F_p)$ and suppose that p has a prime order of n. Then the cyclic subgroup of $E(F_p)$ generated by p is given by: $\langle p \rangle = (\infty, P, 2P, 3P, ...(n-1)P)$. It is to be noted that the prime p in the equation of the elliptic curve, E and the point P and its order, n are collectively referred to as the domain parameters. A private key for the elliptic curve is therefore an integer d that is uniformly selected at random from the interval $[1 : n-1]$, while the corresponding public key would be given by $Q = dP$. However, it is also to be noted that there is a problem known as the discrete logarithm problem (ECDLP) in determining d, given Q and other domain parameters. In summary, the steps for generating the elliptic curve key algorithm include:

(i) Input the elliptic curve domain parameters p, E, P, n. Public key Q and private key d.

(ii) Select $d \in R[1 : n-1]$.

(iii) Calculate $Q = dP$.

(iv) Return (Q, d).

(v) Output elliptic curve public key Q and private key d.

2.3.3.4.3 Elliptic curve Encryption and Decryption Process: The elliptic curve encryption process is started by first representing the plaintext, m, as a point M, and then encrypted by adding it to kQ, where k is an integer that is randomly selected, Q is the recipient's public key. The sender then transmits the point $C_1 = kP$ and $C_2 = M + kQ$ to the receiver who then uses his/her private key d to compute $dC_1 = k(d, P) = kQ$. Thereafter, M is recovered from the expression $M = C_2 - kQ$. The steps for generating the elliptic curve encryption and decryption algorithms can be summarized as follows:

Encryption Algorithm

(i) Input the elliptic curve domain parameters (p, E, P, n), Public key Q and private key d.

(ii) Select $k \in R[1 : n-1]$.

(iii) Compute $C_1 = kP$.

(iv) Compute $C_2 = M + kQ$.

(v) Return (C_1, C_2).

(vi) Output cyphertext (C_1, C_2).

Decryption Algorithm

(i) Input the elliptic curve domain parameters (p, E, P, n), private key d, cyphertext (C_1, C_2).

(ii) Compute $M = (C_2 - dC_1)$.

(iii) Retrieve m from M.

(iv) Return (m).

(v) Output the plaintext, (m).

3 Data Aggregation Schemes in Smart Grid Advanced Metering Infrastructure

A popular research approach for mitigating the effects arising from cyber security threats and privacy violations in a smart grid AMI is the use of data aggregation protocols. In the literature, research solutions proposed on data aggregation based protocols can be categorized as:

(i) Data aggregation solutions for metering data meant for billing purposes.

(ii) Data aggregation solutions for metering data meant for operational purposes and

(iii) Data aggregation solutions meant for the above two purposes.

In the first category, proposed protocol solutions seek to make available smart metering data for billing operations without revealing sensitive information about energy consumers. In the second category, proposed solutions seek to avail smart metering data for control operations such as power quality monitoring without violating the privacy of the consumers by the way the data are obtained, processed or stored. The security threats and privacy violations that could be brought about by the abuse of these data are fundamentally dependent on the frequency at which these data are sampled. Generally, the lower the sampling frequency the less the possibility these data can be attributed to the consumers who own them or the less certainty information concerning these consumers can be deduced. In this survey, the categorization of the data aggregation proposed for the smart grid AMI is dependent on the scheme implemented or the kind of technology implemented. This categorization includes: (1) data aggregation protocols using trusted third parties (TTPs). (2) data aggregation protocols using perturbation technology. (3) protocols using cryptographic algorithms.

3.1 Data Aggregation Protocols using Perturbation Technology

In data aggregation protocols based on the principle of perturbation, the main research idea is to add a kind of randomness to the metering or control data. In such a situation, it would be extremely difficult for the aggregating entity to link the metering data to the smart meters of energy consumers. However, it is expected that aggregated metering data would be calculated with minimal or no error. Typical approaches of data aggregation based protocols using perturbation include: perturbation based on probability distribution (such as binomial or Gaussian), perturbation based on time series using theories such as load signature moderation, theory of rate distortion, and perturbation using orthogonal codes. In [46], Li S. et al encrypted the measured metering data using Walsh orthogonal codes with

ring communication architecture. In the time series based approach using load signature moderation [47], home electrical power routing can be employed to moderate the smart home's load signature so that the appliance usage information could be hidden. On the other hand, the theory of rate distortion which is a sub-field of information theory addresses the problem of lossy compression by analyzing the theoretical fundamentals of determining the bit rate to be communicated over a channel such that the original measured data can be reconstructed at the receiver without or with a negligible distortion error. The taxonomy for data aggregation approaches based on the principle of perturbation is given in Fig. A.9. This taxonomy contains four different categories of research approaches:

1. approach using orthogonal codes [46, 48].
2. approach based on time series theories [47, 49–51].
3. approach based on differential privacy [52–54].
4. approach based on distribution function [52, 53, 55–58].

Notwithstanding the perceived strength of data aggregation protocols based on the principle of perturbation, they generally have some drawbacks. Firstly, the aggregated data (with added randomness) will not be exactly the same as the aggregation of real metering data. Secondly, this approach increases the overhead on computation owing to data processing involved in the recovery of real aggregated metering data. Finally, the certainty of recovering the real aggregated metering data will reduce drastically in the event of one or more smart meters being unable to deliver their randomized metering data to the aggregating entity. In other to overcome these drawbacks, some researchers have resorted to the use of different cryptographic algorithms that can allow entities to aggregate metering data without knowing each smart meter's data.

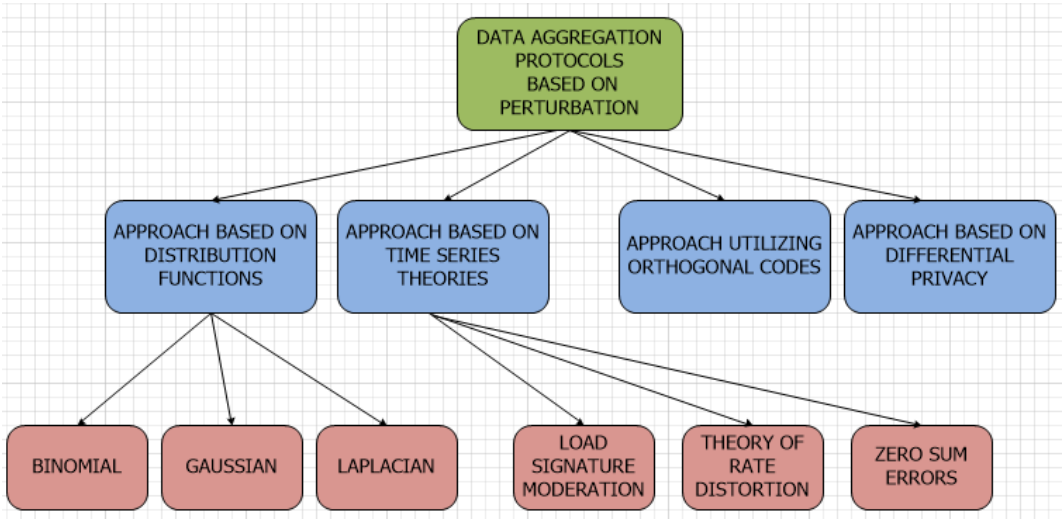

Fig. A.9: Taxonomy of Data Aggregation Protocols Based on Perturbation

3.2 Data Aggregation Protocols using Trusted Third Party

In this data aggregation scheme, a trusted third party is authorized to collect, aggregate and then deliver the metering data to the SG AMI back end systems. Bohli et al [58] introduced this concept of aggregating metering data with a trusted third party acting as an aggregating entity. In order to prevent the violation of end user's privacy by the trusted third party (TTP), the authors suggested that installed smart meters at the consumer's home be incorporated with pseudonyms known only to the data recipients while sending individual meter readings to the TTP. In a similar manner, Kim et al, [51] proposed a related scheme which uses obfuscation function to preserve privacy. This obfuscation function operates on a vector of the smart meter measurements. The measurements are obfuscated in such a way that its original values will be difficult to deduce. In [59], Vetter et al, proposed a hybrid approach which utilizes the cryptographic capabilities of homomorphic encryption and a TTP that manages certificates and cryptographic keys for smart meters. In another development, Fouda et al [60] proposed a lightweight authentication protocol where building area gateway (BA_{GW}) are assumed to be fully trusted and in charge of aggregating metering data from SMs connected to them, encrypts them before forwarding them to their required destination. This proposed aggregation scheme uses Diffie-Hellman key exchange protocol for establishing a secret key shared between a smart meter and its associated gateway. This scheme ensures the confidentiality and integrity of the metering data while transit between the smart meter and the gateway entity.

3.3 Data Aggregation Protocols using Cryptographic Algorithms

In the literature, three popular approaches on data aggregation protocols using cryptographic primitives include:

(i) Schemes using secret sharing.

(ii) Schemes using homomorphic encryption and

(iii) A hybrid scheme combining secret sharing and homomorphic encryption.

An illustration of different schemes for data aggregation protocols using cryptographic algorithms (with references of each scheme as found in the literature) is presented in Fig. A.10. As can be seen from Fig. A.10, there were three references from the reviewed literature for each of the three schemes. Each of the three schemes shown in the above illustration were discussed in more details in the following sub-sections.

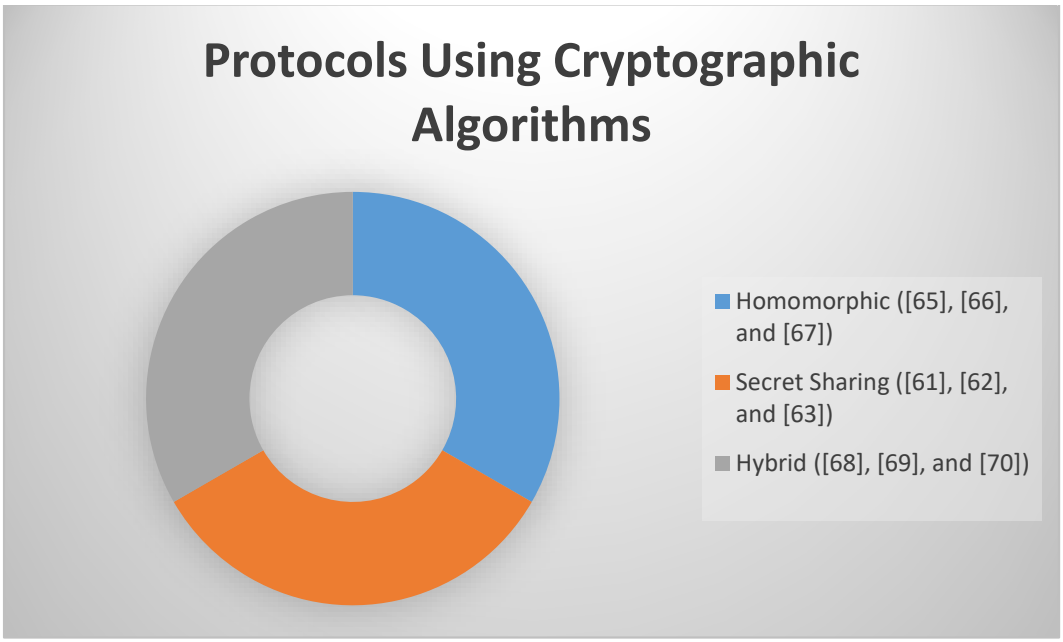

Fig. A.10: Classification of Protocols using Cryptographic Algorithms

3.3.1 Protocols employing the Secret Sharing Schemes

In data aggregation protocols using secret sharing schemes, smart meters split their metering data into different shares. Each share is then sent to a different aggregating node. The aggregating node aggregates the different shares received from different smart meters, and then sends this aggregated value to an authorized data recipient. The final aggregated value of all smart meters is then computed when the data recipient has received the aggregated value from different aggregating nodes. In this way, it will be difficult for the aggregating node and data recipient to compromise the privacy of the metering data since the aggregating nodes receive only a share of the metering data from smart meters while data recipient receives only the aggregated data which consists of shares generated from different smart meters. In [61–63], Rottondi et al proposed a privacy-preserving data aggregation architecture and a communication protocol utilizing the Shamir Secret Sharing encryption scheme. This architecture utilizes gateways placed at the end user's premises which collect data from smart meters. The gateways communicate with each other and with other external entities by means of a public data network. Simulations carried out in this work show that the proposed protocol utilizing Shamir Secret Sharing (SSS) scheme is a feasible solution with regards to the problem of privacy-preserving aggregation of energy consumption data. The major drawbacks in these schemes include (1) high overhead in communication as a result of the secret sharing scheme (2) high overhead in computation at smart meters as each smart meter has to establish communication with each privacy-preserving nodes (PPNs).

3.3.2 Protocols using Homomorphic Encryption

F. Li et al [64] proposed an in-network data aggregation which utilizes smart meters to aggregate user's encrypted metering data. In this scheme, aggregation is done in a distributed manner such that each smart meter node collects data from its children, aggregates them with its own data, and transmits the intermediate result to the parent node. To preserve the privacy of the energy consumption metering data, homomorphic encryption is employed so that inputs and intermediate results are not revealed to smart meters within the aggregation path. It is important to note that this scheme can be utilized to achieve a considerable level of scalability. However, the scheme can only protect end user's consumption data against passive attacks. In [65, 66], an efficient privacy-preserving scheme which uses a homomorphic encryption for demand response in order to realize a privacy-preserving demand aggregation and efficient response was proposed. In this scheme also, an adaptive key evolution technique was examined to guarantee that the users' session keys be forwarded in a secure manner. Finally, Borges et al [67], proposed a protocol that permits authorized data recipients achieve aggregated metering data in a way which preserves the privacy of the data. This protocol illustrates how data aggregation and smart billing can be implemented in a privacy-preserving way. The proposed protocol grants the benefits of data aggregation of the measurements and allows billing with time-based pricing. This proposed work combines homomorphic commitment scheme with a homomorphic encryption scheme. Unfortunately, this scheme also suffers from high overhead in communication and computations. For instance, each smart meter at each time slot, needs to perform $5 + n$ computationally expensive operations made up of one commitment, two signatures, two encryptions and n verifications.

3.3.3 Hybrid Protocols combining Homomorphic Encryption and Secret Sharing

In this section, a brief review of few hybrid protocols that combine both homomorphic encryption and secret sharing. Garcia et al [68] proposed a aggregation scheme which combines the secret sharing scheme and a homomorphic encryption in order to preserve privacy. This protocol defined two roles: (1) first role for the utility company (UC) and (2) second role for customers. In this context, the UC acts as the aggregator. In this protocol, each customer splits their consumption measurements into a random number of shares which is equal to the number of participants. For instance, assuming a three customer participants with smart meters, SM_1, SM_2, SM_3 respectively, and a UC. The customer with SM_1 splits her metering consumption data into three secret shares:

$$Customer_1(SM_1(ss_1) + SM_2(ss_2) + SM_3(ss_3) \mod n \tag{A.8}$$

Where n is a large integer. The customer then keeps the first secret share, $SM_1(ss_1)$ to herself and then encrypts $SM_2(ss_2)$ and $SM_3(ss_3)$ with the public keys of other two customers, and then send the encrypted message to UC. The utility company, UC then receives the encrypted shares from SM_1, SM_2, SM_3, adds the shares which are encrypted with the same public key using the homomorphic encryption properties. The UC then sends the added shares to $Customer_1(SM_1)$. who can now decrypt it using her secret key in order to obtain the plaintexts. Each of the two participants does the same thing after receiving the added shares from UC. Each participant then sends the decrypted plaintexts to UC who then computes the total consumption. This proposed protocol achieves privacy goal as the UC cannot have access to the private measurements from the participant's smart meters. Unfortunately, the reliance of this protocol on secret sharing scheme increases the overheads in computation and communication. Similar protocols utilizing this hybrid approach can be found in [69, 70].

4 Open Issues and Future Challenges on Data Aggregation for SG AMI

In this section, a summary of open issues and future challenges with regards to data aggregation for the SG AMI is presented. From the review carried out in this paper, it can be observed that the state of research in the domain of data aggregation for the SG AMI is really commendable. Notwithstanding, there is obvious indication that research in this area will continue to evolve, considering the popularity of data aggregation approach for providing privacy in SG AMI. Based on this assumption, an analysis based only on the papers reviewed in this work which concentrated on data aggregation protocols is presented. A total of twenty-six research papers were considered in this analysis. Based on this, an important classification has been presented in Fig. A.11. The following observation have been made as a result:

- Many of the privacy-preserving data aggregation protocols were not designed to have a true reflection of smart grid AMI environment.
- As a result of the above point, many of the proposals were designed for a single authorized entity data recipient. In reality, documents from standard bodies for the implementation of the SG AMI show that there are many AMI-back end systems who may be authorized to have access to consumption data for purposes such as billing and other grid control purposes. From Fig.A.11, it can be seen that protocol design in 19 percent of the research papers under consideration, were done without good hint on how many authorized entities are expected to have access to consumption data.

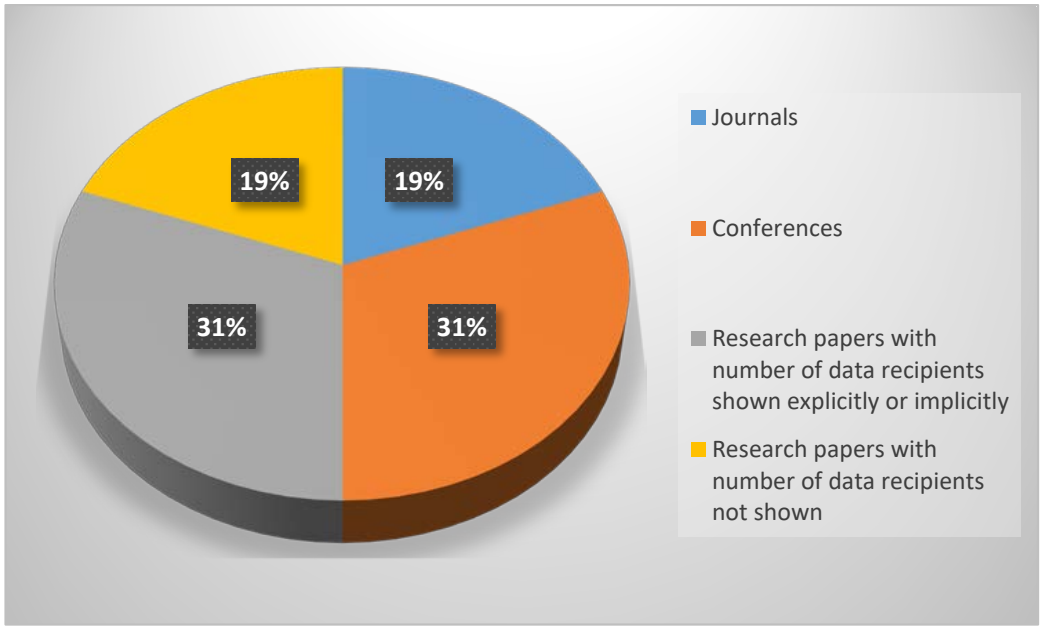

Fig. A.11: Classification of research papers focusing on data aggregation

- From table A.1, it can also be seen that out of the 31 percent of the research proposals that indicated implicitly or explicitly the number of authorized entity data recipients, only 25 percent of these proposals indicate that multiple entities are expected to access user's consumption data.
- It was observed that in all the reviewed proposals on privacy-preserving data aggregation protocols, justifications on overheads in communications and computations were usually based on complexity analysis premised on the number and/or the complexity of primitive operations performed in those protocols.

Based on the above observations, we provide the following future challenges for research in the domain of privacy-preserving data aggregation protocols for SG AMI:

(i) Research proposals on privacy-preserving data aggregation protocols for the SG AMI should be designed to show explicitly the incorporation of multiple data recipients into the proposed models.

(ii) Notwithstanding the popularity of the data aggregation approach towards providing privacy for consumption data, this approach can drive the SG AMI network into a congestive scenario owing to increased overheads on computations and communications. This situation can abate cyber-criminals with the intension of launching data availability attacks against the network. As a result, future researchers ought to justify that these protocols are indeed lightweight.

(iii) It is therefore our considered opinion that the best way of resolving the issue raised in (ii)

above is to implement the designed protocols on wireless communication network standards like Wi-Fi, ZigBee or any wired communication standard which have been recommended for SG AMI communications. In doing this, rather than focus on the complexity analysis of primitive operations utilized in the protocols as justifications for low overheads, quality of service (QoS) performance analysis of the networks ought to be carried out.

(iv) In order to improve the QoS performance of these networks in the presence of these protocols, schemes which tends to minimize congestions in these networks ought to be designed and implemented together with the data aggregation protocols on the wireless or wired communication network of choice.

Table A.1: Analysis on research papers showing the number of data recipients expected

Proposal	Includes multiple entities	Applied to a standard wireless or wired network
Li. S. et al [46]	No	No
Y. Kim et al [51]	No	No
L. Lyu et al [53]	No	No
J. Ni et al [54]	No	No
W. Jia et al [55]	No	No
M. Bohli et al [58]	No	NO
B. Vetter et al [59]	yes	No
C. Rottondi et al [61]	yes	No
C. Rottondi et al [62]	yes	No
C. Rottondi et al [63]	yes	No
F. Li et al [64]	No	No
H. Li et al [65]	No	No
H. Li et al [66]	No	NO
F. Borges et al [67]	No	No
F.D Garcia et al [68]	No	No
Z. Erkin et al [69]	No	No

5 Conclusion

There have been a lot of research efforts directed at improving security and privacy of consumption data and control information against different types of attacks and privacy violations in a SG AMI. The security and privacy issues for the AMI had been compounded in many countries where laws for the protection of private information had been enacted. Research has revealed that consumption data from electricity user's smart meters can be profiled to reveal sensitive information about the customers. It is believed that this sensitive information can be maliciously exploited by cyber-criminals to further launch different ranges of cyber-attacks on consumers. Stakeholders involved in driving the smart grid vision may risk being resisted by final consumers in their efforts towards the mass deployments of the AMI in different countries if these security and privacy issues remain unresolved. Data aggregation of metering data has been a popular research approach towards preserving privacy in SG AMI. In this paper, an extensive review on data aggregation protocols proposed for the SG AMI has been carried out. It was discovered that research in this domain has evolved considerably. However, open issues which can be addressed in order to improve research in this area were highlighted and discussed.

References

[1] 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.

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

[3] B. Botte, V. Cannatelli, and S. Rogai, "The telegestore project in enel's metering system," in *Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*. IET, 2005, pp. 1–4.

[4] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building*. ACM, 2010, pp. 61–66.

[5] F. H. Cate, "The failure of fair information practice principles," 2007. [Online]. Available: https://poseidon01.ssrn.com/delivery.php

[6] M. Lee, O. Aslam, B. Foster, D. Kathan, J. Kwok, L. Medearis, R. Palmer, P. Sporborg, and M. Tita, "Assessment of demand response and advanced metering," *Federal Energy Regulatory Commission, Tech. Rep*, 2013.

[7] M. Calaguas, "South african parliament enacts comprehensive data protection law: An overview of the protection of personal information bill," *Africa Law Today*, vol. 7, 2013.

[8] "Advanced metering infrastructure and customer systems,results from the smart grid investment grant program." [Online]. Available: https://www.smartgrid.gov/files/Final_SGIG_Report_20161220.pdf

[9] W. Zhang and L. Yang, "Sc-fdma for uplink smart meter transmission over low voltage power lines," in *Power Line Communications and Its Applications (ISPLC), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 497–502.

[10] R. Mahmud, R. Vallakati, A. Mukherjee, P. Ranganathan, and A. Nejadpak, "A survey on smart grid metering infrastructures: Threats and solutions," in *Electro/Information Technology (EIT), 2015 IEEE International Conference on*. IEEE, 2015, pp. 386–391.

[11] D. Kundur, X. Feng, S. Liu, T. Zourntos, and K. L. Butler-Purry, "Towards a framework for cyber attack impact analysis of the electric smart grid," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 244–249.

[12] Y. Mo, T. H.-J. 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, 2012.

[13] M. A. Rahman, P. Bera, and E. Al-Shaer, "Smartanalyzer: A noninvasive security threat analyzer for ami smart grid," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 2255–2263.

[14] S. Goel and Y. Hong, "Security challenges in smart grid implementation," in *Smart Grid Security*. Springer, 2015, pp. 1–39.

[15] F. Aloul, A. Al-Ali, R. Al-Dalky, M. Al-Mardini, and W. El-Hajj, "Smart grid security: Threats, vulnerabilities and solutions," *International Journal of Smart Grid and Clean Energy*, vol. 1, no. 1, pp. 1–6, 2012.

[16] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[17] J. Kim and L. Tong, "On topology attack of a smart grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*. IEEE, 2013, pp. 1–6.

[18] E. L. Quinn, "Privacy and the new energy infrastructure," 2009. [Online]. Available: https://ssrn.com/abstract=1370731

[19] S. D. Fugita, F. A. Borges, R. A. Fernandes, and I. N. da Silva, "Methodology based on smart meters applied to the identification of residential loads," in *Electronic System-Integration Technology Conference (ESTC), 2012 4th*. IEEE, 2012, pp. 1–5.

[20] S. William, *Cryptography and network security: principles and practices*. Pearson Education India, 2006.

[21] N. F. Pub, "197: Advanced encryption standard (aes)," *Federal information processing standards publication*, vol. 197, no. 441, p. 0311, 2001.

[22] P. Gallagher and A. Director, "Secure hash standard (shs) fips pub 180-4," *Federal Information Processing Standards Publication, FIPS PUB*, vol. 1, 2012.

[23] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[24] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE transactions on information theory*, vol. 31, no. 4, pp. 469–472, 1985.

[25] R. Cramer and V. Shoup, "A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack," in *Annual International Cryptology Conference*. Springer, 1998, pp. 13–25.

[26] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 1999, pp. 223–238.

[27] R. Lu, *Privacy-enhancing aggregation techniques for smart grid communications*. Springer, 2016.

[28] T. Sridokmai and S. Prakancharoen, "The homomorphic other property of paillier cryptosystem," in *Science and Technology (TICST), 2015 International Conference on*. IEEE, 2015, pp. 356–359.

[29] M. Marwan, A. Kartit, and H. Ouahmane, "Towards a secure cloud database using paillier's homomorphic cryptosystem," in *Multimedia Computing and Systems (ICMCS), 2016 5th International Conference on*. IEEE, 2016, pp. 360–365.

[30] X. Yi, R. Paulet, and E. Bertino, *Homomorphic encryption and applications*. Springer, 2014, vol. 3. [Online]. Available: https://link.springer.com/book/10.1007/978-3-319-12229-8

[31] M. A. Mustafa, N. Zhang, G. Kalogridis, and Z. Fan, "Dep2sa: A decentralized efficient privacy-preserving and selective aggregation scheme in advanced metering infrastructure," *IEEE Access*, vol. 3, pp. 2828–2846, 2015.

[32] W. Mao, *Modern cryptography: theory and practice*. Prentice Hall Professional Technical Reference, 2003.

[33] B. Allen, *Implementing several attacks on plain ElGamal encryption*. Iowa State University, 2008.

[34] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[35] P. Gallagher, "Digital signature standard (dss)," *Federal Information Processing Standards Publications, volume FIPS*, pp. 186–3, 2013.

[36] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the weil pairing," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2001, pp. 514–532.

[37] V. I. Villányi, "Promising one-time signature schemes (survey)," in *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on*. IEEE, 2010, pp. 187–192.

[38] D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and verifiably encrypted signatures from bilinear maps," in *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2003, pp. 416–432.

[39] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of computation*, vol. 48, no. 177, pp. 203–209, 1987.

[40] V. S. Miller, "Use of elliptic curves in cryptography," in *Conference on the theory and application of cryptographic techniques*. Springer, 1985, pp. 417–426.

[41] Z. Tang, "Measurement-device-independent quantum cryptography," Ph.D. dissertation, University of Toronto (Canada), 2016.

[42] A. K. Lenstra and E. R. Verheul, "Selecting cryptographic key sizes," *Journal of cryptology*, vol. 14, no. 4, pp. 255–293, 2001.

[43] D. Bernstein and T. Lange, "ebacs: Ecrypt benchmarking of cryptographic systems, accessed on may 19, 2015." [Online]. Available: https://bench.cr.yp.to/

[44] D. Hankerson, A. J. Menezes, and S. Vanstone, *Guide to elliptic curve cryptography*. Springer Science & Business Media, 2006.

[45] X. Fang, G. Yang, and Y. Wu, "Research on the underlying method of elliptic curve cryptography," in *Information Science and Control Engineering (ICISCE), 2017 4th International Conference on*. IEEE, 2017, pp. 639–643.

[46] S. Li, K. Choi, and K. Chae, "Ocpm: Ortho code privacy mechanism in smart grid using ring communication architecture," *Ad Hoc Networks*, vol. 22, pp. 93–108, 2014.

[47] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 232–237.

[48] Y. Yan, Y. Qian, and H. Sharif, "A secure and reliable in-network collaborative communication scheme for advanced metering infrastructure in smart grid," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. IEEE, 2011, pp. 909–914.

[49] D. Gündüz, J. Gomez-Vilardebo, O. Tan, and H. V. Poor, "Information theoretic privacy for smart meters," in *Information Theory and Applications Workshop (ITA), 2013*. IEEE, 2013, pp. 1–7.

[50] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor, "Smart meter privacy: A utility-privacy framework," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*. IEEE, 2011, pp. 190–195.

[51] Y. Kim, E. C.-H. Ngai, and M. B. Srivastava, "Cooperative state estimation for preserving privacy of user behaviors in smart grid," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*. IEEE, 2011, pp. 178–183.

[52] J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving differential privacy of data disclosure in the smart grid," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 504–512.

[53] L. Lyu, Y. W. Law, J. Jin, and M. Palaniswami, "Privacy-preserving aggregation of smart metering via transformation and encryption," in *Trustcom/BigDataSE/ICESS, 2017 IEEE*. IEEE, 2017, pp. 472–479.

[54] J. Ni, K. Zhang, K. Alharbi, X. Lin, N. Zhang, and X. S. Shen, "Differentially private smart metering with fault tolerance and range-based filtering," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2483–2493, 2017.

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

[56] H.-Y. Lin, W.-G. Tzeng, S.-T. Shen, and B.-S. P. Lin, "A practical smart metering system supporting privacy preserving billing and load monitoring," in *International Conference on Applied Cryptography and Network Security*. Springer, 2012, pp. 544–560.

[57] X. Ren, X. Yang, J. Lin, Q. Yang, and W. Yu, "On scaling perturbation based privacy-preserving schemes in smart metering systems," in *Computer Communications and Networks (ICCCN), 2013 22nd International Conference on*. IEEE, 2013, pp. 1–7.

[58] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Communications Workshops (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[59] B. Vetter, O. Ugus, D. Westhoff, and C. Sorge, "Homomorphic primitives for a privacy-friendly smart metering architecture." in *SECRYPT*, 2012, pp. 102–112.

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

[61] C. Rottondi, M. Savi, D. Polenghi, G. Verticale, and C. Krauß, "Implementation of a protocol for secure distributed aggregation of smart metering data," in *Smart Grid Technology, Economics and Policies (SG-TEP), 2012 International Conference on*. IEEE, 2012, pp. 1–4.

[62] C. Rottondi, G. Verticale, 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, 2013.

[63] C. Rottondi, G. Verticale, and A. Capone, "Privacy-preserving smart metering with multiple data consumers," *Computer Networks*, vol. 57, no. 7, pp. 1699–1713, 2013.

[64] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 327–332.

[65] H. Li, X. Liang, R. Lu, X. Lin, and X. Shen, "Edr: An efficient demand response scheme for achieving forward secrecy in smart grid," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 929–934.

[66] H. Li, X. Lin, H. Yang, X. Liang, R. Lu, 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, 2014.

[67] F. Borges, D. Demirel, L. Böck, J. Buchmann, and M. Mühlhäuser, "A privacy-enhancing protocol that provides in-network data aggregation and verifiable smart meter billing," in *Computers and communication (ISCC), 2014 IEEE symposium on*. IEEE, 2014, pp. 1–6.

[68] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *International Workshop on Security and Trust Management*. Springer, 2010, pp. 226–238.

[69] Z. Erkin and G. Tsudik, "Private computation of spatial and temporal power consumption with smart meters," in *International Conference on Applied Cryptography and Network Security*. Springer, 2012, pp. 561–577.

[70] K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 2011, pp. 175–191.

Chapter III

ZigBee-Based Smart Grid Advanced Metering Infrastructure: Overview of Security Issues

Paper B

ZigBee-Based Smart Grid Advanced Metering Infrastructure: Overview of Security Issues

R.C Diovu and J.T Agee

Accepted and in press

Sensor Letters (American Scientific Publishers)

Abstract

In the last decade, ZigBee technology has emerged among smart grid proponents as a potential candidate for the smart grid advanced metering infrastructure (AMI) applications. The AMI requires reliable communication between the smart meters and the intelligent home appliances that support increased participation of users in the overall energy value chain. One of the key attractions of ZigBee for the AMI applications is its open standard which presents a platform for interoperability among multiple vendors and systems. In this paper, a survey of the current state of research on ZigBee based SG AMI is carried out. However, the focus of this paper is on security and privacy issues surrounding ZigBee based smart grid AMI with the intension of discovering potential areas of vulnerabilities. At the end, open research issues were highlighted and discussed.

1 Introduction

The advanced metering infrastructure (AMI) is the most significant component of the smart grid [1]. A typical architecture of the AMI has been presented in Fig. B.1. With reference to Fig. B.1, the components of the architecture presented in the figure include smart meters, In-Home Displays, Neighborhood Area Network Gateways and AMI concentrators which perform the function of aggregating consumption data transmitted through the NAN GW. Other key components include the AMI master station and the AMI-back-end systems (such as billing systems and meter data management system). It should be noted that consumption data or control information from SM/IHD can be transmitted directly to the NAN GW in the absence of gateways connected to the SM/IHD homes. The advantages derivable from the AMI have motivated utilities to invest huge sums of money for the deployment of the AMI in different countries. In the United States for example, the Smart Grid Investment Grant (SGIG) was launched in 2009 and funded with $7.9 billion to kick-start the realization of the smart grid vision [2]. Approximately 56.2 percent of the fund provided was invested on the advanced metering infrastructure. Fig. B.2 shows the breakdown of how the provided fund was invested. The AMI can be integrated with multiple information and management systems using appropriate communication technology (like ZigBee) in order to unlock new functions that can enhance grid operation efficiency. Fig.B.3 shows the details of the integration of the AMI with key information and management systems for the 60 AMI projects analyzed in the SGIG program. The choice of ZigBee technology for the SG AMI may depend on the topology of the proposed network. As can be seen in this paper, ZigBee may be an ideal choice for the SG AMI for either the HAN or

NAN sections of the network. ZigBee technology is a protocol specification and industry standard for a wireless networking technology otherwise referred to as Low-Rate Wireless Personal Area Network (LR-WPAN) [3]. A typical LR-WPAN is characterized by low power, low cost self organizing wireless devices communicating within a short range in a wireless communication channel. LR-WPAN has been seen by many as an ideal communication network for supporting relatively low throughput applications like smart metering in an AMI network [4]. The network can be organized into different topologies ranging from single-hop star topology to multi-hop mesh topology depending on applications and its design requirements. Fig.B.4 shows the ZigBee technology support for tree and mesh topologies. The main network component featuring in Fig. B.4 include the ZigBee coordinator which normally functions as the parent node to other nodes connected to the network, the router and the end device. The routers and end devices function as the children of the coordinator. It is important to note that the network size in any of the topologies shown in Fig. B.4 can be extended by connecting as much routers as desired. The proponents of ZigBee for SG AMI applications found the technology irresistible and appealing because of its potential for relatively fast transmission, low cost and simplicity of implementation when compared to other traditional wired communication networks. This is particularly true because such wireless network requires minimum cabling infrastructure, permits zero configurations where appliances can be easily removed or added and can inter-connect easily with other wireless based networks [5]. J. Garcia et al [6] reported that there are already large deployments of ZigBee in Mexico for the AMI at the NAN domain. A summary of the deployments for AMI pilot projects for the AMI HAN in Mexico according to [6] is presented in Fig. B.5. Bian et al [7] also captured some selected worldwide SG AMI deployments. A summary of such deployments for the smart meter network (AMI) is presented in Fig. B.6. In a similar way, RF communication technology like ZigBee has been recommended for the AMI front-end communication in South Africa [8]. This evidence of ZigBee based AMI deployments thus far, shows its evolution and the level of attraction of the technology to industry experts. The SG AMI is responsible for the communication of critical information needed for purposes such as billing, pricing schemes, demand response, power quality and other grid control operations. The network could become vulnerable to cyber-attacks like replay attacks, same nonce attacks, spoofing or denial of service attacks [9, 10], if the network is integrated with numerous low cost nodes deployed in insecure locations within the ZigBee network. The same can apply if the network is implemented with poor network configurations. In such circumstance, end user's consumption data needed for different purposes can be stolen or hacked by cyber criminals. This can subsequently reveal sensitive information about consumers such as names, user's address or other personal information which can be utilized to launch various degrees of attack against consumers. This situation is further

compounded in South Africa where a bill known as Protection of Personal Information (POPI) Act has been signed into Law since 2013 [11]. With the near expiration of the grace period for the compliance of this bill, private as well as corporate organizations such as Eskom and municipalities in charge of electric power distributions are expected to comply with this bill in its current deployment of the AMI in the country. In order to avoid or reduce to the barest minimum potential vulnerabilities in the network, the ZigBee specifications [12] have defined standards for ensuring the security of message transmitted over the network. In a nutshell, this standard uses encryption and authentication protocols for providing security requirements such as confidentiality, integrity and availability which are the main objectives of cyber-security for the smart grid according to the National Institute for Standards and Technology (NIST) [13]. A conceptual diagram representing the cyber security objectives and their associated cyber attacks has been presented in Fig. B.7 and B.8 respectively. ZigBee recommends the use of 128 bits mode Advanced Encryption Standard (AES) [14]. The security features defined in ZigBee network spans across the different protocol layers. Details of these features are provided in section II of this paper. The focus of this paper is to review the security features of ZigBee in order to appraise same in terms of their strengths and weaknesses with regards to an SG AMI environment. In addition, this paper introduces ZigBee based smart grid AMI and some research proposals found in the literature that are geared towards the realization of the smart grid vision. To the best of our knowledge, surveys or review works on ZigBee based smart grid AMI having extensive focus on security is very scarce. We provide a brief summary of research works on ZigBee based SG AMI and provide appropriate classifications of such works. Such classifications is helpful in gaining a thorough understanding of what have been done in the subject area before now. Finally, this paper highlights open issues for future research with regards to the security of a ZigBee based SG AMI. It is our hope that this research paper will provide interesting overview about the state-of-the-art of ZigBee based SG AMI which will be beneficial to future researchers, utility companies and standard bodies in power industries in many countries.

2 Overview of the Zigbee protocol Suite

The ZigBee protocol suite is composed of layers which are arranged in a hierarchical manner. Each layer performs a specific service or set of services for layers above it in hierarchy. As presented in Fig.B.9, the lowest two layers were introduced and defined by the IEEE 802.15.4 standard while the remaining upper layers were defined by the ZigBee Alliance. The PHY layer is responsible for modulation, demodulation and for transmission of packets over the air. The MAC layer handles collision issues using CSMA/CA feature during frame transmission. This layer is also responsible for

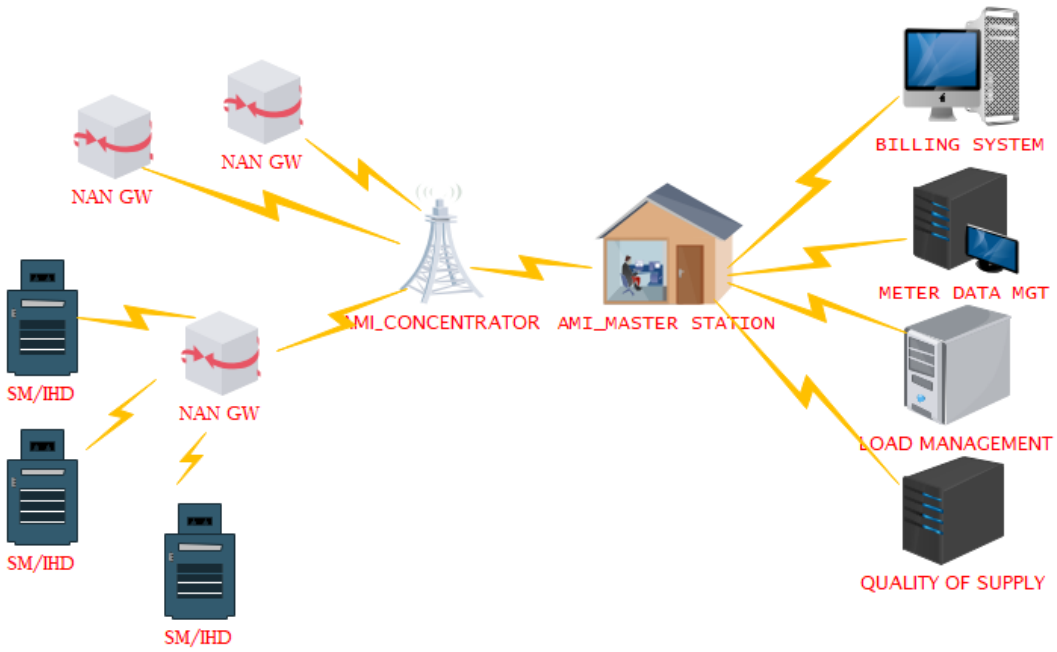

Fig. B.1: A Typical Architecture of Smart Grid AMI Network

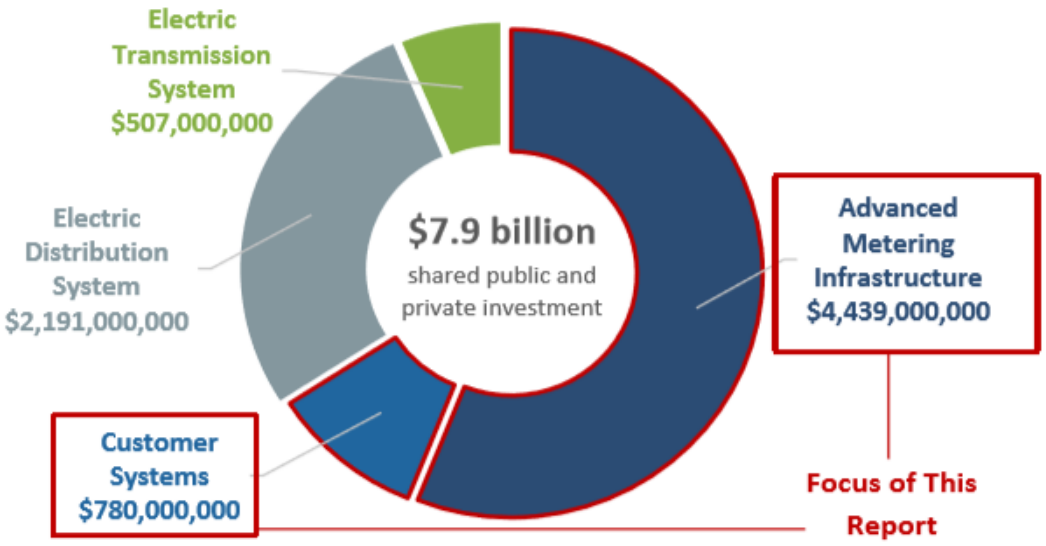

Fig. B.2: Breakdown of 7.9 USD SGIG Investment [2]

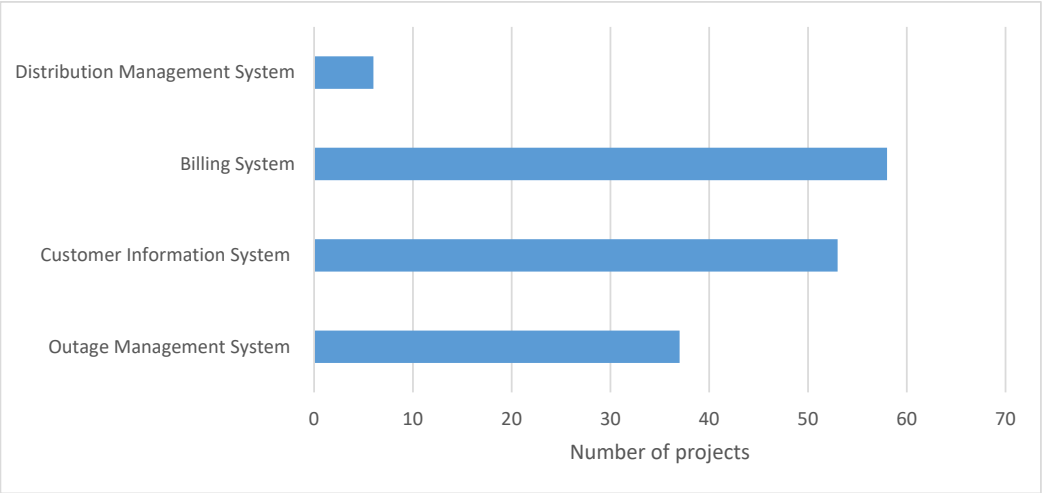

Fig. B.3: Details of AMI projects in the SGIG with regards to key Information Management Systems [2]

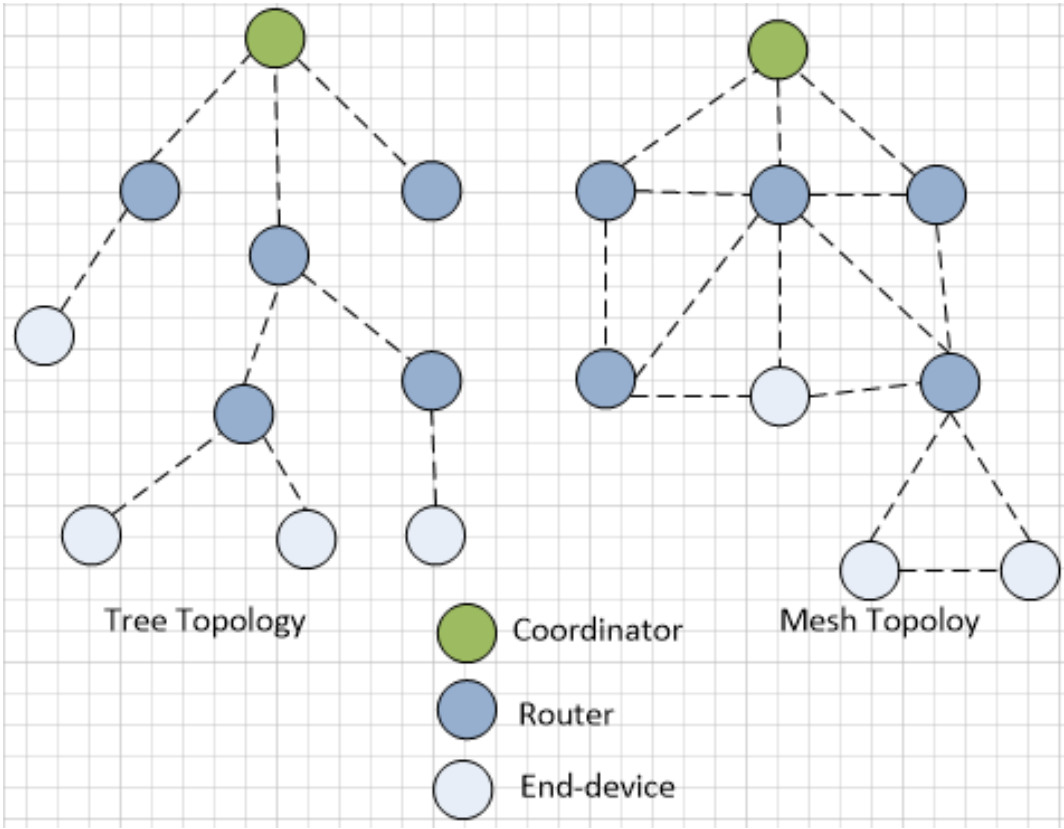

Fig. B.4: Typical Network Topologies supported by ZigBee

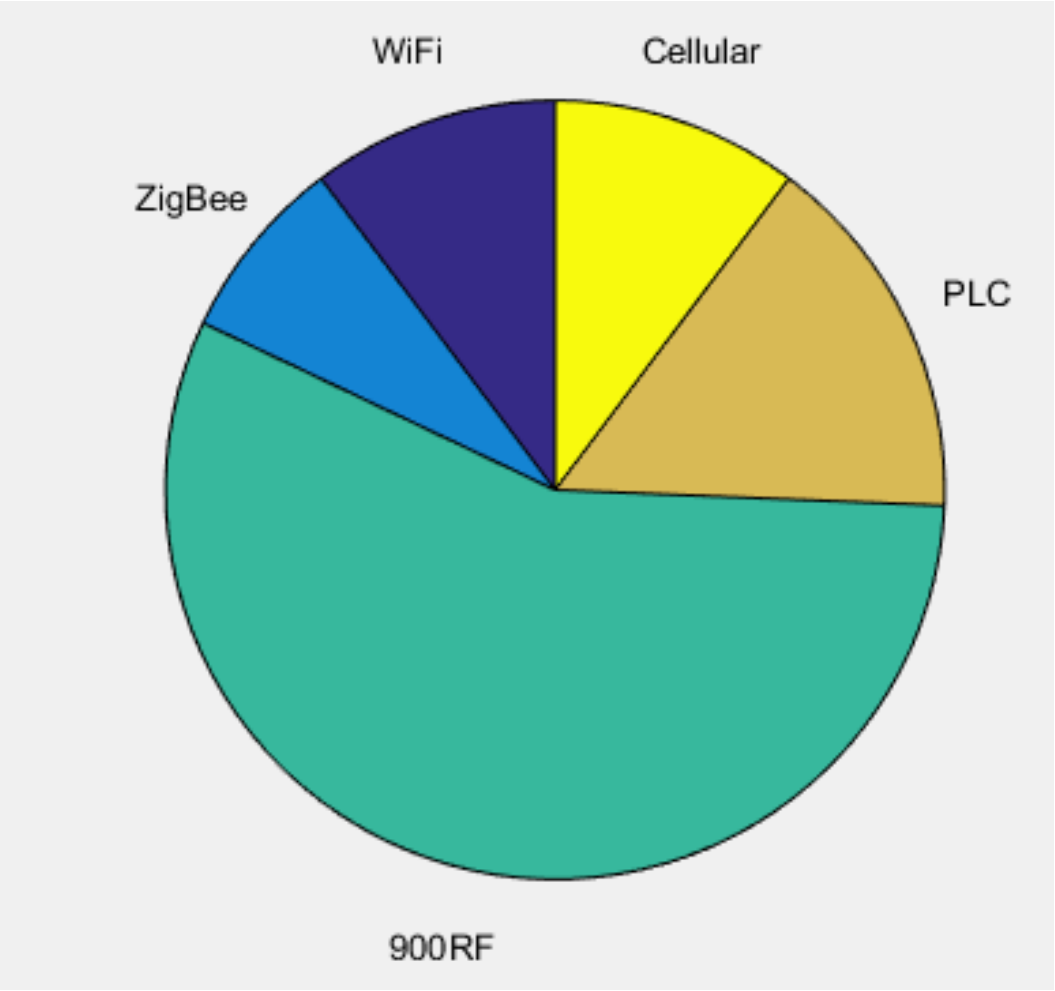

Fig. B.5: A summary of AMI Projects depicting communication technologies for smart meter network

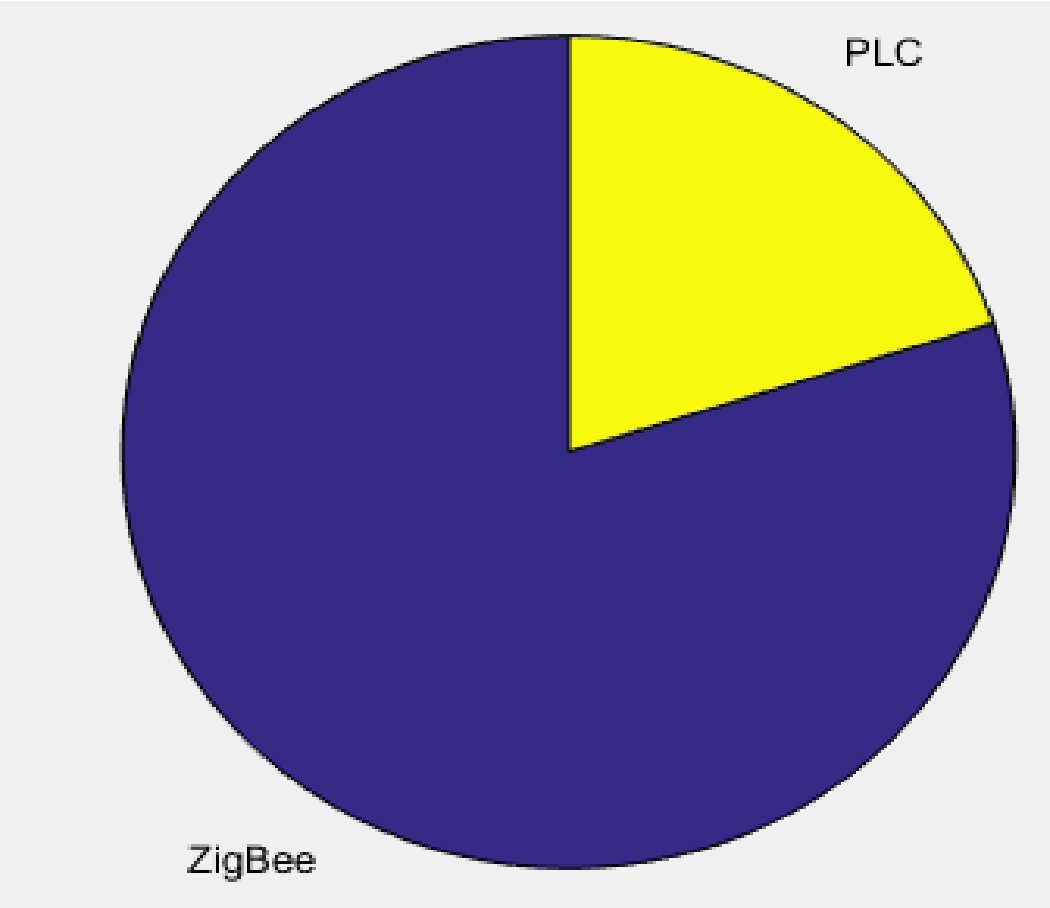

Fig. B.6: A summary of AMI nationwide pilot projects deployment in Mexico for the AMI HAN

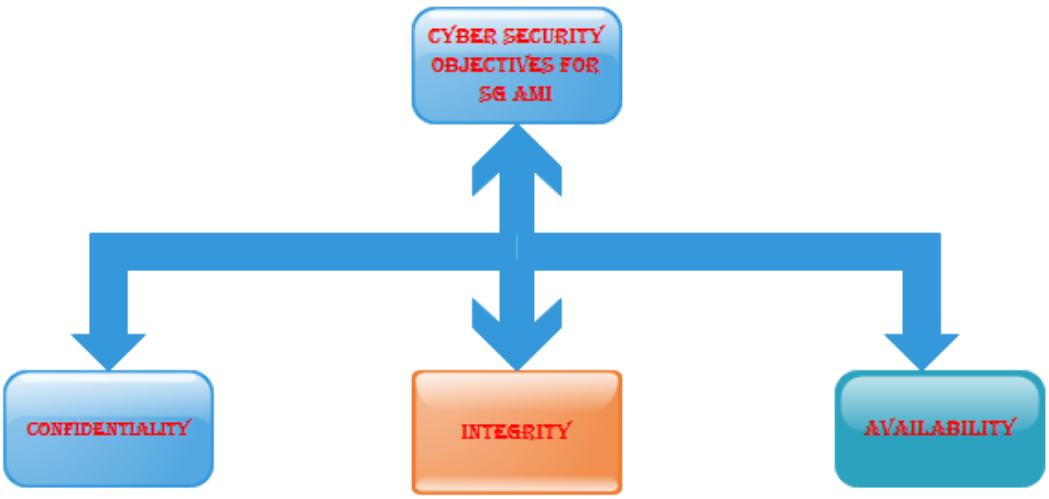

Fig. B.7: A Taxonomy of NIST Cyber-Security Objectives (goals)

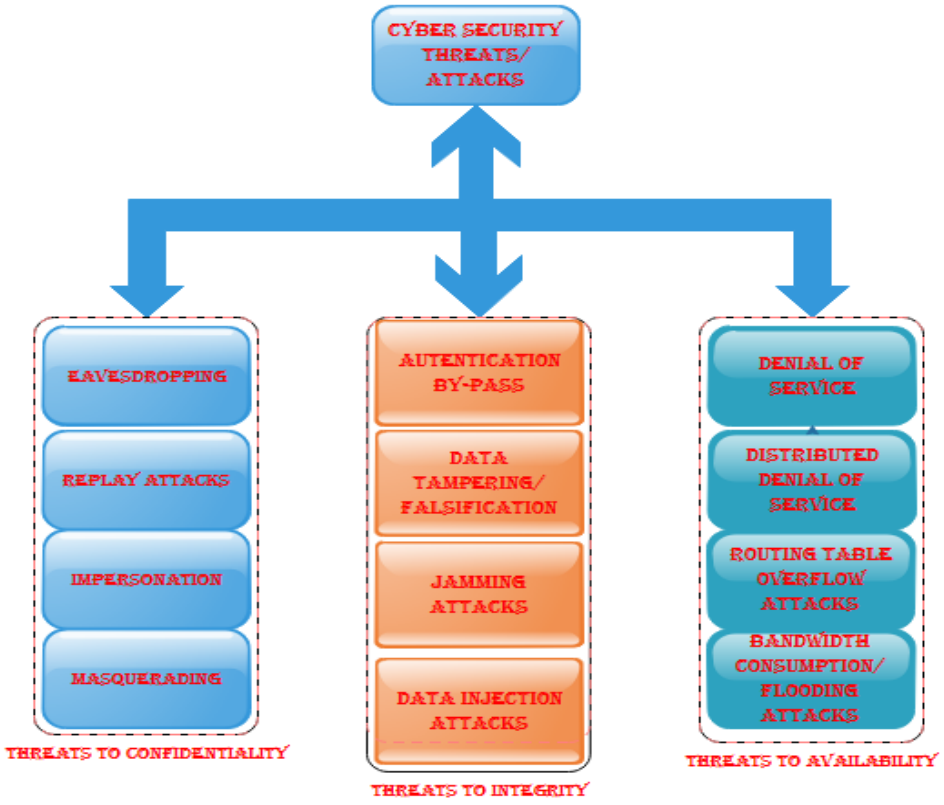

Fig. B.8: Taxonomy of Cyber Attacks with regards to Cyber-Security Objectives

defining frame formats by assigning MAC addresses. The network layer (NWK) on the other hand is responsible for handling routing related tasks and for assigning network addresses to devices in the network. The application (APL) layer is the top most layer in the ZigBee network protocol hierarchy and this layer hosts the application objects. Application objects are developed to customize nodes for various applications. These applications objects manage and control the protocol layers in a ZigBee device. A single ZigBee device can have as many as 240 application objects [15]. In ZigBee networking domain, each of the protocol layer such as MAC, NWK and APS is responsible for providing security for the frames or packets initiated by that layer. By convention, the same security key is used by these layers in a single ZigBee node.

3 Overview of General Research on ZigBee Based SG AMI

In this section, we provide an overview of previous works by researchers on ZigBee based smart grid AMI. The objective of the section is to provide the current state-of-the-art with regards to research on ZigBee based smart grid AMI. H. Shabani et al [16] proposed a smart ZigBee/IEEE 802.15.4 MAC for wireless sensor multi-hop mesh networks. The authors described the construction of a smart

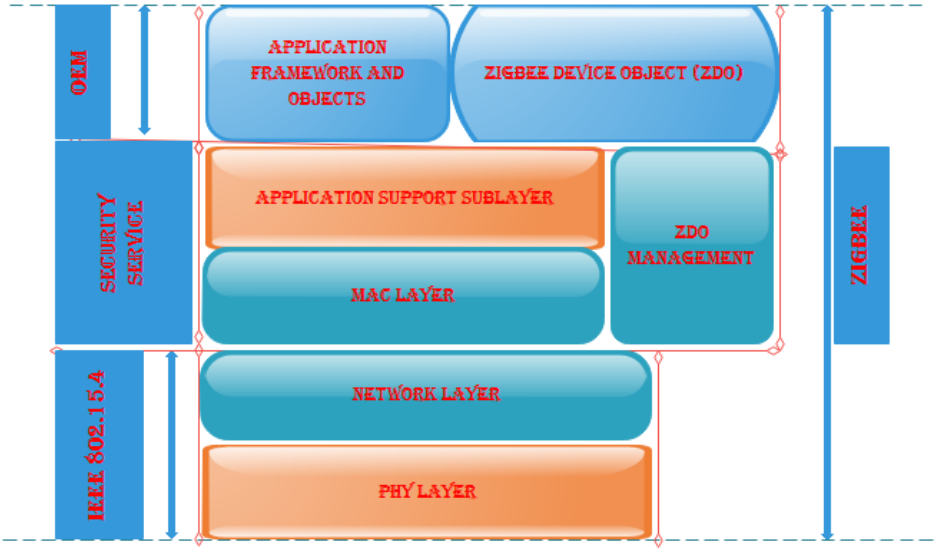

Fig. B.9: ZigBee Protocol Suite

ZigBee-based beacon-enabled mesh network for smart grid auto-metering applications with the goal of improving time delay, duty cycle, energy saving and reliability performance of the network. The authors argued that although, ZigBee beacon-based mode of operation was adopted in IEEE 802.15.4 standard, there was no mechanism or specification for the beacon-mode to be utilized in wireless mesh networks. The authors therefore proposed a modified ZigBee/IEEE 802.15.4 MAC protocol for optimal communication in a wireless mesh smart metering networks. Other research works related to [16] can be found in [17–22]. The objectives of such works include but not limited to the following: (1) Extension of ZigBee super-frame structure so as to increase the number of guaranteed time slots (GTS). (2) Reduction in the wastage of channel bandwidth. (3) Increasing the QoS for multiple devices. In another scenario, Hoi Yantong et al [23] proposed the design of a multi-interface ZigBee building area network (MIZBAN) for delivering high traffic for the advanced metering infrastructure in the presence of high-rise buildings. This design supports both tree and mesh routing in a single network. This research work is an improvement on previous research works proposed in [24, 25] which focused on dynamic channel managements algorithms which could only handle single device interfaces. This implies that such work could not handle adjacent interference originating from multi-interfaces. Hao Ran Chi et al [26] proposed an improvement to their earlier work in [23] by considering interference issues in the design of high traffic for the ZigBee based AMI. In this case, interference-mitigated ZigBee based advanced metering infrastructure is proposed. This design incorporates multi-radios, multi-channels network architecture with interference mitigation features designed by using multi-objective optimization. This work was necessitated by the fact that ZigBee as well as Wi-Fi and Bluetooth operate in the same frequency band known as

industrial, scientific and medical (ISM) band which is prone to interference [27].

On a similar note, Suryatapa Ray et al [28] proposed an advanced metering infrastructure unit for real time load management in a smart grid. In this research work, an embedded controller was utilized for the realization of the AMI unit while ZigBee wireless technology was the choice communication technology for building of the network that connects smart meters to control tower which acts as the base station of the network. In their paper titled "advanced metering and demand response communication performance in ZigBee based HANs," V.Kounev et al [29] provided an overview of SG AMI and the demand response functionalities paying particular attention to the communication requirements with regards to ZigBee and the smart energy profile (SEP). The theoretical performance analysis of the smart energy profile over ZigBee is in line with previous analysis conducted in [30, 31]. However, their analysis was extended to include the effect of Wi-Fi interference on ZigBee transmission. In another development, I. Parvez et al [32] explored the performance of ZigBee/IEEE 802.15.4 frequency bands (868/915 MHz and 2.4 GHz) with respect to the communication of SG AMI at the HAN and NAN domains. In this regard, the authors used the markov chain model to simulate the MAC/PHY layer using unslotted CSMA/CA algorithm. Finally, they investigated the probability of successful transmission, throughput, delay and path loss at these two operating frequencies. A. Mulla et al [33] provided a good overview on the prospects of deploying wireless sensor networks like ZigBee in smart grid applications. Specifically, the authors analyzed the implications and the anticipated challenges of deploying WSN in the AMI. Finally, they compared ZigBee against IPv6 and then focused on the advantages of IPv6 over ZigBee especially for large WSNs for the smart grid.

Finally, A. Kheaksong et al [34] presented a performance evaluation of SG AMI communication featuring ZigBee, Wi-Fi, and Long Term Evolution (LTE) using network simulator version 3 (NS3). The authors specifically carried out the performance of ZigBee for SG AMI network under different locations of the data concentrator unit (DCU). The simulation results reveal that appropriate DCU-AMI distance was 100 meters with maximum AMI node density of 145 nodes. For the purpose of this section, a systematic and fairly exhaustive review of the literature on general research on ZigBee based smart grid advanced metering infrastructure has been carried out. The time frame for this review was from 2007 to 2016. Table.B.1 shows the details of the reviewed papers with a quick description of their main research focus. Similarly, Figs. B.10 and B.11 show the time distributions of the research papers and their appropriate domain classifications respectively. The domain classification and time distribution of the reviewed research papers in this section show that a lot of research work is still needed in the area of ZigBee based SG AMI.

Table B.1: Details of reviewed papers on general research on ZigBee based SG AMI

Research focus	Proposal
Optimal Communication for a ZigBee based wireless Mesh SG AMI	[16], [17], [18], [19], [20], [21], [22], [32] and [34]
Improved Traffic for ZigBee based SG AMI	[23], [24] and [25]
Improved Traffic for ZigBee based SG AMI with interference mitigation	[26]
ZigBee based AMI for load management, demand response and other smart profile applications	[28], [29], [30], [31] and [33]

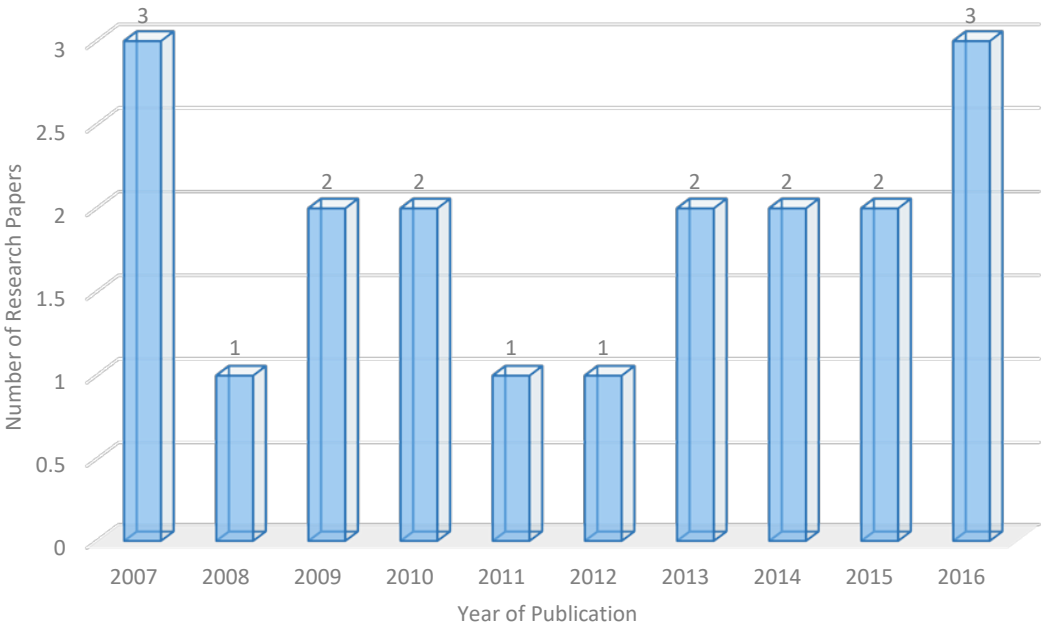

Fig. B.10: Time Distribution of ZigBee Based SG AMI General Research

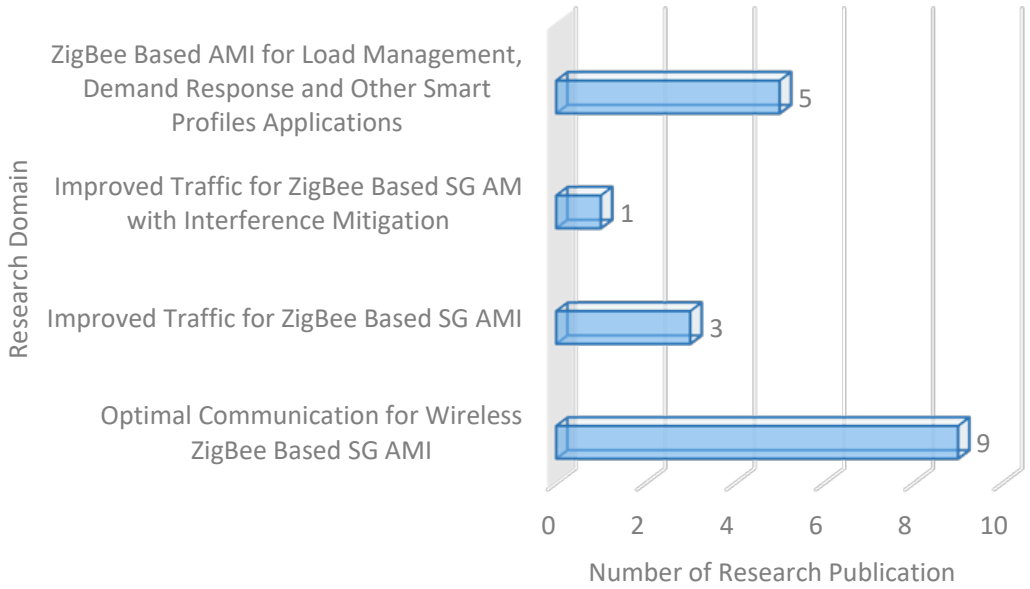

Fig. B.11: Classification of ZigBee Based SG AMI general Research into Domains

4 Overview of the Security Features of ZigBee

In this section, we present general features of ZigBee which could be useful in determining the security strengths and flaws of a ZigBee based AMI network [12], [35, 36]. In general, a transmitted message in a wireless network can be received in normal condition by an authorized nearby device or hijacked by an unauthorized intruder with malicious intentions if appropriate security requirements are not configured in the network. The unauthorized intruder can then gather sensitive information which are supposed to be confidential by listening to the transmitted message. Additionally, the malicious intruder can modify and then replay or resend the messages to legitimate nodes in order to convert them to threat nodes. The first security flaw can be overcome in a ZigBee network by using encryption. However, the intruder can still succeed in carrying out the second threat even if the message is encrypted. In this case, ZigBee uses both device and data authentication to overcome this vulnerability. The following features were found worth reviewing:

(i) ZigBee Data Encryption Feature: Data encryption is an important security feature of a ZigBee network. The ZigBee standard supports the use of symmetric encryption algorithm known publicly as advanced encryption standard (AES). The algorithm if performed on a block of data, results in a cipher-text known as block cipher. ZigBee uses 128-bit block cipher for this purpose. This operation normally protects data from being read by nodes that do not possess the cryptographic key. Considering that the AES algorithm is associated with a security key, the ZigBee standards allow different methods of establishing and sharing security keys

between two or more devices on a ZigBee network. Given that the AES algorithm lies in the public domain and as such known to a potential cyber-criminal, the level of security that can be guaranteed in this network as a result of the encryption process depends on the way and manner the security keys are initiated and distributed.

(ii) Zigbee Authentication Feature: The ZigBee standard uses both data and node authentication procedures in the network for enhanced security. In device authentication procedure, a new node joining the network has to be confirmed authentic. In perspective, the new device must receive a network security key and set its attributes properly in a timely manner in order to be authenticated. On the other hand, data authentication entails that a receiving node should verify that the data from a source node or an intermediary node has not been modified or changed in transit.

(iii) Message Integrity Code (MIC) Feature: In order to ensure a robust and secure data authentication process in ZigBee, a sending node accompanies its frame with message integrity code (MIC) which is generated in a process known by the transmitting and receiving nodes. Thus, the transmitted frame will be considered unauthentic if the MIC value calculated by the receiving node does not match with the one generated by the transmitting node. The level of authenticity for the transmitted data can be improved by increasing the bits in the MIC. The IEEE 802.15.4 and ZigBee standards have MIC options for 32-bits, 64-bits or 128-bits. The MIC is obtained by using an enhanced counter with cipher block chaining message authentication code (CCM*) protocol. This protocol is used in conjunction with the 128-bit AES algorithm and shares the same security key with the algorithm. There are 3 inputs to the AES-CCM*, namely: (i) Data to be transmitted (ii) The security key and (iii) The nonce. The value of the nonce is such that it cannot be the same for two different messages transmitted with the same key. This is because the message counter is incremented in the next transmission time.

(iv) Feature for Frame Freshness Check: The use of nonce in the transmission of frames results in the freshness of the received frame. The security reason behind the use of nonce is to frustrate a cyber-criminal who, even without the required key is capable of receiving a secured message and simply resending same after a calculated period of time. In this context, the resent message can have all the features that could be adjudged as valid security-wise, yet, the counter will indicate that the frame was received previously. As a result, the frame counter can be used to detect and prevent the processing of duplicate frames which makes the network resilient against replay attacks.

(v) Provision for a Trust Center: In the ZigBee network paradigm, the trust centre is a device (usually the ZigBee Coordinator) within the network that all other nodes depend on for security related responsibilities. The trust centre is responsible for the following security roles: (i) Acting as the trust manager in charge of authenticating devices requesting to join the network. (ii) Acting as the network manager in charge of network key distribution. (iii) Acting as a configuration manager responsible for end-to-end security between communicating devices. The trust centre operates in two different modes, namely: (i) Commercial mode and (ii) Residential mode. In the first mode, the trust centre has a lot of responsibilities (maintains a list of devices, master keys, link keys and network keys). In the second mode designed for low security applications, the trust centre maintains only the network key, thereby leaving additional responsibilities as in the former case optional. This implies that as the number of nodes connected to the network increases as in the case of commercial security mode, the memory required by the trust centre increases correspondingly. Table B.2 summarizes the basic security features between the residential and commercial security modes [37].

(vi) Key Distribution and Establishment in ZigBee: The ZigBee standard provides methods for establishing and sharing keys between two or more devices within the network. The keys (link key and the network key) are usually utilized for secure communication. The link key is normally shared between two transmitting nodes while the network key is shared by devices within the entire network. In ZigBee network, a node can acquire a link key by pre-installation, key transport or by key establishment. The key can be embedded in the device in the case of pre-installation method. In the key transport method, the device acquires the security key from the trust centre. On the other hand, key establishment is a method of creating a random key in two different nodes without transmitting the key through a wireless link that may be unprotected. According to the ZigBee standard, the key establishment service is based on the symmetric-key key establishment (SKKE) protocol. It is expected that the devices establishing the key already have the master key. It is noted here, that key establishment is needed primarily to derive a link key (K_L) and not for generating the network key. The SKKE protocol is implemented with the help of two devices which act as the initiator device and the responding device respectively. In this procedure, a link can be established by the initiator device using the master key by transferring specific data to the responding device. The responding device then uses the transmitted data to derive the link key. It is to be noted that the initiator device derives the link key from the same data. The two devices can use the correctly derived link key in symmetric key cryptography. As an illustration of the key transport method, we assume that the two communicating parties, smart meter node 1 (SM_1),smart meter node 2 (SM_2) and the

trust center are involved in this SKKE protocol. The end-to-end Application Key Establishment Protocol is somehow dependent on how the trust center, TC is configured. In this context, the trust center, TC creates the K_L and sends it to the communicating parties. This can be done in such a way that the initiator, and the responder,SM_2 will have no role in the creation of the K_L. Alternatively, the trust center, TC can create a temporary shared key known as Master Key, K_M and sends it to each of the smart meter node. Using this K_M, SM_1 and SM_2 initiate a Symmetric-Key Key Establishment (SKKE) procedure to establish an K_L. This option can allow the two nodes to create an K_L mutually. SKKE is actually a key agreement scheme employed in the ZigBee end-to-end Application Key Establishment mechanism, and its components are defined in [12]. At the end of a successful key initialization procedure, the two ZigBee nodes will be able to establish secure communication using their pairwise encryption key, K_L. This simplified illustration of the end-to-end Application Key Establishment protocol is presented in Fig. B.12 The black solid lines represent the already secure communication paths, labeled by corresponding symmetric encryption keys. On the other hand, the dashed lines represent the resulting secure communication paths after a successful protocol run, also labeled by corresponding encryption keys. Finally, the blue solid lines are the messages in the protocol labeled by their sequence numbers and the encryption keys they deliver. From Fig. B.12, the initiator node, SM_1 begins the procedure of establishing an K_L with the responder node, SM_2 by sending TC the first message, requesting key. This includes the TC address, requested key type (Application Key), and responder address (SM_2) all concatenated into one message. Then TC creates an K_L for the two communicating nodes, and sends it to each communicating party in two similar transport key messages. Since TC is configured to send an K_L directly in this case, the key type value in the last two transport key messages will be Application Link Key $(AppK_L)$. The only difference between these two messages is the boolean value which indicates who among the two communicating nodes is the initiator or the responder in the communication. For example, TRUE indicates that message recipient is the initiator while FALSE indicates that the message recipient is the responder. All the messages in this case are encrypted with a key which the sender/receiver shares with the TC. This key can either be Trust Center Link Key (TCK_L) or Trust Center Master Key (TCK_M), as defined in [12] but for simplicity they will be referred to as KSM_1 for SM_1 node, and KSM_2 for SM_2 node. The illustration of this protocol is summarized below:

(a)

$$SM_1 \rightarrow TC : TC||AppKey||SM_2KSM_1$$

(b)

$$TC \rightarrow SM_1 : SM_1||AppKey||True||K_LKSM_1$$

(c)

$$TC \rightarrow SM_2 : SM_2||AppKey||SM_1||False||K_LKSM_2$$

Table B.2: Basic security features of Residential and Commercial security modes

Features	Residential	Commercial
Provision of Network Layer security with the help of network key	√	√
Provision of APS Layer security with the help of link keys	√	√
Central control and update of keys	√	√
Capacity for deriving link keys between communicating nodes		√
Support for entity authentication		√

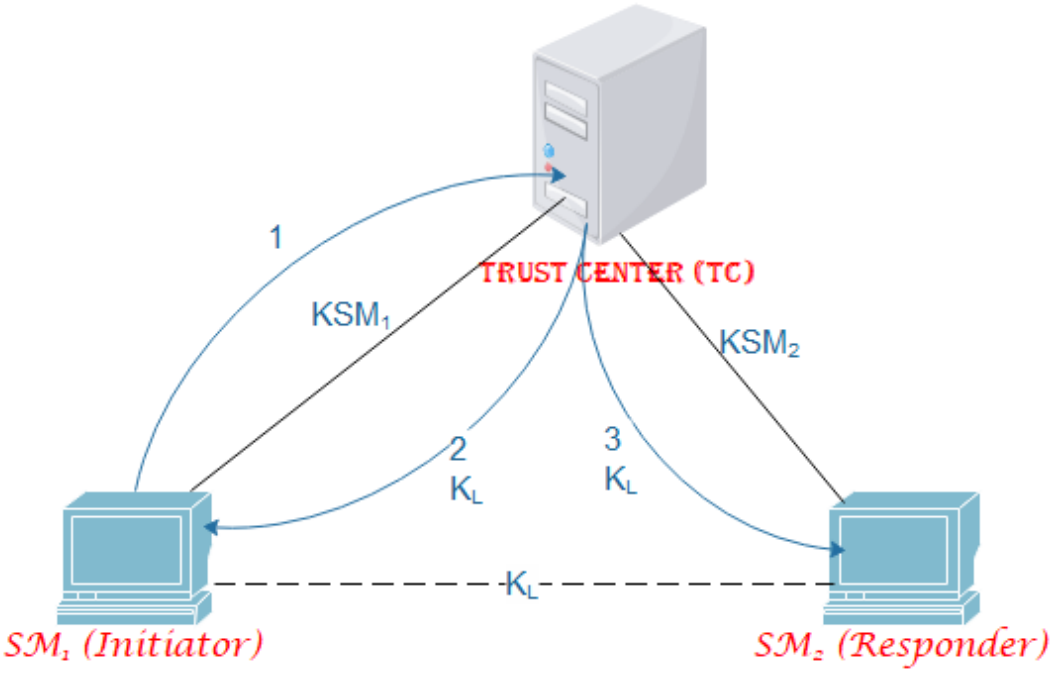

Fig. B.12: An Illustration of Symmetric-Key Key Establishment Protocol in ZigBee

5 Potential Vulnerabilities of ZigBee Security

Notwithstanding the strong security features possessed by ZigBee as we have seen from the above section, there are still concerns about security vulnerabilities in the network. Some of the vulnerabilities are covered in this section [38–40].

1. Same Nonce Attacks: When an AES-CCM* protocol is fed with a nonce value as an input, double encryption which ought to make the system more robust to attacks becomes a point of vulnerability. In this context, the double encryption of the same plaintext will result in two different cipher-texts. This is because the nonce will be different even when the same security key is re-used. In this case, an attacker (eavesdropper) would be able to recover some information about the plaintexts. For example, if the same nonce and key are utilized for the generation of two cipher-texts, the XOR of the two cipher-texts would correspond to the XOR of their corresponding plaintexts [41]. An attacker with this knowledge can succeed in launching this kind of attack on the ZigBee network.

2. Denial of Service Attacks: The main objective of a cyber criminal wishing to lunch a Denial of service attack on any network is to compel network nodes to reject a given percentage or all of its received packets depending on the intensity of the attack [15] This kind of attack can be initiated by a cyber-criminal by first scanning series of network nodes for vulnerabilities. When a vulnerable node is discovered, it can be exploited and maliciously used to scan other nodes for vulnerabilities. A typical illustration of denial of service/distributed denial of service attack is provided in Fig. B.13. The attack can be regarded as distributed denial of service once the handlers and agents are utilized to increase the number of attack sources against the target server. Otherwise, the attack can be regarded as denial of service attack. The handlers can be maliciously used to control a large number of zombies such that the attacker can have a centralized control over the entire network. This kind of attack is geared towards preventing a legitimate network user from using all or some of the network resources [42]. Denial of service attack can be categorized as: attack on bandwidth consumption, consumption of processor time, disruption of network routing information, temporary disruption or destruction of physical components etc. Denial of service attack can be executed at any layer of the network protocol. At the PHY layer, the attack can either be jamming or node destruction attack. At the MAC layer, the attack can be a denial of sleep for end devices. This type of attack can be costly for a ZigBee network which has nodes that are highly constrained by power. Furthermore, denial of service attacks can be a spoofing or replaying attack executed at the NETWORK layer, SYN flood attack at the TRANSPORT layer or reprogramming attacks executed at the APPLICATION layer. Jan Durech et al [43] realized this attack on a ZigBee network by using a program known as SmartRF Studio 7. Their attack model first utilized a packet which they kept resending at a speed of 30 packets per second. The Coordinator in the network was configured to acknowledge every packet or frame received even if the packet is blocked (discarded) at the upper layers because it is a duplicate frame. It was noticed that the

Coordinator stopped responding after 250 seconds of the attack and after the reception of 7442 packets. Xianghui Cao et al [44] also demonstrated that DoS attack can be executed easily in ZigBee network in the following ways: (i) The cyber attacker can send bogus messages geared towards depleting the energy of the victim nodes resulting in disruption of the services offered by these nodes as a result of the power constrained nature of ZigBee nodes. (ii) The attack can also be executed in the network due to the MAC misbehavior. With CSMA/CA protocol configured at the MAC layer of the network, DoS attack can be realized by an attacker by continuously sending traffic to a victim node. This led to the starving of other devices within the interference region of channel access and network services. As a result, each device or node spends so much time sensing and waiting for collision-free channel access thereby resulting in the depletion of node power.

3. Node Compromise: A cyber criminal can launch this type of attack on a ZigBee network by capturing a legitimate node connected to the network and then reprogramming the captured node/s for malicious purposes.

4. Sinkhole and Wormhole Attacks: In this type of attack, a malicious node can attract packets to itself by publishing false routing information, thus, disrupting the network routing table which consequently results in increased packets loss ratio. In wormhole attack, a cyber criminal may receive packets at one point in the network, and then replays them to generate false routing information which can lead to a misdirection of network traffic.

5. Physical Attack: In ZigBee network, physical attack can be made when a desperate cyber criminal gains physical access to nodes/devices connected to the network. In such a situation, security objectives must include the design and implementations of physical monitoring and surveillance systems. In summary, a comprehensive security threats catalogue for a ZigBee based network and their impact on cyber security objectives are presented in Table B.3 and Table B.4 respectively [45–49].

6 Review of Research Efforts on the Security of ZigBee Based SG AMI

In this section, recent research efforts on the security of ZigBee based SG AMI network has been provided. Bashar Alohali et al [50] reported that a node capture attack is possible in ZigBee AMI network. In their work, they demonstrated the possibility of an attacker capturing a given number of smart meter nodes in a targeted personal area network (PAN) and then turning them into threat nodes by extracting maliciously secure keys and measured data from smart meter. They used their attack

Table B.3: ZigBee Based Network Threats Catalogue

Attack Name	Attack Type	Description
Radio Frequency Jamming	Denial of Service	Transmitting noise at the same frequency at the targeted wireless network
Network flooding Attack	Denial of Service	Involves sending of voluminous traffic to vulnerable nodes
Rogue Node Attack	Man-in-the-middle	Involves the introduction of corrupted packets by rogue nodes
Security Parameter Extraction by physical access	Physical Attack	Involves the introduction of corrupted packets by rogue nodes
Sinkhole/Blackhole Routing	Routing Attack	This can involve a network attack where an attacker tries to manipulate routing table to misdirect traffic
Selective forwarding	Routing Attack	This involves the selective forwarding of frames to the next hop
Network Traffic Analysis	Traffic Analysis	Here, the attacker passively monitors transmissions in order to identify communication patterns and participants
False Battery Life Extension	Denial of Service	Here, an attacker can pretend to be in battery extension mode in order to dominate channel access
False Association	Denial of Service	Here, an attacker may send forged association packets thereby depleting the memory of ZigBee coordinator
False Disassociation	Denial of Service	In this case, the attacker may send disassociation packets, thus, causing nodes to be dropped out from the cluster
False Acknowledgment	Masquerading Attack	In this attack, the attacker may send false ACK packets so as to confuse senders into thinking that packets have been successfully received
Plaintext key capture	Eavesdropping	This might involve the overhearing of the network and/or master keys by a cyber attacker
Key capture with SKKE	Eavesdropping	This can happen in network using SKKE where an attacker with master key can guess the pairwise link keys

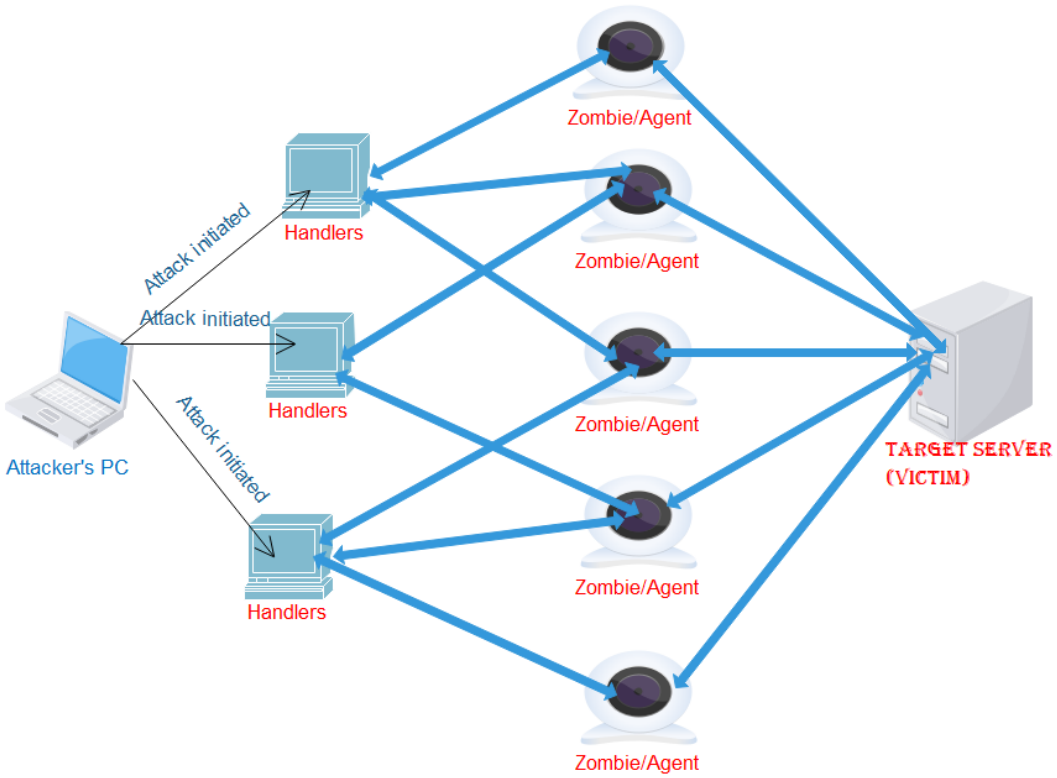

Fig. B.13: Illustration of DoS/DDoS Attack

Table B.4: Impacts of ZigBee Attacks/threats on Cyber Security Objectives

Cyber attack/threats	Impact on Confidentiality	Impact on Availability	Impact on Integrity
Radio Frequency Jamming	Low	High	High
Network flooding Attack	Low	High	High
Rogue Node Attack	Moderate	Low	Moderate
Security Parameter Extraction by physical access	Moderate	Low	Low
Sinkhole/Blackhole Routing	low	High	Low
Selective forwarding	Low	High	Low
Network Traffic Analysis	Moderate	Low	Low
False Battery Life Extension	Low	Moderate	Low
False Association	Low	High	Low
False Disassociation	Low	High	Low
False Acknowledgment	Low	Low	High
Plaintext key capture	High	Low	High
Key capture with SKKE	High	Low	High

model to verify the resiliency of such attack on the different topologies of the ZigBee AMI network. At the end, they concluded that a partial mesh topology is the most resilient to such attacks. The authors in [51] presented a new approach for ZigBee MAC layer design based on security enhancement. The authors stressed the need for a new attribute at the MAC layer which will improve the security of the network. Their major aim was to design or incorporate a feature at the MAC layer that can detect and subsequently block unauthorized nodes at the MAC layer. According to the authors, doing so would make the ZigBee network more secure, reliable, scalable, and more robust against external attacks. Unfortunately, it was not clearly stated how their claimed enhancement differed from the original security features of ZigBee as contained in the ZigBee specifications [12]. Ashraf Khalaf et al [52] investigated the impacts of ZigBee node failure on the Wireless Sensor Network performance against different topologies. According to their findings, the effect of the ZigBee Coordinator failure results in the failure of the entire network when the network is configured with only one ZigBee Coordinator. Their study also noted that the effect of a ZigBee router failure is always more devastating than that of the ZigBee end device failure. They also discovered that mesh and cluster-tree topologies were more resilient to the attack when compared to other topologies. In [53], W. Somkaew et al, presented a performance evaluation of data security implementations over ZigBee based AMI systems. In this study, the implementation and the performance evaluation of the 128-bit advanced encryption standard (AES) utilized in the ZigBee standard was conducted with the help of an advanced security integrated circuit (IC). Specifically, the authors performed a comparative analysis between a hardware-based and software based crypto engines with respect to data encryption and decryption processes. It is to be noted that the research proposed in [50] can bring about an improvement in confidentiality and integrity of data transmitted over the AMI network.

In a similar note, R.K Bhatia [54] proposed a security mechanism for mesh-topology ZigBee based SG AMI which relies on public key infrastructure (PKI) and intrusion detection system for protecting the AMI from internal as well as external threats/attacks. In this proposal, every smart meter and data concentrator unit will be authenticated using the PKI while an intrusion detection module is to be installed between every smart meter and DCU. Furthermore, communication and transfer of data from smart meter to DCU is done by means of ZigBee connection. While this work deserves some commendation, it is to be noted that details of key distribution and management of the PKI in the proposed solution were not provided. Additionally, the authors did not include sufficient analysis of the efficiency of the proposed security in the face of known cyber attacks/threats. Additionally, V. Naboodiri [55] presented a comprehensive outlook for wireless security in a smart grid HAN with potential vulnerabilities identified. However, the authors made occasional references to the security architecture of ZigBee while major part of the research work dealt with wireless networks for the

AMI in a general sense. The authors specifically examined in details, the communication requirements for the AMI HAN scenario with attention given to possible security challenges that need to be addressed. In another instance, M.M Fouda et al [56] introduced a HANid conflict attack in a ZigBee based HAN for the smart grid. In their attack model, the ZigBee coordinator in a HAN is assigned a unique HANid. The authors assumed that if there are more than one HAN coordinator operating in the same operating space, then a HANid conflict can occur. In the event of an occurrence of such conflict, the HAN coordinator may detect this conflict through a received beacon and then perform conflict resolution accordingly. Alternatively, a device or node belonging to the same HAN coordinator can notify the HAN coordinator on receiving a signal from two HAN coordinators with the same HANid. The authors therefore presented a threat model for this scenario where an attacker or adversary device can frequently send forged conflict notification messages to the HAN coordinator, thereby enforcing the HAN coordinator to perform conflict resolution repeatedly. The impact of the HANid conflict attack was demonstrated with a simulation performed in MATLAB. The results generated from this simulation study indicate that the HANid conflict attacks can affect a ZigBee based SG HAN network if left unchecked. Finally, H. Seo et al [57] proposed an attribute based cryptographic security for the ZigBee enabled home automation application. With this proposal, the security of various services in ZigBee based networks for various home automation applications can be enhanced using attributes which reduce the number of security keys needed for secure communication between nodes. The authors performed a comparative analysis of different cryptographic schemes like PKI, IBE, symmetric encryption in the current ZigBee standard and the attribute based encryption scheme. The authors concluded that although attribute based encryption do not offer digital signature, it can nevertheless offer an efficient encryption in ZigBee based network and also offer a reduced overhead on the ZigBee trust centre. In [58], Imen Aquini et al identified vulnerabilities in the join procedure for the admission of new nodes into the network with regards to the ZigBee Alliance's proposal for the smart energy profile. The authors proposed a mechanism that can avoid DoS attack by limiting the number of attempts to be made in order to join a network. The authors concluded that the enhanced join procedure resists DoS attack better than the one proposed by ZigBee Alliance in the smart energy profile (SEP). In summary, the review of important researches already carried out on the ZigBee based security for the SG AMI reveals an appreciable level of research effort on the security of ZigBee based networks. However, this review has shown that some serious cyber attacks/threats are still realizable in a SG AMI as shown in Table B.5. In addition, there are other security issues that must be resolved before the full potentials of a ZigBee based SG AMI can be realized. These security issues have been presented in the next section. Fig. B.14 shows the time distribution of ZigBee based SG AMI security research with concentration only

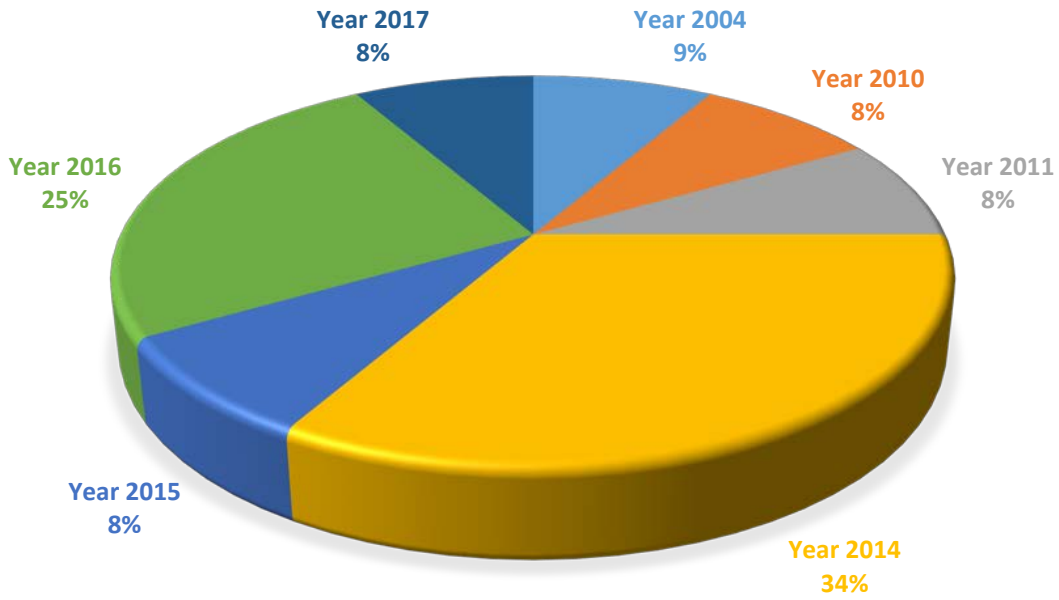

Fig. B.14: Time Distribution of ZigBee Based SG AMI Cyber Security Research

on published papers presented in Table B.5. The summary of the classifications of important security solutions for the ZigBee based SG AMI network with regards to the main cyber security objectives is presented in Fig. B.15. The chart presented in Fig. B.15 (based on publications summarized in Table B.5) can be used to highlight the research gap in cyber security research for ZigBee based SG AMI.

7 Open Research Issues for the Security of ZigBee Based SG AMI

There are great potentials of implementing and deploying ZigBee as a communication technology for the smart grid AMI. These potentials could be limited by myriads of cyber attacks or threats targeting the SG AMI. Studies have showed that these cyber attacks will get more sophisticated in the coming years. As a result, designing security solutions for mitigation and defense against these cyber attacks will not only be burdensome but complicated. This requires a combination of several research methodologies that will help in realizing the major cyber security objectives for the smart grid. It can be discovered from the review carried out in this paper that appreciable level of research effort has been given to the security of the ZigBee based SG AMI. However, the progress made thus far cannot be matched with the potential cyber security vulnerabilities and threats targeting the SG AMI. In the context of the ZigBee based SG AMI, there are still some open research issues that need urgent research attention with regards to cyber security. These issues are discussed briefly:

- Congestion minimizing schemes during data aggregation: An important research approach for preserving the privacy and security of energy consumption data is the use of data aggregation

Table B.5: Summary of security research efforts on a ZigBee based SG AMI

Proposal	Cyber-security objective pursued	Research description
[43]	Availability	Demonstrated the possibility of a node capture attack on a ZigBee based SG AMI
[44]	Availability	Study demonstrated how a DDoS attack can be realized on ZigBee based SG AMI using a program known as smartRF studio 7.
[45]	Availability	Investigated the impact of node failure on a ZigBee based network.
[50]	Availability, confidentiality and integrity	Study developed comprehensive cyber threats catalog for ZigBee based network and their impact on cyber security objectives
[51]	Integrity and Confidentiality	Investigated and implemented data security over ZigBee based SG AMI
[52]	Integrity and Confidentiality	The aim of this research was to design or incorporate a new feature at the MAC layer that can detect and block unauthorized nodes at the MAC layer
[53]	Availability	Investigated the negative impact of ZigBee node failure and their effects on the ZigBee coordinator
[54]	Integrity and Confidentiality	This study carried out a performance evaluation of data security implementations over ZigBee based smart grid AMI
[55]	Integrity and Confidentiality	Study proposed a mesh topology based SG AMI which relies on PKI and intrusion detection system for protecting the AMI against external attacks/threats
[56]	Availability	Study presented an impact analysis of HANid attack on a ZigBee based HAN for the smart grid
[57]	Integrity, Confidentiality and Availability	Study proposed an attribute based cryptography for the security home automation applications.
[58]	Integrity, confidentiality and Availability	Study proposed an enhanced join procedure for ZigBee smart energy profile

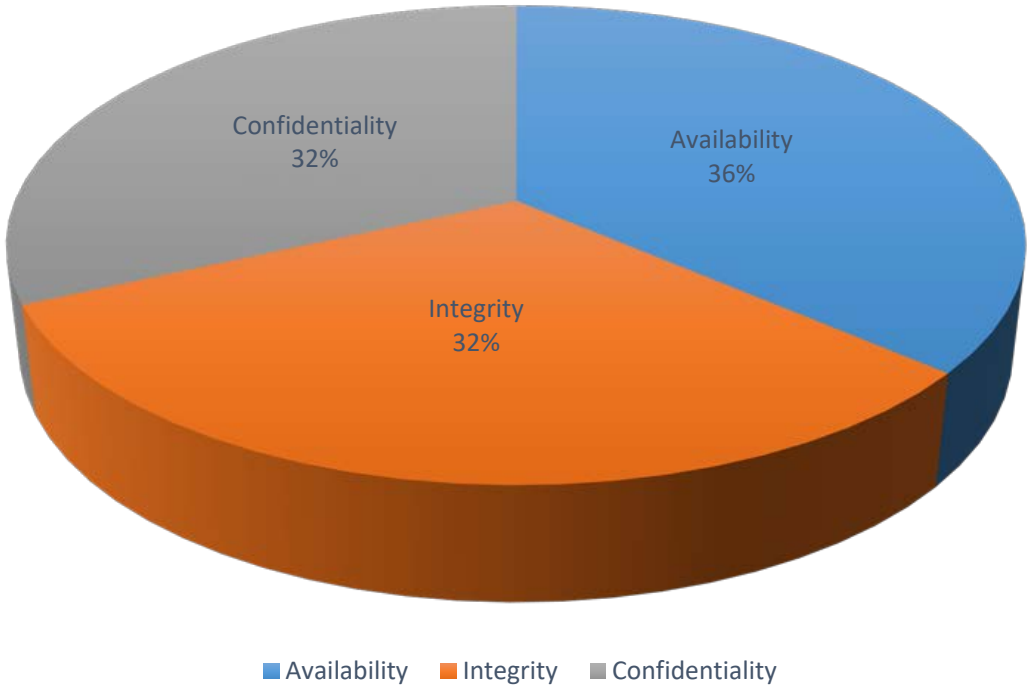

Fig. B.15: Classification of ZigBee Based SG AMI Security Research into Domains

protocols [59–66]. Research on data aggregation for ZigBee based SG AMI is an open research direction as such kind of research is very scare in the literature. Notwithstanding the popularity of this approach, data aggregation brings additional burden on traffic intensity which consequently leads to a congestive network scenario. Unfortunately, the current ZigBee standard has no reliable mechanism for handling congestion management. As a result, an important open research direction can be targeted to research proposals on reliable congestion minimizing schemes for data aggregation for ZigBee based SG AMI.

- Denial of service/distributed denial of service attacks: DoS/DDoS attacks can impact very negatively on the robustness of any network. They are usually aimed at using up in a malicious manner both global and local network resources. This type of attack remains a major security problem for ZigBee based networks as such attacks can be aimed at different layers of the network. Node capture which is a typical DoS attack can lead to malicious exploitation of shared network keys which could consequently lead to other attacks like packet injection and routing-based attacks. This remains an area of research where limited work has been done.

- Security issues resulting from the trust centre: The trust centre is overloaded with three-fold responsibilities of acting as: (i) the trust manager in charge of authenticating devices requesting to join the network. (ii) the network manager in charge of network key distribution and (iii) the configuration manager responsible for end-to-end security between

communicating devices. This will consequently lead to the trust centre being a single point of failure on a ZigBee based SG AMI network. An open research direction will involve the designing of simple key distribution schemes that can lessen the burden on the trust centre. Since the trust centre will be highly constrained by memory especially as the network grows; development of optimization techniques that can impact positively on the memory and energy constrained trust centre node is another open research direction.

8 Conclusion

This review paper has revealed the level of attraction of ZigBee technology as a communication technology of choice for low power, low rate applications like the smart grid advanced metering infrastructure. This is evidenced by the wide deployments of SG AMI pilot projects in different countries as have been revealed by this paper. However, the legislation in different countries concerning the usage of personal information will compel standard bodies in charge of the implementation of the smart grid vision to do a thorough review on the security strengths and weaknesses of any communication technology before its adoption and deployment especially for the SG AMI. The paper has carried out a systematic review of the state-of-the-art research on the ZigBee based SG AMI with a major focus on security. It was discovered that the security features of ZigBee can be adjudged to be strong. However, there are still issues of vulnerabilities which must be urgently addressed in order realize the full potentials of the technology for the SG AMI. These issues have been highlighted in this paper and open research directions which can lead to a more robust ZigBee based SG AMI discussed. It is therefore our believe that this paper will be a vital tool in the hands of standard bodies, policy makers in charge of the electrical energy sector and researchers in the smart grid domain.

References

[1] R. Mahmud, R. Vallakati, A. Mukherjee, P. Ranganathan, and A. Nejadpak, "A survey on smart grid metering infrastructures: Threats and solutions," in *Electro/Information Technology (EIT), 2015 IEEE International Conference on*. IEEE, 2015, pp. 386–391.

[2] U. D. of Energy, "Advanced metering infrastructure and customer systems,results from the smart grid investment grant program." [Online]. Available: https://www.smartgrid.gov/files/Final_SGIG_Report_20161220.pdf

[3] D. Mirzoev *et al.*, "Low rate wireless personal area networks (lr-wpan 802.15. 4 standard)," *arXiv preprint arXiv:1404.2345*, 2014.

[4] C. Bennett and D. Highfill, "Networking ami smart meters," in *Energy 2030 Conference, 2008. ENERGY 2008. IEEE*. IEEE, 2008, pp. 1–8.

[5] V. Aravinthan, V. Namboodiri, S. Sunku, and W. Jewell, "Wireless ami application and security for controlled home area networks," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–8.

[6] J. Garcia-Hernandez, "Recent progress in the implementation of ami projects: standards and communications technologies," in *Mechatronics, Electronics and Automotive Engineering (ICMEAE), 2015 International Conference on*. IEEE, 2015, pp. 251–256.

[7] D. Bian, M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Analysis of communication schemes for advanced metering infrastructure (ami)," in *PES General Meeting| Conference & Exposition, 2014 IEEE*. IEEE, 2014, pp. 1–5.

[8] H. Groenewald, "Nrs049—advanced metering infrastructure (ami) for residential and commercial customers," *ESKOM, Johannesburg, Presentation*, 2009.

[9] W. Meng, R. Ma, and H.-H. Chen, "Smart grid neighborhood area networks: a survey," *IEEE Network*, vol. 28, no. 1, pp. 24–32, 2014.

[10] S.-H. Seo, X. Ding, and E. Bertino, "Encryption key management for secure communication in smart advanced metering infrastructures," in *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*. IEEE, 2013, pp. 498–503.

[11] M. De Bruyn, "The protection of personal information (popi) act-impact on south africa," *The International Business & Economics Research Journal (Online)*, vol. 13, no. 6, p. 1315, 2014.

[12] Z. Specification, "Document 053474r17," *ZigBee Alliance*, 2008.

[13] N. S. Grid, "Introduction to nistir 7628 guidelines for smart grid cyber security," *Guideline, Sep*, 2010.

[14] N. F. Pub, "197: Advanced encryption standard (aes)," *Federal information processing standards publication*, vol. 197, no. 441, p. 0311, 2001.

[15] S. Farahani, *ZigBee wireless networks and transceivers*. Newnes, 2011.

[16] H. Shabani, M. M. Ahmed, S. Khan, S. A. Hameed, and M. H. Habaebi, "Smart zigbee/ieee 802.15. 4 mac for wireless sensor multi-hop mesh networks," in *Power Engineering and Optimization Conference (PEOCO), 2013 IEEE 7th International*. IEEE, 2013, pp. 282–287.

[17] Y.-G. Hong, H.-J. Kim, H.-D. Park, and D.-H. Kim, "Adaptive gts allocation scheme to support qos and multiple devices in 802.15. 4," in *Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on*, vol. 3. IEEE, 2009, pp. 1697–1702.

[18] L. Cheng, X. Zhang, and A. Bourgeois, "Gts allocation scheme revisited," *Electronics Letters*, vol. 43, no. 18, pp. 1005–1006, 2007.

[19] Y. Zhou, Y. Wang, J. Ma, J. Jia, and F. Wang, "A low-latency gts strategy in ieee802. 15.4 for industrial applications," in *Future Generation Communication and Networking, 2008. FGCN'08. Second International Conference on*, vol. 1. IEEE, 2008, pp. 411–414.

[20] A. Van den Bossche, T. Val, and E. Campo, "Modelisation and validation of a full deterministic medium access method for ieee 802.15. 4 wpan," *Ad Hoc Networks*, vol. 7, no. 7, pp. 1285–1301, 2009.

[21] P. Muthukumaran, R. de Paz, R. Spinar, and D. Pesch, "Meshmac: Enabling mesh networking over ieee 802.15. 4 through distributed beacon scheduling," in *International Conference on Ad Hoc Networks*. Springer, 2009, pp. 561–575.

[22] H.-I. Jeon and Y. Kim, "Bop (beacon-only period) and beacon scheduling for meu (mesh-enabled usn) devices," in *Advanced Communication Technology, The 9th International Conference on*, vol. 2. IEEE, 2007, pp. 1139–1142.

[23] H. Y. Tung, K. F. Tsang, K. T. Chui, H. C. Tung, H. R. Chi, G. P. Hancke, and K. F. Man, "The generic design of a high-traffic advanced metering infrastructure using zigbee," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 836–844, 2014.

[24] Q. Yu, J. Chen, Y. Fan, X. Shen, and Y. Sun, "Multi-channel assignment in wireless sensor networks: A game theoretic approach," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.

[25] E. Toscano and L. L. Bello, "Multichannel superframe scheduling for ieee 802.15. 4 industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 2, pp. 337–350, 2012.

[26] H. R. Chi, K. F. Tsang, K. T. Chui, H. S.-H. Chung, B. W. K. Ling, and L. L. Lai, "Interference-mitigated zigbee-based advanced metering infrastructure," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 672–684, 2016.

[27] X. Zhang and K. G. Shin, "Cooperative carrier signaling: Harmonizing coexisting wpan and wlan devices," *IEEE/ACM Transactions On Networking*, vol. 21, no. 2, pp. 426–439, 2013.

[28] S. Roy, B. Bedanta, and S. Dawnee, "Advanced metering infrastructure for real time load management in a smart grid," in *Power and Advanced Control Engineering (ICPACE), 2015 International Conference on*. IEEE, 2015, pp. 104–108.

[29] V. Kounev and D. Tipper, "Advanced metering and demand response communication performance in zigbee based hans," in *Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on*. IEEE, 2013, pp. 31–36.

[30] A. Ghosh, R. Jana, V. Ramaswami, J. Rowland, and N. Shankaranarayanan, "Modeling and characterization of large-scale wi-fi traffic in public hot-spots," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 2921–2929.

[31] W. Yuan, X. Wang, and J.-P. M. Linnartz, "A coexistence model of ieee 802.15. 4 and ieee 802.11 b/g," in *Communications and Vehicular Technology in the Benelux, 2007 14th IEEE Symposium on*. IEEE, 2007, pp. 1–5.

[32] I. Parvez, A. Sundararajan, and A. I. Sarwat, "Frequency band for han and nan communication in smart grid," in *2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, Dec 2014, pp. 1–5.

[33] A. Mulla, J. Baviskar, A. Yerunkar, and R. Sarwadnya, "Convergence of wireless sensor network with smart grid environment based on ipv6 protocol," in *Communication Systems and Network Technologies (CSNT), 2015 Fifth International Conference on*. IEEE, 2015, pp. 153–158.

[34] A. Kheaksong, A. Prayote, and W. Lee, "Performance evaluation of smart grid communications via network simulation version 3," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2016 13th International Conference on*. IEEE, 2016, pp. 1–5.

[35] M. Sun and Y. Qian, "Study and application of security based on zigbee standard," in *Multimedia Information Networking and Security (MINES), 2011 Third International Conference on*. IEEE, 2011, pp. 508–511.

[36] B. W. Ramsey, M. A. Temple, and B. E. Mullins, "Phy foundation for multi-factor zigbee node authentication," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 795–800.

[37] J. Sun and X. Zhang, "Study of zigbee wireless mesh networks," in *Hybrid Intelligent Systems, 2009. HIS'09. Ninth International Conference on*, vol. 2. IEEE, 2009, pp. 264–267.

[38] M. A. B. Karnain and Z. B. Zakaria, "A review on zigbee security enhancement in smart home environment," in *Information Science and Security (ICISS), 2015 2nd International Conference on*. IEEE, 2015, pp. 1–4.

[39] K. Islam, W. Shen, and X. Wang, "Security and privacy considerations for wireless sensor networks in smart home environments," in *Computer Supported Cooperative Work in Design (CSCWD), 2012 IEEE 16th International Conference on*. IEEE, 2012, pp. 626–633.

[40] W. Razouk, "Zigbee security within the framework of iot," in *Service-Oriented Computing and Applications (SOCA), 2014 IEEE 7th International Conference on*. IEEE, 2014, pp. 265–265.

[41] N. Sastry and D. Wagner, "Security considerations for ieee 802.15. 4 networks," in *Proceedings of the 3rd ACM workshop on Wireless security*. ACM, 2004, pp. 32–42.

[42] L. Chhaya, P. Sharma, G. Bhagwatikar, and A. Kumar, "Wireless sensor network based smart grid communications: Cyber attacks, intrusion detection system and topology control," *Electronics*, vol. 6, no. 1, p. 5, 2017.

[43] J. Durech and M. Franekova, "Security attacks to zigbee technology and their practical realization," in *Applied Machine Intelligence and Informatics (SAMI), 2014 IEEE 12th International Symposium on*. IEEE, 2014, pp. 345–349.

[44] X. Cao, D. M. Shila, Y. Cheng, Z. Yang, Y. Zhou, and J. Chen, "Ghost-in-zigbee: Energy depletion attack on zigbee-based wireless networks," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 816–829, 2016.

[45] S. Plósz, A. Farshad, M. Tauber, C. Lesjak, T. Ruprechter, and N. Pereira, "Security vulnerabilities and risks in industrial usage of wireless communication," in *Emerging Technology and Factory Automation (ETFA), 2014 IEEE*. IEEE, 2014, pp. 1–8.

[46] B. Reaves and T. Morris, "Analysis and mitigation of vulnerabilities in short-range wireless communications for industrial control systems," *International Journal of Critical Infrastructure Protection*, vol. 5, no. 3-4, pp. 154–174, 2012.

[47] N. Vidgren, K. Haataja, J. L. Patino-Andres, J. J. Ramirez-Sanchis, and P. Toivanen, "Security threats in zigbee-enabled systems: vulnerability evaluation, practical experiments, countermeasures, and lessons learned," in *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 2013, pp. 5132–5138.

[48] R. Sokullu, I. Korkmaz, O. Dagdeviren, A. Mitseva, and N. R. Prasad, "An investigation on ieee 802.15. 4 mac layer attacks," in *Proc. of WPMC*, vol. 41, 2007, pp. 42–92.

[49] S. S. Jung, M. Valero, A. Bourgeois, and R. Beyah, "Attacking beacon-enabled 802.15. 4 networks," in *International Conference on Security and Privacy in Communication Systems*. Springer, 2010, pp. 253–271.

[50] B. Alohali, K. Kifayat, and Q. Shi, "Impact of topology on service availability in a smart grid advanced metering infrastructure," in *EMERGING 2016: The Eighth International Conference on Emerging Networks and Systems Intelligence*. IARIA XPS Press, 2016, pp. 29–33.

[51] M. J. H. Biddut, N. Islam, R. S. Sultana, A. Sarker, and M. M. Rahman, "A new approach of zigbee mac layer design based on security enhancement," in *Telecommunications and Photonics (ICTP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1–5.

[52] A. A. Khalaf and M. S. Mokadem, "Effects of zigbee component failure on the wsn performance with different topologies," in *Microelectronics (ICM), 2016 28th International Conference on*. IEEE, 2016, pp. 9–12.

[53] W. Somkaew, S. Thepphaeng, and C. Pirak, "Data security implementation over zigbee networks for ami systems," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2014 11th International Conference on*. IEEE, 2014, pp. 1–5.

[54] R. K. Bhatia and V. Bodade, "Defining the framework for wireless-ami security in smart grid," in *Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on*. IEEE, 2014, pp. 1–5.

[55] V. Namboodiri, V. Aravinthan, S. N. 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, 2014.

[56] M. M. Fouda, Z. M. Fadlullah, and N. Kato, "Assessing attack threat against zigbee-based home area network for smart grid communications," in *Computer Engineering and Systems (ICCES), 2010 International Conference on*. IEEE, 2010, pp. 245–250.

[57] H. Seo, C. Kim, and H. Kim, "Zigbee security for home automation using attribute-based cryptography," in *Consumer Electronics (ICCE), 2011 IEEE International Conference on*. IEEE, 2011, pp. 367–368.

[58] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Communications Workshops (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[59] S. Wang, L. Cui, J. Que, D.-H. Choi, X. Jiang, S. Cheng, and L. Xie, "A randomized response model for privacy preserving smart metering," *IEEE transactions on smart grid*, vol. 3, no. 3, pp. 1317–1324, 2012.

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

[61] J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving differential privacy of data disclosure in the smart grid," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 504–512.

[62] C. Rottondi, G. Verticale, and A. Capone, "A security framework for smart metering with multiple data consumers," in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*. IEEE, 2012, pp. 103–108.

[63] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 327–332.

[64] M. A. Mustafa, N. Zhang, G. Kalogridis, and Z. Fan, "Dep2sa: A decentralized efficient privacy-preserving and selective aggregation scheme in advanced metering infrastructure," *IEEE Access*, vol. 3, pp. 2828–2846, 2015.

[65] N. Saxena, B. J. Choi, and S. Grijalva, "Secure and privacy-preserving concentration of metering data in ami networks," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–7.

[66] S. Tonyali, K. Akkaya, N. Saputro, and A. S. Uluagac, “A reliable data aggregation mechanism with homomorphic encryption in smart grid ami networks,” in *Consumer Communications & Networking Conference (CCNC), 2016 13th IEEE Annual*. IEEE, 2016, pp. 550–555.

Chapter IV

Smart Grid Advanced Metering Infrastructure: Overview of Cloud-Based Cyber Security Solutions

Paper C

Smart Grid Advanced Metering Infrastructure: Overview of Cloud-Based Cyber Security Solutions

R.C Diovu and J.T Agee

published

International Journal on Communications Antenna and Propagation (IRECAP)

Abstract

The vision of making the traditional power grid smart has necessitated the incorporation of so many intelligent and smart technologies driven by ICT into virtually all domains of the smart grid. Expectedly, this has opened a lot of routes for vulnerabilities which can be maliciously exploited by today's cyber criminals. The advanced metering infrastructure (AMI) is believed to be the foundation of the smart grid. Malicious exploitations of security vulnerabilities present in the smart grid AMI by cyber criminals would be detrimental to both energy providers and final consumers. This is in view of the huge task involved in the management and storage of data from the numerous intelligent electronic devices (IEDs) that traverse the smart grid's AMI and beyond. A lack of capability in terms of processing, computations and storage of these data can only compound the cyber security problems facing the SG AMI. Cloud Computing is therefore a potential technology that can address the above challenging issues. Previous review works had concentrated mainly on the potentials of cloud computing for the smart grid. None of such works to the best of our knowledge had an extensive focus on the security of the SG AMI. We therefore provide in this paper firstly, a review of the cloud based solutions for the smart grid with appropriate SG domain classification in order to determine the focus points of previous related works. We then provide an elaborate overview on cloud based security solutions for the SG AMI. At the end, future research challenges were highlighted with open research ideas that could help to address the challenges also suggested.

Table C.1: Nomenclature

Symbol	Description
t_n	n^{th} time slot $n = 1...N_t$
$V(t_n)$	Vulnerabilities for a given n^{th} time slot
$P_A(t_n)$	Probability of an attack occurring in a given n^{th} time slot
$P_v(t_n)$	Probability of vulnerability existing for a given n^{th} time slot
$C(t_n)$	Consequence of an attack occurring at a given n^{th} time slot
$R(t_n)$	The system reliability at a given n^{th} time slot
V	Total number of vulnerabilities
$P_T(t)$	Probability of cyber threats assumed at n^{th} time slot
$P_R(t_n)$	Probability of a system risk happening for a given cyber threat at n^{th} time slot

1 Introduction

The smart grid has been considered by many as the future generation power grid. In [1, 2], the electric power grid was defined as electricity of network that can cost-efficiently integrate the behaviour and actions of all users connected to it in efforts towards economic efficiency, energy sustainability and better service delivery to end users. According to Melvin Greeret al [3], smart grid is the term applied to a class of technology designed to modernize the existing utility grid to intelligently and efficiently respond to available power generation, power transmission and consumer demand. The U.S Department of Energy (DOE) has identified seven properties required of the future smart grid and they include: attack resistance, self-healing, consumer motivation, power quality, generation and storage accommodation, enabling markets, and asset optimization [4]. A very important factor that can keep the smart grid stable and efficient is its supporting information infrastructure. Currently, the information supporting infrastructure of the electric power grid includes a system for remote monitoring and control, known as the Supervisory Control and Data Acquisition (SCADA) system, a set of applications used to operate the smart grid called the Energy Management System (EMS), the power system communication infrastructure, and the computational and storage resources [5]. The smart grid incorporates technologies such as the Phasor Measurement Units (PMU), Wide Area Measurement and Monitoring systems, Substation Automation, and Advanced Metering Infrastructure (AMI).

The deployment of technologies such as the AMI will greatly improve the reliability of the grid and reduce costs of power delivery, but they also present new dependencies on ICT resources which are vulnerable to different malicious cyber-attacks [6, 7]. For instance, a compromise of the metering networks may allow an attacker undue access to the control functions that, if corrupted, will threaten the availability of data in the system and consequently violate the integrity of the system. Eraj Khan et al [8], asserts that increase in the number of devices connected to the smart grid via technologies such as the AMI makes it more vulnerable to physical as well as cyber-attacks such as Trojans, Denial of Service (DoS) attacks, viruses and malwares. Indeed, the security threats targeting the smart grid landscape, and in particular the AMI, is continuously increasing. There are already many proposals for enhancing the security of the smart grid AMI in the literature today, but research achievements made already is no match with the sophisticated level of cyber-attacks targeting the AMI.

Incorporating cloud-based technology to the smart grid AMI does not solve out-rightly all the security issues envisaged in the AMI. However, considering the distributed nature of cloud computing and the smart grid, it is almost inevitable that the two technologies will be integrated [9]. On the contrary, non-

integration of cloud computing technology with the smart grid will only compound its security issues. Integrating the two technologies together will present a platform for grid operators or their authorized energy suppliers to satisfy the increased computations and enormous data storage capabilities required in a fully evolved smart grid environment [10–12]. With cloud-based computing system for the smart grid, services and applications can be hosted on a fault-tolerant and consolidated server-centric data centres with optimized QoS and security assurance mechanisms.

This survey introduces the smart grid AMI and the cloud-based solutions that have been proposed in the literature and geared towards the realization of the smart grid vision. There are few review works on the cloud computing and the smart grid and its integration but to the best of our knowledge, none of such research works have an extensive focus on the AMI and its cloud-based security solutions. We provide a brief summary of such research works and provide appropriate classifications of such works that partly dealt with the smart grid AMI and security. Such classifications will be helpful in distinguishing our research focus and what have been done in the subject area before now. It is our hope that this research paper will provide interesting overview about the state-of-the-art of smart grid cloud based security for the AMI which will be beneficial to future researchers, utility companies and standard bodies in power industries in many countries.

2 Overview of the Smart Grid Advanced Metering Infrastructure (AMI)

The smart grid AMI represents a configured infrastructure which incorporates a number of technologies to achieve its design objectives. Examples of infrastructures that could be incorporated to the AMI may include but not limited to: smart meters, data concentrators, customer interface units, back end systems such as meter data management system. A typical architecture of the smart grid AMI is provided in Fig. C.1. Brief descriptions of some of these basic components of the AMI have been provided in this paper. The AMI may also include numerous intelligent electronic devices which are mainly sensor based and used for collection of data and other grid control measurements usually desired depending on application requirements. A fully evolved AMI network could be made up of different sub-communication networks such as Home Area Network (HAN), Neighbourhood Area Network and Field Area Network (FAN). These networks are configured using different communication technology platforms such as: Power Line, ZigBee, Wireless Fidelity, WiMax etc.

The key components of this configured AMI technology is described briefly below [13, 14]:

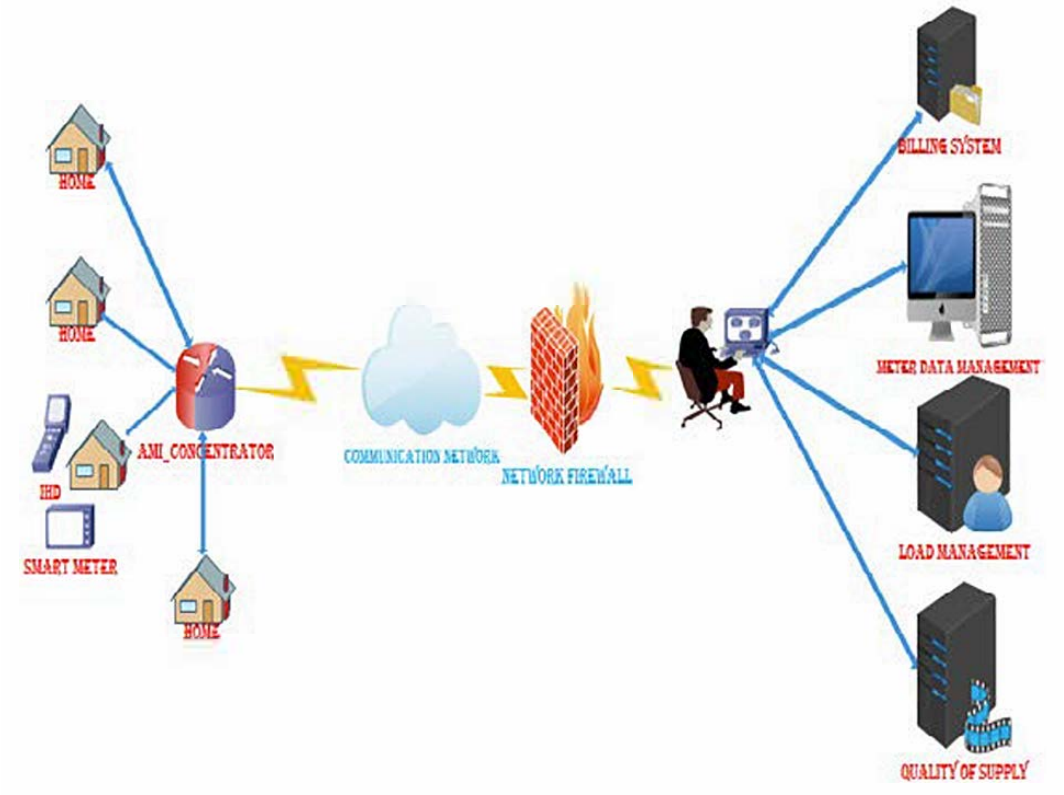

Fig. C.1: A Typical Architecture of a Smart Grid AMI Network

- Smart Meters: These are field devices with remote communication capabilities which provide the basic metering functionalities previously provided by the automatic meter readings (AMR). The smart meters provide additional functionalities such as: time-of-use metering, disconnection/reconnection, interval metering, tamper detection, outage detection, demand management and power quality management.

- Customer Interface Units (In-Home Displays): These are devices usually located at end user's premises to communicate with the smart meters. These devices display the end user's information such as consumption data, rate of consumption, disconnection notices, billing cost, status of load limiting and other important messages.

- AMI Master Stations: The master stations shall have the capability of integrating with the existing utility's management system such as billing, revenue protection and load monitoring systems. In addition they can extract in a secure manner data from the AMI system and shall also have the capabilities of the smart meters and other key components of the system.

- Data Concentrators (Data Aggregators): They facilitate communication between the smart meters, AMI master station and the back end systems. They usually act as a gateway between the backhaul and the field area network.

- Back end Systems: These include utilities and other authorized third parties that perform functions like billing, data storage and management, load management to mention a few.

3 Overview of Cloud Computing

Cloud Computing can be defined as a computing paradigm where systems are pooled together and connected either to a private or public networks so as to provide dynamically and in scalable fashion, infrastructure, application, and software for clients depending on their needs. This technology is irresistible for today's industrial applications as it reduces the cost of application hosting, computations and storage to the barest minimum. Clouds can be categorized as:

- Infrastructure as Service (IaaS).
- Platform as a Service (PaaS) and
- Software as a Service (SaaS).

This categorization is dependent on what scalable virtualization and abstraction that is provided. This may include the computing and data resources, development platform or even working software applications that are deployed. Typical IaaS providers like Amazon make provision for processing, storage, networks and possibly other computing resources like software and operating systems for the consumers. However, the consumers have limited control over selected networking components but no control over cloud physical infrastructure. PaaS providers allow utilities to integrate their customized applications with access control and identity management like active directory to the cloud platform [15]. Clouds can also be categorized as:

- Private cloud: This type of cloud platform is preferable in circumstances where privacy and security has more priority than all other issues like scalability, cost effectiveness and complexity. One disadvantage of this platform is that it limits interdependencies especially in situations where applications and services need to interact with each other.
- Public cloud: In this case, cloud-based services or applications are hosted in a public network like the internet. As a result, this type of cloud service is more vulnerable to cyber-attacks than other cloud platforms. However, it is less complex to implement than other platforms as a result of already deployed robust infrastructure of the internet. As such, it is more cost effective and also offers more promise for scalability.
- Hybrid cloud: As the name implies, this cloud combines both private and public networks in its

provision of cloud-based applications and services. This seems to be the most promising cloud platform for smart grid applications and services.

Some features of cloud computing which can be very beneficial to the smart grid include [16]:

- Agility: Smart grid applications can be reconstructed while communication between SG machines and cloud software can be provided using cloud enabled application programming interface (API) [17].
- Outsourcing and Virtualization: Applications and services can be made available to clients in an abstracted manner that runs as virtual machines instances.
- Flexibility: Cloud based solutions for the SG AMI would afford the grid operators and other third party employees the required flexibility as data/information can be made available either by connecting to office network or from a remote location using virtualized services.
- Elimination or reduction in the cost of hardware and other infrastructural investments: Other elaborate features of cloud computing that offer a lot of promise for the smart grid was summarized and presented in a pictorial form [18] and shown in this paper as Fig. C.2. The authors identified some challenges that could be envisaged towards the integration of cloud computing with the smart grid.

4 Previous Work on Smart Grid and Cloud Computing

The smooth running of most industrial systems like the electric power grid advocates for a reduction in operational expenses. Such reductions are usually supported by the provision of solutions that are capable of providing stability, fault-tolerance and flexibility. One of such solutions for industrial systems is the adoption of cloud based technology for the electric smart grid [19]. Clouds offer features that are well suited for the smart grid software platforms and applications. In this section, we provide previous review works by researchers on cloud computing and its possible integration with the smart grid. The objective of the section is to extract useful information on researches that had focused fully or partly on the cloud based security solutions for smart grid AMI. D.S Markovic [20] examined the opportunities of cloud computing for the different smart grid domains. The authors observed that cloud computing could be leveraged for large scale transmission and storage of data in the smart grid. This can ease off the burden required for real-time computations, transmissions and storage of enormous amount of data generated by smart grid technologies. In addition, they pointed that software and different applications can be outsourced or virtualized using different cloud computing

Fig. C.2: Cloud computing features beneficial for SG integration [18]

models such as SaaS, IaaS or PaaS. In line with their argument, PaaS and hybrid cloud model can be adopted for smart grid applications and services for the AMI, SCADA and EMS depending on the requirement. They noted that outsourcing different applications in the smart grid would help to reduce the burden of energy for computing and storage. From this review work, it is very obvious to note that the research work had concentrated on the opportunities of using cloud computing to impact the smart grid with respect to computations and storage. The work was silent on the potential promise of using the cloud to provide different security services and solutions for the smart grid AMI or even beyond. M. Yigit et al [18] also carried a review of the fundamental issues about cloud computing and the smart grid. Related works on smart grid applications and its integration with cloud platforms were also reviewed. In great details, they analyzed the opportunities and challenges of integrating the smart grid applications with the different cloud platforms. A juxtapose of the challenges and opportunities of cloud computing and smart grid is presented in Fig. C.3. With reference to Fig. C.3, cloud computing opportunities for the SG like security, scalability, cost efficiency, real time responses are contrasted with the challenges of cloud computing for the SG. Some of these challenges include: inefficient policy, compatibility, disaster recovery and data commingling. It is to be noted that data commingling in computing occurs when data of different kinds are stored such that they become easily

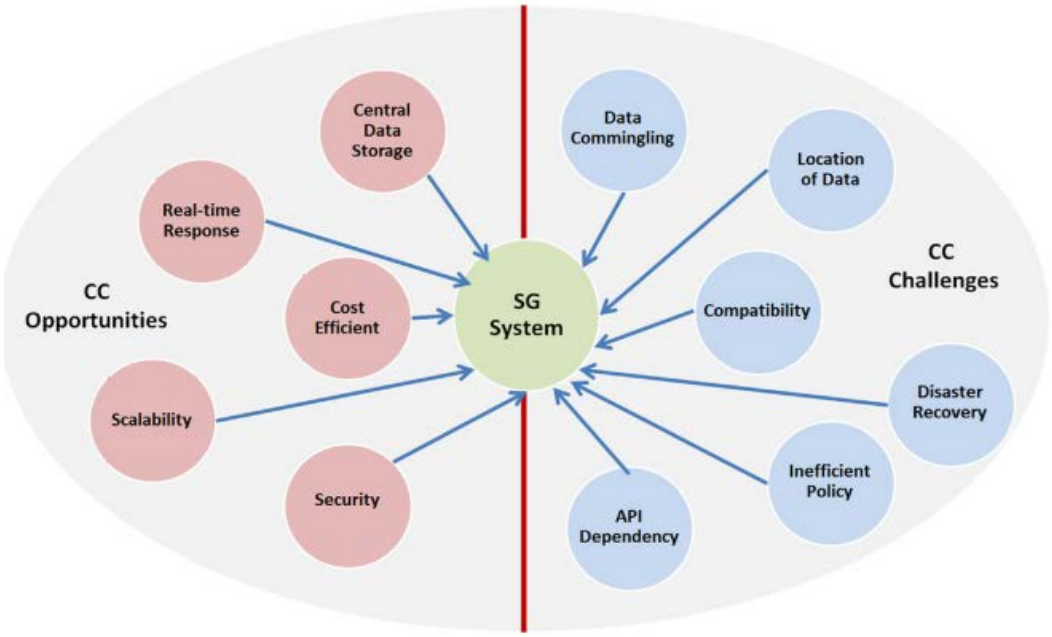

Fig. C.3: A Juxtapose of the Cloud Computing opportunities for the SG with CC Challenges [18]

accessible as an entity instead of being separated. On the other hand, a typical example of the cloud based application for the smart meter is presented in Fig. C.4. In Fig. C.4, SG AMI applications are placed into the smart meter application cloud. Conventionally, these services are developed, uploaded and updated by SG utilities. These services are accessed by smart meter via public communication interface and other intelligent control devices. The proposed framework provides effective solutions for SG AMI in a scalable and reliable manner.

Finally, the authors also reviewed some smart grid applications and their potential advantages to the smart grid. These applications and their advantages to the smart grid were also presented in tabular form. However, this table did not contain a domain by domain classification that could enable one to identify the exact smart grid domain that those applications were designed to impact. As a result, we modified this table to include the smart grid domain classification and then presented in Table C.2 while a summary of other previous review works considered in this section is summarized in Table C.3.

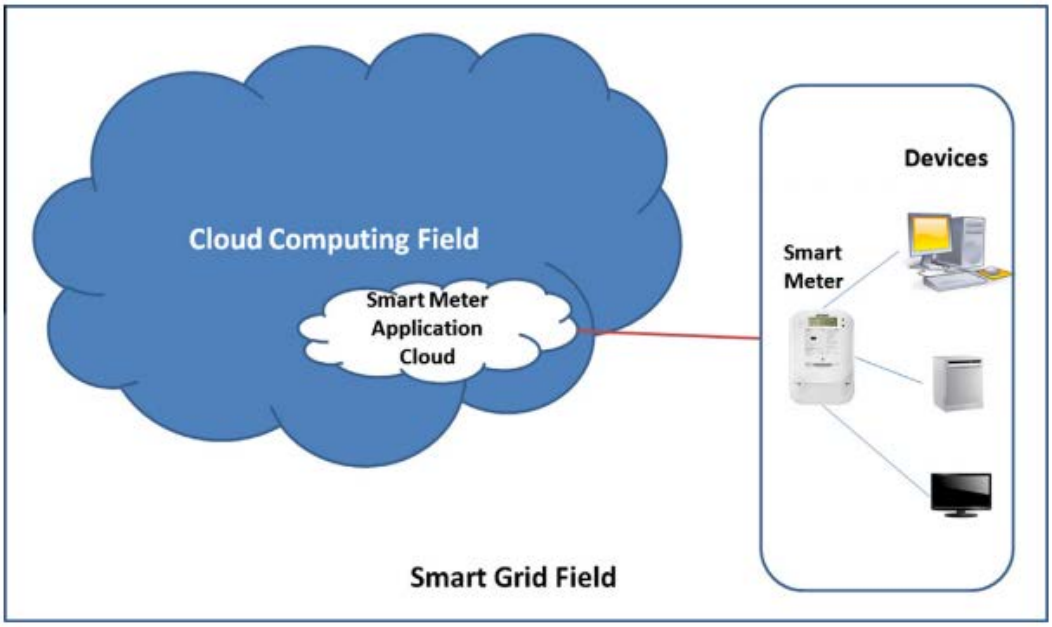

Fig. C.4: A framework for a cloud-based smart meter [21]

Table C.2: Cloud Based Smart Grid Applications/Services

CC Based Application	Domain	Potential Benefit to SG
DR Optimization and scheduling Gen. [22]	AMI	To ease off the burden on data storage
Cloud based smart meter [21]	AMI	For easy integration of new applications to SM
Cyber physical system for SG [23]	WAMMS	Increasing performance and efficiency through renewable resources
Intelligent cloud based energy management system (iCEMS) [24]	WAMMS	For security and load balancing
SG conditional monitoring CC Platforms [25]	WAMMS	For big data analytics and computations
Dynamic internet Data Centres Operations with SG and CC [26]	DR/AMI	Energy management
Netbook Advanced Metering Infrastructure (Net-AMI) [27]	AMI	Efficient communication

Table C.3: Summary of Previous Review Works

Reference	Focus	Comments on domain classifications
[18]	Focus was on SG applications and its integration with CC. Review also looked at fundamental issues about CC and the smart grid	There was no domain classification. However, some works reviewed partly dealt with the security of the SG.
[20]	Focused on opportunities of of using CC to impact the SG with respect to computations and storage.	No domain categorization was provided. The work was silent on the potential promise of using the cloud to provide different security services and solutions for the smart grid AMI or even beyond.
[28]	Focused on the potential Contributions of CC to the Smart grid	There was also no domain classification provided in the review work. However, the review work partly covered the potential contribution of cloud computing to the security of the smart grid.
[29]	Not a review work but the proposal was published at a date later than the above previous review works. Proposed work was titled: SMART FRAME: A Cloud computing based for big data information framework management of the SG	The proposed SMART FRAME was not specifically for the SG AMI. The authors adopted the identity based encryption (IBE) and identity based signature (IBS) security schemes for providing security for their proposed scheme.

The authors of [18] also opined that cloud computing platforms can be utilized for providing security and safety for smart grid applications. In their analysis, a close look at the different cloud platforms indicates that public clouds even though efficient, will be unreliable and insecure for smart grid applications. The following cloud-based security services/applications can improve the security of the smart grid:

- Cloud-based authentication and encryption service.
- A cloud-based information protection management based on well-designed policies.

In another research work, Baling et al [28] also provided a review on potential contributions of cloud technologies like cloud machine, cloud storage, cloud computing and cloud security to the smart grid. They noted that the applications of the above cloud technologies to the smart grid will bring

about changes that will impact positively on the smart grid. With cloud storage, enormous quantities of smart grid generated data can be stored with ease and this can be beneficial to the storage of distributed generated energy data. With cloud machine, energy consumptions can be drastically reduced and services can be seamlessly delivered to the different smart grid terminals. Similarly, cloud computing can impact positively on distributed and parallel processing technologies in order to provide large scale integrated processing capabilities. Finally, the authors showed that cloud security can help to reduce the cyber security risks for the smart grid to an acceptable level, thus, improving the disaster recovery capabilities of the smart grid. While the authors deserve some commendations for the insights provided by this review work, it is to be noted that there was no distinct categorization of the of the different smart grid domains that will be impacted by the above cloud technologies. This categorizations is very fundamental in giving a proper understanding of the impact of cloud computing to the smart grid, as mode of operation, storage, computing and security requirements for the different smart grid domains like SCADA, EMS, WAMMS and AMI would not be exactly the same. As a result, reviews on cloud based solutions for the smart grid ought to be narrowed down to a domain level. A secure cloud computing based framework for big data information management of smart grid otherwise referred to as the smart frame was proposed in [29]. It is important to note that the research work done in [29] was not a review work. However, considering that the research was done recently, it stands to reason that previous reviewers may not have accessed this work as most were done earlier than this work. The idea behind this framework is to set up a three-level hierarchical cloud at the top, regional and end-user levels. According to this proposal, the regional cloud shall be responsible for managing and processing of data from intelligent smart devices while the last level contains end-user smart devices. On the other hand, the top cloud takes care of managing general devices and accumulation of data across the regional level which is below it in hierarchy. In addition, the authors adopted the identity-based encryption (IBE) and identity-based signature security schemes that were firstly introduced in [30, 31] for the security of their proposed framework. It is important to mention here that the security schemes adopted by the authors in this proposed framework can be improved upon despite their advantages as explained by the authors. Recent research conducted in [32] indicates that IBE and IBS as currently implemented have some limitations. One of those notable drawbacks is the inherent key escrow property. This is an arrangement in which the keys needed to decrypt the encrypted data are held in a state referred to as escrow so that under certain conditions, an authorized party may gain access to those keys. This third party is allowed access under a carefully controlled condition. This is a technical security drawback as access to the protected data ought to be provided only to the intended recipient. As a result, researchers are already working on developing IBE and IBS variants that will eliminate or at

least mitigate the escrow feature. Another drawback in implementing the IBE and IBS schemes is the issue of high level of assurance required in the private key generator (PKG). Since the PKG holds all private keys, it requires a high level of assurance and availability as provided by certificate authority (CA). While CAs may if desired be disconnected from the network, the PKG must be available to furnish users with their private keys. This may increase the vulnerability of the system to cyber-attacks. As a result, future research should focus on developing PKG that are secured beyond the level of security currently obtainable from CAs. The observation that can be made from the review carried out in this section is that many proposals on the cloud based solutions were not made with security as the main focus. In addition, very few of the proposals were designed with specific intension of impacting the SG AMI. In reality, considering the distributed nature of the smart grid, it is unlikely that any cloud-based solution can benefit all the domains of the smart grid at same time as each domain has a different computing, storage, security and privacy requirements. However, cloud based solutions for the smart grid should be designed with scalability in mind, and in such a way that it will not interfere with the smooth functioning of the domains that were designed a priori to interface with other domains. It is therefore our considered opinion that future research on this subject should focus on giving a detailed analysis of how a particular cloud based solution will impact a given domain of the smart grid. Finally, this section has shown that cloud based security solutions for the smart grid in general (and even for the AMI) is very much at its infancy.

5 Cloud Based Security Solutions for the Smart Grid AMI

In this section, we focus exclusively on the review of cloud based security solutions for the smart grid advanced metering infrastructure (AMI). The reason for focusing on the AMI is not far-fetched. In the first instance, there are a lot of interdependencies between the AMI and the rest of the smart grid network. As such, a security compromise of one smart grid domain could result in a compromise of other domains with a catastrophic effect which is dependent on the cascading effect of the cyber-attack launched against it [33]. Secondly, cyber security challenges to be addressed for the traditional power grid in its evolution to the new smart grid would be entirely different from the challenges of cyber security in the conventional telecommunication and other information technology industries. In the case of the telecommunications and other IT industries, the rate of tolerance to cyber-attacks such as service disruption would be higher than that of the smart grid where the interdependencies on one domain on another will span across the different phases of generation, transmission and distribution of electricity to final consumers. In other words, the rate of tolerance of cyber-attacks to the smooth functioning of the smart grid is relatively low. Finally, the AMI is regarded as the building block and

the foundation of the new smart grid [34, 35], as such, lack of capability in handling security issues in the SG AMI could cripple the business for the energy markets. This can result in black-out for several thousands or millions of electricity consumers, thus, leading to colossal economic loss [36]. This makes the research on the cyber security of the SG AMI as compelling as it is urgent. For easy comprehension of this work, this section has been categorized into different cloud-based security solutions for the SG AMI based on the different objectives of cyber security. Before proceeding with this section, a quick overview of the objectives of cyber security for the smart grid has been provided.

5.1 Smart Grid Cyber Security Objectives

The legislation against violations of privacy and security of end user's consumption data in different countries today [37, 38] has placed a serious burden on how the smart grid vision ought to be pursued in such a way as to comply with these legislations. In addition, according to the guidelines for the smart grid cyber security released by the National Institute for Standards and Technology (NIST) [39], the Cyber Security Working Group has identified three major objectives for the smart grid security. It is therefore imperative that justifications for cloud based security solutions for the smart grid AMI be based on these cyber security objectives. In other words, the objective of the cyber criminal is to use malicious ways or by brute force to by-pass the access control schemes configured in a given cyber physical system. Fig. C.5 shows a contrast between cyber security objectives for the AMI and the criminal objectives of a cyber-criminal. The three objectives include:

- Confidentiality: Cyber security solutions for the smart grid should be designed to prevent unauthorized disclosure of information.
- Integrity: Integrity as an objective of cyber security for the smart grid entails that security solutions should be designed in order to prevent unauthorized modification or destruction of information or data.
- Availability: On the issue of availability as an objective of cyber security for the smart grid, security solutions designed to solve the problem of availability ought to prevent the disruption in the form of delay or outright denial of the use of data or information as and when needed.

It is therefore very important to understudy the cyber security risks of cyber physical systems like the SG AMI. If accurate models can be derived for the vulnerabilities, threats and attacks targeting such systems, then an equation for the cyber security risk can be given as [40]:

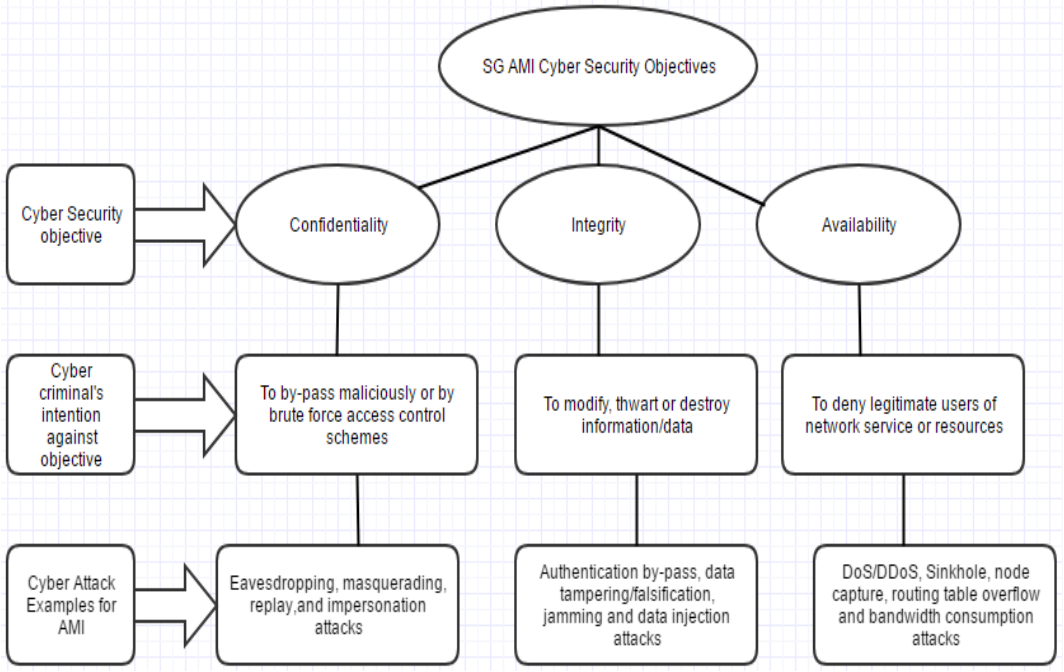

Fig. C.5: A conflicting objectives for the cyber-criminal and smart grid cyber security

$$Risk = SystemVulnerability * Threats * AttackConsequence \tag{C.1}$$

Models for system vulnerability, cyber threats and cyber attack consequence can be represented in equations C.2, C.3 and C.4 respectively [40]. System vulnerability, V at n^{th} time slot includes system configuration errors, or faults which can cause serious harm to the system. Therefore, the probability of system vulnerability for a given threat at n^{th} time slot can be given as:

$$P_v(t_n) = \frac{V(t_n)}{V} \tag{C.2}$$

where $V(t_n)$) is system vulnerability at n^{th} and V is the total number of system vulnerabilities in the system. Note that eqn. (C.2) is predicated on the assumption that vulnerability and threats are independent to each other. If vulnerability and threats represent two different events in probability theory which are independent to each other, it can be stated with clarity that the probability of one event happening does not in any way affect the probability of the other event happening. Otherwise, the probability would have to be calculated following the rules of conditional probability. Similarly, a cyber attack is an action maliciously intended to cause harm to the system. According to [40], the consequence of an attack occurring at n^{th} time slot is given by:

$$C(t_n) = \frac{P_A(t_n)}{P_T(t_n)} \tag{C.3}$$

where $P_A(t_n)$ is the probability of an attack occurring at n^{th} time slot and $P_T(t_n)$ is the probability of cyber threats assumed at n^{th} time slot. It can be noted that consequence here means the outcome or

effect of a malicious action. Attack on the other hand is an intuitive entity likely to cause harm. If the inter-arrival rate of the attack follows a poison distribution, then:

$$P_A(t_n) = e^{-\lambda t_n} * (e^{-\lambda t_n})^n$$

where n is the number of attacks occurring in a time interval, t_n. The consequence model is also dependent on the probability of threats assumed in the system in a given time t_n . This probability of threat $P_T(t_n)$ is defined in eqn. C4 for clarity. Finally, a cyber threat can be any adversary with sufficient motivation to exploit the vulnerabilities existing in a cyber system [41]. Therefore the probability of a cyber threat for a given system at n^{th} time slot can be given by:

$$P_T(t_n) = P_v(t_n) * R(t_n) \tag{C.4}$$

where $P_v(t_n)$ is the probability of the vulnerability for a given threat at n^{th} time slot, $P_T(t_n)$) is the probability of a threat assumed at n^{th} time slot, and $R(t_n)$ is the system reliability at n^{th} time slot. It can therefore be stated that the probability of a system risk happening for a given cyber threat at n^{th} time slot can be written as:

$$P_R(t_n) = P_v(t_n) * P_T(t_n) * C(t_n) \tag{C.5}$$

Where the terms $P_V(t_n,)$, $P_T(t_n)$ and $C(t_n)$ have been defined previously. The risk model which is represented in eqn. C5 can be used to predict or estimate the likelihood of a hazardous event happening to the system.

Summarily, the overall objective of cyber security for the smart grid ought to be pursued in such a way that risks arising from cyber threats and attacks targeting the smart grid be highly mitigated and reduced to the barest minimum. This entails that when any defense point is compromised or circumvented, the single point of vulnerability ought to be detected in a timely manner and stopped as quickly as possible in order to reduce the cascading effect of the attack failure.

5.2 Cloud Based Security Solutions for the Smart Grid AMI

Mikhail Simonov et al [42] proposed a cloud based software framework which integrates smart meters with the AMI using service-oriented architecture of the future Internet public-private partnership. In their approach, the authors provided a logical abstraction of the metering data providers who gather enormous data from different energy meters at the configured data rates. The key innovation in this proposal is the process-oriented and the opportunity to account for energy analytically. This proposal is a solution that seeks to improve privacy concerns that the data of energy

consumers could be compromised as data analytics would no longer be done in the public domain. In another development, K. Billewicz proposed a cloud based architecture for the smart grid AMI [43]. In this model, each smart meter will be equipped with a Wi-Fi and Power Line communication modules. In context, the smart meter would have a regular connection (once a day or once in every one hour) with the cloud based application. Alternatively, the smart meter can be constantly left in an online mode with connection to the cloud based application. In the design, the smart meter can send consumption information or events and would verify that the cloud based application does not have any requests related to the smart meter's configuration. The author argues that configuration commands such as change in tariff and change in billing method are not signals that are needed to reach the smart meter instantaneously. In this context, a delay associated with the smart meter connecting to the cloud service is reduced to an acceptable level. In summary, the implementation of this proposal will guarantee the uploading of real-time data to the cloud and this approach increases the safety and availability of those data. It is to be noted that this cloud based solution for the smart grid AMI is completely different from the majority of proposals found in the literature. While most proposals have concentrated on improving the computation, storage capabilities of the AMI or improving or safeguarding the integrity of information within the AMI, this approach tends to ensure the availability of data in the AMI. However, there was no analysis of this approach to prove the availability of data in the face of data availability attacks like denial/distributed denial of service attacks. Another interesting cloud based solution for the smart grid AMI was proposed by Md. Mahmud Hasan et al [44]. The authors proposed a framework called Encryption as a Service for the smart grid AMI (ES4AM). According to the authors, the AMI deals with and conveys an enormous amount of sensitive information which needs to be protected from unauthorized access or modification. The architecture for this framework is presented in Fig.C.6. The framework presented in Fig.C.6 features three levels of communication. In the first level, consumption information can be collected from smart meters and other intelligent sensor-driven electronic devices. Communication from consumers' homes to the next level in the SG AMI network is done through the Home Area Network Gateway (HAN GW) while Building Area Network Gateway (BAN GW) performs similar functions for the buildings connected to the network. The AMI concentrator is responsible for the aggregating of data from the HAN GW and BAN GW. In the third level, the aggregated data from the AMI concentrators is transmitted to the SG control center. Encryption is a candidate technology with primary security objective of preventing unauthorized access to information. Unfortunately, encryption brings about an added burden to the system in the form of overheads in computations and communications respectively. Thus, moving such a service or application to the cloud is not only innovative but also rewarding. The ES4AM was designed to deliver service for the AMI at the third

level of the AMI architecture where AMI concentrators aggregate and deliver aggregated data to the control center. The details of the key management and other cryptographic operations have been excluded from this work. As explained earlier, this kind of security solution can impact on the security and privacy of information within the smart grid domain. Furthermore, Bashar Alohali et al [45] proposed a cloud of things (CoT) based security for smart grid Home Area Network (HAN). With the CoT security based solution, a collection of appliances that will use real-time data can be enabled. The CoT is a virtualized internet of things (IoT) which also has the capability of providing monitoring and control. The paper explained how a device which is connected at the level of the HAN of the SG AMI can be serviced from the CoT security solution. In other words, the devices connected to the HAN are mapped to a virtual layer on a cloud infrastructure where the security solution is domiciled. The paper also illustrates how the security requirements of the network are administered from the CoT infrastructure. The security solution provided in this paper was designed for the prevention or mitigation of data confidentiality, availability and integrity attacks. The authors also included in their proposal, a symmetric key cryptography with a secure key management scheme for achieving confidentiality and integrity during end-to-end communications. The details of the group key distribution and management for their proposed solution is beyond the scope of this paper. This proposed work by the authors deserves some commendations especially in the performance analysis of the security solution against confidentiality and integrity attacks. However, the authors could not show clearly how their proposed solution could be used to prevent or mitigate data availability attacks such as denial of service or distributed denial of service attacks (DoS/DDoS) as claimed in their paper. Finally, Neetesh Saxena et al [46] proposed a lightweight cloud based authentication protocol for providing security solutions among HAN environment, energy providers, gateways, and advanced metering infrastructure network. This cloud based solution is a centrally controlled, distributed authentication protocol that provides authentication services among smart grid communication entities. According to the authors, the protocol provides security solutions such as privacy preservation, forward secrecy and protection against identity theft. The protocol also provides protection against cyber-attacks such as man-in-the-middle attack, impersonation attack, redirection attack, and denial of service attacks. Specifically, the cloud based authentication protocol provides authentication between the following pair of entities:

- Energy providers and the smart meters.
- Smart meters and the HAN gateway.
- BAN gateway and NAN gateway.
- NAN gateway and the cloud computing infrastructure.

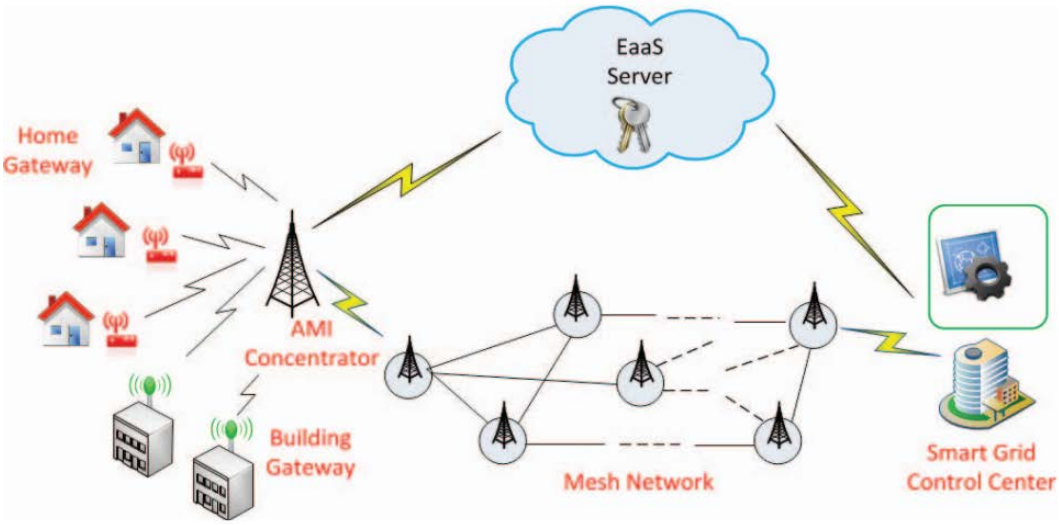

Fig. C.6: A framework for Encryption as a Service for smart grid AMI (ES4AM) [42]

In context, the authentication protocol provides a secure security solution which enables energy consumers to easily switch between energy providers seamlessly. It is to be noted that this security solution uses cloud based hierarchical trusted authorities (TAs) for key distribution and management with no key exchange or PKI complications. A schematic diagram for the hierarchy of the TAs in this solution is provided in Fig.C.7. According to Fig. C.7 above, the central cloud computing center is located at the first tier of the hierarchical architecture while the distributed cloud computing centre is located at the second tier. Finally, the smart meters, HAN, BAN, and NAN gateways, and the energy providers are located at the third tier of this architecture. This cloud based solution is designed such that various cloud computing services can be deployed at each trusted authority (TA). For instance, infrastructure-as-a-service (IaaS) can be deployed to aid in data/information collection, processing and storage. Similarly, platform-as-a-service (PaaS) can be deployed for developing and integrating different cloud based security applications while software as a service (SaaS) can be deployed for enhancing the optimization of energy usage. While a cloud based authentication security solution would be highly desirable for preventing or mitigating data integrity and confidentiality attacks for the SG AMI network, it is to be noted that it may not be very effective in combating data availability attacks as claimed by the authors. Simply put, security solutions seeking to mitigate or prevent data integrity or confidentiality attacks are designed primarily to control network users' access to the network and to define dynamically user's profiles required for granting access to network resources. Thus, authentication protocols only may not be sufficient for combating data availability attacks. In other words, there is urgent need to design and implement cloud based security solutions that have the capability of inspecting network packets and traffics with the intension of filtering or dropping illegal packets or traffics while at the same time allowing legitimate traffic or packets to pass through

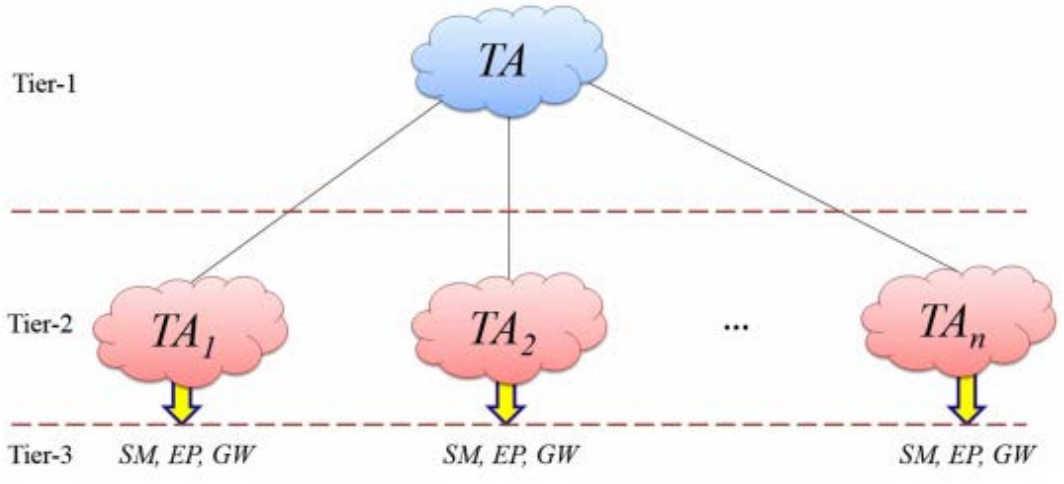

Fig. C.7: A hierarchical structure for the TA_s as implemented in [46]

to their destinations. Research has shown that today's cyber criminals have the technical skills needed for by-passing a lot of authentication protocols. Considering the dependencies of the SG AMI on the other SG domains, security solutions that can mitigate against data availability attacks in the event of a successful by-pass of authentication protocols should be designed and deployed at the right places within the SG AMI network. This is even more precarious given the different means of launching data availability attacks. For instance, distributed denial of service (DDoS) which is an example of data availability attack exploits numerous attack sources that are spread across so many hosts on a network thereby making the defense against it to be more complicated and its effect to be more damaging. In a client-server networks for instance, the cyber-criminal can initiate this kind of attack by trying to circumvent traditional security defenses by mimicking regular web traffics; and consequently initiating requests that cannot be detected by traditional firewalls and gateway securities.

5.3 Summary of Cloud Based Security Solutions for the Smart Grid AMI

In this section, a brief summary of the review of proposed cloud based security solutions for the smart grid AMI has been presented. There is no gain saying the fact that research in this area is still in its infancy. This is not to underscore in any way the contributions made by researchers in the proposed solutions reviewed in this paper. It is to be noted that all proposals aimed at preventing or mitigating data availability attacks failed to give an analysis of the strength of their proposed solution in the face of data availability attacks. A summary of the comparison based on cyber security objectives for the proposed security solutions have been presented in Table C.3 and Fig.C.8. Based on the analysis presented for proposals seeking to enhance data availability, their perceived impact is summarized in

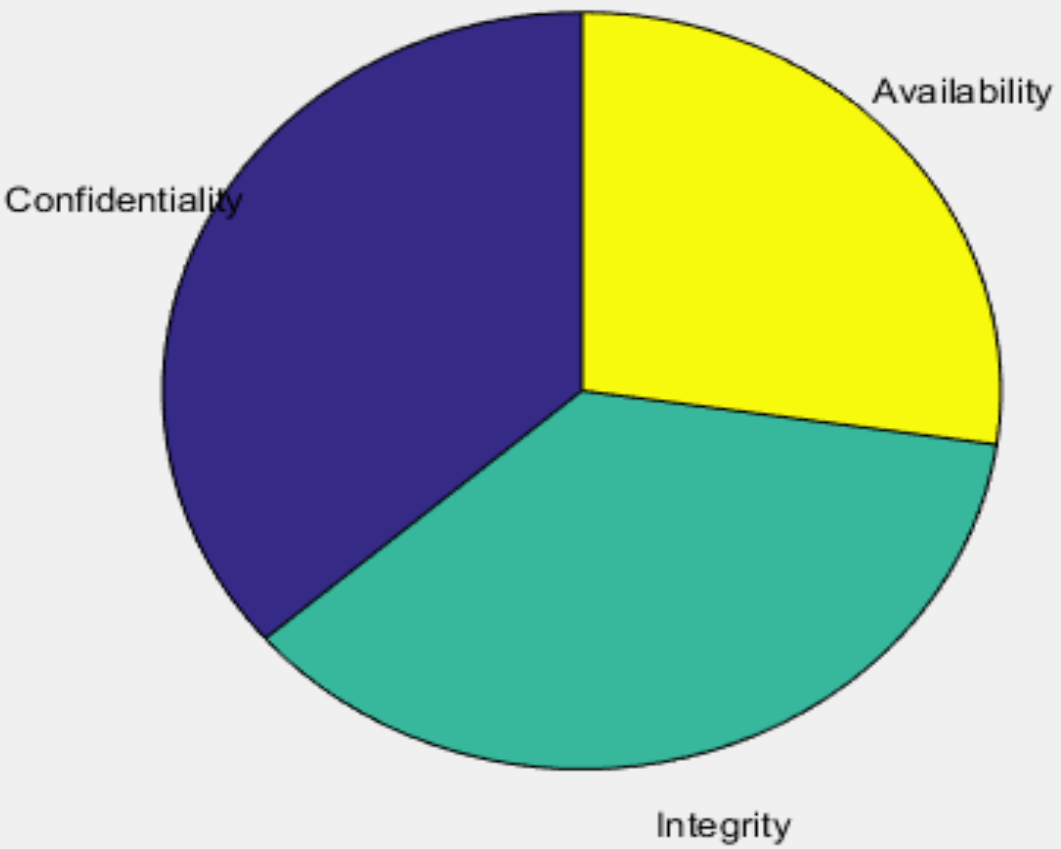

Fig. C.8: Summary of Cloud Based Security Solutions with regards to Cyber Security Objectives

Table C.5. In the key to rating for the perceived impact analysis, rating 1 stands for very low, rating 2 stands for low, rating 3 stands for moderate, while rating 4 and 5 stand for high and very high respectively. Most of the proposals on data integrity/confidentiality prevention or mitigation relied on security mechanisms such as authentication, access controls and cryptographic algorithms. Such security measures can bring about the needed data integrity/confidentiality for the SG AMI. However, caution must be applied in deploying such security mechanisms on the SG AMI in order not to limit the availability of data to authorized parties. This is in view of the overhead on computations and communications that can be incurred as a result of such protocols and algorithms. In other words, some security solutions can either be inapplicable, non-viable, insufficiently scalable, incompatible or simply inadequate in addressing the challenges posed by a highly complex environment such as the smart grid [47]. As a result, the SG interoperability panel and cyber security working group has advocated that research aimed at producing cost effective, attack resilient architectures for SG AMI and which are necessary for system level survivability and resiliency to be given utmost priority [48].

Table C.4: Comparison of the cloud based security solutions for the SG AMI

Proposed solution	Methodology	Cyber security objective pursued
[42]	software based design	Integrity and confidentiality
[43]	Application/Software-based design	Availability
[44]	Cryptography	Integrity and confidentiality
[45]	Authentication and cryptography	Integrity, confidentiality and availability
[46]	Authentication	Integrity, confidentiality and availability

Table C.5: Perceived impact of the cloud based security solution on data availability for the SG AMI

Proposed solution	Analysis in the face of attack	perceived impact of proposed solution
[43]	No	3
[44]	No	2
[45]	No	3

6 Future Challenges and Open Research Issues

It is to be noted that cyber threats targeting the smart grid AMI will continue to evolve in the coming years in terms of their technology, sophistication levels and dynamics. As a result, research solutions seeking to prevent or reduce to the barest minimum vulnerabilities arising from cyber threats against the SG AMI, would require continuous effort. Stronger security controls requiring the combinations of different security methodologies would be highly desirable. Such solutions ought to guarantee instant detection, mitigation and protection against vulnerable points on the SG AMI network. A security solution for the future SG AMI ought to include scalability in the heart of its design given the distributed nature of the SG AMI which will inevitably result in periodic expansion of the network. Finally, security solutions for the future SG AMI should be designed to be robust; considering the enormous amount of data that transverse the entire SG AMI network and beyond. Therefore, a lack of capability for handling potential vulnerabilities within the network could lead to an increase in data confidentiality, integrity and availability attacks. Most cloud based security solutions for the SG AMI have been designed using cryptographic techniques or authentication schemes or both. It is to be noted that many of the solutions are not suitable for addressing the above challenges. Specifically, most solutions available in the literature are not very suitable for mitigating or preventing data availability attacks like distributed denial of service (DDoS) attacks. Therefore, an untapped research direction

for future research on cloud-based security solution for the SG AMI is the use of firewalls. Security firewall remains one of the most important security defense mechanisms to be implemented for the SG AMI network. The technological innovation that resulted in the use of firewalls happened in the early 1990s [49] but its efficacy cannot be doubted even till date. Detailed information about the history of firewalls can be found in [50, 51]. Simply put, a firewall is a collection of security measures that are designed to prevent unauthorized access to a networked environment [52]. Firewalls can also be designed to prevent data leakage, to prevent access to data or information or to enforce rules or policies as to which traffic or packet should be stopped, rejected or passed to its destination [53, 54]. A cloud based firewall for the SG AMI can be designed to address the following challenges:

- To eliminate the need for the procurement of on-premises security hardware in the event of network expansion.
- To protect SG AMI back-end servers with no single point of failure.
- To filter the ingress and egress traffic of SG AMI back-end servers, thus providing a secure segmentation of all application traffic.
- To guarantee automatic, instant detection, mitigation solutions and protection against flood based DDoS vulnerable points on the SG AMI.
- To mitigate against volumetric flood-based attacks (up to 200 Gbps and above) which cannot be handled easily by traditional firewalls and other gateway securities.
- To provide SG AMI server load balancing to mask servers and applications in real time.

7 Conclusion

The enormous amount of data traversing the SG AMI has necessitated the incorporation of cloud computing technology into the security solutions designed for the SG AMI. This huge amount of data would require a lot of capabilities in terms of computations and storage. Although cloud computing as a technology has its own vulnerability to cyber-attacks, it is our argument in this paper that non-incorporation of cloud computing to the security solutions designed for the SG AMI, would only aggravate the security vulnerabilities of the SG AMI. These vulnerabilities would consequently result in an increased success rate for cyber-attacks against the integrity, confidentiality and availability of data/information. In this review paper, a review of the cloud based solutions for the SG with a focus on the cloud based security solutions for the SG AMI has been provided. It was discovered that most cloud based security solutions proposed for the SG AMI may not address the future security

challenges facing the SG AMI. It was therefore noted that future security challenges for the SG AMI can be reasonably resolved by leveraging on cloud computing based firewall security solutions.

References

[1] K. O. Aduda, W. Zeiler, G. Boxem, and T. Labeodan, "On defining information and communication technology requirements and associated challenges for 'energy and comfort active'buildings," *Procedia Computer Science*, vol. 32, pp. 979–984, 2014.

[2] C. Clastres, "Smart grids: Another step towards competition, energy security and climate change objectives," *Energy Policy*, vol. 39, no. 9, pp. 5399–5408, 2011.

[3] M. Greer and M. Rodriguez-Martinez, "Autonomic computing drives innovation of energy smart grids," *Procedia Computer Science*, vol. 12, pp. 314–319, 2012.

[4] U. NETL, "A systems view of the modern grid," *White Paper, Jan*, 2007.

[5] S. Meiling, T. Steinbach, T. C. Schmidt, and M. Wählisch, "A scalable communication infrastructure for smart grid applications using multicast over public networks," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. ACM, 2013, pp. 690–694.

[6] 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, 2012.

[7] E. B. Rice and A. AlMajali, "Mitigating the risk of cyber attack on smart grid systems," *Procedia Computer Science*, vol. 28, pp. 575–582, 2014.

[8] E. Khan, B. Adebisi, and B. Honary, "Location based security for smart grid applications," *Energy Procedia*, vol. 42, pp. 299–307, 2013.

[9] A. Califano, E. Dincelli, and S. Goel, "Using features of cloud computing to defend smart grid against ddos attacks," in *10th Annual symposium on information assurance (Asia 15), ALBANY*, 2015, pp. 44–50.

[10] E. Brynjolfsson, P. Hofmann, and J. Jordan, "Cloud computing and electricity: beyond the utility model," *Communications of the ACM*, vol. 53, no. 5, pp. 32–34, 2010.

[11] M. T. Khorshed, A. S. Ali, and S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing," *Future Generation computer systems*, vol. 28, no. 6, pp. 833–851, 2012.

[12] L. Zheng, S. Chen, Y. Hu, and J. He, "Applications of cloud computing in the smart grid," in *Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011 2nd International Conference on*. IEEE, 2011, pp. 203–206.

[13] H. Groenewald, "Nrs049—advanced metering infrastructure (ami) for residential and commercial customers," *ESKOM, Johannesburg, Presentation*. [Online]. Available: http://www.ameu.co.za/Portals/16/Conventions/

[14] K. Subramoney and H. Madhoo, "Value proposition of advanced metering infrastructure- internal stakeholder communication," *Eskom Presentation Report, version 4*, 2013.

[15] Y. Simmhan, B. Cao, M. Giakkoupis, and V. K. Prasanna, "Adaptive rate stream processing for smart grid applications on clouds," in *Proceedings of the 2nd international workshop on Scientific cloud computing*. ACM, 2011, pp. 33–38.

[16] J. Baek, Q. H. Vu, J. K. Liu, X. Huang, and Y. Xiang, "A secure cloud computing based framework for big data information management of smart grid," *IEEE transactions on cloud computing*, vol. 3, no. 2, pp. 233–244, 2015.

[17] J. POPEANGĂ, "Cloud computing and smart grids," *ERP and E-Business Application Deployment in Open Source Distributed Cloud Systems*, vol. 3, pp. 57–66, 2012.

[18] M. Yigit, V. C. Gungor, and S. Baktir, "Cloud computing for smart grid applications," *Computer Networks*, vol. 70, pp. 312–329, 2014.

[19] L. Yu, T. Jiang, and Y. Zou, "Real-time energy management for cloud data centers in smart microgrids," *IEEE Access*, vol. 4, pp. 941–950, 2016.

[20] D. S. Markovic, D. Zivkovic, I. Branovic, R. Popovic, and D. Cvetkovic, "Smart power grid and cloud computing," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 566–577, 2013.

[21] X. Fang, S. Misra, G. Xue, and D. Yang, "Managing smart grid information in the cloud: opportunities, model, and applications," *IEEE network*, vol. 26, no. 4, 2012.

[22] Y. Simmhan, S. Aman, B. Cao, M. Giakkoupis, A. Kumbhare, Q. Zhou, D. Paul, C. Fern, A. Sharma, and V. K. Prasanna, "An informatics approach to demand response optimization in smart grids," Tech. Rep., 2011. [Online]. Available: https://www.semanticscholar.org/paper/

[23] J. Byun, Y. Kim, Z. Hwang, and S. Park, "An intelligent cloud-based energy management system using machine to machine communications in future energy environments," in *Consumer Electronics (ICCE), 2012 IEEE International Conference on*. IEEE, 2012, pp. 664–665.

[24] X. Jin, Z. He, and Z. Liu, "Multi-agent-based cloud architecture of smart grid," *Energy Procedia*, vol. 12, pp. 60–66, 2011.

[25] H. Bai, Z. Ma, and Y. Zhu, "The application of cloud computing in smart grid status monitoring," in *Internet of Things*. Springer, 2012, pp. 460–465.

[26] P. Wang, L. Rao, X. Liu, and Y. Qi, "D-pro: Dynamic data center operations with demand-responsive electricity prices in smart grid," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1743–1754, 2012.

[27] K. Nagothu, B. Kelley, M. Jamshidi, and A. Rajaee, "Persistent net-ami for microgrid infrastructure using cognitive radio on cloud data centers," *IEEE Systems Journal*, vol. 6, no. 1, pp. 4–15, 2012.

[28] B. Fang, X. Yin, Y. Tan, C. Li, Y. Gao, Y. Cao, and J. Li, "The contributions of cloud technologies to smart grid," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1326 – 1331, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032116000629

[29] J. Baek, Q. H. Vu, J. K. Liu, X. Huang, and Y. Xiang, "A secure cloud computing based framework for big data information management of smart grid," *IEEE Transactions on Cloud Computing*, vol. 3, no. 2, pp. 233–244, April 2015.

[30] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology*, G. R. Blakley and D. Chaum, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985, pp. 47–53.

[31] D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in *Advances in Cryptology — CRYPTO 2001*, J. Kilian, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 213–229.

[32] D. Kalyani and R. Sridevi, "Survey on identity based and hierarchical identity based encryption schemes," *International Journal of Computer Applications*, vol. 134, pp. 32–37, 01 2016.

[33] E. Liu, M. L. Chan, C. W. Huang, N. C. Wang, and C. N. Lu, "Electricity grid operation and planning related benefits of advanced metering infrastructure," in *2010 5th International Conference on Critical Infrastructure (CRIS)*, Sept 2010, pp. 1–5.

[34] D. Wei, Y. Lu, M. Jafari, P. Skare, and K. Rohde, "An integrated security system of protecting smart grid against cyber attacks," in *2010 Innovative Smart Grid Technologies (ISGT)*, Jan 2010, pp. 1–7.

[35] W. Boyer and S. A. McBride, "Study of security attributes of smart grid systems: Current cyber security issues," 05 2018. [Online]. Available: https://www.smartgrid.gov/files/Study_Security_Attributes_Smart_Grid_Systems_Current_Cyber_200903.pdf

[36] D. G. Hart, "Using advanced metering infrastructure to realize the smart grid," *IEEE Power and Energy Society General Meeting*, pp. 1 – 2, August 2008.

[37] P. Schwartz and J. R. Reidenberg, *Data privacy law: a study of United States data protection*. LEXIS law, 1996.

[38] M. Calaguas, "South african parliament enacts comprehensive data protection law: An overview of the protection of personal information bill," *Africa Law Today*, vol. 7, 2013.

[39] N. S. Grid, "Introduction to nistir 7628 guidelines for smart grid cyber security," *Guideline, Sep*, 2010.

[40] M. A. Khan and M. Hussain, "Cyber security quantification model," in *Proceedings of the 3rd international conference on Security of information and networks*. ACM, 2010, pp. 142–148.

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

[42] M. Simonov, G. Daltoe, G. Zanetto, and R. Conti, "Smart meters using the architecture of future internet," in *PowerTech, 2015 IEEE Eindhoven*. IEEE, 2015, pp. 1–6.

[43] K. Billewicz, "The use of cloud computing in ami system architecture," in *Modern Electric Power Systems (MEPS), 2015*. IEEE, 2015, pp. 1–6.

[44] M. M. Hasan and H. T. Mouftah, "Encryption as a service for smart grid advanced metering infrastructure," in *Computers and Communication (ISCC), 2015 IEEE Symposium on*. IEEE, 2015, pp. 216–221.

[45] B. Alohali, M. Merabti, and K. Kifayat, “A cloud of things (cot) based security for home area network (han) in the smart grid,” in *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2014 Eighth International Conference on*. IEEE, 2014, pp. 326–330.

[46] N. Saxena and B. J. Choi, “Integrated distributed authentication protocol for smart grid communications,” *IEEE Systems Journal*, 2016.

[47] Y. Mo, T. H.-J. 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, 2012.

[48] U. NIST, “Guidelines for smart grid cyber security (vol. 1 to 3),” *NIST IR-7628, Aug*, 2010.

[49] N. Amalina, R. Alsaqour, M. Uddin, O. Alsaqour, and M. Al-Hubaishi, “Enhanced network security system using firewalls,” *ARPN Journal of Engineering and Applied Sciences*, vol. 8, No. 12, 2013.

[50] R. Braden, D. Clark, and S. Crocker, “C. huitema," report of iab workshop on security in the internet architecture,” RFC 1636, USC/Information Sciences Institute, MIT, Trusted Information Systems, INRIA, Tech. Rep., 1994.

[51] A. Muffett, “Proper care and feeding of firewalls,” in *In Proceedings of the UKERNA Computer Security Workshop. United Kingdom Education and Research Networking Association, Atlas*. Citeseer, 1994.

[52] W. R. Cheswick, S. M. Bellovin, and A. D. Rubin, *Firewalls and Internet security: repelling the wily hacker*. Addison-Wesley Longman Publishing Co., Inc., 2003.

[53] G. R. Barme and S. Ye, “The great firewall of china: At isps, internet cafes, even state censorship committees, we meet the wired of china,” *WIRED-SAN FRANCISCO-*, vol. 5, pp. 138–149, 1997.

[54] M. J. Ranum, “A network firewall,” in *Proceedings of the World Conference on System Administration and Security, Washington, DC*, 1992.

Chapter V

Data Aggregation in Smart Grid AMI Network for Secure Transfer of Energy User-Consumption Data

Paper D

Data Aggregation in Smart Grid AMI Network for Secure Transfer of Energy User-Consumption Data

Remigius Chidiebere Diovu and John Terhile Agee

Published

International Journal of Engineering Research in Africa

Abstract

In a secured smart grid AMI environment, congestion management during data aggregation with security encryption for privacy preservation is a challenging issue. By introducing data communication network schemes into the Advanced Metering Infrastructure (AMI), network traffic congestion and service rates can be improved while preserving user's privacy from the grid operator's end. In this paper, a resilient architecture called Ring Triangulation Communication Architecture (RTCA) for data aggregation and user privacy protection is proposed. To preserve privacy as well as for reducing traffic congestion in the architecture, DMF homomorphic encryption algorithms were formulated for local concentrators while using a global concentrator to check for anomalies in the AMI server clusters. With TCP/IP protocol and IEEE 802.11 MAC/PHY on the network, TCP message flooding was contextualized for congestion scenario. Stochastic TCP congestion management schemes with wired equivalent privacy (WEP) and the Data Minimizing Function (DMF) scheme were compared. Our proposed architecture significantly reduced transmission congestion and cryptographic overheads incurred during message aggregation. The results of the performance of the DMF Homomorphic encryption scheme incorporated into our proposed architecture for the SG AMI were discussed. These include service rate and other QoS metrics which are negatively affected by a congestive network condition.

1 Introduction

The advanced metering infrastructure (AMI) makes it possible to realize a multi–way communication between electricity users and its service providers [1]. A reliable smart grid model allows energy consumers to participate in the overall energy value chain. Hence, there is an interaction between the content providers and utility operators [2]. The exchange of information by consumers helps them to smartly adjust their consumption profile in the most efficient manner [3]. Usually, these smart meters are used to extract user-centric details both from residential and nonresidential contexts. It records energy used in the time range of cycles, amounting to 60 secs per cycle or less and feeds back the service provides for predictive analysis and billing profiling. With this, the utility providers can use point estimation to ascertain the overall energy demands and its flow [3]. Preserving the privacy of energy consumption data can be achieved by designing robust and resilient communication architectures for the AMI. Such architectures ought to incorporate state-of- the-art cryptographic algorithms and data aggregation protocols that will preserve the privacy and security of end user's

energy information in the face of different cyber-attacks. Communication of user data is done through wireless medium depending on the network architecture. Most models use a multi-hop architecture so as to relay received data to a concentrator or middleware gateway [4]. This is usually applied when multiple meters are connected together so as to route their data with a dedicated access point. With respect to data storage and cluster server processing, data aggregation in AMI network for secure transfer of energy user-consumption data must take cognizance of congestion control and user data privacy. Such networks ought to have common characteristics like latency, throughput, resource utilization, abstracted security of end users. When congestion occurs in AMI network, processing delay could lead to the saturation of network resources or the eventual grounding of the network. Secure user details which have cryptographic overheads can add to the delays. As a result of this kind of processing delay, services in smart grid AMI may suffer oscillations and packet dropping probability may increase. Data aggregation for security and privacy preservation in AMI network has attracted appreciable level of research attention as a consequence. Three major categories of non-trivial data aggregation techniques have been highlighted in the literature, viz: (1) data aggregation protocols utilizing trusted third parties (TTPs) (2) protocols using perturbation and (3) protocols using cryptographic algorithms. Aggregation protocols that rely on TTP [5, 6] have one major drawback. This is because the scheme must rely on one entity which must be assumed to be trustworthy. With such arrangement, consumers would always be worried that their consumption data can be compromised. In order to overcome the above drawback, aggregation by perturbation was proposed [5, 7–11]. The main idea behind this aggregation scheme is to add noise or some means of randomness to the metering data of each smart meter so that the aggregating node or entity does not infer the metering data of each individual SM. Aggregation protocols by means of perturbation suffers from two major drawbacks. Firstly, aggregated metered noisy or fake data will not amount to exactly the aggregated real metered data due to the added randomness to each SM data. Secondly, if some smart meters do not deliver all their noisy metering data successfully to the aggregating node or entity, authorized entities would find it difficult to obtain a good approximation of the real aggregated data. The above drawbacks can be overcome by using cryptographic data aggregation based protocols.

Data aggregation schemes using cryptographic algorithms can be categorized as: (1) protocols using secret sharing schemes and (2) protocols using homomorphic encryption. Aggregation protocols using secret sharing schemes [12–18] generally have two major drawbacks which include: (1) high computational costs at the SMs as each SM needs to securely connect with the respective aggregating nodes. (2) They suffer from high communication overheads in the system as each SM needs to send a number of shares of its metering data to different entities. The above drawbacks can be reduced by

utilizing homomorphic encryption algorithms for privacy preservation. Unfortunately, most of the existing aggregation protocols utilizing homomorphic encryption have one major drawback. This is because they were designed for a single entity [19] recipient system which is not ideal for a real SG environment. In addition, congestion scenarios resulting in delays incurred from aggregation schemes have not been thoroughly studied. This paper uses our proposed Ring Triangulation Communication Architecture (RTCA) to overcome the above mentioned drawback as it is designed for a multi-party recipient system model. In addition, our model was designed to transmit data packets using the TCP/IP protocol which is crucial for reliable services [20] and contextualized herein this paper for congestion scenarios. However, in using TCP message flooding as shown in Fig. D.1 and Fig. D.2, the metered data of a user can become a target for malicious attackers when transmitted. A reliable security must therefore be applied to re-enforce integrity. Furthermore, our proposed scheme assumes a comprehensive version of the AMI which should have the capability of interconnecting numerous intelligent electronic devices (IEDs) driven by sensors [21]. As such, a SM can serve as a gateway for variety of real time signals such that profiling via smart-metering and other grid control metrics are transmitted through a robust wireless sensor architecture [22]. Finally, our scheme incorporates a DMF algorithm which is used not only to preserve sensed metering or control operational data but to preserve privacy of the data without an additional message exchange (overhead) during aggregation. The proposed scheme also incorporates a Hilbert-curve based technique [23, 24] which makes it difficult to spoof/hijack the actual sensed data by cyber attackers. In this context, even if an attacker tries to overhear it, the sensed data is changed by the homomorphic based DMF algorithm. The effect of packet aggregation during a congestion scenario on the security overhead of the transmitted payload was sufficiently considered.

This follows the discovery made from literature review carried out in section 2 (and also in chapter 2 of this thesis) which showed that data aggregation provides privacy for user's consumption data but can drive the SG AMI to a congestive scenario because of increased overheads in communication and computations. Justifications for proposed data aggregation schemes in the literature had always been done by analyzing the complexity of cryptographic primitive operations contained in such protocols. One open issue that had remained in this research domain is the implementation of the already proposed data aggregation protocols on communication network standards recommended for the SG AMI; and then analyzing the quality of service (QoS) that are likely to be affected in a congestive network scenario. Rather than concentrating on making justifications on overheads due to computations and communication, this chapter will consider applying our proposed data aggregation scheme to IEEE 802.11 network standard and then analyzing the performance of important QoS parameters. Since congestion is the main problem associated with data aggregation protocol, some

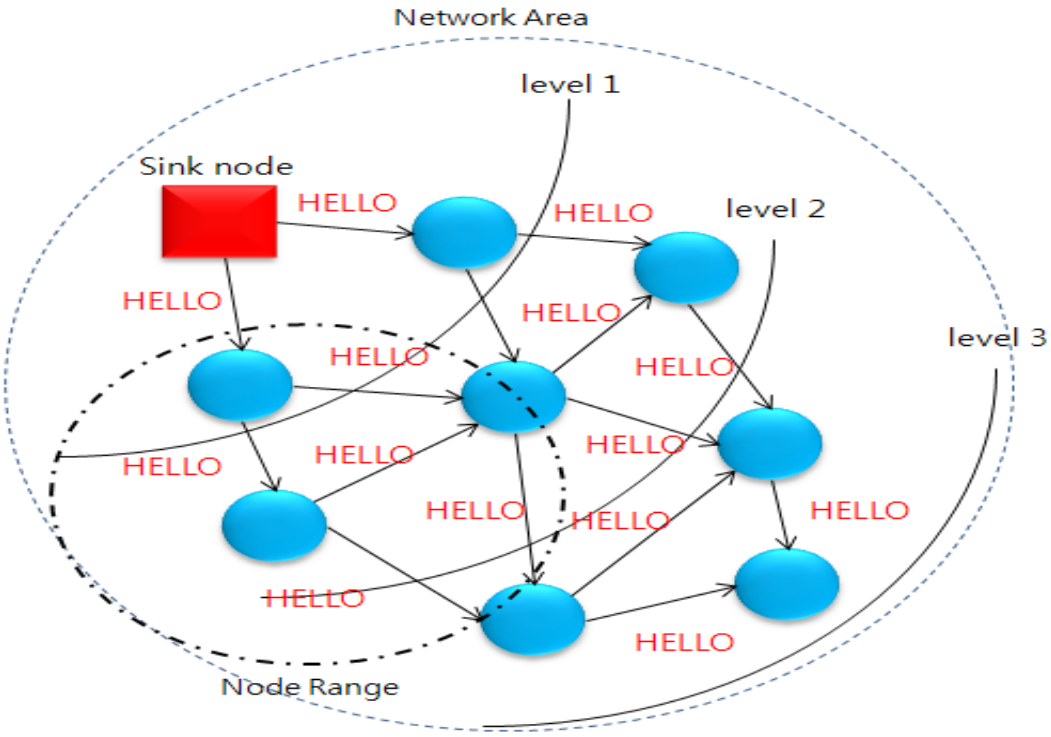

Fig. D.1: Message Flooding [24]

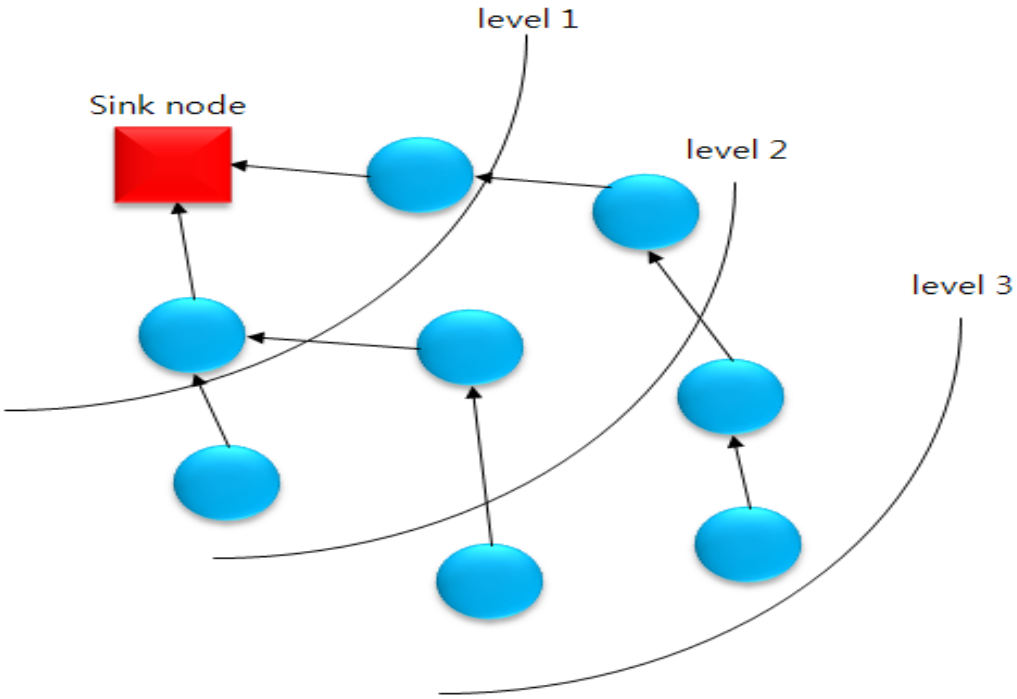

Fig. D.2: Network construction [24]

network QoS performance of the proposed scheme will be compared with the QoS guaranteed by TCP congestion control schemes which have been proposed in the literature, presumably tested, prototyped and already implemented in the IEEE 802.11 wireless standard. These TCP control schemes include- TCP Tahoe, TCP Reno, TCP New Reno, TCP Sack and TCP Vegas. After several reports of congestion collapses in October 1986, the first version of the TCP control scheme known as TCP Tahoe was proposed by Van Jacobson [25] in 1988 and was modified by the same author in 1990 as new variant known as TCP Reno [26]. TCP Reno was later modified by Brakmo et al in 1995 and Mathis M et al in 1999 into TCP Vegas and TCP Sack respectively [27].

The rest of this research work is arranged as follows: In the next section, privacy preservative data aggregation (PPDA) approaches relating to our proposed scheme is discussed. This is followed by the methodology which includes the system formulation, design, implementation and validation. Finally, scenario evaluation and result analysis were presented, while our conclusion represents the last part of this paper.

2 Related Research Efforts

There are existing data aggregation proposals which relates to our scheme. These are used in addressing privacy and integrity of critical data. This section presents efforts in this area. The work in [28] presented a secure aggregation scheme in which SM information is uniquely mapped leading to new singular packet aggregation. Their approach preserves the uniqueness of every SM. The processing in their scheme introduces expensive overhead owing to congestion. To address the limitation, the authors in [16, 29, 30] proposed H- encryption that focused on data aggregation in order to reduce its computation overheads and still secure locally generated but consumed data. The authors opine that there is no need for sequential decryption considering the core concentrator such that the base data is uniformly generated. The work in [31] introduced a scheme for the SG AMI network that optimizes network traffic as well as the service rate time. This also, retains the privacy of energy user details. To account for computational overhead in smart grid applications, the network characteristic was modeled. However, their network communication structure has a defective formation for their contextual PECA scheme. Also, the two topologies used failed to account for congestion management at the event of over subscription at the local concentrators. A cryptographic compression layer which could reduce overhead at the local concentrator was not discussed. In [19], the work proposed a Ring Communication Architecture (RCA) which leverages on the capability of an orthogonal code mechanism for privacy protection of the customer's data. Their system focused

on privacy issues based on RCA. This is realized via RCA network measurement and payment processes involving complex mathematical formulations. In most of these works, the security of the control center is not guaranteed since the utility control center is exposed to external denial of service attacks.

Other privacy preserving data aggregation schemes [24, 32–34] which are based on wireless sensor networks are presented here briefly. W.B He et al [32], presented a Cluster-based aggregation (CBA) scheme for network optimization. From the work, CBA scheme suffers from high communication overhead owing to the amount of communication needed to perform optimal data fusion. The work done in [32] presented SMART method for preserving the privacy of data using a data slicing technique. In this work, the node selection process introduces errors into the data transmission and reception from neighbors. It literally suffers from high communication overhead arising from the burden of a node division of neighbor load sharing [24].

Conti et al [33], proposed a data preservation scheme otherwise referred to as Twin-key. This scheme is robust to data loss particularly during congestion scenario as it can prevent the leakage of the sensed data during data aggregation. The design incorporates the setting up of a Twin-key during a ring circuit like cluster construction [34]. In this arrangement, two neighboring nodes could provision at least one common key mapping to a given hash value. As a result, data aggregation is performed twice resulting in each node adding its sensed value to the partial aggregate value. Similarly, this scheme like other schemes mentioned earlier suffers from high communication overhead, in this case, generated from the process of physical announcement and logical data aggregation. In [35], the authors advocated for a generic aggregation and privacy preserving scheme known as GP2S. Generic Privacy Preservation Scheme (GP2S) which handles data aggregation for composite queries because it maintains both individual data and aggregate data. In their design, each network node is loaded with a simple (one-way) hash function which is secure. This on-demand hash function maps a bit string for parametric optimization. The source and sink nodes uses this function to validate route distribution. Again, it can be noted that the GP2S scheme has issues which borders on accuracy and congestive distortion. From the literature review, there is a need to contextualize and implement a highly secure, and available data aggregation scheme (with Homomorphic Encryption algorithms for network construction, encryption data construction, and data aggregation phases), which supports data privacy, reduces encryption/decryption overhead (via message size), secures the AMI server machine clusters, reduces latency considerably, and enhances other quality of service metrics. The proposed scheme focuses on computational efficiency which factors-in metrics of resource utilization, energy consumption, delay (latency), and precision of aggregated result as well as the

system throughput.

3 Methodology

This involves the phases of system formulation, design, implementation and validation. An application for SG AMI network was developed based on network construction, encryption data construction and data aggregation. Applications in the network by our assumption operate within the latency requirement of 15 minutes periodic metering. Phase composite technique (PCT) was adopted. A smart grid AMI network was designed by introducing phase of data encryption. After the testbed setup, the respective smart meter computes time-based token data to be transmitted to the local concentrators. During this process, an elliptic-curve key exchange encryption algorithm is used to encapsulate its data for integrity test. Fig. D.3 demonstrates the sequence of events for the algorithm:

1. Private keys (e.g: pSender and pReceiver) are set for the source AMI node and its base node (receiving node).
2. Respective AMI nodes compute the result T by using an abstracted constant key (having a public elliptic curve) to multiply an estimated point (K).
3. The AMI node then sends this result (R) to the base node (either a local or global concentrator). It then computes the time-based token data by multiplying the value of R with its private constant key.

The time-based token data is the aggregation of current data (corresponding to $i - coordinate$) and voltage data (corresponding to v-coordinate) since the elliptic curve is two-dimensional model. With elliptic key exchange algorithm, its feasibility to transfer AMI data with a fair cryptographic overhead is guaranteed. The data content is then hidden from a malicious cyber-criminal during transmission. Even when there is congestion, the network delay is optimal.

A homomorphic token is utilized for abstracting the legitimate data from a malicious attacker. The idea is that the legitimate data can be changed by extracting with a token value while being sent to the base AMI core. Consequently, the tokenized sensed data is hidden while exchanging through the local concentrator. During the network construction, TCP hello message is used to ascertain the liveliness of the AMIs. Algorithm 1 shows a simple network construction algorithm using TCP message Flooding. It depicts a secured processing of network construction phase. With reference to algorithm 1, DMF is enabled once there is a smart meter (AMI source node) in a subnet group. If a sink node is located within the communication range of a smart meter node, the smart meter node

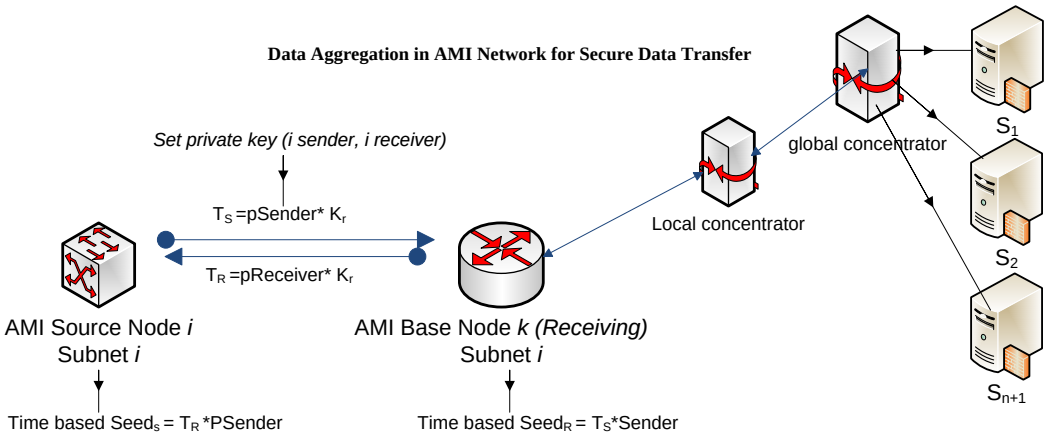

Fig. D.3: Secure Data Aggregation RTCA model

receives the TCP hello message (broadcast from the AMI base hub) and sets the sink node as the cluster parent node. Following the node level setting shown in Fig. D.2, a smart meter node can verify whether the message is from the sink node or not. In this architecture, nodes at higher levels can communicate across different clusters while nodes at the lower levels can only communicate with nodes within the same level and with nodes acting as cluster head (cluster parent node). In this way, each node is able to determine its sibling nodes, parent node, and child nodes. Secured network data can then be constructed by configured cluster parent nodes from all source nodes present in all the AMI subnet clusters.

To service an end user metering data transmission, known parent AMI base node aggregates available data obtained from its SM CUI terminals. The AMI base terminal packages the tuned-aggregated result into two-dimension while being encrypted using Hilbert curve as found in [24]. The Hilbert curve is then used for localization and privacy preservation. The Hilbert curve ensures tighter security encryption by ensuring that the homomorphic curve level is met. Once verified, the data image is then sent to the local concentrator for overhead reduction. Algorithm 2 illustrates the encryption data construction for TCP message flooding. With reference to algorithm 2, each node needs to find neighboring nodes by exchanging key seed. After this step, each node generates a Hilbert curve direction and curve level based on the new value. The next stage in this algorithm involves each node encrypting its data by Hilbert curve. Lastly, each node packs the encrypted data in preparation for sending to a local concentrator.

Algorithm 1: Tokenized and Secured Network Construction

Input: Call AMI Tokenized Network_Construction (Message msg, MsgType msgType) {

Output: Send to Sink (Cluster Servers){

1. For N(i) = 1 to K_{n+1}
2. If (A node is Source node) Then (Set Homomorphic Encryption (DMF) ==1)
3. {
4. Flooding (initLevel, base_stationID); exit}
5. Call (AMI_base hub);
6. Wait until TCP HELLO message is received ;
7. Child.nodes = ClusterParent_ID (True);
8. If (AMI_device n+1 acknowledges and Receives token-Message msg from a Smart Meter) {
9. If (token Mode = MegType_HELLO) {
10. Configure ClusterParent_ID, from Message msg ;
11. Set Counter_Ready i ++ ;
12. If (Message msg HopCount >initial HopCount + 1) Reduce Message msg;
13. else return;
14. For i (ClusterParent_ID > 0
15. Start ( )
16. Message msg_Flooding (NodeID) ; }
17. If (Token = Encryted) Then Allow {
18. If (child.node < 1)
19. Search Cluster.AMI = All_Child.nodes_ID ;
20. else check Next Cluster.AMI;
21. locate Child.node (N) = Message msg (Token);
22. }
23. End if
24. End if
25. End if
26. End if
27. End if
28. End if

29.End // terminate Algorithm

Algorithm 2: Encryption Data Construction

Input: Call AMI Tokenized Network_Construction (Message msg, MsgType msgType) {

Output: Send to Sink_AMI local concentrator{

1. For N(i) = 1 to K_{n+1}
2. Call command EncryptData (Message msg, MegType msgType) { AggregateNode = gather_neighbor(random);
3. Send (AggregateNode, KeySeed_data) to AMI cluster_Node;
4. Wait until response message from AMI cluster_Node;
5. newValue = ComputeKey (Received_KeySeed_data from AMI.cluster_Node, Keytoken_data);
6. directionValue = makeDirection(newValue)
7. setCurveLevel = currentCurveLevel(newValue)
8. encryptedData = Homomorphic_Curve(direction, curveLevel, newValue)
9. packing(encData)
10. Wait ( )
11. Ready to Send to Local concentrator ( )
12. }
13. If construction = Ready Then 14. Send ( )
15. Else Return ( )
16. End Algorithm

4 Data Transmission Phase

During the encrypted data transmission stage, each smart meter transmits hash tagged (encrypted) data to the subnet AMI base core. This then processes encrypted data received from tokens. Once the encryption direction of the base SM is different from its assigned one, it cannot relay the energy data to the local concentrator. In this case, the encryption curve direction must match. As such, the local concentrator (sink gateway) aggregates all the received data and reduces the overhead before sending to the global concentrator in an orderly hierarchy. Essentially, the DMF in the concentrators carries out cryptographic message compression before relaying the message to the global concentrator. Fig. D.3 illustrates typical data flow in a secure data aggregation RTCA model. Packet loss in the network is reduced using Time Division Multiple Access (TDMA) scheme [36] was considered for data transmission.

Theory 1: Assuming smart meter nodes $J_1, J_2 \ldots J_C$ where the number of cluster site smart meter nodes is C; the start time of the data transmission, i.e., Start Time t_0, for J_{ith} smart meter node J_i is determined by:

$$AMI_{start}time(J_{ith}) = \sum_{i}^{k}(i-1) * \langle \frac{LifeTimeofSendSection + LifeofReceptionSection}{LifeTimeofuserQuery} \rangle \tag{D.1}$$

The above theory explains the principle that decides the transmission time for smart meter nodes. In other words, each smart meter sends its encrypted data at its own transmission time. This theory assumes that users can direct a query request to a storage-enabled node (for example, a base station). The query request can be an aggregation request which is dependent on an aggregation function specified in the query. In this context, a basic requirement is that the node in reception of this request synchronizes its clock according to the timing information contained in the user request.

Algorithm 3 shows secured data aggregation by TCP message flooding. The process starts with the smart meter nodes. In the smart grid mode, an intermediate node called the AMI base node receives the data from the CIUs and then re-encrypts the data with its own data. In this way, the encrypted data of source nodes reaches the global concentrator where encryption with sufficient security key lengths is performed before sending the aggregated data to the AMI master stations. The AMI nodes must be enabled before the encrypted message is sent to the sink node. The local and global concentrators must also satisfy the message encryption and decryption requirements in order to propagate upstream and downstream messages to the sink. The sink node then receives the event message before calling

the local concentrator subroutine (DMF). Data captured from the local concentrator from the edge nodes are equally encrypted before pushing to the Grid OpenFlow firewall node. At any time, the sink node is instantiated, anomaly detection is verified through Filter TCP encrypt scheme. Legitimate traffic is then decrypted upon being received by the user AMI master node while illegitimate traffics are discarded accordingly.

5 AMI Modeling for System Analysis

A message appending AMI network is considered whereby the event based data obtained is logically sent to the sink received data. It preserves the integrity of the transmitted information. The local concentrator uses the DMF to reduce the high computational overhead before sending the summarized data to the global concentrator. In this section, various TCP based AMI algorithms for smart grid application scenarios were modeled. Network delay, throughput, utilization, latency, access delay etc., were considered. The RTCA topology would be analyzed in relation to the various TCP workloads. Since processing delay is very essential; a reliable data rate could be used. In a context domain, the time it takes to execute data aggregation process was observed to rely on the composite network delay and its security processing overhead. The RTCA topology is modeled to account for the network latency/delay computations in a smart grid domain considering a reliable transmission approach. In this regard, TCP algorithms and IEEE 802.11 MAC/PHY were introduced for the characterization of the SG AMI local and global concentrators. The proposed deterministic data aggregation payload (i.e., Data Minimizing Function- DMF) combines these to derive the network delay. The results of the DMF Homomorphic encryption in the Smart grid AMI will be discussed.

The TCP behavior for round-trip time (RTT) dynamics and IEEE 802.11 MAC/PHY link capacity are key elements of the Stochastic TCP congestion management variants with wired equivalent privacy (WEP). It is a common knowledge that the performance of a TCP flow is affected by its round-trip time (RTT). Therefore, the RTT of a TCP flow can be defined as the time interval between the point a packet is sent by an AMI source node and the point that an ACK is received by the same node. This interval includes the propagation delays, delays from links connecting AMI nodes to the channel, queuing delays and other delays from connected hosts.

Therefore, to derive the round trip time for SG AMI with respect to the RTCA topology, it is assumed that m number of TCP flows from AMI source nodes J_i, $(i = 1, 2, 3, ..., m)$ passes through a local concentrator via the AMI base hub.

Let $R_i(t)$ be the RTT at the time ACK is received at the source node,

Algorithm 3: Secured Data Aggregation

Input: Call AMI Network_Construction (Message msg, MegType msgType){

Output: Send to Sink_AMI (AMI meters)

1. For N(i) = 1 to K_{n+1}
2. Call function Data Aggregation (Message msg, msgType msgType)
3. If (a metering node is Set) then Send Message(encData) to Sink AMI_Node
4. Else{
5. If (a AMI_node receive Event.message(encData) from Rx.sensor node { Then
6. Call local concentrator (DMF)
7. If (the local concentrator node receive Event.message(encData)) { Then
8. Stores encData from msg (Homomorphic);
9. decryptedData = decryption(encData)
10. aggregatedData += decryptedData;
11. Call Global Concentrator ()
12. newEncData = Homomorphic Curve(direction, curveLevel, aggregatedData);
13. If (all data is received from local concentrator.childNode)
14. SendMessage(encData) to Grid OpenFlow Firewall.ParentNode; }
15. If (a node is SinkNode){ Then
16. Filter TCP encrypt.Anomaly
17. Store encData from msg;
18. decryptedData = decryption(encData);
19. Send Message(decryptedData) to User AMI Master_Stations; }}}
20. End if
21. End if
22. End if
23. End if
24. End if
25. End // terminate Algorithm

$\delta_{l,i}(t)$ be the link (path) delay for packets from flow i

$\delta_{p,i}(t)$ be the corresponding propagation delay

$\delta_{q,i}(t)$ be the queuing delay

Then

$$R_i(t) = \sum_{i=1}^{m} (\delta_{l,i}(t) + \delta_{p,i}(t) + \delta_{q,i}(t)); t \geq 0, 1, ..., m. \tag{D.2}$$

From eqn. D2, it can be assumed that in addition to the queuing delay dynamics, the propagation and link delays contribute to the RTT with a factor that is fairly constant during each network cycle.

But queuing delay is related to queue length and link capacity by the expression [37]:

$$\delta_{q,i}(t) = \frac{Q_i(t)}{C_i}$$

where $Q_i(t)$ is the queue length at AMI source node i, and C_i is the link capacity of node i.

Therefore, the round trip time will be given by:

$$R_i(t) = \sum_{i=1}^{m} \frac{Q_i(t)}{C_i} + \delta_{l,i}(t) + \delta_{p,i}(t) \tag{D.3}$$

In order to derive the flow throughput for the TCP flows from the AMI source nodes J_i with respect to the RTCA proposed in this chapter, it is important to note that all TCP flows from the AMI source nodes compete for access to the wireless channel. This is facilitated by the distributed coordination function (DCF) wireless access mechanism which is based on CSMA/CA. In this case, all AMI nodes have equal contention parameters and such, have equal opportunity to access the channel. In this access mechanism, an AMI source node that is ready to transmit its packet first monitors the channel activity. If the channel is perceived to be idle for a period of time equal to a distributed interframe space (DIFS), the node transmits its packet. If on the other hand the channel is sensed busy either immediately or even during the DIFS, the node persists in monitoring the channel until it is measured idle for DIFS. At this point, the node generates a random backoff interval before transmitting. With this collision avoidance mechanism, the probability of collision from AMI source nodes is reduced. It is assumed in this analysis that each AMI source node always has a packet in its transmission queue in readiness for transmission in line with the DCF access mechanism. It is to be noted that congestion control in the mechanism is also facilitated by considering additive increase multiplicative decrease (AIMD) feature. With the AIMD in place, a TCP sender additively increases its rate when it perceives

that the end-to-end path (link) is free from congestion and multiplicatively decreases its rate when it perceives that the channel is congested.

Let τ be the normalized throughput which is the fraction of time the wireless channel is used to transmit successfully payload bits. This can be derived by considering how the transmission can happen in randomly chosen slot time. Thus, we assume:

P_{tr} to be the probability that there is at least one transmission within the considered time slot.

P_s to be the probability that the transmission is successful (ie, packet is delivered successfully).

P to be the data payload.

Therefore, τ is given by:

$$\tau = \frac{E[payload\ \ information\ \ transmitted\ \ in\ \ a\ \ slot\ \ time]}{E[length\ \ of\ \ slot\ \ time]} \tag{D.4}$$

If $E[P]$ is considered to be the average packet payload size, then the average amount of payload information that is transmitted successfully in a slot time is given by $P_{tr}P_sE[P]$.

The above expression is valid on the condition that a successful transmission occurs in a slot time with probability of $P_{tr}P_s$. On the other hand, the average length of slot time can be obtained by considering that with the probability of $1 - P_{tr}$, the slot time would be idle or empty. Now, considering that a successful transmission will happen with a probability of $P_{tr}P_s$, it can be deduced that a collision can happen with a probability of $P_{tr}(1 - P_s)$ (this is in consideration of three possible events which include: (i) slot time is empty or idle (ii) there is a successful transmission and (iii) there is a collision) . It is to be noted at this point that the event of having an idle or empty slot time will depend on a variable ξ which is the duration of an empty or idle slot time. Similarly, the event of having a successful transmission will be affected by a variable T_s which is the average time that the channel is sensed busy. Finally, the event of a collision happening will be affected by a variable T_c which is the average time the channel is sensed busy by each transmitting AMI source node.

Therefore, the average length of a slot time will be equal to:

$$(1 - P_{tr})\xi + P_{tr}P_sT_s + P_{tr}(1 - P_s)T_c$$

Throughput, τ will then be given by:

$$\tau = \frac{P_sP_{tr}E[P]}{(1 - P_{tr})\xi + P_{tr}P_sT_s + P_{tr}(1 - P_s)T_c} \tag{D.5}$$

It is to be noted that the above expression for throughput can be used when there is no need to make specific reference to a particular access mechanism. However, the throughput for the considered RTCA network is based on the DCF access mechansism. In this situation, the variables T_s and T_c must be specified such that throughput will now be given by:

$$\tau = \frac{P_s P_{tr} E[P]}{(1 - P_{tr}) + P_{tr} P_s T_s + P_{tr}(1 - P_s) T_c} \tag{D.6}$$

In the TCP model, link capacity C is vital to determining RTT, queue length and congestion window size. Since, C is likened to link bandwidth, B_w, it is necessary to determine the total link capacities connecting all the AMI source nodes to the wireless channel. This is given by:

$$T_{LC} = B_w * \tau \tag{D.7}$$

This leads to equation D.8:

$$T_{LC} = \frac{(P_s P_{tr} E[p]) B_w}{(1 - P_{tr}) + P_{tr} P_s T_s + P_{tr}(1 - P_s) T_c} \tag{D.8}$$

The above expression refers to the total link capacity used to carry the traffic overhead in RTCA topology. Given that data processing D_p involves Service rate S_r and arrival rates A_r, for the RTCA topology, D_p is then given by:

$$D_p = S_r + A_r \tag{D.9}$$

The RTCA topology has service rate, latency, delay, queuing workload and throughput behavioral analysis under traffic congestion and security encryption. This work assumed the number of smart meters to be 20 for each subnet cluster with 1024 bytes message sizes with a propagation delay of 0.001s.

6 Network Simulation and Modeling

6.1 Preliminaries

This section describes important features of Riverbed Modeler simulation engine used for the study carried out in this chapter. Other important information that helped in the design of the RTCA testbed is also included in this section. Riverbed modeler has interface feature for easy integration of external objects files and libraries with numerous collection of different types of protocols. It also features a development environment which makes it easier for researchers to model different types of network devices and protocols of their choice. Developing network models for a simulation study could be

easy and straight-forward by using standard network objects incorporated in the modeler. However, developing new or modifying existing network models, devices or protocols can be an unpleasant experience. This normally involves introducing new protocols/algorithms into the models/devices or modifying the existing ones. Riverbed modeler has capability for performing discrete event simulation (DES).

With DES, simulation progresses as scheduled events are executed, and simulation time updated. To ensure efficiency of the simulation, the simulation kernel processes the scheduled events iteratively in an ordered manner. A pseudo-code representation of DES is shown in Fig. D.4 [38]. This comprises three main phases which include: initialization, simulation kernel operations, and the concluding operations. During the initialization phase, all state variables (such as simulation time, statistics, memories, event lists, etc.) are populated with initial state values. In the simulation kernel phase, simulation runs in the main loop until a simulation termination condition (such as simulation finish time) is satisfied. After exiting the simulation kernel phase, the simulation processes conclude by taking important actions such as: writing statistical records to files, freeing up memories, etc.

Discrete event simulation in Riverbed allows the introduction of important node objects modelled by the user. Each of the nodes can contain several modules which are separated by logical functionalities. Each module represents a functionality in the given node. For example, a module can be used to send or receive packets, process data, store data, etc. A module on the other hand can contain processors, queues, transmitters, and receivers. These can be utilized to implement the logic or functionality (process domain) represented by a module. A logic can comprise several state variables where transition from one state to another can be triggered by certain conditions. The processors are fully programmable through their process model while queues can be used to control data packets. The process domain has support for writing C++ codes within a state that can perform some desired operations. In other words, codes that implement important algorithms and protocols are incorporated at this level.

In this simulation, important objects such as Cluster Parent (AMI Tx), AMI base Hub (CIU), Data Aggregator (AMI local Concentrator), and Global Concentrator were modeled and introduced into the RTCA architecture. In developing the process node models, Conjugate batch Gradient was used instead of stochastic gradient descent to build the Cluster Parent (AMI Tx), since the later does not require knowledge of a complete training of CIU and AMI node sets. This procedure provides a step-wise random projection onto available node sets (hyper-node planes) until convergence. The method starts with an arbitrary initial vector CIU node ϑ_0 and at every iteration i, the algorithm shown in Fig.D.5 randomly selects a CIU row $i(k) \in i \in 1, ..., n$ of the layered cluster as the parent nodes.

```
1  void main( )
2  {
3   //initialization
4  initialize_variables ( );
5  allocate_memories ( );
6  ...
7  //simulation kernel operations
8  while(simulation_time < finish_time)
9  {
10 current_event = pop_next_event_from_list ( );
11 process_event(current_event);
12 update_simulation_time ( );
13 ...
14 }
15 //finishing operations
16 write_records_statistics_to_file ( );
17 free_memories ( );
18 ...
19 }
```

Fig. D.4: Pseudo-code Representation of Discrete Event Simulation

It then uses the sampled data streams to calculate the gradient based on the local loss function, i.e $\delta f_i(x_i, y_i)$. The parameter ϑ^k is then updated by moving a small step size along the flow gradient as shown in Algorithm IV (Fig.D.6). The CIU and AMI are updated after every sampling period provided the buffer limits are not exceeded.

Further, the AMI local aggregators in Algorithm VI (Fig. D.7) essentially models the node process and carries out multiple rounds of Conjugate batch Gradient in parallel on node $i \in 1, \ldots, n$ using the initial parameter ϑ^k obtained from the T_{th} iteration. The algorithm for the device reduces all the gathered operation that computes ϑ^{k+1} in Algorithm V. This is done via the sample averaging of all the gathered data streams from all the AMI nodes. The AMI local concentrator reduces operation using synchronization to achieve data aggregation. By applying AMI local concentrator scheme on the edge network, it improves performance with increase in the node size. The data stream queue system occurs in such a manner that stream arrivals reach infinite concentrator queued (L_{ag}) at time (t). In the algorithm, a read function obtains incoming streams and creates the linked list representing the corresponding input arrivals. All the localized but aggregated data are synchronized further and recycled to the Grid global concentrator for Open Firewall processing.

Algorithm IV: CIU input terminals (destination address, source address, queue size, link Information) procedure for user connection.

---

1 **CIU Data:** Source address, destination address, queue size, link Information
2 **CIU Output**: DataStream hits local aggregator La_g
3 **Begin ( )**
4 CIU_ Polyadd_CIU & AMI Dest (Input, Output)
5 Initialize CIU = 0; CIU > 0; i ++)
6 Call Cluster Parent (AMI Tx);
7 **End**

Fig. D.5: Algorithm for the process Node Modelling of CIU input terminals

---

Algorithm V: Cluster Parent (AMI Tx)/ AMI_base Hub (CIU) (Conjugate batch Gradient)

---

1 #Define CIU as Cluster Parent (AMI Tx)
2 Initialize: i, iterations T, $\vartheta^k \leftarrow 0$
3 **for** (i = 0; i > n; i++) // convergence or maximum iteration i
4 **Do** AMI $\leftarrow 0$ until ($AMIn_o$, $AMIn_1$, $AMIn_2$, $AMIn_3$, $AMIn_4$, $AMIn_5$, $AMIn_{n+1}$) $\xrightarrow{\Delta}$
5 convergence 5 or maximum iteration i
6 Int AMI_base Hub (CIU);
7 **for** i = $AMIn_0$ to $AMIn_{n+1}$
8 Return Link;
9 **Draw** i ∈ {1, ... n}
10 Map i ← AMI_base Hub (CIU)
11 **Upper** Limit Buffer ← Set ETXthersh;
12 **Update** $\vartheta^{k+1} = \vartheta^{0k} = \partial fi\,(xi, yi)$.
13 **end**
14 **end**
15 **Return** ϑ^k(CIU & AMI)

Fig. D.6: Algorithm for the process Node Modelling of Cluster Parent

---

Algorithm VI: AMI local concentrator/ local aggregator

---

1 **Define** (AMI local concentrator)
2 **Input:** local ID, destination ID, queue size, link Information
3 **Output**: Gather up streams and dispatch the infinite queues to Global sink
4 **Draw** AMI_local Connection ( ) ; AMI_Global Connection j ( )
5 Initialize: i, iterations T, $\vartheta^{k} \leftarrow 0$; ϑ^{k}(CIU & AMI)
6 Map AMI data (individual nodes)
7 While all data (i) not converged (Buffer) do
8 **For** all i (AMI) $\in \{0, \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \mathrm{n}-1\}$ **do** read (i);
9 **For** i:=0 to N-1 **do** read (φ[i]);
10 **For** i:=0 to N-1 **do** read (∂[i]);
11 **For** i:=0 to N-1 **do** read (n_{k+1}[i]);
12 **For** j:=0 to N-1 **do** read ϑr[i]: = φ([i]);+ ∂([i]);+……(n_{k+1}[i]);
13 **For** i:=0 to N-1 **do** Write (j[i]);
14 $\vartheta^{k+1}(\mathrm{CIU}\ \&\ \mathrm{AMI}) = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} \vartheta^{k}$
15 **End;**
16 **End**
17 **End**
18 **End**
19 **End**
20 **End**
21**End**

Fig. D.7: Algorithm for the process Node Modelling of Local/Global Concentrator

6.2 Simulation Design Details

In the design in Fig.D.8, the smart meters/AMI local/global concentrators are driven by TCP/IP protocol and IEEE 802.11 MAC/PHY layers. This assumption is widely used in AMI researches [24]. Using the TCP/IP model assumptions, smart meter TCP model was generated to generate round-trip-time. Beside the analytical models presented above, Riverbed Modeler is used to analyze the performance of the RTCA procedural-algorithms for data aggregation including user privacy protection. The simulation considered the Homomorphic encryption algorithms (i.e., Network construction, Encryption data construction, Data aggregation) formulated for local concentrators while using a global concentrator to check for anomalies into AMI server clusters. TCP/IP message flooding was contextualized for congestion scenario taking cognizance of IEEE 802.11 MAC/PHY. Stochastic TCP congestion management variants with wired equivalent privacy (WEP) and our formulated deterministic data aggregation payload (i.e., Data Minimizing Function-DMF) were compared in the Smart grid AMI. Before discussing the simulation results, the design setup is hereby presented. It was assumed that 40 AMI are evenly distributed in subnet clusters having 500m radius. The smart meter CIU and the AMI RF base core interact on instantaneous basis. Two subnet clusters with AMI CIDR IP_subnet $= 172.16.1.\frac{1}{24}$ running on IEEE 802.11n MAC/PHY metering nodes was introduced. A local concentrator was placed at the edge of RTCA design. The MAC layer configuration is based on IEEE 802.11n MAC/PHY with simulated data rate of 54 Mbps was used at the network layer for communication via the IP cloud to the global concentrator. This implements eDDOS_grid OpenFlow firewall which serves as a load balancer while protecting the type-1 virtual machine server clusters. TCP data minimizing function (DMF) for homomorphic encryption in the RTCA topology was implemented on the concentrators. After enabling the encryption schemes, as for the transport layer, TCP_Tahoe_WEP, TCP_Reno_WEP, TCP_NewReno, TCP_SACK_WEP, and TCP_Vegas_WEP were all used while comparing with TCP_DMF_WPA_3. The default retransmission timeout (RTO) was set to a 30secs for the RTCA topology. In each AMI in the subnet clusters, the homomorphic encryption is enabled. Each meter was configured to generate $2^8 - 1$ data every 1hour. For the proposed TCP_DMF_WPA-3, the size of the default payload of the energy data is set to 1024bytes due to the message encryption scheme. This is active and increases as the aggregation takes place. The AMI transmission power in the presence of the concentrators is 0.0002W. The simulation run time in this case lasted for a maximum of 3600secs.To evaluate the security processing delay; the proposed homomorphic algorithm is subjected to traffic workloads. This processing delay is determined by the number of active smart AMI nodes as well as the topological layout.

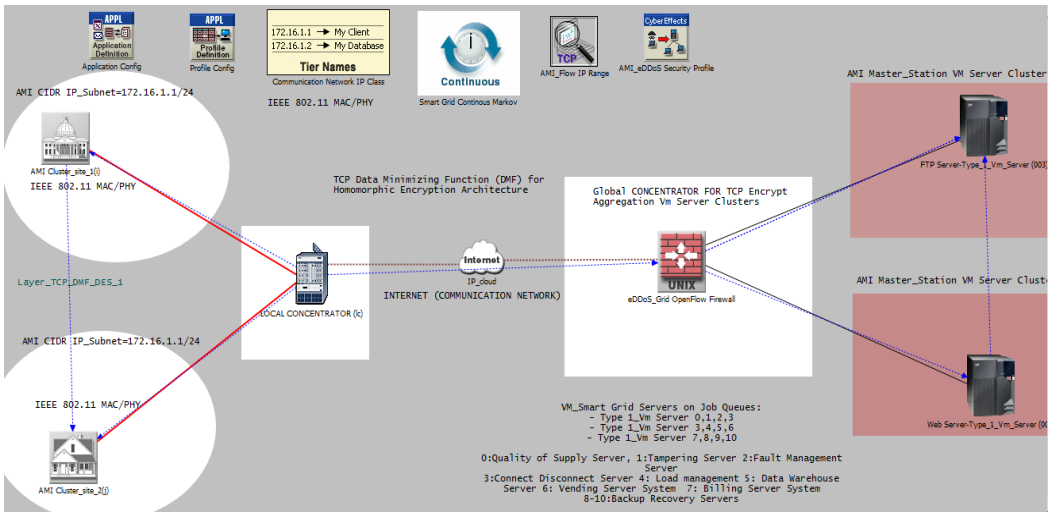

Fig. D.8: Experimental testbed (RTCA topology)

7 Scenario Evaluation

Various levels of configurations were carried out while evaluating the performance of the RTCA topology for demand response, dr. The demand response in smart grid model was used to improve SG efficiency by allowing customers to adjust their energy consumption pattern while maintaining privacy. The AMI instantaneously updates the local concentrators with the energy consumed data. This concurrently updated on the AMI_server clusters. At time t, the SG utility load center of the subnet sites broadcasts demand response dr details such as energy cost, dr strategy to the local concentrators, l_c. The TCP l_c now broadcasts the received data to the global concentrators gc. The AMI_Master station server clustering through the gc now receives the energy data. This determines the actual consumption profiles for each subnet cluster using the received (collected) data and profiles of the results for each energy user on-demand. Fig. D.8 illustrates the RTCA topology where data aggregation payloads (such as Data Minimizing Function- DMF) were analyzed. It focuses on DMF for homomorphic encryption in the Smart grid AMI as previously explained. The essence is to observe how the DMF can control AMI network data traffic and the computational overhead while declaring customers' privacy at all times from the AMI. The AMI master stations have their servers logically isolated using Type-1 virtualization for optimal performance.

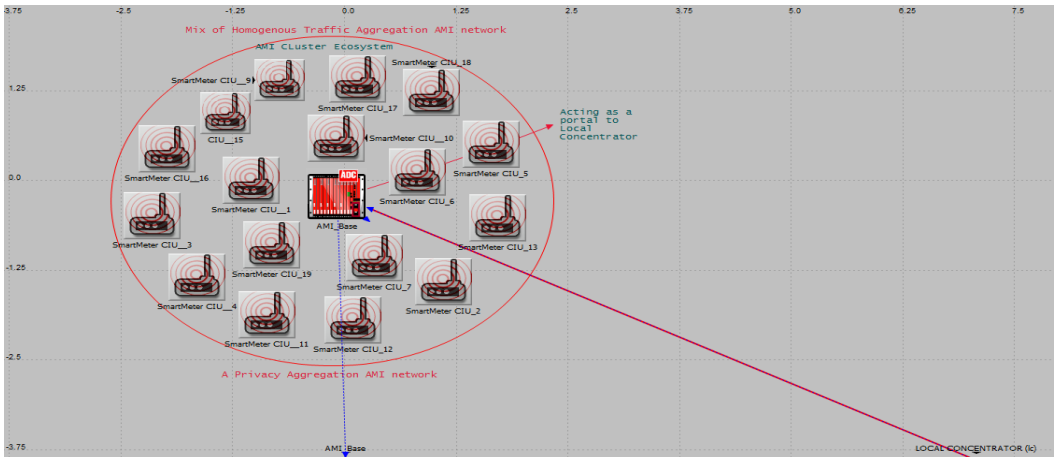

Fig. D.9: Smart meter/AMI nodes in subnet cluster (RTCA topology)

8 Results Analysis

This section presents an analysis of the results of the proposed RTCA in comparison with the different TCP congestion control algorithms implemented in the IEEE 802.11 network standard. The QoS metrics considered in this analysis include: service rate, queuing length, latency, throughput and media access delay.As mentioned already in section 6, the TCP congestion control algorithms implemented in this study were configured with Wired Equivalent Privacy (WEP) which is a security algorithm introduced in the IEEE 80.11 for providing data confidentiality. Fig. D.9 shows the aggregation of smart meters based on IEEE 802.11n MAC/PHY. Security processing delay is relatively low despite the traffic overhead. On the other hand, Fig.D.10 shows the validation analysis of RTCA service rate for data aggregation and user privacy protection. With data aggregation workload, the RTCA was used to relieve the transmission congestion with cryptographic overheads incurred during message aggregation. Homomorphic Encryption algorithms formulated at the phases of network construction, data encryption and data aggregation at the concentrators and the checks for anomalies into AMI server clusters constitute a huge traffic workload. With TCP message flooding on the network, the proposed deterministic data aggregation payload (i.e., Data Minimizing Function- DMF) was compared with other TCP encryption variants. The DMF homomorphic encryption in the Smart grid AMI received the most satisfactory service rate (36.04 %) compared with others as shown in Table D.1. The implication is that traffic workloads from various users will receive fair share of computations regardless of the overhead. This is very significant for environments that have many users in the AMI.

Similarly, Fig. D.11 shows the observing processor queue length in the RTCA topology. The RTCA could have a sustained processor queue with rising processor usage. It ensures that the processor

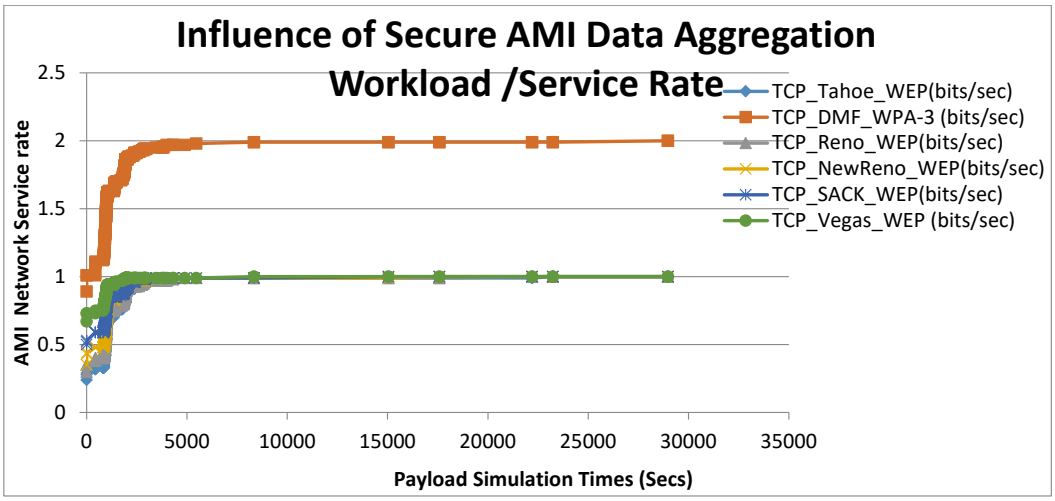

Fig. D.10: Validation of RTCA service rate behavior

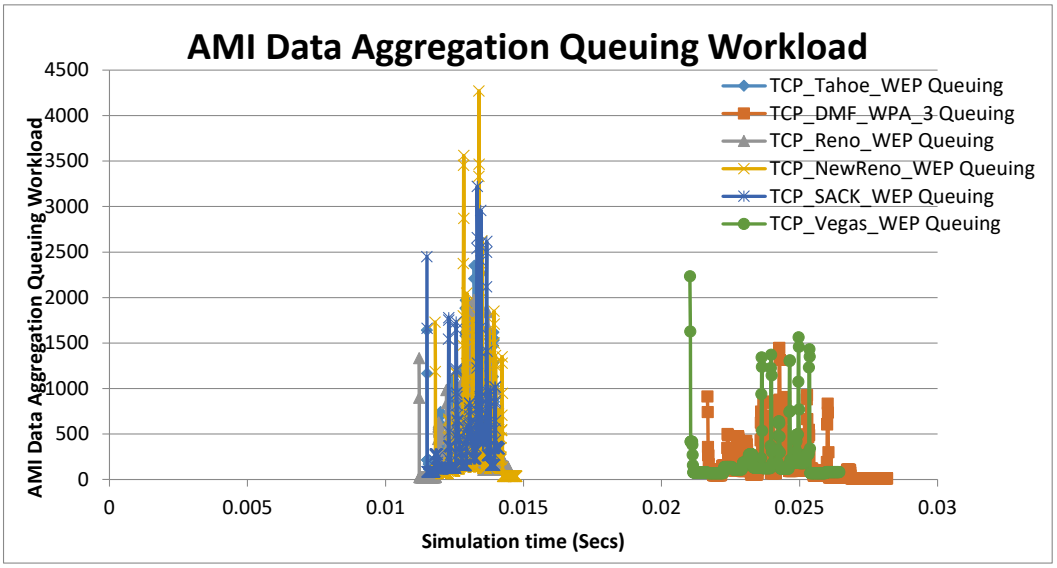

Fig. D.11: Validation of RTCA Queuing workload behavior

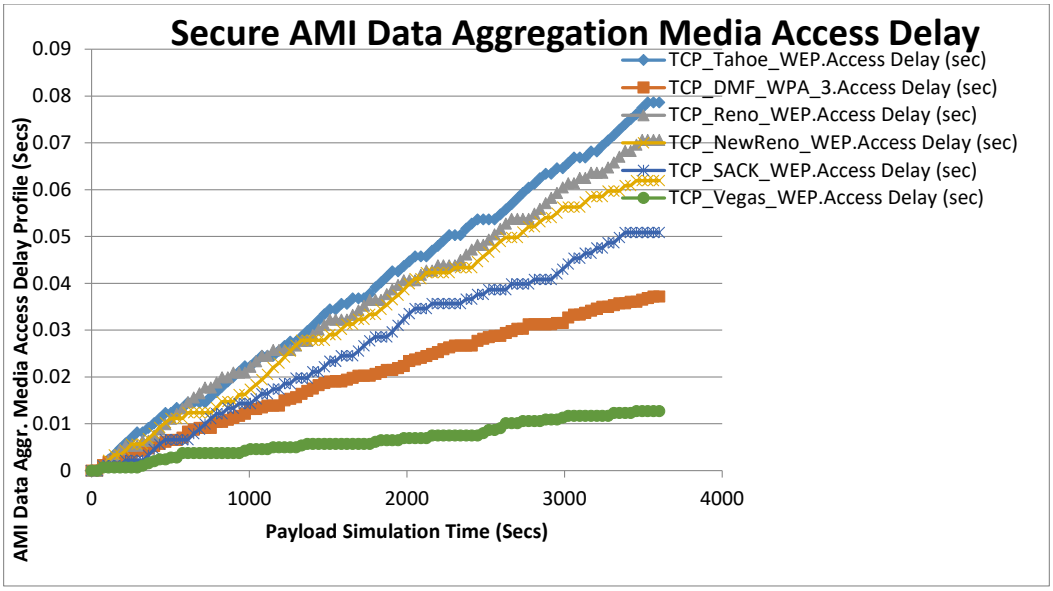

Fig. D.12: Validation of RTCA Media Access Delay behavior

queue is served at 100% service rate. It can even have a sustained processor queue with maximum processor usage or maintain a saturated queue state with data aggregation workload. Fig. D.11 illustrates how a processor bottleneck interferes with AMI data aggregation workload. It could be deduced that when the RTCA summarized CPU cycle is already at 100% utilization; instantiating another virtual machine process achieves little or no work. A processor queue is a summarization and gathering of ready threads (which may be single or multiple) which cannot run on the processor due to a running, simultaneous active thread. In context, the visible indication of CPU saturation in RTCA is controlled by concurrent multiple threads executing the service rates for tasks/jobs. Clearly, the workload queues literally manifests when the CPU is very busy with its service rates, this can occur when utilization is well below 90%. It usually happens when arrival requests for processor time t appears randomly as well as when the job-thread workload pushes oscillatory time slots from the CPUs. From Fig. D.11, the proposed DMF offered the least queuing workload (9.97%) relative to others. This is as a result of the 36.04% service rate profile making for more processing of cryptographic overhead when in existence.

Finally, Fig. D.12 shows the measured media access delay in RTCA running various TCP workloads. It depicts the time from when the AMI-data transverses the MAC-physical layer of the AMI nodes until successful transmission from the wireless medium to the local concentrator is established. The essence of analyzing average access delay is explained by the fact that the Smart grid AMI offers a real-time metering behavior having acceptable delay for AMI data propagation. Therefore, it is expedient to have smaller delays for AMI real-time flows. It was observed that the proposed DMF

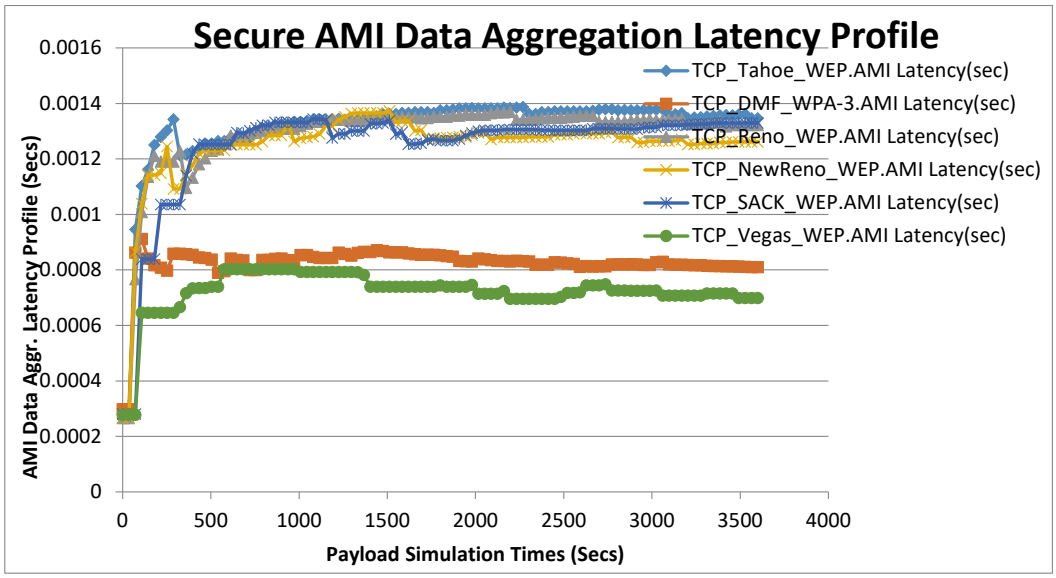

Fig. D.13: Validation of RTCA latency behavior

had 12.14% media access delay which is comparatively reasonable considering other TCP schemes. This has implication in the RTCA. It could be deduced that the medium of communication affects the latency profile. With a reliable two-way AMI, MAC access delay sets the tolerable and optimal rate of energy information transmitted. Hence, the lower the MAC access delay the more optimal the data aggregation workload propagation. Table D.1 shows the summarized variations. From Fig. D.13, latency response from AMI server clusters is quantified by the RTT feedback. It is measured in a directional one-way or round trip delay time. Round-trip latency is determined from a single hub. It does include the time a destination system (sink) spends processing the packet. However, in the RTCA network, energy data packet is sent via gateways in form of concentrators. Absolute throughput is established once all the data is completely received. RTCA latency is the sum of the link latencies, including link delays and the forwarding latency of each gateway/concentrators. Queuing and processing delays adds to the minimum latency. Queuing delay happens if a concentrator receives multiple packets from different subnet clusters. It was observed that the proposed DMF in RTCA has 11.90% latency profile which is as a result of lower queuing length and compact network layout. Again, Fig. D.14 shows the validation of RTCA throughput behavior where the proposed DMF had 25.69% which is relatively better than other TCP variants. This implies that the RTCA has the ability to deliver successfully the real time data captured from the end user AMI. Table D.1 shows the summarized variations when contextualized in Fig. D.3.

Table D.1: Result Summary of the Data Aggregation workload in Smart grid RTCA

System metrics	TCP-Tahoe -WEP	TCP-DMF -WPA-3	TCP-Reno -WEP	TCP-New Reno	TCP-SACK -WEP	TCP-Vegas -WEP
Avg. service rate	8.11%	36.04%	13.51%	9.01%	15.32%	18.01%
Avg. Queuing length	15.95%	9.97%	13.29%	23.92%	21.59%	15.28%
Avg. Media Access Delay	25.24%	12.14%	22.36%	19.81%	16.29%	4.16%
Avg. latency profile	20.83%	11.90%	19.20%	18.60%	19.35%	10.12%
Average throughput	20.78%	25.69%	19.62%	15.15%	12.41%	6.35%

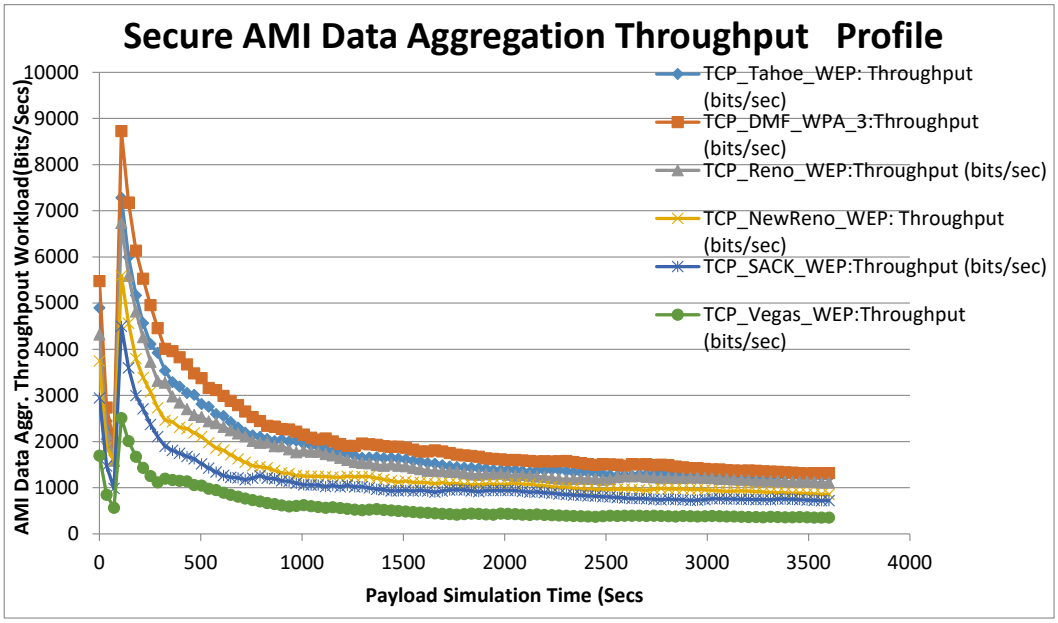

Fig. D.14: Validation of RTCA Throughput behavior

9 Conclusion

In this paper, a novel communication architecture for data aggregation in a smart grid advanced metering infrastructure (SG AMI) network is proposed. The architecture relies on homomorphic

based encryption algorithms which seeks to improve network traffic and service rates while supporting the security and privacy of end user's consumption data during data aggregation. From the result analysis and evaluation, the proposed Ring Triangulation Communication Architecture (RTCA) can be used to relieve transmission congestion and reduce the cryptographic overheads incurred during data aggregation. Additionally, the paper has shown that the scheme can be very reliable and efficient in delivering state-of-the-art QoS in terms of resource utilization, propagation latency/delay, and system throughput, during data aggregation as depicted in Table D.1. The QoS metrics evaluated show that the scheme can reduce latency considerably and also secure the AMI server clusters.

References

[1] 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.

[2] H. Farhangi, "The path of the smart grid," *IEEE power and energy magazine*, vol. 8, no. 1, 2010.

[3] S. Misra, P. V. Krishna, V. Saritha, H. Agarwal, and A. Ahuja, "Learning automata-based multi-constrained fault-tolerance approach for effective energy management in smart grid communication network," *Journal of Network and Computer Applications*, vol. 44, pp. 212–219, 2014.

[4] T. Khalifa, A. Abdrabou, K. B. Shaban, M. Alsabaan, and K. Naik, "Transport layer performance analysis and optimization for smart metering infrastructure," *Journal of Network and Computer Applications*, vol. 46, pp. 83–93, 2014.

[5] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Communications Workshops (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

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

[7] X. He, X. Zhang, and C.-C. J. Kuo, "A distortion-based approach to privacy-preserving metering in smart grids," *IEEE Access*, vol. 1, pp. 67–78, 2013.

[8] H.-Y. Lin, W.-G. Tzeng, S.-T. Shen, and B.-S. P. Lin, "A practical smart metering system supporting privacy preserving billing and load monitoring," in *International Conference on Applied Cryptography and Network Security*. Springer, 2012, pp. 544–560.

[9] H.-Y. Lin, S.-T. Shen, and B.-S. P. Lin, "A privacy preserving smart metering system supporting multiple time granularities," in *Software Security and Reliability Companion (SERE-C), 2012 IEEE Sixth International Conference on*. IEEE, 2012, pp. 119–126.

[10] S. Wang, L. Cui, J. Que, D.-H. Choi, X. Jiang, S. Cheng, and L. Xie, "A randomized response model for privacy preserving smart metering," *IEEE transactions on smart grid*, vol. 3, no. 3, pp. 1317–1324, 2012.

[11] E. Shi, H. Chan, E. Rieffel, R. Chow, and D. Song, "Privacy-preserving aggregation of time-series data," in *Annual Network & Distributed System Security Symposium (NDSS)*. Internet Society., 2011.

[12] C. Rottondi, M. Savi, D. Polenghi, G. Verticale, and C. Krauß, "Implementation of a protocol for secure distributed aggregation of smart metering data," in *Smart Grid Technology, Economics and Policies (SG-TEP), 2012 International Conference on*. IEEE, 2012, pp. 1–4.

[13] C. Rottondi, G. Verticale, and A. Capone, "A security framework for smart metering with multiple data consumers," in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*. IEEE, 2012, pp. 103–108.

[14] R. Cristina, G. Verticale, and A. Capone, "Privacy-preserving smart metering with multiple data consumers," *Computer Networks*, vol. 57, no. 7, pp. 1699–1713, 2013.

[15] C. Rottondi, M. Savi, G. Verticale, and C. Krauß, "Mitigation of peer-to-peer overlay attacks in the automatic metering infrastructure of smart grids," *Security and Communication Networks*, vol. 8, no. 3, pp. 343–359, 2015.

[16] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *International Workshop on Security and Trust Management*. Springer, 2010, pp. 226–238.

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

[18] S. Salinas, M. Li, and P. Li, "Privacy-preserving energy theft detection in smart grids: A p2p computing approach," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 257–267, 2013.

[19] S. Li, K. Choi, and K. Chae, "Ocpm: Ortho code privacy mechanism in smart grid using ring communication architecture," *Ad Hoc Networks*, vol. 22, pp. 93–108, 2014.

[20] S. Tonyali, K. Akkaya, N. Saputro, and A. S. Uluagac, "A reliable data aggregation mechanism with homomorphic encryption in smart grid ami networks," in *Consumer Communications & Networking Conference (CCNC), 2016 13th IEEE Annual*. IEEE, 2016, pp. 550–555.

[21] C. Gopi and V. Lalu, "Sensor network infrastructure for ami in smart grid," *Procedia Technology*, vol. 24, pp. 854–863, 2016.

[22] Y. Liu, "Wireless sensor network applications in smart grid: recent trends and challenges," *International Journal of Distributed Sensor Networks*, vol. 8, no. 9, p. 492819, 2012.

[23] A. R. Butz, "Alternative algorithm for hilbert's space-filling curve," *IEEE Transactions on Computers*, vol. 100, no. 4, pp. 424–426, 1971.

[24] M. Yoon, Y.-K. Kim, and J.-W. Chang, "A new data aggregation scheme to support energy efficiency and privacy preservation for wireless sensor networks," *International Journal of Security and Its Applications*, vol. 7, no. 1, pp. 129–142, 2013.

[25] V. Jacobson, "Congestion avoidance and control," in *ACM SIGCOMM computer communication review*, vol. 18, no. 4. ACM, 1988, pp. 314–329.

[26] J. Van, "Modified tcp congestion avoidance algorithm," Tech. Rep., 1990. [Online]. Available: ftp://ftp.ee.lbl.gov/email/vanj.90apr30.txt

[27] O. Ait-Hellal and E. Altman, "Analysis of tcp vegas and tcp reno," *Telecommunication systems*, vol. 15, no. 3-4, pp. 381–404, 2000.

[28] A. Bartoli, J. Hernandez-Serrano, M. Soriano, M. Dohler, A. Kountouris, and D. Barthel, "Secure lossless aggregation for smart grid m2m networks," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 333–338.

[29] S. Ruj, A. Nayak, and I. Stojmenovic, "A security architecture for data aggregation and access control in smart grids," *arXiv preprint arXiv:1111.2619*, 2011.

[30] P. Deng and L. Yang, "A secure and privacy-preserving communication scheme for advanced metering infrastructure," in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*. IEEE, 2012, pp. 1–5.

[31] M. Bae, K. Kim, and H. Kim, "Preserving privacy and efficiency in data communication and aggregation for ami network," *Journal of Network and Computer Applications*, vol. 59, pp. 333–344, 2016.

[32] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher, "Pda: Privacy-preserving data aggregation in wireless sensor networks," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*. IEEE, 2007, pp. 2045–2053.

[33] M. Conti, L. Zhang, S. Roy, R. Di Pietro, S. Jajodia, and L. V. Mancini, "Privacy-preserving robust data aggregation in wireless sensor networks," *Security and Communication Networks*, vol. 2, no. 2, pp. 195–213, 2009.

[34] H. Choi, S. Zhu, and T. F. La Porta, "Set: Detecting node clones in sensor networks," in *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*. IEEE, 2007, pp. 341–350.

[35] W. Zhang, C. Wang, and T. Feng, "Gp2s: Generic privacy-preservation solutions for approximate aggregation of sensor data," in *Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2008, pp. 179–184.

[36] S. Madden, M. J. Franklin, J. Hellerstein, and W. Hong, "Tinydb: An acquisitional query processing system for sensor networks," *ACM Transactions on Database Systems*, vol. 30, pp. 122–173, 03 2005.

[37] V. Misra, W.-B. Gong, and D. Towsley, "Fluid-based analysis of a network of aqm routers supporting tcp flows with an application to red," in *ACM SIGCOMM Computer Communication Review*, vol. 30, no. 4. ACM, 2000, pp. 151–160.

[38] G. S. Fishman, *Discrete-event simulation: modeling, programming, and analysis*. Springer Science & Business Media, 2013.

Chapter VI

Congestion Minimizing Scheme for Enhanced Data Aggregation in ZigBee-Based SG AMI Network

Paper E

Congestion Minimizing Scheme for Enhanced Data Aggregation in ZigBee-Based SG AMI Network

R.C Diovu and J.T Agee

Published

International Journal on Communications Antenna and Propagation (IRECAP)

Abstract

It is envisaged that the multi-directional communication of end user's energy consumption data in a smart grid advanced metering infrastructure (AMI) network needed for billing and other grid controlled purposes will raise concerns about the potential violations of the privacy of end users' personal data. Preserving the privacy of consumers can be achieved by designing robust and resilient AMI communication architectures with state-of-the-art cryptographic algorithms and data aggregation protocols. However, research has shown that these aggregation protocols can result in a congestive network scenario in the absence of a proper congestion management scheme. Unfortunately, there is no efficient mechanism for congestion management in the current ZigBee standard. In this paper, a well-designed Ring Triangulation Communication Architecture (RTCA) is utilized in our congestion minimizing scheme for enhancing data aggregation in a ZigBee based SG AMI network. The RTCA is implemented with a well-constructed ZigBee data and its aggregation algorithms. The average throughput and latency results from our scheme indicates a significant reduction in transmission congestion as well as traffic overheads incurred during message aggregation relative to the congestion control schemes (ZCCF and ZTCC) proposed in the literature.

1 Introduction

The smart grid advanced metering infrastructure (AMI) is expected to connect end users, grid/third party operators and an evolved market [1]. A typical architecture of the smart grid AMI network is provided in Fig. E.1. The AMI is principally responsible for communication of end user's consumption data for different purposes. Concerns have been raised by customers over possible violations of their sensitive personal information. One approach for preserving the privacy of the end user's consumption data is the use of secure data aggregation based protocols. Given that the aggregated metering data of groups of consumers are sufficient for authorized operators or third party recipients to perform their duties, allowing the recipients receive only the aggregated metering data is an important way of preserving the privacy of such data.

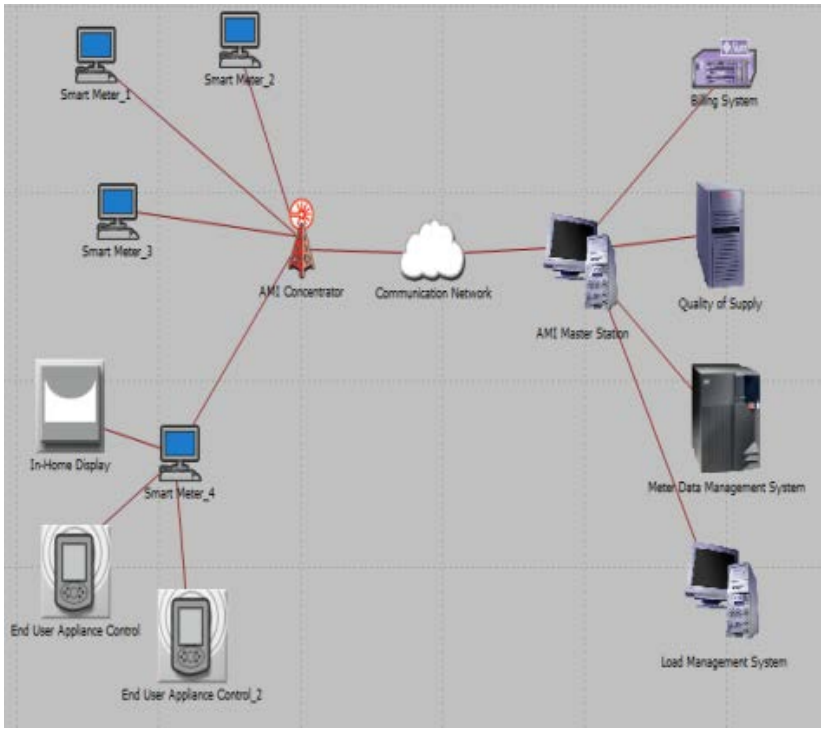

Fig. E.1: A Typical Architecture of a Smart Grid AMI Network

In this context, it becomes difficult for those recipients to decompose user's metering data from the aggregated data. The objectives for data aggregation can be realized better with resilient and robust communication technologies. The choice of ZigBee technology as a communication platform for the AMI has some obvious advantages. They include but not limited to the following [2]: support for low cost, low power applications, open standard which gives room for multiple vendors and systems to be interconnected, its ability to utilize the free industrial, scientific and medical (ISM) spectrum bands, to mention this few. However, data aggregation brings an increased burden on traffic intensity which results in a congestive network scenario. Unfortunately, the current ZigBee standard has no reliable mechanism for handling congestion management. Even though ZigBee provides re-transmission and acknowledgment (ack) to prevent data loss, but this provision is insufficient for dealing with congestion in the network. Instead, repeated retransmissions and ACK can aggravate the network situation and drive it to congestive scenario [3]. A scheme which seeks to minimize congestion incurred during data aggregation in a ZigBee enabled smart grid AMI network is proposed in this paper. In pursuance of this research objective, a well-designed ring triangulation communication architecture (RTCA) was configured in line with the IEEE 802.15.4 PHY/MAC protocol specifications as utilized in the ZigBee technology. With regards to the research goals of this paper, this architecture is designed to reduce the overhead incurred during data aggregation in a ZigBee enabled smart grid AMI network. For the rest of this paper, this RTCA shall be referred to as RTCA

Congestion Minimizing Scheme (RTCA_CMS). This RTCA_CMS scheme provides an easy-to-use wireless data connectivity with reliable and secure wireless network architectures. Some of its features are support for multiple subnet topologies such as point-to-point, point-to-multipoint and mesh topologies. Other features include smaller duty cycle which leads to prolonged battery life, low latency and a promise for scalability. The RTCA_CMS scheme was designed with support for mesh topology networking. In this case, the nodes are interconnected with other nodes such that multiple link states connect each node. The links between these nodes are optimized and dynamically updated through complex, built-in mesh lookup/routing table. Since, the mesh networks are decentralized in structure; individual nodes have the potential for self-discovery considering the lookup table. Once any node leaves the network, the design topology automatically reconfigures routing table and paths based on the latest state. When the nodes have more arrivals than the service rates of the mesh topology, stability is lost during these changing conditions. Hence, node failure becomes obvious. In a congested ZigBee network, either of these happens when a node attempts sending data to the sink: (1) The source node receives a busy signal or an indication that the signal cannot be effectively processed at that time. (2) A message is queued and delivered according to specified parameters. (3) A transmitted data is returned, rejected, or completely lost. When buffered data queues become excessively long or frequency of busy signals becoming excessive, the network is said to be in a condition of high-loss. One major aim of traffic management is to minimize or eliminate high-loss situations via a reasonable throughput threshold. For an efficient design, the number of rejected data or failed data should be as close to zero as possible while ensuring lower latency and higher throughput response. A review of related research effort geared towards solving the issue of congestion in ZigBee networks has been provided in section III of this paper. However, our proposed RTCA_CMS architecture is utilized for minimizing congestive network problems during data aggregation in the AMI network. It is to be noted that similar researches reviewed in section III above have not been extended to congestion issues related to data aggregation in a SG AMI context. The comparative analysis between our proposed scheme and two important congestion control schemes (ZigBee congestion control frame and ZigBee time congestion control) proposed in [4] was based on latency and throughput performance. The choice of throughput and latency is obvious. This is because of the fact that the two quality of service (QoS) parameters are usually affected by congestive network scenarios. The RTCA_CMS scheme out-performed the ZCCF and ZTCC schemes in terms of latency and throughput profiles. The rest of this paper is organized as follows: Section II contains an overview of ZigBee and the CSMA-CA algorithm employed in ZigBee network for congestion management while a review of related work is provided in section III. In section IV, the design methodology of our proposed RTCA congestion minimizing scheme is described while section V

contains the simulation results and analysis. Finally, section VI contains the conclusion for this study.

2 Overview of ZigBee and its CSMA-CA Algorithm

In an ideal ZigBee network, typical nodes needed for different design configurations include: the coordinator, the router and the end device nodes. Ideally, the ZigBee coordinator and the router are associated with the end devices in a relationship referred to as parent-child relationship. This parent-child relationship is established during the network joining procedure and this helps to define the topology of the network either as a tree or a cluster depending on the choice of the network designer. The ZigBee coordinator sets the number of routers (R_m) and the maximum number of end devices (D_m) such that each router may have children with fixed maximum depth (L_m). The size $A(d)$ of the range of the network addresses assigned to a router node at a depth $d < L_m$ is defined by this recurrence relation [2], [5]:

$$A(d) = \begin{cases} 1 + D_m + R_m & \text{if } d = L_m - 1. \\ 1 + D_m + R_m A(d+1) & \text{if } 0 \leq d \leq L_m - 1. \end{cases} \tag{E.1}$$

The above recurrence relation can be solved easily by each configured router and utilized by the router to assign addresses to its children. In [6], L. Bai et al reported that for a cluster topology, the allocated address A_n of n^{th} router-capable child node for Aparent node address and network depth d, should be given by:

$$A_n = A_{parent} + A(d) * (n-1) + 1 \tag{E.2}$$

Whereas the allocated address A_m of an m_{th} end device is given by eqn. E.3:

$$A_m = A_{parent} + A(d) * R_m + m \tag{E.3}$$

The mechanism employed in ZigBee for the management of congestion related issues is referred to as the Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA). When this algorithm is configured, a device wishing to transmit will first perform a Clear Channel Assessment (CCA) to ensure that the channel is not in use by any other device. This algorithm has two variants referred to as slotted and unslotted CSMA-CA respectively. The slotted CSMA-CA is utilized when there is a superframe structure in place. Ideally, a superframe divides the active transmission period into 16 equal time slots. On the contrary, the unslotted CSMA-CA is utilized when there is no superframe structure in place. In this case, there is no need for back-off alignment. Generally, a non-beacon

enabled ZigBee network uses the unslotted CSMA-CA algorithm for channel access. If the Clear Channel Access (CCA) operation returns an indication that the channel is busy, the transmitting device will back-off for a random period of time after which a re-trial can take place. The random back-off period for both slotted and unslotted CSMA-CA is an integer multiple of the unit back-off period (a unit back-off = aUnitBackoffPeriod) [2]. The CSMA-CA algorithm uses three important variables which include:

- Back-off exponent which determines the range of the back-off period
- Number of back-offs and
- Contention window length

The back-off is given by the expression:

$$Back{-}off = (A\ random\ integer\ between\ (0\ to\ 2^{BE}{-}1)){*}aUnitBack{-}offPeriod\ {+}IFS \tag{E.4}$$

Where IFS = Inter frame spacing. If the device enters a blocking state, it stops its operation and waits until the next contention access period (CAP). In the next CAP, the device generates another number and then repeats its previous steps. This is disadvantageous, as the energy-constrained device is depleted in the process. In addition, this leads to an increase in the queuing delay for the frame. Another limitation of this algorithm which requires improvement for network designers is that it leads to poor utilization of the channel. M. Baz et al [7] observed that this type of transmission wastes about 90% of the next time slot thereby making other devices to deplete their energy by always sensing a busy channel. Finally, an average delay experienced which is related to the back-off interval and the number of CCAs will likely increase exponentially in the event of the device failing to send a frame, thereby increasing the probability of frames being dropped.

3 Review of Related Work

Generally speaking, congestion is a major problem in wireless sensor networks (WSNs) like ZigBee. It usually results in the degradation of the throughput, increased latency and the depletion of the energy constrained sensor nodes. [8]. As a result, congestion control is a challenging issue for all WSNs [9, 10]. As it is with most WSNs, congestion control in ZigBee based networks (which utilizes the CSMA-CA algorithm) is done by simply trying to regulate the transmission rate at the transport layer in order to relieve congestion. With this algorithm configured, a considerably low latency and high throughput which can result in a stable network cannot be guaranteed. A review of

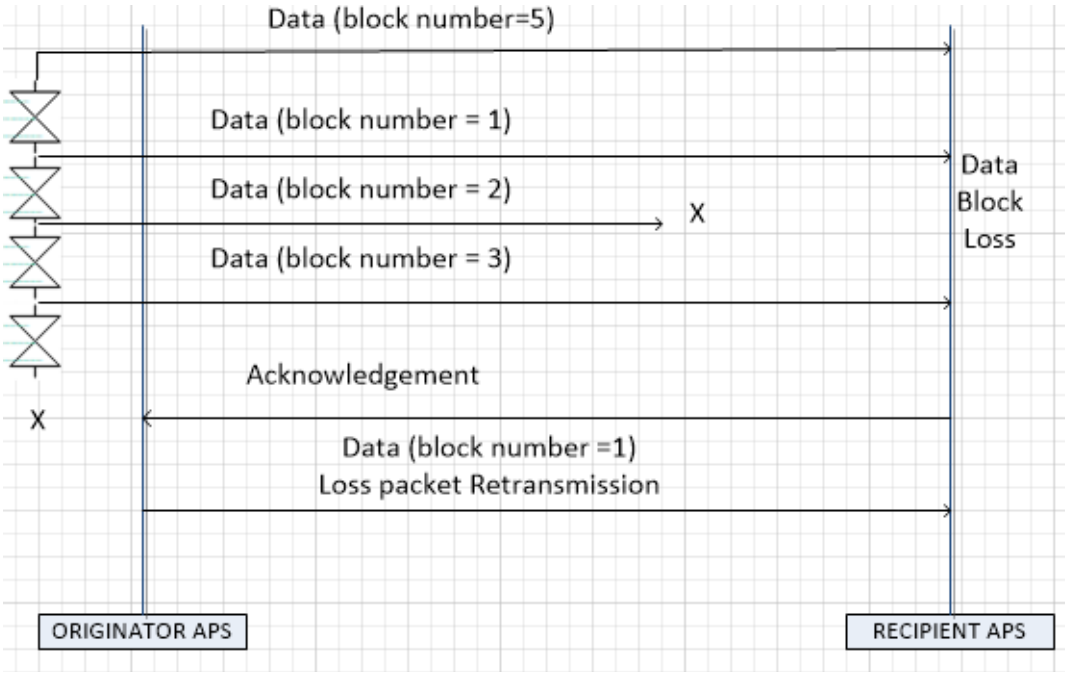

Fig. E.2: A flow diagram for ZTCC [4]

different research strategies geared towards controlling congestion in ZigBee based networks is provided in this section. Woo Sub Jeong [4], proposed ZigBee APS layer in a many-to-one communication application for efficient congestion control. According to their proposed ZigBee time congestion control (ZTCC), the moment a node originator recognizes congestion, it will control time duration for data transmission as shown in Fig.E.2. In the ZigBee protocol, time value is added to application support sub-layer (APS) information base. If that time value continues to increase, data transmission time will become enormously long. Fig. E.3 on the other hand shows their proposed ZigBee congestion control frame (ZCCF) algorithm for detecting congestion. Just like in Fig. E.2, Fig. E.3 still illustrates the mechanism on how the originator and the recipient detect the network congestion. In Fig. E.3, the ZigBee originator considers the situation as the loss of the ACK frame and recognizes congestion. If it fails to receive the ACK in specified period after data transmission, ZigBee originator then activates send 'Congestion Inform Request Command' to the recipient announcing that congestion has occurred. The recipient sends 'Congestion Reply Command' to the originator. The ZigBee APS Command id field tags the APS Command with 1 byte. The network status field is used to announce which data frame caused congestion. The moment the ZigBee originator receives 'Congestion Reply Command,' it controls time duration for data transmission [4]. The ZigBee APS layer, other variants were implemented in NS2 for comparison and analysis.

M. Ouadoua et al [11] discussed IEEE 802.15.4 ZigBee sensor networks while enhancing its topology using Minimum Spanning Tree (MST) in terms of energy consumption and network's lifetime. Other

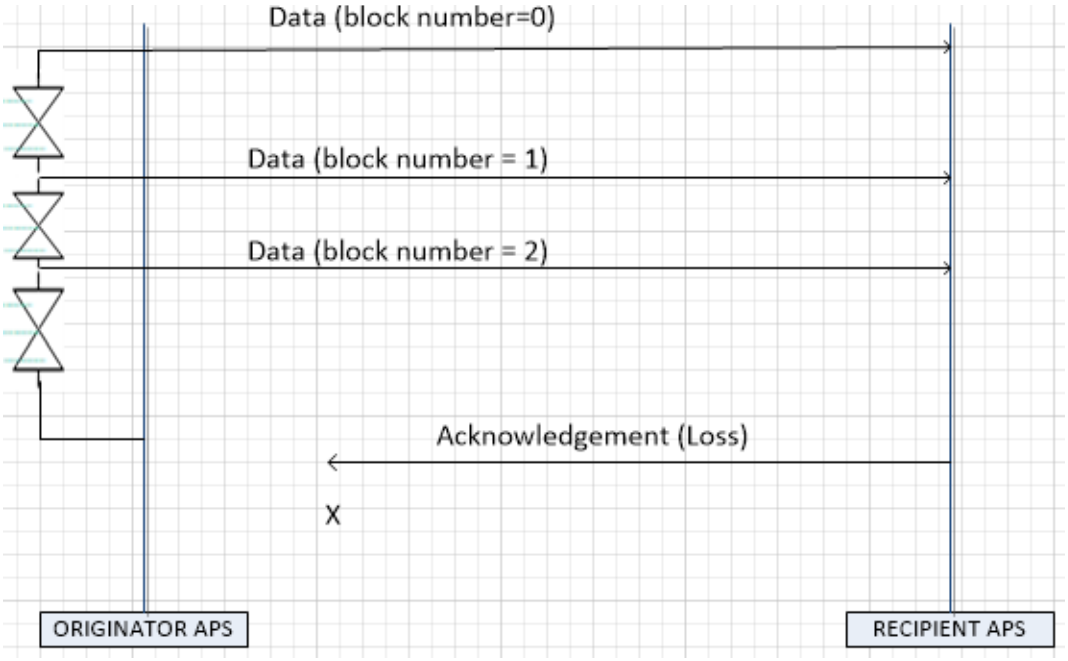

Fig. E.3: A flow diagram for ZCCF [4]

works on ZigBee data communication networks vulnerable to congestion were studied in [12, 13]. The work in [14] discussed congestion control in ZigBee networks but their approach will be unsuitable for a smart grid AMI. Finally, Baz et al [7] proposed a time-based frame aggregation and selective frames algorithms for improving the slotted CSMA-CA algorithm used in IEEE 802.15.4. The time-based frame aggregation was designed to reduce the waiting times resulting from differences in frame sizes. In this case, multiple frames are combined to reach a multiple of a slot duration of about 10 octets. This reduces the overhead incurred in transmission, increases throughput and channel utilization and reduces delay. On the other hand, the selective frames scheme was proposed to resolve issues that could lead to a blocking state for a device in its attempt to transmit a frame longer than its allocated transmission period. This enables the device to select an alternative frame which can be transmitted within the remainder of the period. From the literature, it can be noticed that congestion management in ZigBee enabled networks has received very little research attention. More so, the few research works dealing on the congestion management has not been applied to data aggregation in the context of the smart grid advanced metering infrastructure (AMI) network. To guarantee a reliable transmission and network efficiency in ZigBee network, an effective and reliable data transmission is yet to be found in ZigBee networks. This paper, proposes an RTCA_CMS to mitigate possible congestion in ZigBee based AMI networks. This can control and reduce congestion scenario in real-time. In this context, a message flooding is contextualized for congestion scenario. The ZigBee congestion time control (ZTCC), and ZigBee congestion control frame ZCCF [4] are compared with the RTCA_CMS. This work seeks to significantly reduce and prevent transmission congestion considering the traffic

overheads during message aggregation.

4 RTCA_CMS Design Methodology

The congestion minimizing scheme proposed in this work was designed and implemented using a robust Ring Triangulation Communication Architecture (RTCA). The node process modelling of important network objects introduced into the RTCA testbed had already been explained in chapter 5 of this thesis. This architecture was designed to conform to a typical SG AMI architecture (as shown in Fig. E.1) which has support for multi-party AMI back-end systems. On the design environment, the application support configuration offers a well-defined interface and control services for the network design. It works as a bridge between the network layer and the other elements of the interface. The advanced metering infrastructure (AMI) is designed as a multi-way communication between electricity users and its service providers. A reliable smart grid model is developed which allows energy consumers to participate in AMI value chain. The exchange of information by consumers is done via the ZigBee nodes which were properly configured using the IEEE 802.15.4 MAC/PHY layers. On the AMI, smart adjustment of consumption profile is done via the AMI gateway which can be connected via the Internet. Fig.E.4 shows a conceptual model for basic data aggregation by the ZigBee nodes in the network. On the other hand, Fig. E.5 shows feasible data aggregation using our RTCA_CMS scheme while Fig. E.7 shows the smart grid AMI nodes in the designed RTCA architecture.

The algorithms 1 and 2 are utilized for the compressed data payload generation and data aggregation respectively. A local concentrator was placed at the edge of RTCA design while using a global concentrator to check for anomalies in the AMI server clusters. The algorithms 1 and 2 were then converted to C++ trace files in order to enable the simulation runs. It is to be noted that the above algorithms utilized symmetric based encryption and decryption schemes based on AES. The activation of Grid Openflow firewall in algorithm 2 enables full-duplex transmission which makes easy for illegitimate traffic to be isolated from legitimate traffic. Finally, the default retransmission timeout (RTO) was set to a 30secs for the RTCA topology while simulation run time was set to 3600secs. The RTCA_CMS design was carefully done in this research work in order to achieve a lower delay and high throughput at all times.

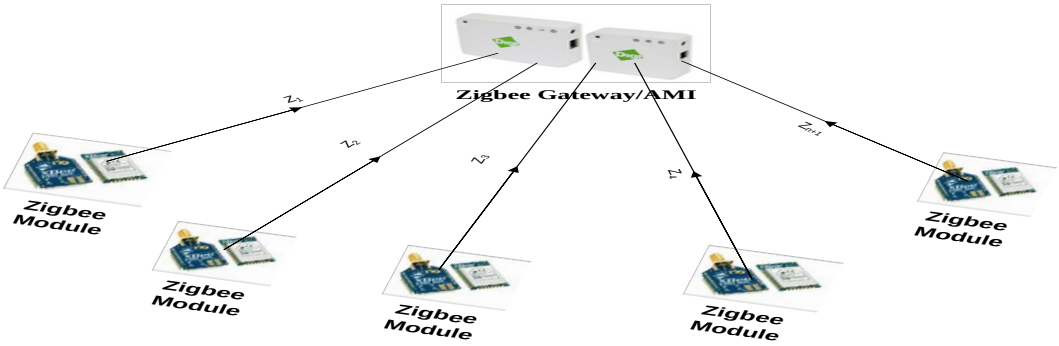

Fig. E.4: Basic Data aggregation by ZigBee Nodes

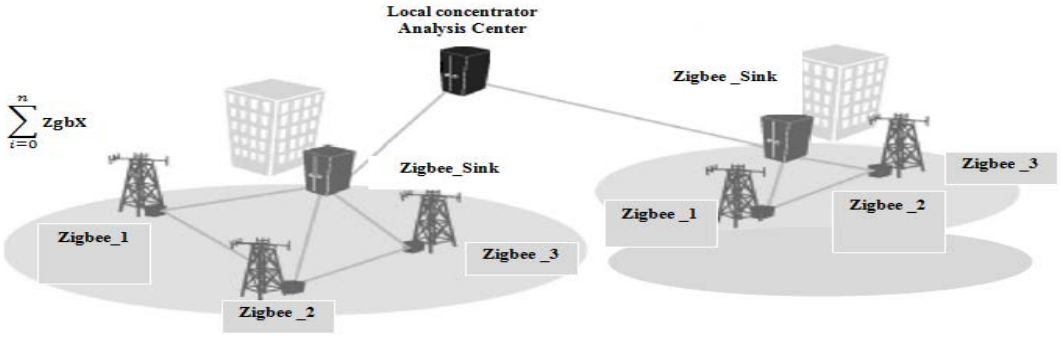

Fig. E.5: ZigBee Data Aggregation clusters used for the RTCA_CMS

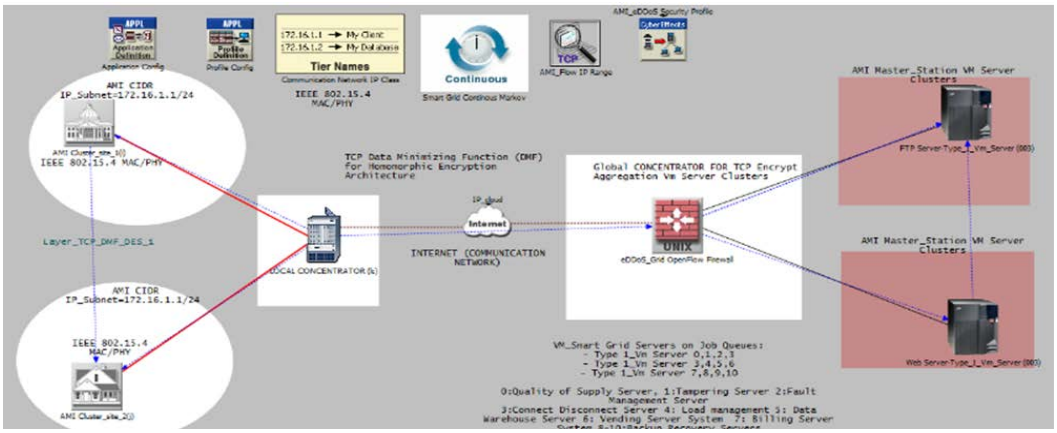

Fig. E.6: ZigBee Experimental RTCA topology based on the IEEE 802.15.4 PHY/MAC

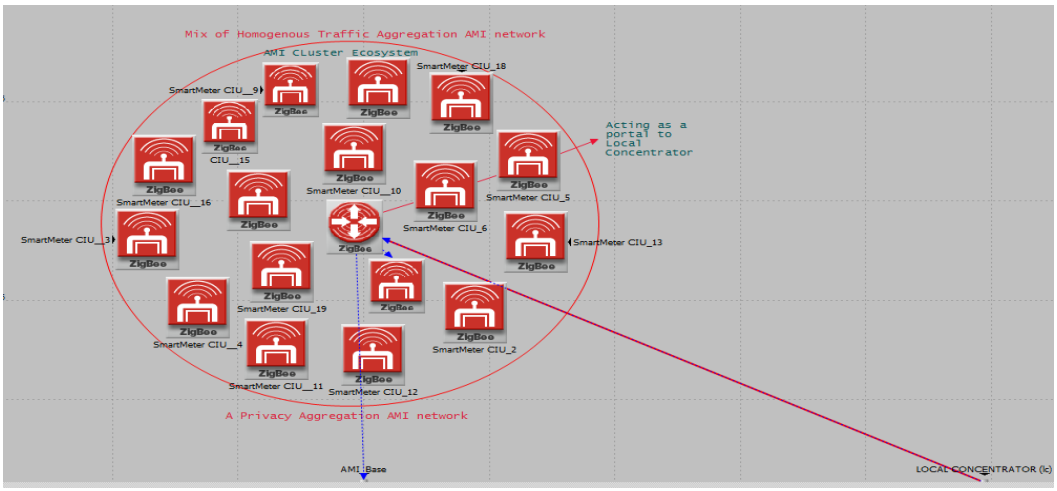

Fig. E.7: ZigBee Smart Meter/AMI Nodes in RTCA Topology

5 Simulation Results and Analysis

With our RTCA_CMS, traffic processing delay is relatively low despite the traffic overhead. Using Fig. E.6 for validation analysis of RTCA for data aggregation workload, our proposed scheme was used to relieve the transmission congestion during message aggregation. Data encryption and data aggregation at the concentrators constitute a huge traffic workload. The RTCA_CMS scheme was compared with ZCCF and ZTCC schemes, in a smart grid AMI case study scenario with regards to latency and throughput as shown in Table E.1. With our proposed scheme, traffic workloads from various users will receive fair share of computations regardless of the overhead. This is very significant for environments that have many users in the smart grid AMI. From Fig. E.8, latency response from ZigBee AMI clusters is quantified by the RTT feedback. It is measured as a directional one-way or round trip delay time. Round-trip latency is determined from a single hub. It does include the time a destination system (sink) spends processing the packet.

However, in the RTCA network, energy data packet is sent via gateways in the form of concentrators. Absolute throughput is established once all the data is completely received. RTCA latency is the sum of the link latencies, including link delays and the forwarding latency of each gateway/concentrator. Queuing and processing delays adds to the minimum latency. Queuing delay results of a concentrator receives multiple packets from different subnet clusters. It was observed that the proposed RTCA_CMS had the least latency profile (5.96%) which is as a result of lower queuing length. The ZCCF and ZTCC have 10.71% and 83.33% latencies respectively. Again, Fig. E.9 shows the validation of RTCA_CMS throughput behaviour which has 39.28% which is relatively better than other congestion control schemes. The ZCCF and ZTCC had 26.79% and 33.93% respectively. This

implies that the RTCA_CMS has the ability to deliver successfully the real time data captured from the end user AMI. Table E.1 shows the summarized variations of the different schemes under consideration.

Algorithm 1: Zigbee Data Construction Algorithm

Input: Call AMI ZigBee Network_Construction (Message msg, MsgType msgType) {

Output: Send to Sink_AMI local concentrator{

1. For N(i) = 1 to K_{n+1}
2. Call command EncryptData (128 AES EncrytedData){
3. AggregateNode = gather_neighbor(random);
4. {
5. KeyExpansion (CipherKey, ExpandedKey)
6. AddRoundKey (State, ExpandedKey)
7. AddRoundKey (State, ExpandedKey)
8. For (i =1, i<Nr; i++)
9. Round (State, ExpandedKey + Nb*i)
10. FinalRound (State, ExpandedKey + Nb * Nr)
11. Ready to Send to Local concentrator ( )
12. }
13. If construction = Ready Then Send ( )
14. Else Return ( )
15. End Algorithm

Algorithm 2: Secured Data Aggregation Algorithm

Input: Call AMI Network_Construction (Message msg, MsgType msgType) {

Output: Send to Sink_AMI (AMI meters)

1. Begin ( )
2. For N(i) = 1 to K_{n+1}
3. Call function Data Aggregation (Message msg, msgType msgType)
4. If (a metering node is Set) then Send Message(encData) to Sink AMI_Node
5. Else{
6. If (a AMI_node receive Event.message(encData) from Rx.sensor node { Then
7. Call local concentrator (RTCA_CMS)
8. If (the local concentrator node receive Event.message(encData)) { Then
9. Stores encData from msg (AESCipher_DataBlock);
10. decryptedData = decryption(encData)'
11. aggregatedData += decryptedData;
12. Call Global Concentrator ()
13. If (all data is received from local concentrator.childNode)
14. SendMessage(encData) to Grid OpenFlow Firewall.ParentNode; }
15. If (a node is SinkNode){ Then
16. Store encData from msg;
17. decryptedData = decryption(encData);
18. Send Message(decryptedData) to User AMI Master_Stations; }}}
19. End if
20. End if
21. End if
22. End if
23. End if
24. End // terminate Algorithm

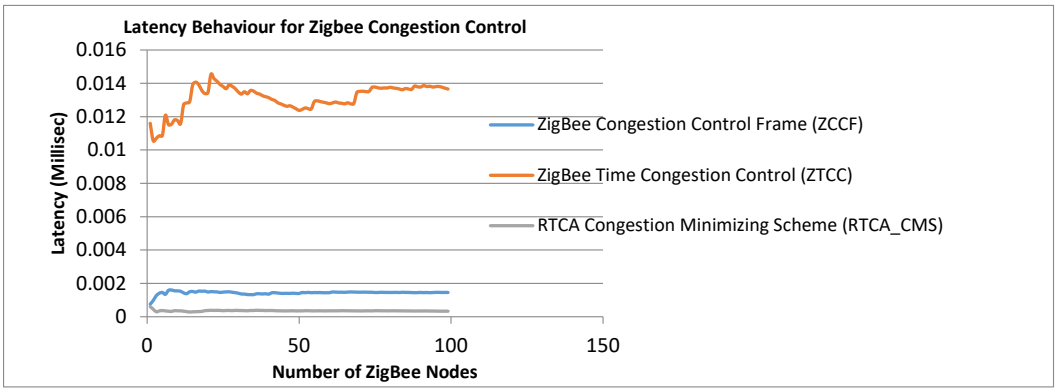

Fig. E.8: Validation of ZigBee RTCA_CMS latency behaviour

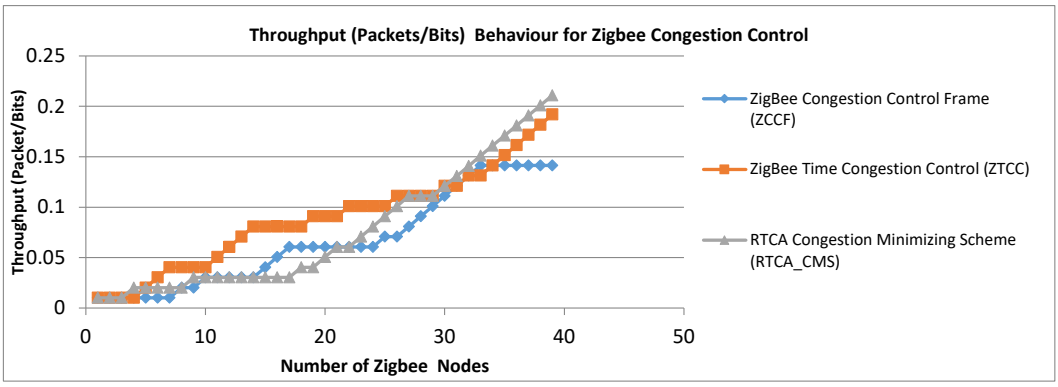

Fig. E.9: Validation of ZigBee RTCA Throughput behavior.

Table E.1: Result Summary for data aggregation workloads

System metrics	ZCCF	ZTCC	RTCA_CMS
Average latency profile	10.71%	83.33%	5.96%
Average throughput	26.79%	33.93%	39.28%

6 Conclusion

In this paper, ZigBee RTCA congestion minimizing scheme in smart grid AMI for enhanced data aggregation is presented. The scheme seeks to reduce network traffic leading to congestion scenario during data aggregation. From the result analysis, the proposed RTCA_CMS can be used to relieve transmission congestion and reduce overheads incurred during data aggregation. QoS metrics such as network latency and system throughput, during data aggregation were analyzed for ZigBee congestion control frame (ZCCF), ZigBee Time congestion control (ZTCC), in relation to the proposed ZigBee RTCA_CMS. The QoS metrics evaluated show that our scheme can reduce latency

considerably at the aggregation point in the subnet cluster of RTCA. Its throughput behaviour shows significant improvement for the smart grid AMI network.

References

[1] M. M. Hasan and H. T. Mouftah, "Encryption as a service for smart grid advanced metering infrastructure," in *Computers and Communication (ISCC), 2015 IEEE Symposium on*. IEEE, 2015, pp. 216–221.

[2] Z. Specification, "Zigbee standards organization," *Document 053474r17, Jan*, vol. 17, p. 26, 2008.

[3] J. F. Kurose, *Computer networking: A top-down approach featuring the internet, 3/E*. Pearson Education India, 2005.

[4] W. S. Jeong and S. H. Cho, "Congestion control for efficient transmission in zigbee networks," in *Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on*. IEEE, 2009, pp. 1–4.

[5] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, and Y. F. Hu, "Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and zigbee standards," *Computer communications*, vol. 30, no. 7, pp. 1655–1695, 2007.

[6] L. Bai, Y. Liu, S. Qian, and S. Zhang, "Zigbee hybrid routing algorithm for network energy optimization based on node cluster label," in *Software Engineering and Service Science (ICSESS), 2016 7th IEEE International Conference on*. IEEE, 2016, pp. 726–729.

[7] M. Baz, P. D. Mitchell, and D. A. J. Pearce, "Improvements to csma-ca in ieee 802.15.4," in *2012 IEEE 14th International Conference on High Performance Computing and Communication 2012 IEEE 9th International Conference on Embedded Software and Systems*, June 2012, pp. 1549–1554.

[8] G. Wu, F. Xia, L. Yao, Y. Zhang, and Y. Zhu, "A hop-by-hop cross-layer congestion control scheme for wireless sensor networks," *arXiv preprint arXiv:1201.0207*, 2011.

[9] B. Hull, K. Jamieson, and H. Balakrishnan, "Mitigating congestion in wireless sensor networks," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*. ACM, 2004, pp. 134–147.

[10] W.-w. Fang, J.-m. Chen, L. Shu, T.-s. Chu, and D.-p. Qian, "Congestion avoidance, detection and alleviation in wireless sensor networks," *Journal of Zhejiang University SCIENCE C*, vol. 11, no. 1, p. 63, 2010.

[11] M. Ouadou, O. Zytoune, D. Aboutajdine, Y. El Hillali, and A. Menhaj-Rivenq, "Improved cluster-tree topology adapted for indoor environement in zigbee sensor network," *Procedia Computer Science*, vol. 94, pp. 272–279, 2016.

[12] Z. Yi, H. Hou, Z. Dong, X. He, Z. Lv, C. Wang, and A. Tang, "Zigbee technology application in wireless communication mesh network of ice disaster," *Procedia Computer Science*, vol. 52, pp. 1206–1211, 2015.

[13] O. Okoro, E. Muoegbunam, A. Ekene, K. Okafor, and A. Uzoamaka, "The impact of traffic congestion on wireless sensor networks; a case for traffic distribution on telosb rf transceiver device"," *Engineering and Technology*, vol. 2, no. 2, pp. 35–47, 2015.

[14] D. S. Yun and S. H. Cho, “A data transmission method in zigbee networks using power efficient device,” in *Advanced Technologies for Communications, 2008. ATC 2008. International Conference on*. IEEE, 2008, pp. 162–165.

Chapter VII

A Cloud-Based OpenFlow Firewall for Mitigation Against DDoS Attacks in Smart Grid AMI Networks

Paper F

A Cloud-Based OpenFlow Firewall for Mitigation Against DDoS Attacks in Smart Grid AMI Networks

R.C Diovu and J.T Agee

Published

IEEE Xplore (Presented at the 2017 IEEE PES PowerAfrica)

Abstract

Recent architectures for the advanced metering infrastructure (AMI) have incorporated several back-end systems that handle billing and other smart grid control operations. The non-availability of metering data when needed or the untimely delivery of data needed for control operations will undermine the activities of these back-end systems. Unfortunately, there are concerns that cyber attacks such as distributed denial of service (DDoS) will manifest in magnitude and complexity in a smart grid AMI network. Such attacks will range from a delay in the availability of end user's metering data to complete denial in the case of a grounded network. This paper proposes a cloud-based (IaaS) firewall for the mitigation of DDoS attacks in a smart grid AMI network. The proposed firewall has the ability of not only mitigating the effects of DDoS attack but can prevent the attack before they are launched. Our proposed firewall system leverages on cloud computing technology which has an added advantage of reducing the burden of data computations and storage for smart grid AMI back-end systems. The openflow firewall proposed in this study is a better security solution with regards to the traditional on premises DoS solutions which cannot cope with the wide range of new attacks targeting the smart grid AMI network infrastructure. Simulation results generated from the study show that our model can guarantee the availability of metering/control data and could be used to improve the QoS of the smart grid AMI network under a DDoS attack scenario.

1 Introduction

The smart grid is believed by many to be the electric power grid for the future. In [1, 2], the electric power grid was defined as electricity network that can cost-efficiently integrate the behaviour and actions of all users connected to it for better economic efficiency, energy sustainability and service delivery to end users. A very important factor that can keep the smart grid stable and efficient is its supporting information infrastructure. The advanced metering infrastructure (AMI) is regarded as the foundation of the smart grid [3]. The AMI includes smart meters, communication networks, and data management systems. The AMI also incorporates in its design a means of collecting data into software applications and interfaces [4]. Smart meters connected to the AMI network can provide end user's with control over their energy consumption and also can facilitate the integration of distributed energy resources. On the other hand, smart grid utilities will benefit from being able to perform core operational functions like outage detection, remote meter readings and options like

prepaid or time-of-use metering solutions. The deployment of technologies such as the AMI to the electric power grid will increase the reliability of the grid and reduce costs of power delivery, but they also bring about dependencies on cyber resources which may be vulnerable to attack [5, 6]. A compromise of the metering networks may allow an attacker undue access to the control functions that, if corrupted, will threaten the availability of the data in the system and consequently violate the integrity of the system. Two major research questions that must be addressed in order to enhance the wide-spread adoption and implementation of smart grid AMI concept include: (1) How can the real-time data from smart meters and other numerous intelligent electronic devices (IEDs) that transverse the smart grid's AMI networks be managed efficiently and intelligently? (2) How can appropriate cyber security solutions against wide range of vulnerability attacks or cyber threats targeting the AMI network be designed and deployed at the right places? Storing and processing the huge data needed for AMI's applications such as meter data management, billing system and load management is an arduous task which can reveal sensitive personal information of consumers if appropriate security is not put in place. This issue can be minimized by hosting some AMI's applications on a fault tolerant and consolidated Server-centric data centres with QoS and security assurance mechanisms. Standard documents released by standard bodies for AMI architectural implementations in different countries are already incorporating data management systems/data centres into the AMI architecture [7, 8]. Califano et al [9], opines that it is almost inevitable that the smart grid and cloud computing technology will be integrated given their distributed nature. Integrating cloud computing into the smart grid presents a new platform for energy suppliers to satisfy the enormous data storage and computation capabilities required in an evolved SG environment [10–12]. On the other hand, dealing with the issue of vulnerabilities arising from cyber threats against the AMI network will require a continuous research effort as cyber crimes are constantly evolving. Bou-Harb [13] categorized these cyber threats as: (1) connection-based and (2) device-based threats or vulnerabilities. According to the study [13], the connection-based attacks exploit the vulnerabilities existing within the communication channels and protocols. Examples include: jamming, eavesdropping, and message injection or modification attacks. The device-based attacks on the other hand usually exploit flaws and vulnerabilities on the devices to perform malicious activities and they include: denial of service (DoS/DDoS), man-in-the-middle, and metering data tampering/falsification. The focus of this study is on mitigating the effect of distributed denial of service attack (DDoS) on an AMI network. The rest of the paper is organized as follows: Section II contains an overview of DDoS attacks while section III has a brief review on related research works. The design methodology of the grid openflow firewall is described in section IV while section V contains the simulation results and analysis. Finally section VI provides the conclusion for our study.

2 Overview of Distributed Denial of Service (DDoS) Attacks

In contrast to the traditional DoS attack which uses a single source of attack, DDoS attack exploits numerous attack sources that are spread across multiple hosts thereby amplifying its damaging effect and making defense more complicated. DDoS attacks have become the weapon of choice by cybercriminals as their impact on targeted organizations can result to several hours, days or even weeks of network downtime. Many a times, DDoS attacks in order to circumvent traditional network defenses, target web applications and application databases of organizations. These DDoS attacks do so by trying to mimic regular web traffic thus initiating requests that are too slow to be detected by traditional firewalls and other network gateway securities. A diagram for the conceptual model of distributed denial of service (DDoS) attack in AMI network was given in [14]. This model represented in Fig. F.1 shows a DDoS attack initiated by an attack agent targeting to exploit system vulnerabilities that will affect the availability of network service and devices. According to the model, the attacker's main objective is to exhaust the bandwidth and processing power of its vulnerable victims which can be smart meters or other edge devices connected to the AMI network. The discovery of vulnerable devices is the first stage in launching a DDoS attack in AMI network. This stage is then followed by an attack stage. In an AMI network, three types of attack have been identified. They include: attacks targeting network protocols, attack on network infrastructures, and attacks on network bandwidth. Asri [15] noted that a cybercriminal can launch TCP SYN flooding attack in order to disable the network service of the AMI back-end applications. For a packet exchange oriented AMI network utilizing special infrastructures like routers, an attacker may decide to maliciously disorganize the routing tables in order to affect the packet delivery success [16]. In the case of bandwidth attack, the attacker's main goal is to saturate the available bandwidth by sending volumetric traffic to its agents (victims) which could result in increased packet loss ratio and blockage of legitimate traffic [17].

3 Review of Related Work

In this section, a quick review of the researches related to our proposed study is provided. Yonghe et al [14] have studied the impact of bandwidth consumption attack on the AMI network using network simulator 3. From the study, it was discovered that when 60% of the total smart meters connected to the AMI network have been compromised and used as agents, the impact of this attack results in almost 50% increase in packet loss. In addition, the average and maximum end-to-end delays were

13.5 and 5.48 times higher than their values in normal network operating condition. A similar study was conducted in [15] using network simulator 2. In this case, the target server was flooded with UDP attack. It was discovered that at a simulation time of 300 ticks (atomic discrete time unit), a sudden increase in the number of packets forwarded to the server was observed. The server was finally grounded as the attack continued at a simulation time of 500 ticks resulting in the dropping of all the subsequent incoming packets. In [18], a risk assessment tool called SecAMI was introduced to function at the NAN of the AMI network. The aim of the research was to study how quickly an attack can spread at the NAN and how fast it can be detected using their proposed tool. They also used the proposed tool to evaluate the average response time needed to mitigate such an attack. Finally, in [19], the authors also studied the impacts of DDoS attack on the AMI network using OMNET ++ simulator.

From the review, it was discovered that previous research efforts dealing with DDoS in AMI network had focused mainly on studying the impact and detection of DDoS on the network. One major drawback from all the research work reviewed is that they all have stopped at the analysis of the impact of such attacks on the AMI network. Our proposed study have gone a step higher by proposing a security solution to such attacks and showing how it can be used to prevent or mitigate the effects of the attack and improve the QoS of an AMI network in an attack scenario. Our proposed study is unique when compared to existing related studies in the following ways. Firstly, it is a hybrid cyber security solution that combines the recent technological advances in cloud computing technology with the traditional infrastructural security capability. Secondly, our proposed grid openflow firewall (GOF) has the ability of not only detecting and mitigating the effects of DDoS attacks but can prevent the attack before they are launched. Thirdly, our proposed firewall can handle 250 Gbps DDoS volumetric attack which cannot be managed by the traditional gateway securities. A firewall can be described as a set of security measures primarily designed to prevent unauthorized users access to a network. A firewall can be implemented in the form of a hardware or a software or a combination of both. It has the ability of monitoring and/or controlling traffic between nodes connected to the network. Prevention of unauthorized access to data, information or network resources can be done by enforcing rules or policies on which traffic or network packets can be stopped or allowed to pass to its destination. It is our belief that our proposed GOF will open more routes for several infrastructural-based firewall security researches for the smart grid AMI which is at the moment very scarce in the literature.

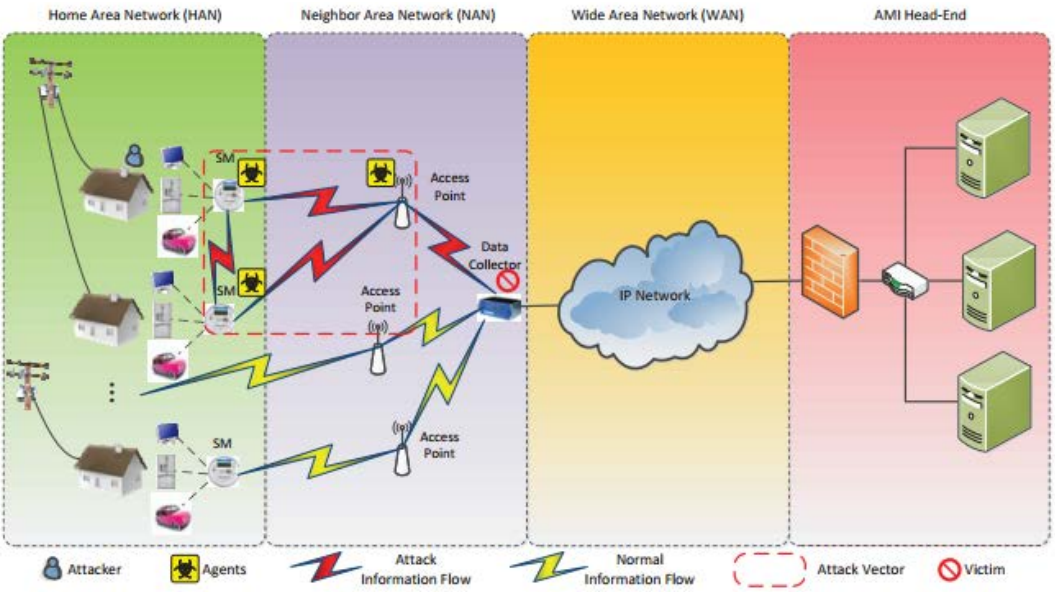

Fig. F.1: DDoS in AMI network [14]

4 Design Methodology of Grid OpenFlow Firewall (GOF)

Our proposed grid openflow firewall is a cloud based AMI firewall network which is designed to mitigate and dynamically handle the changing threat landscape of distributed denial of service (DDoS) attacks. This proposed cloud-based firewall service allows only legitimate traffic from the AMI to reach the cluster backed severs while stopping illegitimate traffic at the edge, before it hits the utility energy server clusters. In this network, DDoS mitigation is focused on keeping the TCP/IP web characteristics constant on a twenty four hours basis regardless of the network condition. In this research work, a well designed cloud based network was developed using the parameters in Table F.1 to understudy and validate the influence of the proposed grid openflow firewall (GOF) on the AMI based server infrastructure.

The GOF was designed leveraging on the emerging potentials of software defined networking (SDN) paradigm. Software defined networking (SDN) enhances programmability and makes it easier to configure the network by decoupling the control plane from the data plane. The control plane is the logic infrastructure that controls the way traffic is forwarded from the data plane of a networking device such as a router or a switch. The openflow protocol utilized for the design of the proposed firewall uses an openflow enabled switch which can isolate traffic by the application of virtual local area networks (VLANs) to the normal process of the switch. With the openflow protocol, message structure can be defined which can enable the controller to manipulate (to add or modify) the flow entries in the openflow logical switch flow tables. The flow table can be used to decide how each data

flow should be processed. This is typically done by associating an action with each flow entry. In the proposed openflow firewall design, segmented logical virtualization was achieved through component instantiation. This means that the network has a layer two services captured in the list record of the OpenFlow table map services. In its design, services are binded and bundled as services into a virtualized node. This was done in the form of services consolidation. With the introduction of linked lists, the function then maps these services into the OpenFlow controller node. As the list table increases, the firewall adapts by making more subroutine calls on new resources instances. Algorithm I shows the openflow firewall logical instantiation while Algorithm II represents the openflow firewall services routine map. The two algorithms are shown in F.2 and F.3 respectively. It is to be noted that the choice of the pointer operator (up arrow) used in the two algorithms is associated more with PASCAL than C++. This deliberate choice was made just to avoid confusion of using one symbol ($*$) to represent both multiplication and pointer operators. By way of definition, a variable of the pointer type contains a memory address where data of another variable can be stored.

Algorithm I: OpenFlow Firewall logical instantiation

Begin *( )*
Create Instantiation map ( )
For *node = I to* N_{k+1}
OpenFirewall Function add $F(I_0, J_1, K_1, N_{k+1}$*:link): link ;*
Var ℓ:link; i = integer;
    Begin
      ℓ: = β*; // OpenFlow load balancer self connection*
    recycle loop;
      new (ℓ $\uparrow$*. next); ℓ: = ℓ* $\uparrow$*: next;*
      P:P$\uparrow$ *next; Q:=Q*$\uparrow$*. next; R*$\uparrow$*. next;* N_{k+1} $\uparrow$ *next*
      Do Until *(P=* β*), (Q=* β*),(R=* β*) and (*N_{k+1} *=* β *);*
      (ℓ $\uparrow$*. next):=* $\mathbb{Z}$*; add:* β $\uparrow$*next F;*
End

Fig. F.2: OpenFlow Firewall logical instantiation

Algorithm II: OpenFlow Firewall Services C (*VLAN Switch-*$V_{sw,}$ *Firewall-*$f_{w,}$ *SLB-*$S_{slb,}$ *VBB-*$V_{bb,}$ *….Xn+1*) ,(destination address, source ID, Queue size, link tabel entries) routine map

```
1:  Inputs:       OpenFlow monitorCallSchedule Openflow firewall services bundling
                  Source address, destination address, queue size, link Information
                  α1&β2 // initialValue legitimate & trendPosteriorValue, illegitimate traffic
    Output:       OpenFlow_DES
    Parameters:       OpenFlow _weight←Empty; // OpenFlow weighted Moving Average
                      weight←0; weightedMoving←0; totalWeight←0;
                      fiboA←0; fiboB←1;
                      OpenFlow _weight ContainerhistoryItem← null; // QoS Pattern
    int i←0;
    While i < OpenFlow _monitorCallSchedule d do
        ListItem←HistoryList.get ( HistoryList.Size() – OpenFlow_monitorCall – i )
        OpenFlow _weight ←fiboA1 + fiboB2;
        OpenFlow weightedMoving←weightedMoving+ (ContainerhistoryItem * weight);
        total α1&β2 _weight ←total α1&β2 _weight + α1&β2 Sweight;
        i ++;
    end while
    OpenFlow Firwall_DES ← α1&β2 weightedMoving / total α1&β2_Weight;
    // Calculate event_filtering(α1&β2 ) & execute dynamic network balancing
    Legitimate_initialValue ←(α1 * OpenFlow_weight) * (pastInitialValue +
    trendPosteriorValue);
    illegitimateTrendPosteriorValue←β2 * (initialValue - pastInitialValue)+ (β2) *
    pastTrendPosteriorValue);
    OpenFlow_DES ←initialValue+ trendPosteriorValue; (α1&β2)
    //Result:  Services bundling and load dispatch L_M to cluster servers
    //Define OpenFlow load balancer-Services C
    Var ℓ:link; c: real; j = integer;
        Begin
         Type link ℓ: = ↑ node; // controller devices (OpenFlow)
         Node = record c: real; j: integer; next: link end;
         Function listadd (ℓ: link, c: real; j: integer): link ;
         Function listadd F(I_0, J_1, K_1, N_{k+1}:link;
            If (Sensedevent =1) then Enable C // Set services if OpenFlow gateway Enabled,
               new (ℓ ↑. next); ℓ: = ℓ ↑: next;
               ℓ ↑. c : = c;ℓ ↑. j: = j;
               listadd: =N; //Create another instance of virtual node on OpenFlow
                  Forj = 1: n  // Optimize flow table do
                     listcontrol: =N.c // Services control➔delete, merge, modify, read;
           Endifs;
    End;
    Return OpenFlow_DES
```

Fig. F.3: OpenFlow Firewall Services Routine Map

The Openflow firewall list node essentially employs non-zero terms for the services, so that the list node contains values of C (services) and its being controlled by the variable j. This is shown in Algorithm II and was used in the openflow firewall design. In this algorithm, after bundling the services, they remain activated (i.e. firewall, Service Load Balancer, Volume Billing Batcher, Virtual Local Area Network, flow table operations, among others) as long as the Openflow firewall gateway is active. Legitimate and illegitimate connections are concurrently mapped into these services each time a connection is made. The listadd function creates new virtual node and gives it the relevant

specified fields (attributes), and links it into controller interface c_i. Now, the control list routine is used to optimize the flow table such that services can be removed, merged or modified. From the OpenFlow model above, an instantiated dummy node β is used to hold the link which points to the first node as well as connections on the list. This entry list is then constructed for dispatch by a load balancer (Service Load Balancer). After the list generation, β is set to the link implicitly. The essence is to make it feasible for the connected entities to be read and written to form an infinite queue.

In the analysis of the design firewall, volumetric DDoS attack flood from a malicious hacker was focused on the target network. This was done with data packets (ping of death traffic from six hackers as can be seen from the testbed) that have the capability of completely saturating the available network bandwidth. The attack used was 250 Gbps which generated very high volumes of traffic congestion, overloading the targeted smart grid AMI network and its server subsystems. Essentially, this can cause an extensive service disruption for legitimate location based users trying to gain access via web http/TCP/IP service. This volumetric attack (250 Gbps DDoS) which can literally last for a longer duration in a production environment can hijack and possibly shutdown the entire smart grid network within a few minutes in the absence of robust security mechanism. This attack is very pronounced at the network layer (layers 3 and 4) as they can overwhelm the internet link of a server, network resources, and network nodes that are not able to absorb the increased traffic volumes. Fig. F.4 shows the designed experimental test-bed based on the network parameters of Table F.1. This work made comparison between two major DDoS attack mitigation scenarios which include our proposed grid openflow firewall algorithm and the conventional non grid openflow firewall algorithm. With compatible C++ library, a simulation with Riverbed Modeler Engine 17.5 was carried out. Various layers of integrations were satisfied while using the external libraries to populate and build the network map shown in Fig.F.4. On the other hand, Fig. F.5 and Fig.F.6 represent the successful trace file engine build work design for the network and successful compilation of simulation trace file respectively.

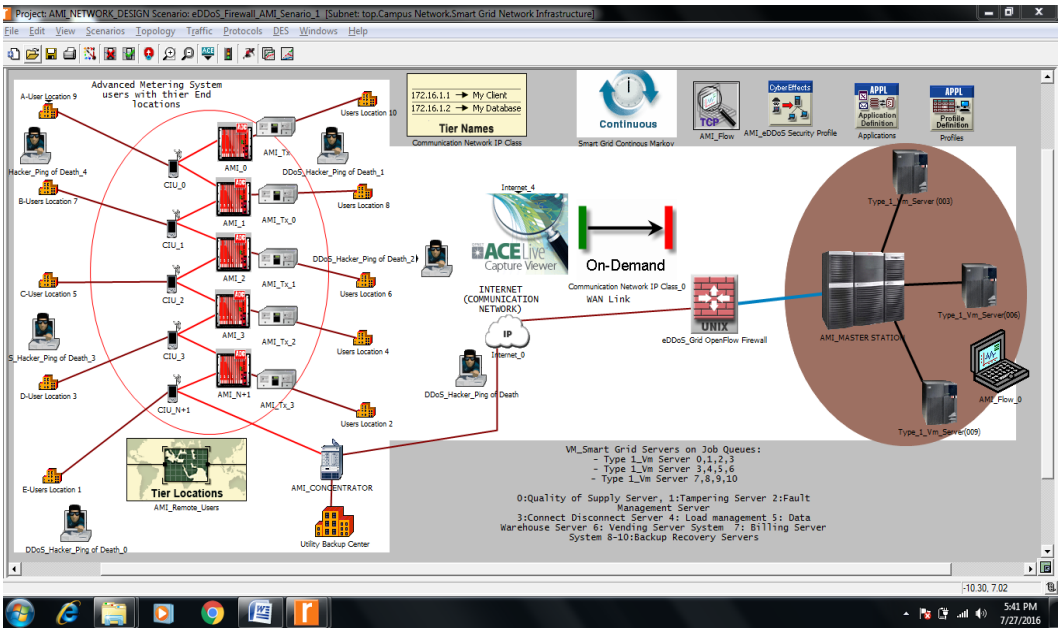

Fig. F.4: Designed cloud-based AMI experimental test-bed

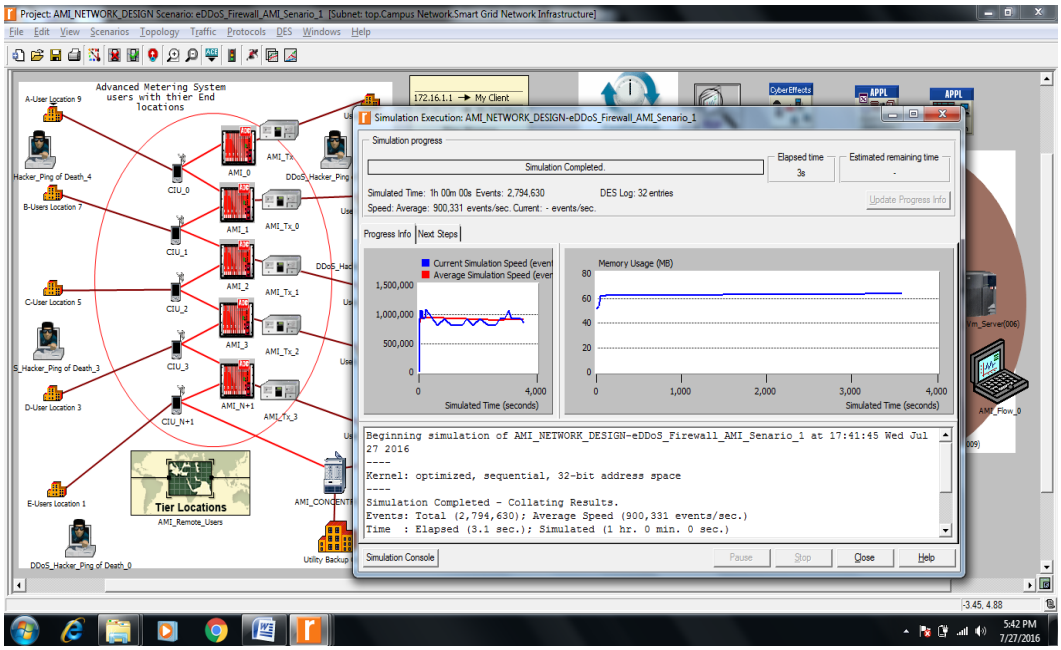

Fig. F.5: Successful trace file engine build work design for the network

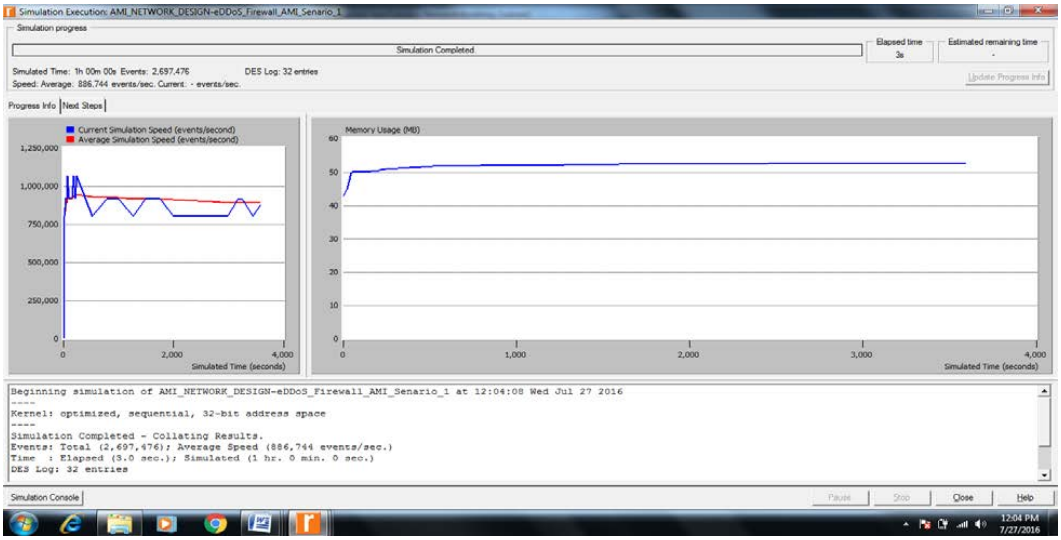

Fig. F.6: Successful compilation of simulation trace file

Table F.1: Network Design Parameters

Parameter/specifications	Values/ Type
Number of AMI network units	5
Number of AMI CIU	5
No. of AMI Concentrator gateway	1
No of Utility datacenter backup	1: 6 Server Clusters
No of Grid OpenFlow Firewall	1 (Cisco Nexus 9000 firewall with support for Virtual DDoS protection in the DCCN threat mitigation design)
No of AMI Master Station	1:3 Server clusters
Type of Payload	TCP/IP
No of Type-1 Virtual Machine Instance	12 Virtual Machines
No of Servers	9:SunUltra 10:333 MHz;1 CPU;1 Core (Simple CPU Mode)
Attack Vector Traffic	DDoS 250 Gbps
AMI Traffic	On-demand DB Query
Ethernet Technology	PPT1 (40 Gbps)

5 Simulation Results and Analysis

Query Response Time

An access to the AMI server to dispatch the query result value can be measured by the local response time. Slow query log provides exact energy information about queries executed. At the remote utility AMI server, there is a large number of queries that could take time to execute. This feature in the AMI server provides a tool for analyzing the information by counting and displaying the number of queries according to the length of time they took to execute. After the server has finished the processing of bypass sensing, data are then collected. Fig. F.7 shows the plot of eDDoS for two considered scenarios of open flow and non-open flow data captures. From the plot, no visible behaviour was observed until at 2000 secs. From the simulation time of 2000 secs to 3500 secs, a gradual gradient is shown. From the simulation time lapse of 3500 secs to 4000 secs, the query response time of openflow AMI firewall is slightly lower (52,000 Msec) than that of the non-openflow AMI firewall (55,000 Msec). By using the openflow firewall, an online analytical processing (OLAP) cache on the AMI server will experience optimal performance. In this case, the response time will show tremendous improvement by accessing the query result set data from the server OLAP cache as well as from its database. As shown in Fig. F.4, the AMI OLAP cache is architected to store energy theft query result sets as highly compressed cluster data. It gives the utility or third party operator (admin entity) application server access to the result sets. If the same query (or a subset) is then executed by another user, the subsequent query request can be filled by accessing the result set already stored in the AMI OLAP cache. In this design, the OLAP cache offers significant performance gains in the server. Storage options are available for query results and can be optimized to best meet the needs of the system as well as for the individual queries. As shown in Fig. F.7, the presence of the openflow firewall created a slight differential with a non-openflow firewall. This plot shows that the AMI will have slightly lesser attack vulnerability on the network when compared to the non-openflow firewall and its access to stored data will be faster. In the real world scenario, prolonged delay can result when there is an attack which cannot be managed by the gateway firewalls.

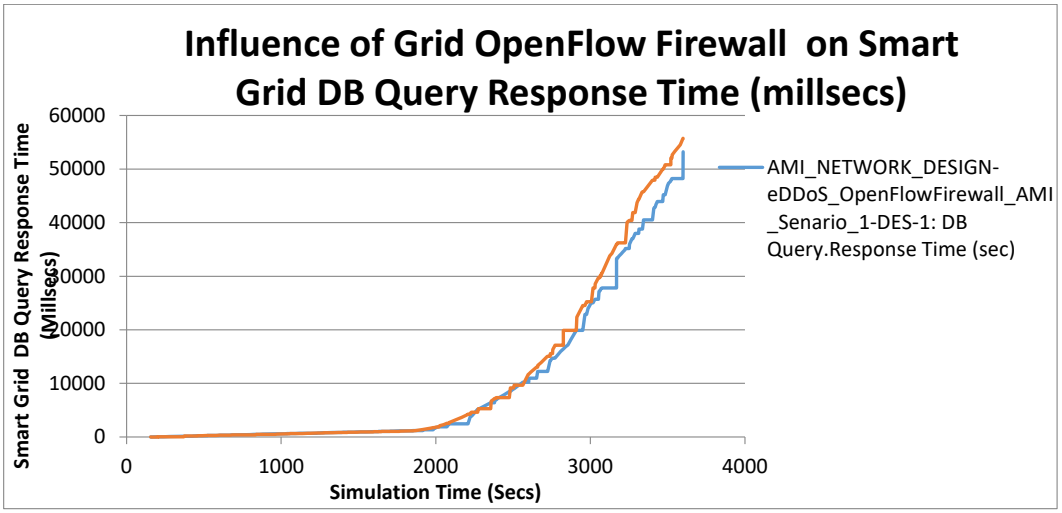

Fig. F.7: Influence of GOF on SG AMI DB Query Response Time (milli-sec)

Resource Utilization Response

This work used resource utilization as another fundamental QoS metric for the smart grid AMI network. By first examining the traffic data used to study the link utilization and possible packet loss at the network core, the temporal patterns of data traffic using packet tracers were then characterized. The observations were then used to evaluate various firewall schemes for traffic analysis in AMI base station servers. The firewall devices in the AMI network are organized into openflow and non-openflow multiple layers. These devices in two scenarios have different physical capabilities. Characterizing the link utilization and packet loss of the scenarios will determine how the security module will benefit from traffic engineering. Fig. F.8 shows that link utilization is over four times lower in the openflow security aggregation layer than in the non-openflow which has about 95^{th} percentile utilization. This is likely because in a generic AMI network topology, there are many users connecting with the AMI base station with aggregation links in most cases. With a connection request by end users, feasible regions of resource allocation are first established. As shown in Fig. F.4, enterprise servers are mapped as virtual machine instances for server consolidation and virtualization.

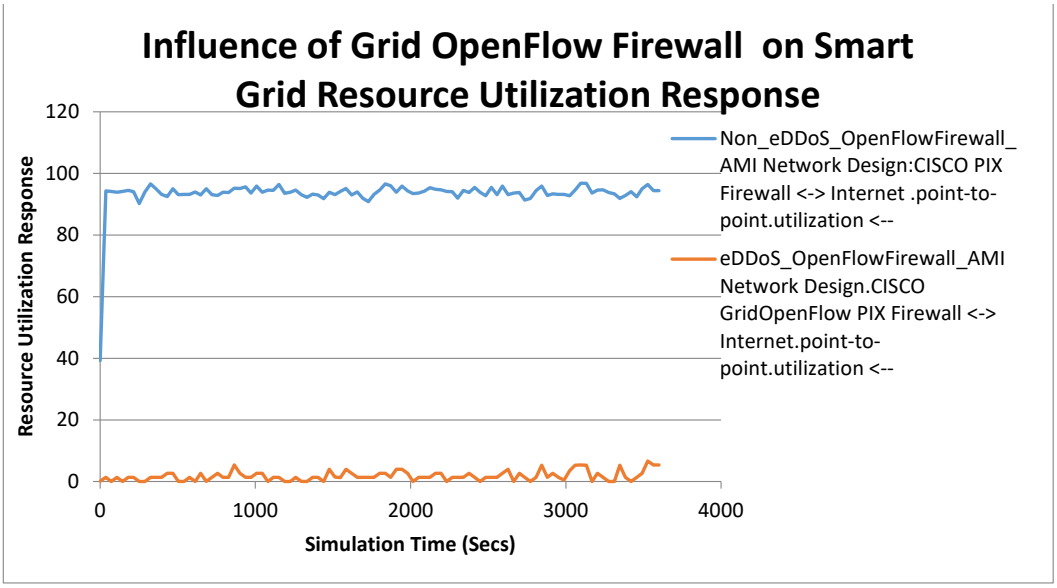

Fig. F.8: Influence of GOF on SG Resource Utilization Response

Network Throughput Response

An interesting QoS aspect of the AMI is the throughput behaviour of the network. The network throughput exponentially remains stable at about 250000 bits/secs. This follows a characteristic behaviour of the low network resource utilization and efficient query response times. Under any attack scenario, the openflow algorithm carefully detects such anomaly and creates an isolation mechanism. This usually impacts positively on the network throughput as shown in Fig. F.9. The converse is the case in Fig. F.10 under non-openflow security scheme where any disruptive attack on the AMI network will normally affect the network throughput. By introducing the attack vector on the network, Fig. F.10 shows the degraded throughput scenario. The more the attack, the more unstable and unreliable the network becomes until it is finally grounded.

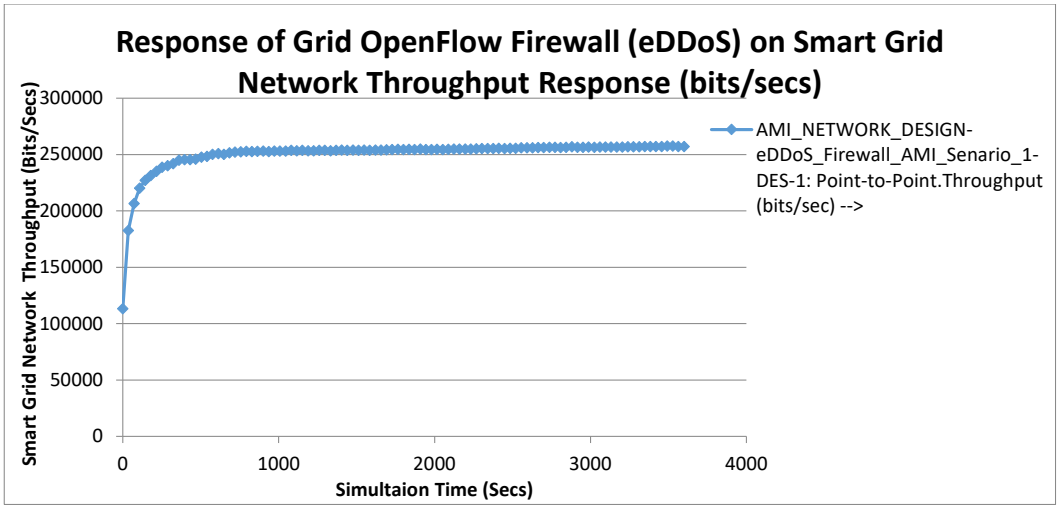

Fig. F.9: Response of GOF on SG AMI network throughput (bits/secs)

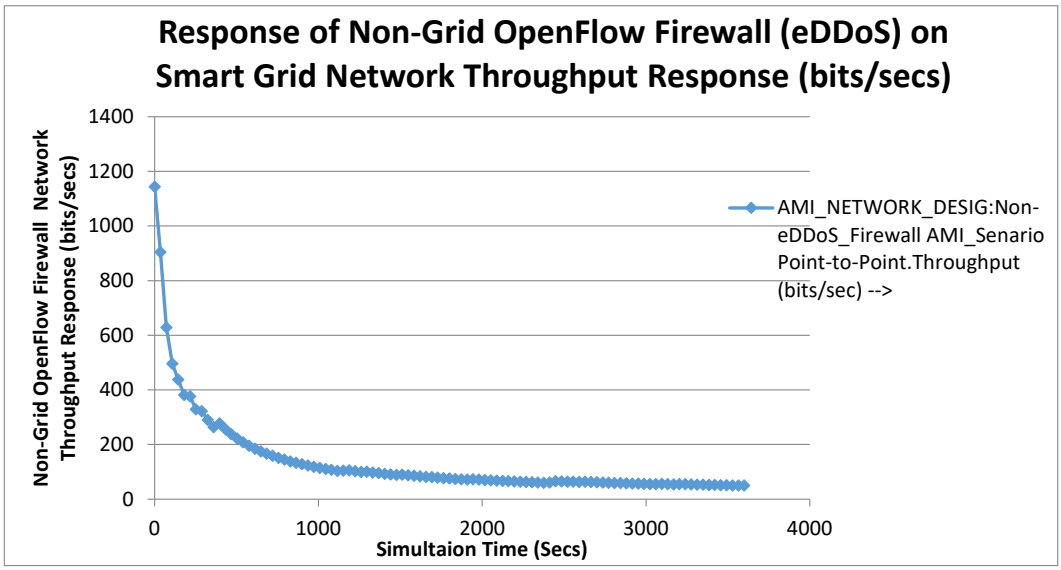

Fig. F.10: Response of non-GOF on network throughput (bits/secs)

Network Latency Response

In the smart grid AMI network infrastructure, when a DDoS attack is detected by openflow firewall algorithm, the system disconnects the victim AMI server from the network for quick remediation. Network latency is a QoS metric which expresses the delay incurred owing to queuing, congestion, and intrinsic traffic processing from the source device to the target AMI base servers. In cases involving DDoS attacks, resources such as CPU power, bandwidth, and memory will be blocked and processing time on the paths that lead to the targeted AMI server system will be affected. The main goal of the AMI openflow algorithm for DDoS defence mechanism is to detect DDoS attacks as soon as possible

and stop them as near as possible to their sources. As can be observed from Fig.F.11, the openflow dense architecture offers a lower point to point delay (0.00052 secs) than the non-openflow architecture (0.0018 secs). The lower latency is partly due to efficient resource management and openflow service processing.

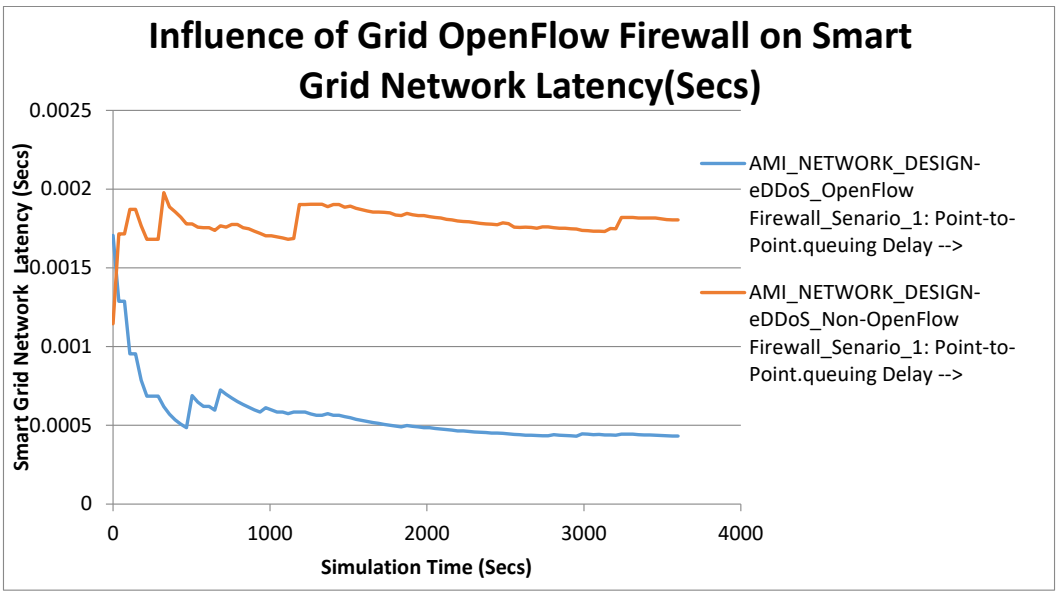

Fig. F.11: Influencene of GOF on SG AMI network latency (secs)

6 Extended Comparative Analysis

In this section a brief extended analysis of the performance analysis of the proposed firewall with two major classes of firewalls that have been proposed in the literature is carried out.The two major classes of firewalls include: IP Gateway based firewall [20], [21] and VLAN based Access Firewall [22], [23], [24]. This time around, the comparison is based on two important parameters which are availability response and overheads in the face of the DDoS attacks. By introducing a DDOS attack on the network, cryptographic overhead as well as availability responses are observed. The impact of the SG AMI OpenFlow Firewall Algorithm contributed to lower overhead and high availability response in the presence of volumetric DDoS attack against the network. The Cloud-based OpenFlow firewall was observed to have the best effort overhead at various traffic workload times compared with VLAN based Access Firewall, and IP Gateway firewall as can be observed from Fig. F.12. Similarly, availability was observed to be very satisfactory compared with other schemes as can be observed from Fig. F.13. This is feasible because of the instantiation algorithms which supports services bundling for mitigating attack vectors and its payloads. A light weighted density OpenFlow firewall will always give satisfactory cryptographic overhead against attacks while supporting more legitimate connections

and filtering illegitimate ones accordingly.

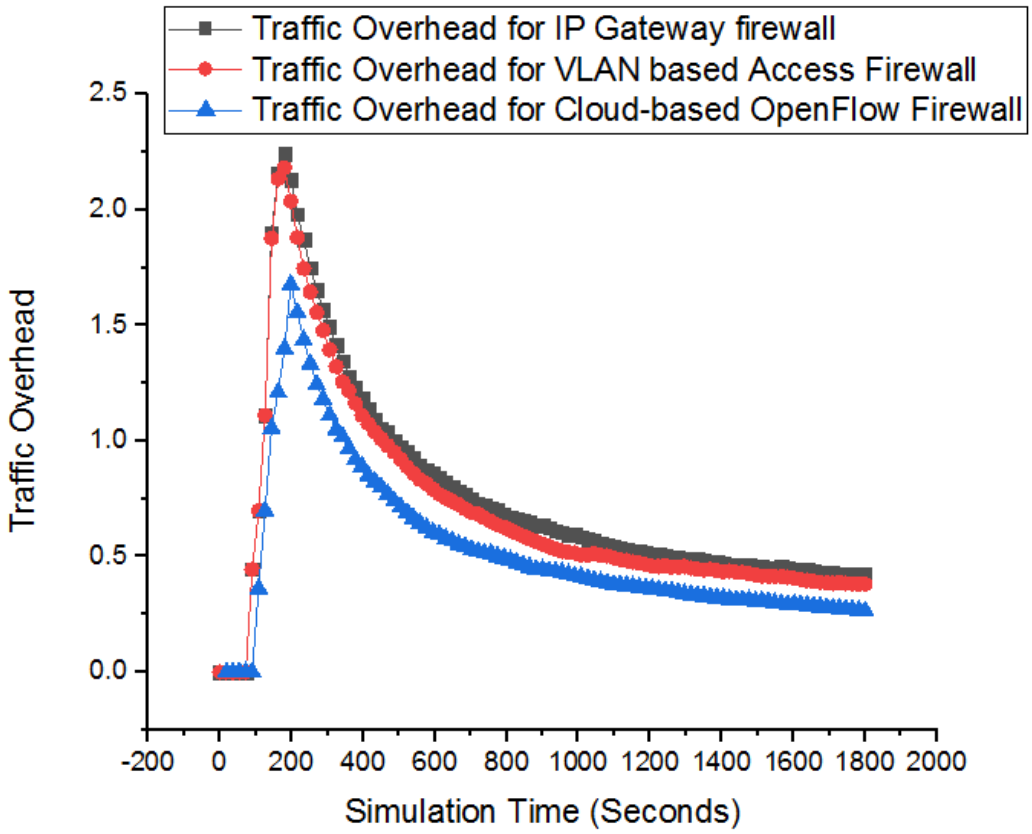

Fig. F.12: Cloud-based OpenFlow firewall with least cryptographic overhead

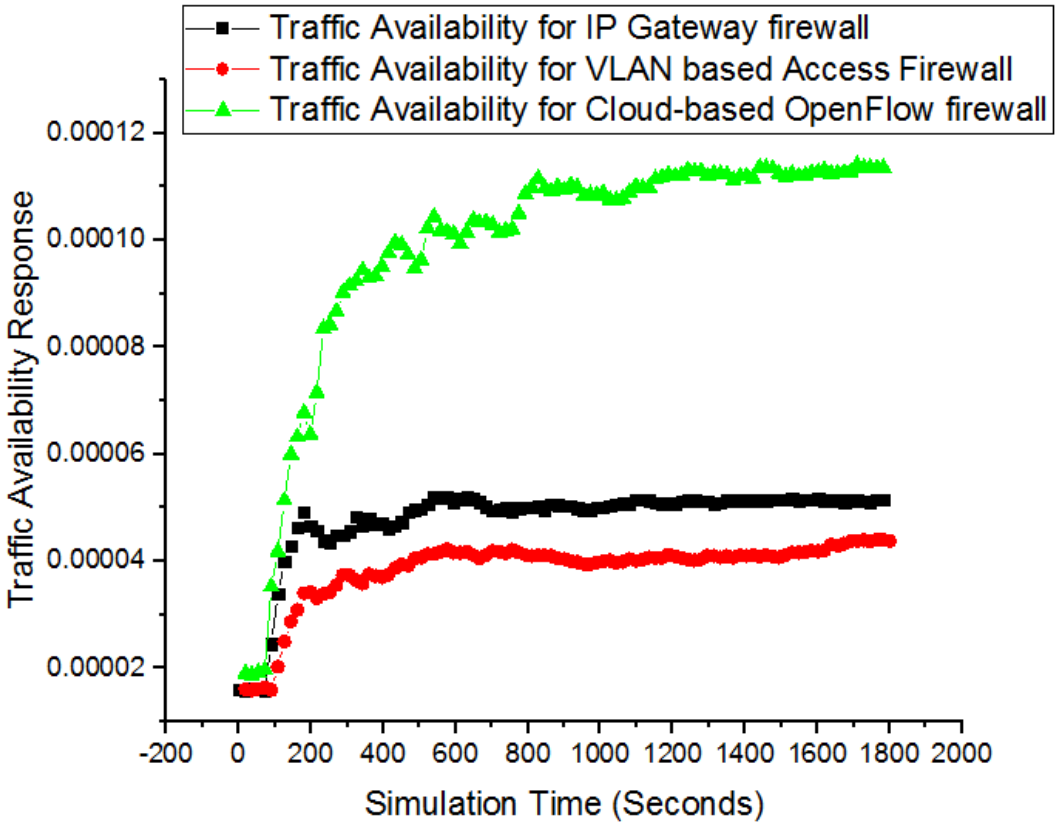

Fig. F.13: Cloud-based OpenFlow firewall with significant availability trend pattern

7 Conclusion

Volumetric attacks by cyber criminals targeting the smart grid AMI network will most likely continue to increase in the near future. This requires stronger security controls within the SG AMI network. Since DDoS attacks are evolving progressively in terms of their technology, sophistication levels and dynamics, leveraging on our proposed grid openflow will seriously mitigate the effects of such attacks. In this paper, we have shown from the analysis of the simulation results, that our proposed firewall algorithm is an improved offer over the traditional reliance on conventional gateway for server back-end protection. Our proposed grid openflow firewall for DDoS mitigation could be used to improve the quality of service (QoS) of the smart grid AMI network in an attack scenario. Our cloud-based security approach to DDoS attacks on the AMI network can guarantee an automatic, instant detection, mitigation solutions and protections against DDoS vulnerable points on the network. In the event of network expansion, it can be very scalable and can save a lot of cost as it eliminates the need for procuring on-premises security hardware.

References

[1] K. O. Aduda, W. Zeiler, G. Boxem, and T. Labeodan, "On defining information and communication technology requirements and associated challenges for 'energy and comfort active'buildings," *Procedia Computer Science*, vol. 32, pp. 979–984, 2014.

[2] M. S. Jiménez, ""european commission's on smart grids: from innovation to deployment," 2011.

[3] N. M. G. Strategy, "Acompendium of smart grid technologies," 2009.

[4] N. M. P. O. 21st Century Economy, "Advanced metering infrastructure," *US Department of Energy Office of Electricity and Energy Reliability*, 2008.

[5] 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, 2012.

[6] E. B. Rice and A. AlMajali, "Mitigating the risk of cyber attack on smart grid systems," *Procedia Computer Science*, vol. 28, pp. 575–582, 2014.

[7] P. Gallagher and N. Framework, "Roadmap for smart grid interoperability standards," *NIST Special Publication 1108R2*, 2012.

[8] H. Groenewald, "Nrs049—advanced metering infrastructure (ami) for residential and commercial customers," *ESKOM, Johannesburg, Presentation*, 2009.

[9] A. Califano, E. Dincelli, and S. Goel, "Using features of cloud computing to defend smart grid against ddos attacks," in *10th Annual symposium on information assurance (Asia 15), ALBANY*, 2015, pp. 44–50.

[10] E. Brynjolfsson, P. Hofmann, and J. Jordan, "Cloud computing and electricity: beyond the utility model," *Communications of the ACM*, vol. 53, no. 5, pp. 32–34, 2010.

[11] M. T. Khorshed, A. S. Ali, and S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing," *Future Generation computer systems*, vol. 28, no. 6, pp. 833–851, 2012.

[12] L. Zheng, S. Chen, Y. Hu, and J. He, "Applications of cloud computing in the smart grid," in *Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011 2nd International Conference on*. IEEE, 2011, pp. 203–206.

[13] E. Bou-Harb, C. Fachkha, M. Pourzandi, M. Debbabi, and C. Assi, "Communication security for smart grid distribution networks," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 42–49, 2013.

[14] Y. Guo, C.-W. Ten, S. Hu, and W. W. Weaver, "Modeling distributed denial of service attack in advanced metering infrastructure," in *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society*. IEEE, 2015, pp. 1–5.

[15] S. Asri and B. Pranggono, "Impact of distributed denial-of-service attack on advanced metering infrastructure," *Wireless Personal Communications*, vol. 83, no. 3, pp. 2211–2223, 2015.

[16] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Detection of invalid routing announcement in the internet," in *Dependable Systems and Networks, 2002. DSN 2002. Proceedings. International Conference on*. IEEE, 2002, pp. 59–68.

[17] D. Dittrich, P. Reiher, and S. Dietrich, *Internet Denial of Service: Attack and Defense Mechanisms*. Pearson Education, 2004.

[18] T. Shawly, J. Liu, N. Burow, S. Bagchi, R. Berthier, and R. B. Bobba, "A risk assessment tool for advanced metering infrastructures," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*. IEEE, 2014, pp. 989–994.

[19] K. I. Sgouras, A. D. Birda, and D. P. Labridis, "Cyber attack impact on critical smart grid infrastructures," in *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*. IEEE, 2014, pp. 1–5.

[20] S. H. M. Zarch, H. Soltani, and M. Yazdani, "Attacks simulation on computer networks by simulator," *Journal of Engineering Computers Applied Sciences(JECAS)*, vol. 3, No 6, pp. 12–17, 2014.

[21] B. Rababah, S. Zhou, and M. Bader, "Evaluation the performance of dmz," *International Journal of Wireless and Microwave Technologies*, pp. 0–13, 2018.

[22] C. Udeze, K. Okafor, C. Okezie, and O. Okeke, "Performance evaluation of openflow vlan strategy in data center switched ethernet," *African Journal of Computing ICT*, vol. 6, no. 3, pp. 63–72, 2013.

[23] S. A. Abdullah, "Simulation of virtual lans (vlans) using opnet," *Journal of Electronics and Com munication Engineering (IOSR-JECE)*, vol. 11, Issue 6, p. 67—80, 2016.

[24] A. AlQhtani, E. Aloboud, R. Altamimi, and H. Kurdi, "Impacts of vpns and firewalls on public cloud performance," in *2017 European Modelling Symposium (EMS)*. IEEE, 2017, pp. 197–201.

Chapter VIII

Enhancing Cloud-Based Smart Grid AMI Network Security by Leveraging on Quantum Key Distribution Features

Paper G

Enhancing Cloud-Based Smart Grid AMI Network Security by Leveraging on Quantum Key Distribution Features

R.C Diovu and J.T Agee

Published

Transactions on Emerging Telecommunications Technologies

Abstract

In the last decade, there have been many research proposals on the potentials of incorporating cloud computing technology to the electrical power grid as it evolves rapidly to the smart grid. The advanced metering infrastructure (AMI) is one domain of the smart grid that can be greatly impacted by cloud computing technology with regards to an improved capability for data storage and computations which will be a serious burden to the AMI in an evolved smart grid environment. In our previous study, a cloud-based OpenFlow firewall for the mitigation of distributed denial of service (DDoS) attacks in an SG AMI was proposed. In that study, it was clearly demonstrated that cloud computing technology, in addition to improving the storage and computation capability of the AMI, can also be leveraged on to enhance the security of the SG AMI in an attack scenario. In this study, we propose a cloud-based SG AMI system model utilizing quantum key distribution (QKD) cryptography for enhancing the security of a cloud-based SG AMI. The proposed model has been designed to complement the security features of the previously designed OpenFlow firewall to guarantee not only the availability but also the confidentiality and integrity of data traversing the AMI network and beyond. Results from the simulation of the QKD (BB84) protocol shows that QKD can be leveraged on to provide information theoretic security for the cloud-based SG AMI.

1 Introduction

The smart grid advanced metering infrastructure (AMI) is a key technology on which the whole of the smart grid vision is anchored on. As such, its deployment will greatly increase the reliability of the traditional grid and reduce costs of power delivery; but they present a new dependence on ICT resources which are vulnerable to numerous malicious and exploitative cyber-attacks [1, 2]. The SG AMI is expected to be connected to numerous intelligent electronic devices (IEDs) transmitting data within the AMI and beyond. It follows that the mass deployment of the AMI will obviously put more pressure on grid operators as they will experience an accelerated increase in the volume of data which are needed for billing and other grid operations like load management, meter data management and quality of supply.

There are different kinds of cyber-attacks against the SG AMI. Cyber-attacks such as data modification, falsification or injection attacks can compromise the integrity of the AMI system. Similarly, attacks like eavesdropping, masquerading, or replay attacks can compromise the confidentiality of energy consumption or grid control data in the AMI. Attacks targeting the integrity

and confidentiality of data within the AMI network can have serious consequences. On the other hand, data availability attacks like denial of service or distributed denial of service (DoS/DDoS) attacks can impact more negatively on the AMI system as they can compound issues with regards to the overall management of the voluminous data traversing the SG AMI for grid operators and other AMI back-end operators. In addition, data availability attacks will undermine the activities of the grid operators and other AMI back-end systems. For instance, billing operations will be badly affected while control decisions cannot be taken in a timely manner when data needed for these activities are maliciously delayed or denied.

It is therefore imperative that SG AMI network should be designed to be robust and resilient against attacks that compromise the main objectives of cyber security such as integrity, confidentiality, and availability. In this paper, it is believed that these three objectives of cyber-security for the SG AMI can be realized using cloud computing technology. Califano et al [3] argue that it is almost inevitable for cloud computing and the smart grid to be integrated given their distributed nature. On the contrary, non-integration of cloud computing technology with the smart grid will only compound the SG security issues. Integrating the two technologies together will present a platform for grid operators or their authorized energy suppliers to satisfy the increased computations and enormous data storage capabilities required in a fully evolved smart grid environment [4, 5]. With cloud-based computing system for the smart grid, services and applications can be hosted on a fault-tolerant and consolidated server-centric data centers with optimized QoS and security assurance mechanisms. Aside from the added advantage of increased capacity for storage and computations, cloud-based security solutions can also be designed and deployed at the right places to enhance the security of the SG AMI system.

Some previous researchers on this domain had focused on the potential contributions of cloud computing to the smart grid and/or the possible integration of smart grid applications with cloud computing [6–8]. However, some of these papers had no focus on security and they were not targeted specifically for the SG AMI. As can be seen from the review of related work in section II, some of the cloud-based security solutions proposed for the SG AMI focused mostly on the design of authentication and other cryptographic protocols. Such security measures can enhance data integrity/confidentiality in the SG AMI network. However, in deploying such security solutions for the SG AMI, caution must be applied in order not to limit the availability of data to grid operators and other AMI back-end systems who are authorized to have access to the data. This is because of the overhead on computations and communications that can be incurred as a result of such protocols and algorithms. In other words, some security solutions can either be inapplicable, incompatible or

simply inadequate in addressing the security challenges of the SG AMI.

There is, therefore, an urgent need for a cloud-based security solution for the SG AMI that would be capable of mitigating the effects of attacks against integrity, confidentiality, and availability in the AMI system. In previous paper, we designed a cloud-based OpenFlow firewall for the mitigation of distributed denial of service (DD0S) attacks against the SG AMI [9]. It was demonstrated in that paper that QoS requirements can be guaranteed in the face of 250 Gbps volumetric DDoS attack. In this paper, a more comprehensive cloud-based security solution for the SG AMI is proposed. This security solution extends the work presented in [9] by incorporating some lightweight authentication/cryptographic protocols into the proposed cloud-based architecture. In order to guarantee a quantum-safe AMI network, the incorporated protocols relied on security keys securely distributed using quantum key distribution (QKD). With this proposed security solution, the three cybersecurity objectives can be enhanced. In addition, the SG AMI will have the required capacity to deal with the burden of an anticipated increase in storage and computations requirements owing to the increased volume of data traversing the AMI. The proposed security architecture is an advancement on previous cloud-based security solutions for the SG AMI because of the following features:

- Security keys needed for authentication and other cryptographic operations are distributed using quantum key distribution (QKD) which can be utilized to provide information theoretic security (ITS).
- All authentication schemes and cryptographic protocols incorporated into the proposed solution were achieved using symmetric keys generated using the secure QKD scheme.
- Unlike similar proposals which rely on public key infrastructures, the advent of quantum computers is not a threat to the security of the protocols utilized in this paper as their security keys were generated using the QKD scheme. In other words, the security of some similar protocols (shown in section II) are dependent on their computational complexity and with the advent of quantum computers, such protocols will be broken very easily.
- Since the protocols utilized in this paper were implemented using symmetrical based cryptosystem, their overhead on computations and communication will be lighter than most similar solutions based on asymmetric cryptosystem.

This paper therefore provides the following contributions:

- The design of a cloud-based SG AMI system model using the security features of quantum key distribution. This design was done seamlessly to overcome the limitation of QKD which

conventionally allows key transmission between two parties only.

- The design incorporated lightweight authentication protocols at the upper cloud layers and a data aggregation scheme based on homomorphic encryption at the root cloud to preserve the privacy of high-frequency data needed for profiling or for other grid control purposes.
- The incorporated protocols together with the previously designed OpenFlow firewall were utilized to provide confidentiality, integrity, and availability for the cloud-based SG AMI.
- An extensive simulation of the QKD protocol was carried out to discover key factors that affect the length of the QKD generated key. The simulation done in this work was very useful since the proposed cloud-based systems depend on symmetric cryptosystem which is usually designed to offer security that is equal to the length of secret keys.

The rest of this paper is structured as follows: section II contains the review of related research work. While section III contains an overview of quantum key distribution (QKD), section IV describes the proposed cloud-based AMI system model. Section V describes the simulation design and analysis of the BB84 QKD protocol. Finally, the conclusion of the study is presented in section VI.

2 Review of related work

In this study, we focus on cloud-based security solutions for the SG AMI which are related to our study. K. Billewicz et al proposed a cloud-based architecture for the smart grid AMI [10]. In their proposed model, each smart meter will be equipped with a Wi-Fi and Power Line communication modules. In context, the smart meter would have a regular connection with the cloud-based application. Alternatively, the smart meter can be constantly left in an online mode with a connection to the cloud-based application. In the design, the smart meter can send consumption information or events and would verify that the cloud-based application does not have any requests related to the smart meter's configuration. The key idea behind this proposal is to reduce the delay associated with the smart meter connecting to the cloud service to an acceptable level. According to the authors, this proposed solution will guarantee the uploading of real-time data to the cloud and this approach increases the safety and availability of those data.

In [11], Md. Mahmud Hasan et al proposed a framework called Encryption as a Service for the smart grid AMI (ES4AM). This proposed solution was necessitated by the nature of sensitive and voluminous amount data which needed to be protected from unauthorized access or modification. Encryption is a candidate technology with primary security objective of preventing unauthorized

access to information. Unfortunately, encryption brings about an added burden to the system in the form of overheads in computations and communications respectively. Thus, moving such a service or application to the cloud is not only innovative but also rewarding. The ES4AM was designed to deliver service for the AMI at the third level of the AMI architecture where AMI concentrators aggregate and deliver aggregated data to the control center. This kind of security solution can enhance the integrity and confidentiality of AMI data.

In addition, Onoshakpor et al [12] proposed a smart grid convoluted network (SGCN). With this proposed SGCN, activities like query requests, data processing and monitoring can be implemented using fog computing layer 3 devices. The proposed network helps to mitigate the effects of cyber attacks which affect the computational requirements of smart grid applications by using Fourier Predictive Cyber Monitor (FPCM). The SGCN can guarantee some level of security against hackers and malicious malware. However, the authors did not focus specifically on the smart grid AMI network. In addition, an encryption scheme using good hash function algorithm can be designed for mitigating the effects of eavesdropping attacks on the SGCN.

In a similar development, Bashar Alohali et al [13] proposed a cloud of things (CoT) based security for smart grid Home Area Network (HAN). With the CoT security-based solution, a collection of appliances that will use real-time data can be enabled. The CoT is a virtualized internet of things (IoT) which has the capacity of providing monitoring and control. The paper illustrates how the security requirements of the network are administered from the CoT infrastructure. The security solution provided in this paper was designed for the prevention or mitigation of data confidentiality, availability and integrity attacks. The authors also included in their proposal, symmetric key cryptography with a secure key management scheme for achieving confidentiality and integrity during end-to-end communications. The performance analysis of the security solution against confidentiality and integrity attacks carried out in this paper is commendable. However, the authors could not show clearly how their proposed solution could be used to prevent or mitigate data availability attacks as claimed in their paper.

Also, in [14], Neetesh Saxena et al proposed a lightweight cloud-based authentication protocol for providing security solutions among HAN, energy providers, gateways, and the AMI. This cloud-based solution provides distributed authentication services among smart grid communication entities. Specifically, the cloud-based solution provides authentication between the following pair of entities: energy providers and the smart meters, smart meters and the HAN gateway, BAN gateway and NAN gateway, and NAN gateway and the cloud computing infrastructure. The security solution proposed by the authors uses a hierarchical cloud-based trusted authority (TAs) for key distribution

and management based on public key infrastructure (PKI). In this proposal, a cloud-based authentication security solution was utilized for mitigating data integrity and confidentiality attacks for the SG AMI network. However, it may not be very effective in combating data availability attacks as claimed by the authors. Thus, authentication protocols only may not be sufficient for combating data availability attacks. This drawback was resolved in our proposed paper by incorporating our previously designed OpenFlow firewall in other to make the whole security solution more comprehensive. Another drawback in [14] is the reliance on public key infrastructure for the design of the authentication protocols. Ideally, most security solutions based on PKI rely on the mathematical complexity of their encryption algorithms; as such, would take a very long time for their security to be broken. Unfortunately, the speed at which quantum algorithms can solve such complex problems possesses a severe challenge to the above assumption. On the other hand, symmetric ciphers (such as AES) are considered to be quantum safe because they have the ability to adapt to a quantum attack just by increasing the key size in order to rectify the vulnerability that may have been introduced by the quantum algorithm.

3 Overview of Quantum Key Distribution (QKD)

Quantum key distribution is an evolving security technology that can be exploited positively to achieve information-theoretic security (ITS) [15]. QKD can be utilized for the generation and distribution of shared secret key between two parties conventionally referred to as Alice and Bob. The security of quantum key distribution is primarily based on the fundamental principles of quantum mechanics. A basic QKD protocol model involving two parties has been shown in Fig. G.1. As can be observed from Fig. G.1, the key exchange between the two parties can be completed using both classical and quantum channels. In the QKD protocol model, no limit is placed on the computational power of an adversary (eavesdropper). However, fundamental features of QKD ensures that any eavesdropping attempt on the key exchange process will introduce detectable errors. An unconditionally secure key can be transmitted if the errors introduced by the eavesdropper is below a defined threshold. Otherwise, the protocol is aborted, and the process re-started.

As of today, there are many variants of the QKD protocol. The first version of the QKD protocol was proposed in 1984 by Charles Bennett and Gilles Brassard [16]. This protocol which is simply referred to as BB84 (named after author's names and the year of publication) is based on the Heisenberg's uncertainty principle. According to this principle, only one property of a pair of conjugate properties in a quantum system can be measured with certainty. Another QKD protocol known as Eckert's protocol

[17] is based on the principle of quantum entanglement. According to this principle, it is possible for two quantum properties to become entangled in such a way that a particular property can be measured on one particle while observing the opposite state on an entangled particle instantaneously. The QKD protocol utilized in this research work is the BB84 protocol which is described in further details in the next sub-section.

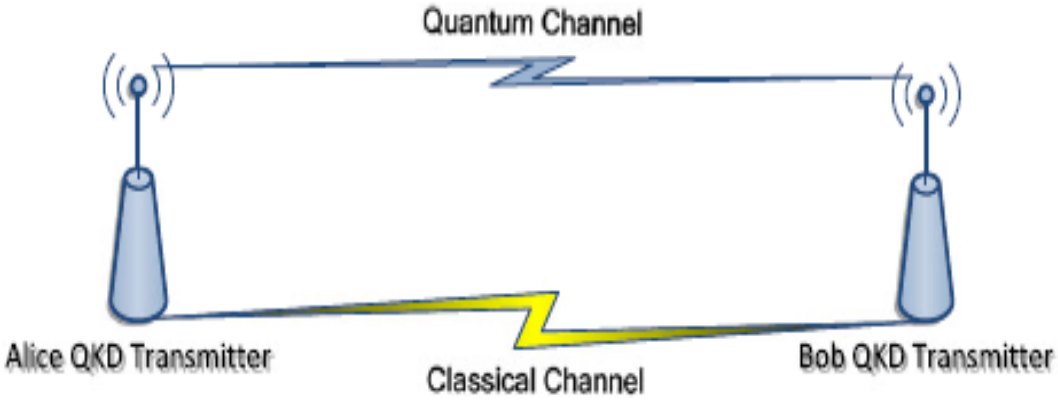

Fig. G.1: Structure of a QKD link

BB84 Protocol

According to BB84 protocol, the sending party (Alice) prepares a sequence of qubits and sends them to the receiving party (Bob) over a quantum channel. The sending party then randomly chooses a basis for each qubit with either rectilinear ($\oplus$) or diagonal polarization ($\otimes$). A qubit can be represented in either of the following quantum states [18]:

$$|\psi(0,\oplus)\rangle = |0\rangle$$

$$|\psi(1,\oplus)\rangle = |1\rangle$$

$$|\psi(0,\otimes)\rangle = |+\rangle$$

$$|\psi(1,\otimes)\rangle = |-\rangle$$

Fig. G.2 also shows how a qubit can be encoded in the polarization of a photon in the BB84 protocol. The receiving party then randomly selects a basis for the measurement of the received polarized photons. If his selected basis matches with the sending party's basis, he will measure the encoded bit with 100% accuracy. If on the other hand, their basis does not match, he can still measure the encoded bit with 50% probability. This is possible because a photon polarized in one basis yields a random value when the same photon is measured in its conjugate basis. The idea behind encoding of classical bits in two conjugate bases in the BB84 is that it makes it possible for keys to be transmitted

between two parties such that any eavesdropper would destroy information, and thus, gets detected. While the quantum channel is used for the exchange of quantum encoded qubits, the classical channel is used for error reconciliation during the key exchange process.

Other important steps in the BB84 protocol include: sifting, error estimation and reconciliation, and privacy amplification. During sifting stage, the receiving party announces on a public classical channel the qubits which he managed to measure successfully. The sending party and receiving parties then reveal and exchange the bases which they had used. At this point, the bits corresponding to photons measured in different bases between Alice and Bob are discarded. The two parties should now have identical string of bits which is called the sifted key. In the error estimation and reconciliation process, the two parties perform error correction scheme using the public classical channel in order to locate and correct erroneous bits in their sifted bit strings. It can be noted at this point that an eavesdropper can gain partial information about the sifted key during the reconciliation process. Privacy amplification is therefore employed to reduce and eliminate effectively an eavesdropper's knowledge of the sifted key to an arbitrary low value. This can be achieved by using a universal hash function chosen at random from a publicly known family of hash functions. In the BB84 simulation carried out in this paper, a universal hashing scheme based on Toeplitz matrices is employed. In QKD post-processing, Toeplitz matrices can be used for three purposes: error verification, authentication and privacy amplification. Toeplitz matrices are Boolean matrices with special structure that can be given as:

$$M = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots & a_{-m+1} \\ a_1 & a_0 & a_{-1} & \ddots & \\ a_2 & a_1 & \ddots & & \vdots \\ \vdots & \ddots & & & \\ a_{l-1} & & \dots & & a_{l-m} \end{pmatrix} \tag{G.1}$$

where $a_i = 0, 1; M_{i,j} = a_{i,j}$ and $M_{i,j}$ is the (i, j) element of M [19]. It is to be noted from here that fully random Toeplitz matrices specified by $m + l - 1$ random bits can be used for privacy amplification while Toeplitz matrices specified by smaller number $(2l)$ of random bits can be used for error verification and authentication. In the QKD BB84 protocol simulation carried out in this paper, sifting authentication was done using Linear Feedback Shift Register (LFSR) universal hashing based on Toeplitz matrices. The LFSR-based Toeplitz is derived based on the universal hashing construction proposed by Krawczyk [20].

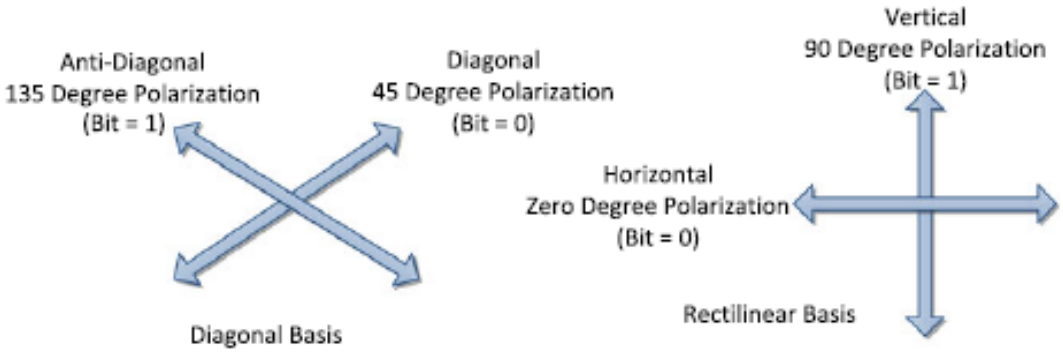

Fig. G.2: Illustration of qubits encoding using different polarization states [15]

4 Proposed System Model

A. Overview

The proposed model shown in Fig. G.3 is composed of a hierarchical cloud structure [14] with a key distribution scheme that can enhance greatly the security and privacy of consumption data or control information traversing the AMI network. This hierarchical cloud-based AMI network is made up of:

- Root cloud: The root cloud comprises the smart meters, In-Home-Displays (IHDs), intelligent electronic devices (IEDs), root cloud key management server, and data aggregators which also act as gateways for the Home Area Networks connected to the root cloud.
- Regional cloud: The regional cloud is made up of the Neighborhood Area Network Gateways (NAN GWs), and the regional key management server.
- Top cloud: The top cloud comprises the top cloud server, the OpenFlow firewall, and AMI master station which is considered in this model as a trusted third party. The OpenFlow firewall has been incorporated in this model to help in the detection and mitigation of distributed denial of service (DDoS) attacks against the AMI network. The OpenFlow firewall can also be used in filtering illegitimate traffics moving from the lower clouds to the top cloud.
- Communication channels: As can be seen in Fig. G.3, the proposed system model comprises different types of communication channels. There are two kinds of quantum channels incorporated in this model. The blue thick line represents quantum channels utilized for a quantum key generation for communication between the root cloud key management servers and the regional key management servers. In the case of communication between the regional

clouds and the top cloud, messages are to be encrypted with QKD key generated using quantum channels represented by black thin dotted lines. In addition to quantum channels that are used for QKD key generation, the model also includes classical channels (public channels) which are also needed for completing the QKD key generation protocol. Messages encrypted with desired encryption keys can be transmitted using the classical channels.

B. Threat Model

The threat landscape for the smart grid advanced metering network can be large, however, the proposed system model is designed to address the following threats:

- Privacy violations of end user's consumption data: The privacy of end user's consumption can be hijacked by cybercriminals and maliciously utilized to launch various degrees of cyber-attacks. This kind of violation can happen easily when end user's consumption data are collected at high frequency.
- Breaches from authorized third parties: Authorized third parties otherwise referred to as AMI-back ends may decide to violate the confidence reposed in them for pecuniary gains. For instance, if an authorized energy supply obtains by malicious means the energy consumption data of consumers from other suppliers, the supplier can use this information to their advantage in an open liberalized energy market.
- Denial of service/distributed denial of service (DoS/DDoS) attacks: Cybercriminals can target the availability of end user's consumption data or control information. In this context, the operations of AMI back-end system responsible for billing will be seriously undermined. In addition, timely control decisions could be affected, a situation, which can lead to the entire system failure.
- Replay attack: An experienced cybercriminal can intercept message (consumption data or control information) from Home Area Network (HAN) or their associated gateways (in this system model, the data aggregators) and then replay those messages to the NAN GWs.
- Message Injection Attack: As in the case of replay attack, the adversary can inject a fake message into an intercepted message, and then send it to the NAN GWs.

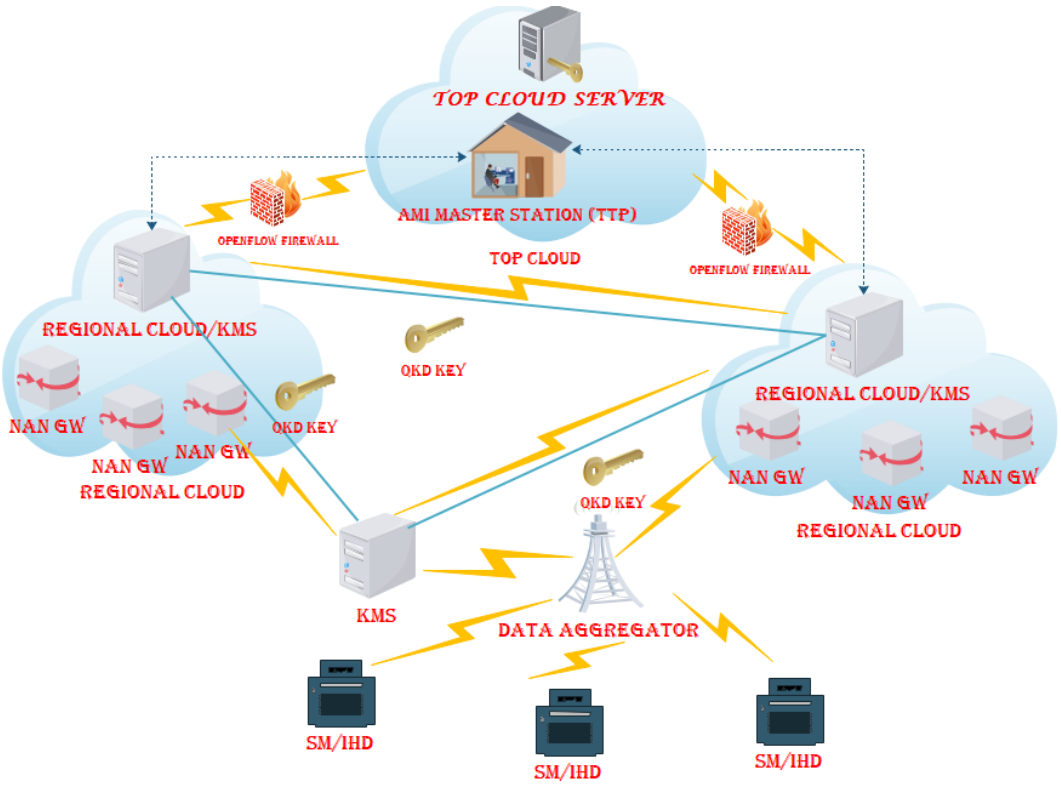

Fig. G.3: A Cloud-Based SG AMI Network Model using Quantum Key Distribution System

C. Design Goals

The system model proposed in this paper have the following design goals:

- Message Authenticity/Integrity: The recipient of a message should be assured that the message has not been altered during transit, is fresh and is indeed from the claimed source.
- Confidentiality: End-users' consumption or metering data should not be disclosed to unauthorized entities.
- Data Availability: End user's consumption data or control information should be made available to AMI-back ends as and when needed.
- Authorization: Authorized AMI-back end entities and third parties should have access only to the metering data that meets the requirements for the fulfillment of their contractual obligations. As a result of this requirement, the AMI master station is assumed to be fully knowledgeable about the needs of the authorized entities and they should be granted access rights to the end user's consumption data based on those needs.
- Low Computation and Communication Overheads: The key distribution and authentication

schemes in this proposed work were designed to be lightweight. As a result, this paper relies mostly on the use of symmetric cryptographic algorithms for these protocols.

D. Cryptographic Building Blocks utilized in the proposed Model

In this section, the building blocks utilized in this proposed system model is briefly presented. They include:

Homomorphic Encryption

Homomorphic cryptosystem possess properties which can perform a set of operations on the ciphertext without revealing the encrypted message (plaintext). This is done in such a way that when the ciphertext is decrypted, the decrypted value will be equal to the plaintext that was obtained when the same set of operations are performed on the plaintext. Homomorphic encryption has two basic operations which are addition and multiplication. More formally, homomorphic encryption on addition and multiplication operations are defined below:

- Additive Homomorphic Encryption: This homomorphic property deals with the addition of encrypted data.

  Definition 1: Let's consider m_1 and m_2 to be two plaintext messages which are encrypted to $Enc(m_1)$ and $Enc(m_2)$ respectively. It then follows that the application of additive homomorphic encryption on the encrypted plaintexts would yield:

  $$Enc(m_1).Enc(m_2) = Enc(m_1 + m_2) \tag{G.2}$$

  In other words, the multiplication of the encrypted messages results in the ciphertext of the sum of the messages.

- Multiplicative Homomorphic Encryption: This property deals with the multiplication of encrypted plaintext messages.

  Definition 2: Let's consider two plaintext messages, m_1 and m_2 which are encrypted to $Enc(m_1)$ and $Enc(m_2)$ respectively.

  It also follows that the application of homomorphic encryption on the encrypted messages would yield:

  $$Enc(m_1).Enc(m_2) = Enc(m_1.m_2) \tag{G.3}$$

Claude Castellucci Symmetric Homomorphic Encryption [21]

Unlike most homomorphic encryption schemes which use asymmetric keys, Claude Castellucci's homomorphic encryption scheme uses a symmetric key for encryption. The major steps in the encryption and decryption process in this scheme are provided below:

- Encryption

  Step 1: Represent plaintext message as integer, $m \in [0, M-1]$ where M is large integer.

  Step 2: Let k be a keystream generated randomly such that $k \in [0, M-1]$.

  Step 3: Compute $C = Enc(m, k, M) = m + k(ModM)$.

- Decryption

  Step 1: $Dec(c, k, M) = C - K(ModM)$.

  Addition of ciphertexts

  Step 1: $Letc_1 = Enc(m_1, k_1, M)$ and $c_2 = Enc(m_2, k_2, M)$.

  Step 2: $Fork = k_1 + k_2, Dec(c_1 + c_2, k, M) = m_+m_2$.

Hashed Message Authentication Codes (HMAC)

The hashed message authentication code (HMAC) is a key-dependent one-way hash function which is used to provide message integrity and origin authentication for the message transmitted between two communicating entities. The main input parameters for HMAC algorithm are the message and the secret key which is known between the message sender and the intended receiver. The HMAC is normally created by concatenating the message with a secret key and hashing the result with a cryptographic hash function. The standard for the creation of HMAC has been defined in RFC2104 [22]. Given a message, m, the secret key, k, the HMAC for the message can be given by:

$$HMAC((K_0, m) = H(k_0 \oplus opad)||H((k_0 \oplus ipad)||m)) \tag{G.4}$$

where H is the cryptographic hash function, $ipad$ and $opad$ are distinct padding constants while $||$ and $\oplus$ are the concatenation and Exclusive-OR operators respectively.

One-Time-Pad Encryption In the one-time-pad encryption scheme, communicating parties share a random bit string, as large as any message they wish to transmit. The string, is the symmetric key.

In [23], it was shown that in order to compute the ciphertext from the message, the sending party computes:

$Sending\ \ party \rightarrow c \oplus k$

Since

$$c \otimes k = (M \otimes k) \otimes k \tag{G.5}$$

$$= M \otimes (K \otimes K)$$

$$= M \otimes 0$$

$$= M$$

Where 0 denotes the bit string of all 0's in the same length as M.

E. System Model Details

This section provides the details on important communications and cryptographic computations included in this scheme which are needed for enhancing the security and privacy of the AMI network. Table G.1 shows the meaning of important notations used in the proposed model. Without loss of generality, we assume that data should be released by the AMI Master station to the different AMI back-end systems (such as billing system) depending on the functions that they perform. In other words, high-frequency data ought not to be sent to billing system since low-frequency data is sufficient for them to perform their duties. Other interesting details have been presented under the following sub-headings:

Message Upload from Root Cloud to the Upper Cloud

Case 1: Data needed for billing

This kind of data can be collected at a very low frequency [24] such as once per month. From a privacy-preserving perspective, this kind of data does not pose serious privacy threats to the end users. In this paper, the use of Trusted Platform Module (TPM) built into the metering system [25, 26] is recommended. Isolating the calculation of customer's bill within the TPM would ensure the integrity and confidentiality of the system. However, for this to be realistic, the TPM must be assumed to be secure. Additionally, the trusted computing software to be implemented for the TPM must be verified by the AMI master station which should be an unbiased umpire in ensuring that the codes utilized in

Table G.1: Basic Notations and their definitions

Notations	Definition
K_{QKD}	Symmetric secret key by QKD BB84 Protocol
m_i	Metering data from smart meters
ID_i	Identity of smart meters, i= 1, 2, ...
$\|\|$	Concatenation operator
C_{m_i}	Encrypted metering data
AGG Value	Aggregated value of metering data
NAN_{GW}	Neighborhood Area Network Gateway
$KMS_{REGIONAL\_CLOUD}$	Key management server for the regional cloud
M_i	Time-based message from NAN_{GW}
$HMACK_{QKD}$	A keyed-hash based message authentication function using K_{QKD}
TS^{t_n}	Time stamp
$\otimes$	Exclusive-OR operator
DATA AGG	Aggregated function used for concatenating the ciphertexts of metering data with smart meter identities

the design of the TPM do not violate the privacy of end users. This category of data can be uploaded to the AMI master station without passing through the process of data aggregation.

Case 2: Data needed for profiling or other grid control purposes

In this second case study, data collected at high frequency should be aggregated before they are uploaded to the upper cloud in order to preserve the privacy of energy consumers. The data aggregation introduced in this proposed system model uses a variant of homomorphic cryptosystem proposed by Claude Castelluccia et al [21]. In their work, data aggregation was performed using enhanced homomorphic encryption. Their enhanced homomorphic encryption scheme was peculiar in two ways. Firstly, the exclusive-OR operation usually utilized in most cryptosystem was replaced with modular addition operator. Secondly, unlike most systems that use asymmetric keys, their homomorphic encryption scheme was implemented with a symmetric master key shared between the source nodes and the sink nodes This encryption was proven to be secure. The difference between their scheme and the one proposed in this paper is the method of symmetric key generation. In our scheme, the symmetric key used for the homomorphic encryption is generated by running the BB84

QKD protocol between the AMI master station and a given root cloud KMS server. In addition, an identity-based attribute from the smart meters was incorporated into the aggregation value to facilitate recoverability of metering data at the target node. The generated K_{QKD} key is then shared between the root cloud server and the smart meters connected within its domain. To preserve the integrity and confidentiality of the consumption data, the generated K_{QKD} key should not be shared with the data aggregator. This is important since homomorphic encryption makes it possible for encrypted metering data to be aggregated by the data aggregator without first decrypting the encrypted metering data. This will help to resist data injection or modification attacks at the root cloud. This requirement will be particularly important in the event that the root cloud is serviced by a public cloud provider. The details of important processes leading to a privacy-preserving data aggregation have been provided below:

- Every smart meter encrypts its metering data m_i, to obtain the corresponding ciphertext. For i_{th} number of smart meter nodes connected to a given aggregator, the symmetric encryption key, K_{QKD} is generated. The generated key is shared with the closest NAN_{GW} to the data aggregator. The identity of each smart meter, ID_i, is concatenated with ciphertext computed in eqn. G.6 and shown in eqn.G.7:

$$C_{m_i} = Enc(m_i, K_{QKD}, M) \tag{G.6}$$

$$ID_i||C_{m_i} \rightarrow DATA\ \ AGG \tag{G.7}$$

  Where M is a message space (set of all possible messages) which needs to be large enough to prevent overflow.

- The smart meter then forwards C_{m_i} to the aggregator which then aggregates C_{m_i} for all smart meters within its domain by performing the addition of modulo M. The aggregated value will then be forwarded to the nearest NAN_{GW}:

$$AGG\ \ Value \rightarrow NAN_{GW} : C_{m_i} = \sum_{ID-1}^{n} C_{m_i}(mod\ \ M) \tag{G.8}$$

  where n is the number of ciphers added.

- Given C_{m_i} and K_{QKD}, m_i can be recovered from the decryption algorithm given in eqn. G.9.

$$m_i = Dec(C_{m_i}, K_{QKD}, M) \tag{G.9}$$

  Using the smart meters identities, ID_i, the respective metering data can be attributed to the smart meters by the NAN_{GW}. For onward transmission or upload of data from NAN_{GW} to upper cloud, additional security measures are then applied by the NAN_{GW}.

Upload of Message from Regional Cloud to the Top Cloud

Upload of Message from Regional Cloud to the Top Cloud: In this section, a description of how message received from the root cloud is uploaded to the top cloud is provided. Included in this section are explanations on cryptographic computations associated with authentication and key distributions utilized in this message upload process. This description begins from the point when consumption data or control information is received by the NAN_{GW_s} located at the regional cloud. Important details with regards to this section are provided below:

- Key Generation: From the system model shown in Fig. G.3, it is assumed the key used for providing information theoretic security (ITS) is generated by executing the BB84 quantum key distribution protocol between two regional cloud servers or one regional cloud server and a root cloud server [16]. The QKD generated key, (K_{QKD}) will then be used for one-time-pad encryption.

- Message authentication: It is noteworthy to point out that the authentication that is performed during the execution of the BB84 protocol is crucial for providing security for some or all of the classical messages exchanged during the public discussion during the protocol execution. In addition to this authentication protocol, our scheme incorporates another authentication which is executed between any regional cloud server that wishes to transmit message to the AMI master station. This very process is very crucial because the AMI master station is responsible for relaying these data to the AMI back-end systems who need them for billing and other control operations. We assume in this model that the regional cloud server (KMS) has unrestricted access to all the NAN_{GW_s} in its domain while the NAN_{GW_s} have authenticated access to KMS server within their domain. An illustration of message transmission with appropriate authentication from NAN_{GW} to $KMS_{REGIONAL\_CLOUD}$ and then to the AMI _Master _Station is provided below and summarized in Figs. G.4 and G.5 respectively:

  - Transmission of M_i from NAN_{GW} to an associated regional cloud server, $KMS_{REGIONALCLOUD}$: To ensure the integrity of the consumption data/control information (simply referred to as message), the NAN_{GW} generates message authentication code by using the generated K_{QKD} key. The time-based message, M_i containing the consumption data from smart meters or control information for a given period of time is passed to $HMACK_{QKD}$ which then generates message authentication code for the message, M_i. The NAN_{GW} then concatenates this message, M_i, time stamp, TS^{t_n} and $HMACK_{QKD}$, encrypts the whole packet by using AES algorithm.

The NAN_{GW} then sends this encrypted packet to the $KMS_{REGIONAL\_CLOUD}$. It is to be noted that the time stamp, TS^{t_n}, was incorporated to the above process in order to frustrate a cyber-criminal with the intension of launching a replay attack on the message, M_i. The whole process is represented below:

$$NAN_{GW} \rightarrow KMS_{REGIONAL\_CLOUD} : Enc(M_i||TS^{t_n}||HMACK_{QKD}(M_i)) \tag{G.10}$$

When the $KMS_{REGIONAL\_CLOUD}$ receives the encrypted packet from the NAN_{GW}, it first decrypts the packet using the K_{QKD} key in the process represented below:

$$Dec(M_i||TS^{t_n}||HMACK_{QKD}(M_i)) \tag{G.11}$$

The $KMS_{REGIONAL\_CLOUD}$ will then check the freshness of the time stamp, and if the stamp is fresh, the $KMS_{REGIONAL\_CLOUD}$ verifies the integrity of the message, by calculating $HMACK_{QKD}$ of the message and then forwards the authenticated message to the AMI master station. The session is aborted if $KMS_{REGIONAL\_CLOUD}$ is unable to verify the freshness of the time stamp.

- One-Time-Pad Encryption: After $KMS_{REGIONAL\_CLOUD}$ had successfully decrypted the message received from NAN_{GW_s} in the process described above, it performs one-time-pad encryption before uploading the decrypted message to the AMI master station, AMI_Master_Stn. One-time-pad is preferred here because it remains the only encryption scheme for which ITS can be proven. It can be noted that using keys generated by QKD to perform one-time-pad encryption will lead to an unconditionally secure message transmission protocol [27]. The K_{QKD} generated key, the decrypted Message, M_i, and the ciphertext are the basic components of the one-time-pad encryption process. Given the K_{QKD} and M_i the ciphertext computation and recovery of M_i in the one-time-pad process is given below:

$$Ciphertext : C = M_i \otimes K_{QKD} \tag{G.12}$$

$$Message : M_i = C \otimes K_{QKD} \tag{G.13}$$

F. Security and Privacy Analysis

In this section, analysis of the security and privacy properties of our proposed model is presented. Following the design goals presented in section IV of this paper, this section focuses on how the proposed model can be used to enhance the privacy, integrity, confidentiality, and availability of critical information within the SG AMI.

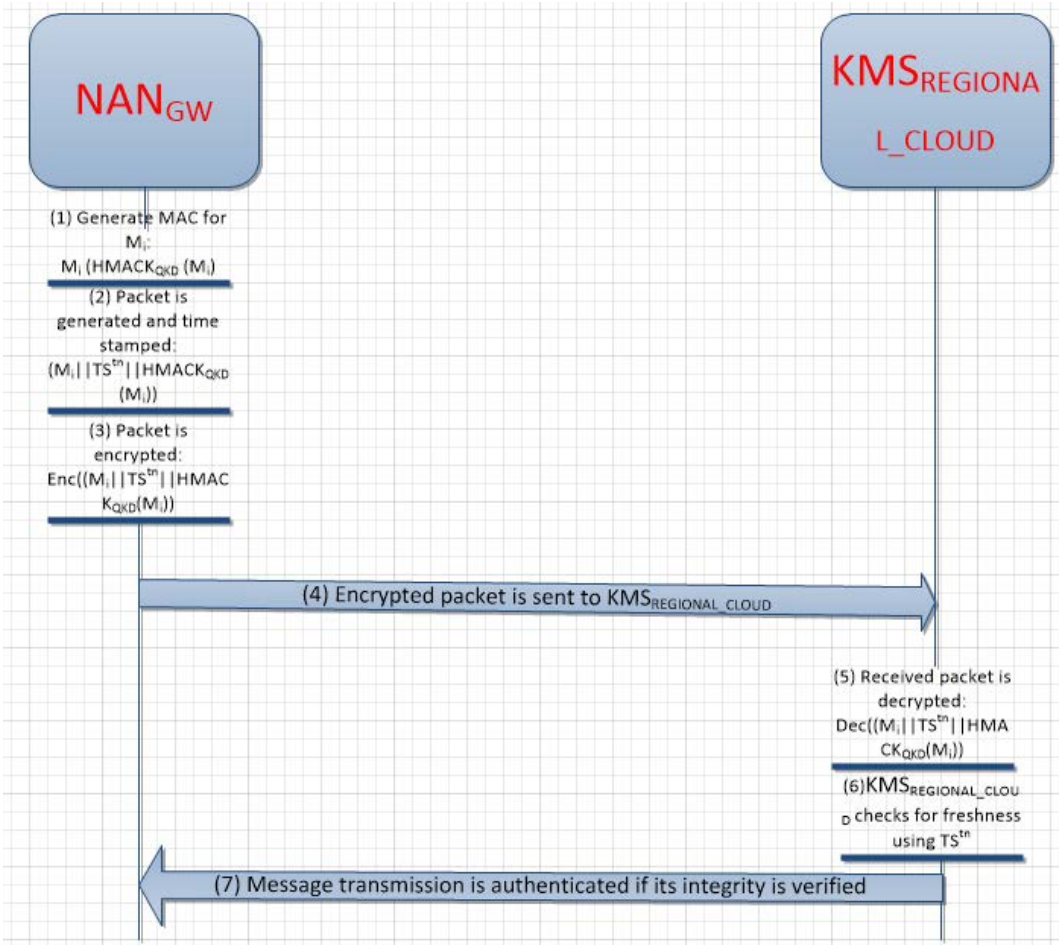

Fig. G.4: Trasmission and Authentication between NAN_{GW} and $KMS_{REGIONAL\_SERVER}$

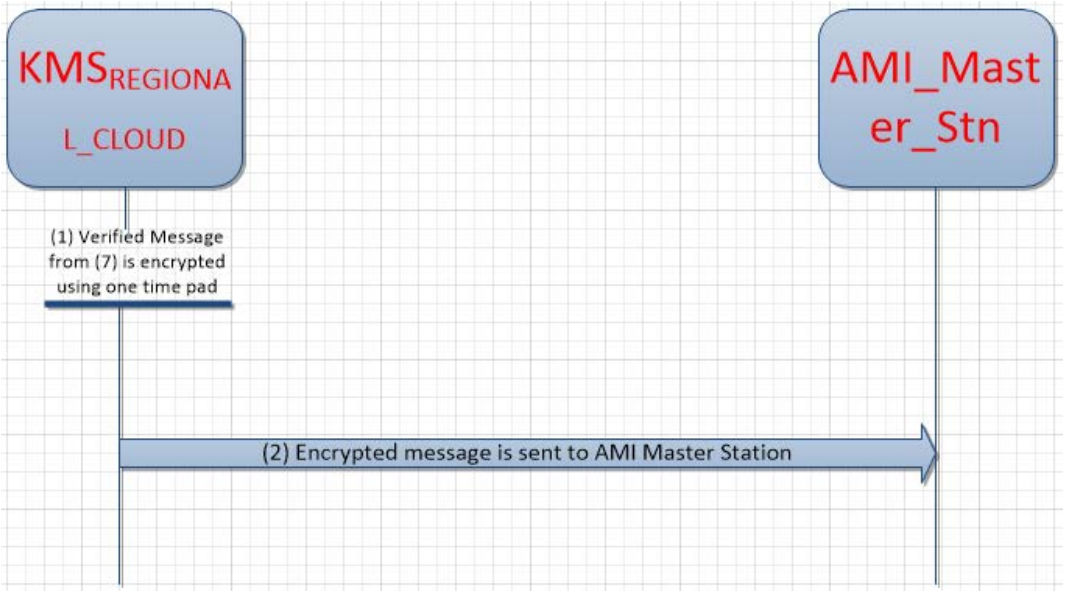

Fig. G.5: Message Transmission between $KMS_{REGIONAL\_SERVER}$ and AMI_Master_Stn

Privacy

To preserve the privacy of metering data which can be exploited to the detriment of consumers, both low and high-frequency data are never sent directly to the AMI back-end systems. While low-frequency data may be sent without being aggregated, high-frequency data must be aggregated before being uploaded to the AMI master station. These high-frequency data are aggregated using the security properties of homomorphic encryption such that even the data aggregators at the root cloud are unable to decrypt the aggregated data.

Integrity and confidentiality

To ensure that integrity and confidentiality of critical information are enhanced in the proposed model, important security measures are enforced on the data before they are uploaded from the regional cloud to the upper cloud. Firstly,

The regional cloud server authenticates the NAN_{GW} from the same regional cloud which is in reception of an aggregated metering data from a data aggregator at the root cloud. This is done to ensure that the $KMS_{REGIONAL\_CLOUD}$ is dealing with an entity within its regional domain. Another security measure ensures that the decrypted aggregated metering data is time stamped and then passed through a keyed-hash based authentication function $(HMACK_{QKD})$ using the QKD generated symmetrical key before encrypting the resultant packet using an AES algorithm. Finally, the encrypted packet when received by the $KMS_{REGIONAL\_CLOUD}$ is further encrypted by passing it through a one-time pad encryption algorithm in order to provide information theoretic security for the information. The time stamp was embedded in the encrypted packet so that each of the communicating party can verify the freshness of the time stamp to check whether it is the same time stamp present in the encrypted packet. The above security measures combine to ensure that data modification and injection attacks are frustrated. In addition, the way and manner the QKD key is generated and distributed makes it difficult for attacks like eavesdropping and replay attacks to succeed.

Availability

The authentication protocols incorporated into the proposed model can be useful in frustrating cyber-attacks targeting the critical information in the AMI network. In other words, malicious messages being transmitted by unauthorized nodes within the network will eventually be detected by the key

management servers at the different cloud levels since these servers are assumed to have unrestricted access to nodes within their domain. Such messages, if detected, will be discarded, thus, frustrating the efforts of such attacks. In the realization that there are many cyber-criminals who have the skills of by-passing many authentication protocols and access rules in wireless networks, the previously designed openFlow firewall was also incorporated into the proposed model to complement in the mitigation of data availability attacks against the network.

5 Performance Evaluation

In this section, an analysis of the communication and computation costs of the protocols incorporated into the proposed system model is provided. These two parameters can be very crucial in determining the performance of any smart grid communication. As such, the communication and computation costs of protocols designed for the smart grid AMI ought to be as low as possible. On the contrary, protocols with high costs of computation and communication can lead to a congestive network scenario which can make data availability attacks against the network to thrive. In this analysis, we compare our protocol with an integrated distributed authentication protocol for smart grid communication proposed by Neetesh et al [14] and a novel Identity-Based key establishment (NIKE) protocol for SG AMI proposed by Amin M et al [28]. The NIKE protocol proposed in [28] has another variant tagged NIKE+. It is to be noted that NIKE and NIKE+ are equivalent security-wise. However, they differ in the manner of their construction which explains the differences in the communication costs especially at the different domains of the network. Table G.2 shows the symbols used in this analysis. On the other hand, table G.3 shows the comparison of the computation costs of the proposed protocol with that proposed by Neetesh et al and Amin M et al. It is to be noted that the total computation costs of NIKE and NIKE+ are equal. This is the main reason why NIKE and NIKE+ were not presented differently in table G.3. It can be seen clearly from table G.3 that the protocol proposed by Amin M et al has a lower computation cost than our proposed protocol. However, it is easy to observe from the same table that the protocol proposed by Amin M et al had fewer entities involved in their protocol than our own protocol. For the protocol proposed by Amin M et al, the entities involved includes smart meters and AMI-head end (AHE). In contrast, the entities involved in our protocol includes smart meters, home appliances (like In-home displays, intelligent electronic devices), home area network gateways (HAN GW), trusted third party (TTP), neighborhood area network gateways (NAN GW), and regional key management servers (RKMS). In all fairness, the computation cost of our proposed protocol can truly be compared with that proposed by Neetesh et al because they have almost the same number of entities as can be seen in table G.3. In

that regard, it is easy to discover from table G.3 that the computation cost of our proposed protocol is far less than the protocol proposed by Neetesh et al. This is as a result of so many operations involved in the protocol proposed by Neetesh et al which are majorly dependent on public key cryptography that are generally adjudged to be computationally intensive. On the contrary, the number of operations involved in our protocol are far less than that proposed by Neetesh et al. In addition, the operations utilized in our proposed protocol are mainly symmetric based cryptographic operations which are less computationally intensive than that of public key based cryptographic operations. In the case of communication cost, NIKE and NIKE+ were treated differently because their total communication costs are different. The communication cost in bits of the different operations utilized in the analysis for the proposed protocol has been presented in table G.4. It is to be noted from table G.4 that the communication cost for time stamping operation is equal to the cost for an entity identity operation, ie, $CC_{TS} = CC_{ID} = 32$ bits. Similarly, the cost of computing the aggregation function is 64 bits [29]. Before presenting the comparison of the communication costs of our protocol and other protocols considered in this paper, a detailed calculation of the communication cost of our protocol has been broken down into different communication domains which include SM to HAN_{GW}, HAN_{GW} to NAN_{GW}, NAN_{GW} to $KMS_{REGIONAL-CLOUD}$ and $KMS_{REGIONAL-CLOUD}$ to $AMI_{MASTER-STATION}$ (TTP). The communication costs between the entities in each domain are calculated as follows:

$$SM \rightarrow HAN_{GW} : 1 * CC_{QKDG} + 1 * CC_{ES} + 1 * CC_{ID}$$

$$\Rightarrow SM \rightarrow HAN_{GW} = 1 * 128 + 1 * 128 + 1 * 32 = 288 \ \ bits.$$

Where CC_{QKDG}, CC_{ES}, CC_{ID} are the communications costs for QKD key generation operation, symmetric encryption operation and entity identification operation respectively. Similarly,

$$HAN_{GW} \rightarrow NAN_{GW} : 1 * CC_{AGG} + 1 * CC_{DS} + 1 * CC_{ID}$$

$$\Rightarrow HAN_{GW} \rightarrow NAN_{GW} = 1 * 64 + 1 * 128 + 1 * 32 = 224 \ \ bits.$$

Where CC_{AGG} and CC_{ES} are the communications costs for QKD key generation operation and symmetric decryption operation respectively. On the other hand, the communication cost between NAN_{GW} and $KMS_{REGIONAL-CLOUD}$ is given by:

$$NAN_{GW} \rightarrow KMS_{REGIONAL-CLOUD} = 1 * CC_{HMAC} + 1 * CC_{ES} + 1 * CC_{DS} + 1 * CC_{TS}$$

$$\Rightarrow NAN_{GW} \rightarrow KMS_{REGIONAL-CLOUD} = 1 * 160 + 1 * 128 + 1 * 128 + 1 * 32 = 448 \ \ bits.$$

Where CC_{HMAC} and CC_{TS} are the communications costs for hashed message authentication code operation and time stamping operation respectively. Finally, the communication cost between

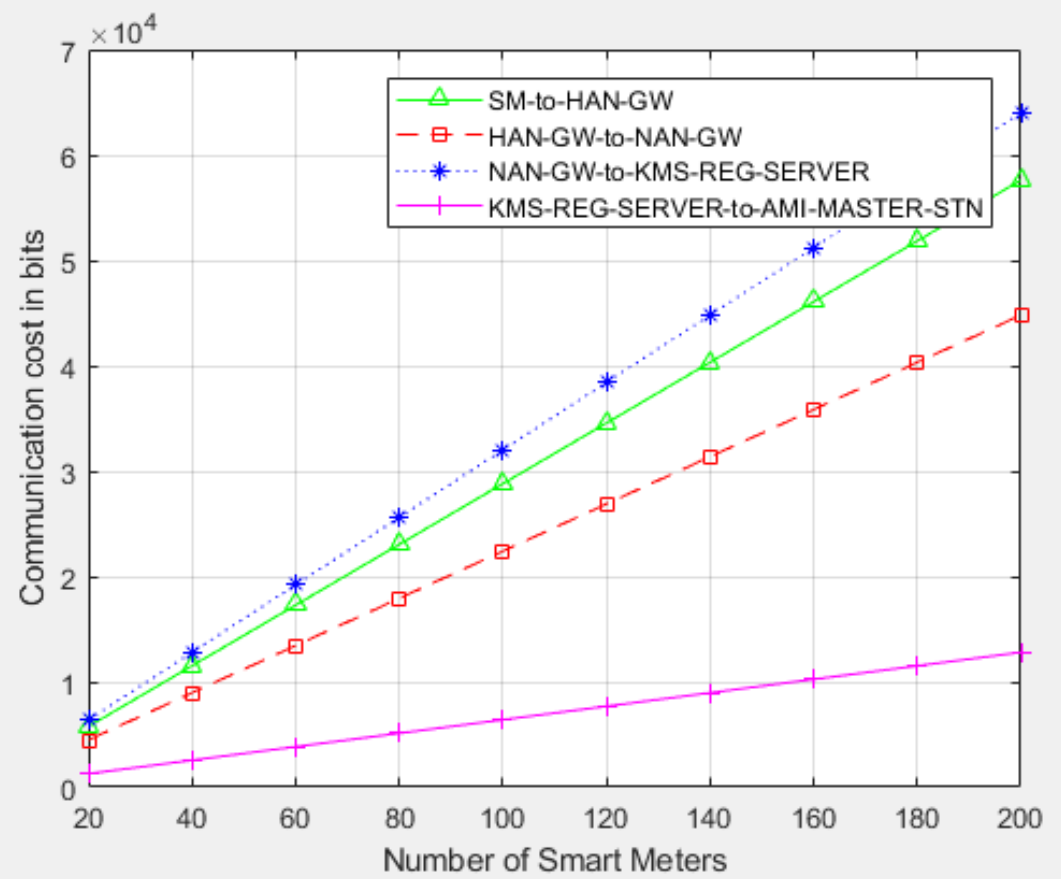

Fig. G.6: Communication cost for different AMI domains

$KMS_{REGIONAL-CLOUD}$ and $AMI_{MASTER-STATION}$ (TTP) is given below:

$$KMS_{REGIONAL-CLOUD} \rightarrow TTP = 1 * CC_{ES}$$

$$\Rightarrow KMS_{REGIONAL-CLOUD} \rightarrow TTP = 1 * 128 = 128 \ bits.$$

Fig. G.6 shows the communication cost of the proposed protocol for the different domains of the network. It can be observed from the figure that as the number of smart meters increases to 200, the domain with the most expensive communication cost ($SM \rightarrow NAN_{GW}$) will experience a communication overhead of approximately 64 kilo bits. On the other hand, Fig. G.7 shows a comparison of the of the communication cost of the proposed protocol with those proposed in [14] and [28]. It can be observed from the graph that the proposed protocol has the lowest communication cost when compared to other protocols. As the number of smart meters increases to 200, the communication cost for the proposed protocol is less than 100 kilo bits.

6 Simulation of the BB84 Protocol

Table G.2: Symbols (and their descriptions) utilized for computation cost analysis

Symbols	Descriptions
C_{RN}	Computation latency of generating a random number
C_M	Computation latency of point scalar multiplication operation
C_{HMAC}	Computation latency for a hashed message authentication code operation
C_{ES}	Computation latency for symmetric encryption operation
C_{DS}	Computation latency for symmetric decryption operation
C_{QKDG}	Computation latency for QKD key generation
C_{TS}	Computation latency for time stamping operation
C_{ID}	Latency of computing the identity of an entity
C_{AGG}	Computation latency of consumption data aggregation operation
C_{EP}	Computation latency for public key encryption operation
C_{DP}	Computation latency for public key decryption operation
C_A	Computation latency for scalar addition operation
C_{XOR}	Computation latency for exclusive bit-wise operation
C_{EMUL}	Computation latency for elliptive curve multiplication operation
C_{ESUB}	Computation latency for elliptive curve subtraction operation
C_{hash}	Computation latency for hash function operation

Table G.3: Comparism of Computation cost of different proposals

Proposal	Computation cost	Number of entities involved in computation
Neetesh et al	$7C_{EP} + 4C_{EMUL} + 7C_{DP} + 1C_{ESUB}$ $+5C_{hash} + 4C_{XOR} + 14C_{HMAC}$ +$1C_M + 1C_A$	EP, SM, HA, HAN GW, BAN GW, NAN GW
Amin M et al (NIKE and NIKE+)	$2C_{RN} + 5C_M + 7C_{hash}$	SM, AHE
Proposed	$3C_{ES} + 2C_{DS} + 1C_{QKDG} + 1C_{AGG}$ $+C_{HMAC} + 2C_{TS}$	SM, HA, HAN GW,TTP, NAN GW, RKMS

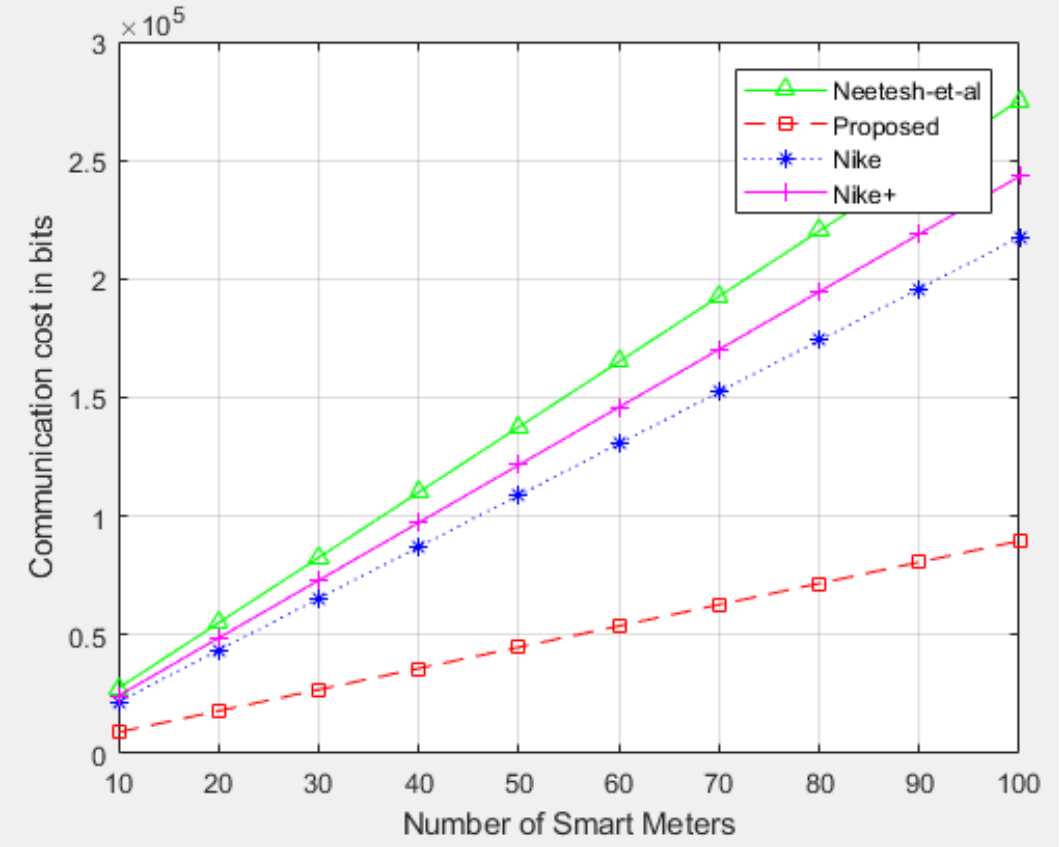

Fig. G.7: Communication cost for different proposals

Table G.4: Operations and their corresponding communication costs in bits

Communication cost for protocol operation	Size in bits
CC_{RN}	128
CC_{M}	128
CC_{HMAC}	160
CC_{ES}	128
CC_{DS}	128
CC_{QKDG}	128
CC_{hash}	256
CC_{TS}	32
CC_{ID}	32
CC_{AGG}	64

6.1 Preliminaries

All the simulation of the BB84 protocol carried out in this study was done using an online web-based QKD simulator [30]. This simulation toolkit is unique and capable of customizing a wide range of parameters. BB84 protocol and indeed many QKD protocols consist of several processes which end up reducing the length of the final key needed as inputs to the required encryption algorithms. As earlier mentioned, the eavesdropper can always sniff on the quantum channel and even the classical channel in order to gain partial information about the sifted key. On the quantum channel, Eve may decide to intercept a photon sent from Alice, perform measurement in a randomly chosen basis and re-send a new photon to Bob according to her measured results. It is to be noted at this point that before a transmission is initiated between Alice and Bob, Alice has to randomly choose to use either rectilinear or diagonal basis to encode the photon. The polarization of each photon is also chosen randomly from a set of polarization options as discussed in section 3. Therefore, it is practically impossible for Eve to determine the polarization state of a photon without first knowing the encoding basis chosen by Alice. If Eve uses for example a polarization beam splitter to project the input photon into either horizontal or vertical polarization (which is a measurement in rectilinear basis), then she will destroy information encoded in diagonal basis. This is a fact since a 45 degree or 135-degree polarized photon has the same chance of being projected into either horizontal or vertical polarization state. As a result, any attempt by Eve to randomly choose the basis and perform the measurement will introduce some errors, and these errors can be statistically calculated by Alice and Bob. In order to verify this, a biased estimation plot generated from our simulation and shown in Fig. G.8, shows that increase in the rate of eavesdropping increases the error introduced in diagonal basis measurements but decreases the error introduced in rectilinear basis measurements. However, the average of the two error plots can be utilized effectively by Alice and Bob to determine the presence of an eavesdropper. In addition, QKD sifting plot shown in Fig. G.9 can provide a lot of insight on how to handle the activities of an eavesdropper. This plot contains important statistics like basis match and mismatch, measurement match and mismatch before and after sifting. Since the basis are always compared by Alice and Bob, any photon processed using a non-matching basis is discarded and then removed from the raw key material. Therefore, the mismatch measurement outcomes can be useful to the legitimate communicating parties in determining Eve's attack strategies. This will subsequently lead to an improved key rate. On the other hand, the partial knowledge of the sifted key which can be gained by the eavesdropper during reconciliation and error correction can lead to the reduction of some bits in the key material. Unfortunately, most symmetric key protocols were designed such that their security is equal to their key length. In order words, key length can be used as a measure for

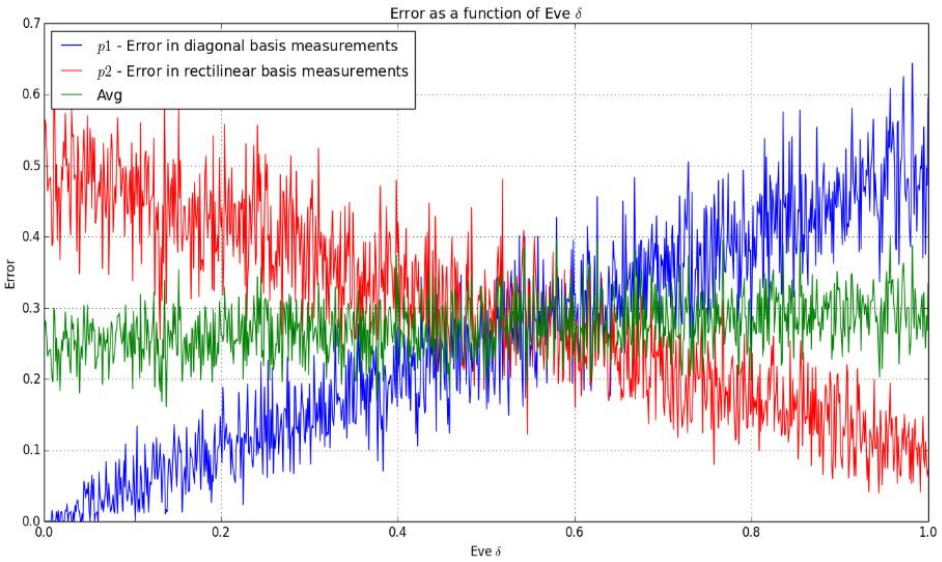

Fig. G.8: QKD Biased estimation plot

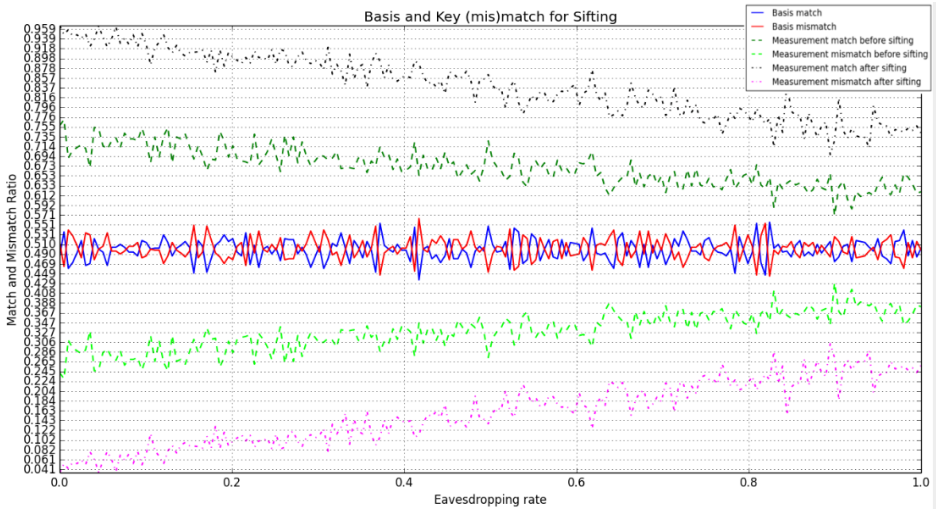

Fig. G.9: QKD sifting plot

the security that can be attained by an encryption algorithm. Miralem Mehic et al [31] states that if the final key rate is not long enough, encryption solutions like the one-time-pad might not be feasible because of lack of key material. It is, on this note that further simulation studies and analysis is carried out in the following sub-section. The main aim of this further simulation study is to discover the key factors affecting the length of the final key generated by the BB84 protocol.

6.2 Investigating the Effects of Basis Bias Ratio on the Key Length

In furtherance of earlier simulation carried out in the previous section, several simulations were carried out by varying some important parameters while keeping one or two parameters constant.

This approach was adopted in order to observe the effects of the varied parameters on the key length. In doing this, biased error estimation and eavesdropping were enabled during the simulation runs. Enabling eavesdropping makes it possible for the adversary (Eve) to launch specific intercept-resend attacks on the public classical channel. It is to be noted that the initial number of quantum bits (qubits) sent over the quantum channel was set to 500 qubits throughout the simulations. On the other hand, biased error estimation was enabled all through the simulation for countering the effect of eavesdropping attack on the public channel. It was discovered that among all the tested varying parameters, only basis choice bias for Alice (referred to as basis choice bias delta in the simulation kit) and the basis choice bias for the Eavesdropper (Eve) were the key parameters that affected the key length. As a result, the simulation experiments presented herein this sub-section were done after investigating the effects of the varying simulation parameters. These were done in three categories:

1. Firstly, the basis choice bias for Alice and Eavesdropper (Eve) were jointly varied while error estimation sampling rate, error tolerance, and eavesdropping rate were kept constant. It is to be noted that the basis choice bias probability for selection of qubits was varied in the simulation runs from 0.1 to 0.9. On the other hand, the error tolerance is a threshold for the quantum bit error rate (QBER) to be tolerated. The error tolerance and the eavesdropping rate were set to a value of 0.05 throughout the experiments.

2. In the second experiment, the basis choice bias for Eve was varied while keeping the basis choice bias for Alice constant along with other constant parameters as in (1) above.

3. The last experiment was conducted by varying the basis choice bias for Alice while keeping the basis choice bias for Eve constant along with other constant parameters defined in (1) above.

The configuration parameters for this simulation has been summarized and provided in Table G.5.

Table G.5: Summary of Simulation Configuration Parameters

Simulation Category	Property Qubit Count	Basis Choice for Alice	Basis Choice for Eve	Eavesdropping	Basis Error Estimation
1st Stage	500	Varied (0.1-0.9)	Varied (0.1-0.9)	Enabled	Enabled
2nd Stage	500	Constant (0.1)	Varied (0.1-0.9)	Enabled	Enabled
3rd Stage	500	Varied (0.1-0.9)	Constant (0.1)	Enabled	Enabled

6.3 Simulation Results

The simulation results generated in the first simulation category show that the bases choice bias for Alice and Eve had obvious effects on the key rate. In this simulation, there was a downward decrease in the key rate as the basis choice bias probability for selection of qubits for Alice and Eve were increased from 0.1 to 0.5. However, a further increase in the basis choice bias probability from 0.6 to 0.9 resulted in an increase in the key rate (key length) as illustrated in Fig.G.10 It is to be noted that a comparative analysis can be made between the key lengths before error correction and the final key length as shown in all the simulation results in Fig. G.10 to Fig.G.12.

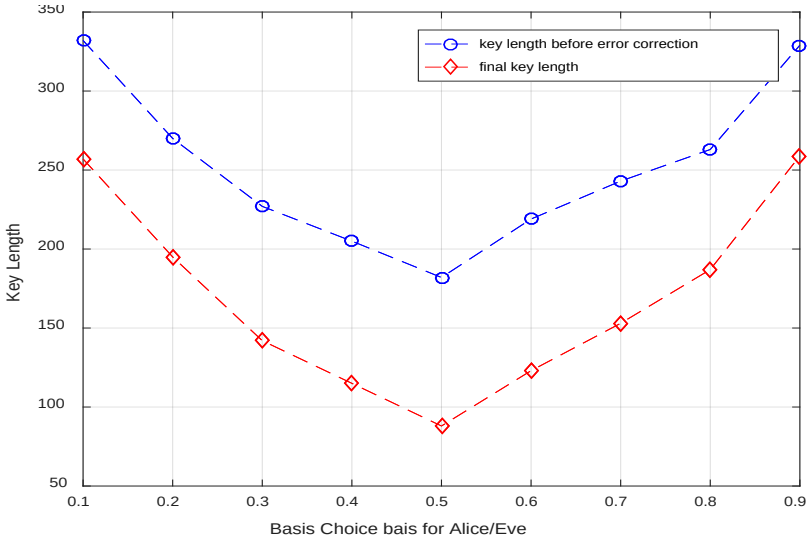

Fig. G.10: Plot of final key length against bases bias ratio for Alice and Eve

From the graphs, it can be seen that there is a decrease in the key material, a trend that agrees with theoretical knowledge, which has been previously explained. The trend shown for the final key length in Fig. G.10 can obviously be utilized by the sending party in making a good choice for the selection of basis for encoding qubits to be transmitted to the receiving party. However, it was very important to further verify which basis choice bias (Alice or Eve) affected the key length specifically since the sending party (Alice) has no direct control over the adversary (Eavesdropper). In the second simulation category, it was discovered that there was a deviation from the trend recorded in the first simulation category. This time around, varying the basis choice bias probability for the eavesdropper (Eve) from 0.1 to 0.9 did not show a defined trend in the key length. Instead, Fig. G.11 showed an intermittent decrease and increase in the key length. It can, therefore, be concluded that basis choice bias probability for Eve was not the main parameter affecting the key length. In the last simulation category with the result shown in Fig. G.12, varying the basis choice bias probability for Alice affected the key length with a trend almost the same as the one shown in Fig.G.10. It can then be concluded

that the basis choice bias probability for Alice was the main parameter affecting the final key length. In other words, the optimal basis choice bias probabilities (bias ratios) to be utilized by the sending party (Alice) to generate final key length that could be suitable for symmetric encryption algorithms utilized for the cloud-based SG AMI proposed in this paper lies in the range of 0.1-0.2 and 0.8-0.9 respectively. The implication of this is that the sending party with basic understanding of how to select the optimal basis bias ratio can succeed in generating symmetric keys with the final key length that is sufficient for the secure implementations of the proposed protocols.

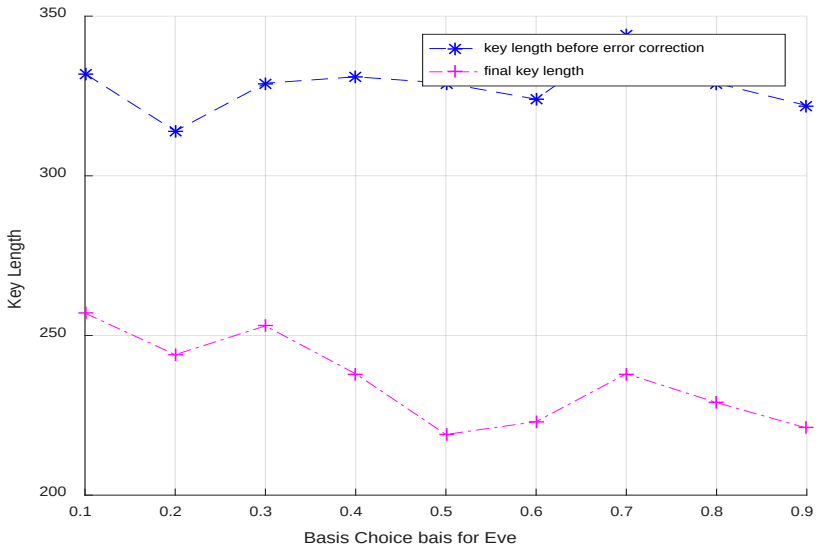

Fig. G.11: Plot of final length against basis bias ratio for Eve

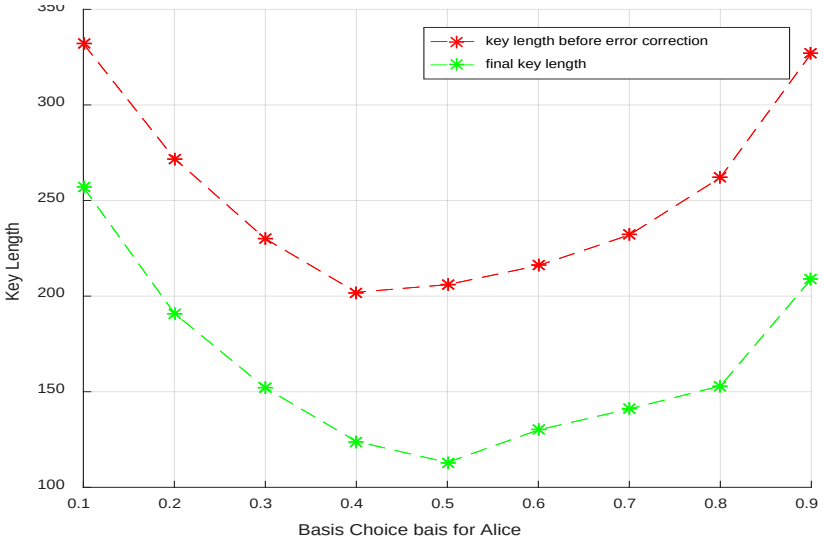

Fig. G.12: Plot of final key length against basis bias ratio for Alice

7 Conclusion

There is a general belief that cyber-attacks targeting cyber-physical systems would get more sophisticated in the coming years. This is as a result of the tremendous skills possessed by these criminals. Research has also revealed that cyber-criminals across the globe have evolved their communication network to become more collaborative than their victim organizations. This trend would only get worse with the availability of quantum computers especially in commercial quantity. Quantum computing will render very vulnerable, many cryptographic protocols that are adjudged to be secure today. With quantum algorithms, such cryptographic keys can be broken quickly, and this can allow eavesdroppers to listen to messages, replay them or even modify them. In this paper, a cloud-based SG AMI system model that can be used to enhance security and privacy in the network is proposed. The proposed system model leverages on the security features of quantum key distribution (QKD). The cryptographic protocols incorporated into this model are all based on symmetric cryptosystem in contrast to similar protocols found in the literature which depend on public key cryptosystems which would be vulnerable to quantum attack. In this paper, we have shown through simulation, that QKD is a cryptographic primitive that can offer information theoretic security (ITS) guaranteed by the laws of physics. Consequently, QKD provides the means for a secure distribution of secret keys that can be used with quantum safe symmetric key protocols like advanced encryption standard (AES) and one-time pad (OTP) encryption for a cloud-based SG AMI.

References

[1] 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, 2012.

[2] E. B. Rice and A. AlMajali, "Mitigating the risk of cyber attack on smart grid systems," *Procedia Computer Science*, vol. 28, pp. 575–582, 2014.

[3] A. Califano, E. Dincelli, and S. Goel, "Using features of cloud computing to defend smart grid against ddos attacks," in *10th Annual symposium on information assurance (Asia 15), ALBANY*, 2015, pp. 44–50.

[4] E. Brynjolfsson, P. Hofmann, and J. Jordan, "Cloud computing and electricity: beyond the utility model," *Communications of the ACM*, vol. 53, no. 5, pp. 32–34, 2010.

[5] M. T. Khorshed, A. S. Ali, and S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing," *Future Generation computer systems*, vol. 28, no. 6, pp. 833–851, 2012.

[6] M. Yigit, V. C. Gungor, and S. Baktir, "Cloud computing for smart grid applications," *Computer Networks*, vol. 70, pp. 312–329, 2014.

[7] D. S. Markovic, D. Zivkovic, I. Branovic, R. Popovic, and D. Cvetkovic, "Smart power grid and cloud computing," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 566–577, 2013.

[8] B. Fang, X. Yin, Y. Tan, C. Li, Y. Gao, Y. Cao, and J. Li, "The contributions of cloud technologies to smart grid," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1326–1331, 2016.

[9] R. Diovu and J. Agee, "A cloud-based openflow firewall for mitigation against ddos attacks in smart grid ami networks," in *PowerAfrica, 2017 IEEE PES*. IEEE, 2017, pp. 28–33.

[10] K. Billewicz, "The use of cloud computing in ami system architecture," in *Modern Electric Power Systems (MEPS), 2015*. IEEE, 2015, pp. 1–6.

[11] M. M. Hasan and H. T. Mouftah, "Encryption as a service for smart grid advanced metering infrastructure," in *Computers and Communication (ISCC), 2015 IEEE Symposium on*. IEEE, 2015, pp. 216–221.

[12] R. Onoshakpor and K. Okafor, "Cyber security in smart grid convolution networks (sgcns)," in *Electro-Technology for National Development (NIGERCON), 2017 IEEE 3rd International Conference on*. IEEE, 2017, pp. 392–399.

[13] B. Alohali, M. Merabti, and K. Kifayat, "A cloud of things (cot) based security for home area network (han) in the smart grid," in *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2014 Eighth International Conference on*. IEEE, 2014, pp. 326–330.

[14] N. Saxena and B. J. Choi, "Integrated distributed authentication protocol for smart grid communications," *IEEE Systems Journal*, 2016.

[15] M. Grimaila, J. Morris, and D. Hodson, "Quantum key distribution: A revolutionary security technology," *The Information System Security Association (ISSA) Journal*, pp. 20–27, 2012.

[16] H. Bennett Ch and G. Brassard, "Quantum cryptography: public key distribution and coin tossing int," in *Conf. on Computers, Systems and Signal Processing (Bangalore, India, Dec. 1984)*, 1984, pp. 175–9.

[17] A. K. Ekert, "Quantum cryptography based on bell's theorem," *Physical review letters*, vol. 67, no. 6, p. 661, 1991.

[18] B. Huang, Y. Huang, J. Kong, and X. Huang, "Model checking quantum key distribution protocols," in *Information Technology in Medicine and Education (ITME), 2016 8th International Conference on*. IEEE, 2016, pp. 611–615.

[19] C.-H. F. Fung, X. Ma, and H. Chau, "Practical issues in quantum-key-distribution postprocessing," *Physical Review A*, vol. 81, no. 1, p. 012318, 2010.

[20] H. Krawczyk, "Lfsr-based hashing and authentication," in *Annual International Cryptology Conference*. Springer, 1994, pp. 129–139.

[21] C. Castelluccia, E. Mykletun, and G. Tsudik, "Efficient aggregation of encrypted data in wireless sensor networks. 3rd intl," in *Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Sensor Networks, Italy*, 2005.

[22] H. Krawczyk, R. Canetti, and M. Bellare, "Hmac: Keyed-hashing for message authentication," 1997.

[23] M. T. Goodrich and R. Tamassia, *Algorithm design: foundation, analysis and internet examples*. John Wiley & Sons, 2006.

[24] S. Finster and I. Baumgart, "Privacy-aware smart metering: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1088–1101, 2015.

[25] R. Petrlic, "A privacy-preserving concept for smart grids," *Sicherheit in vernetzten Systemen*, vol. 18, pp. B1–B14, 2010.

[26] M. LeMay, G. Gross, C. A. Gunter, and S. Garg, "Unified architecture for large-scale attested metering," in *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*. IEEE, 2007, pp. 115–115.

[27] R. Alléaume, C. Branciard, J. Bouda, T. Debuisschert, M. Dianati, N. Gisin, M. Godfrey, P. Grangier, T. Länger, N. Lütkenhaus *et al.*, "Using quantum key distribution for cryptographic purposes: a survey," *Theoretical Computer Science*, vol. 560, pp. 62–81, 2014.

[28] A. Mohammadali, M. S. Haghighi, M. H. Tadayon, and A. Mohammadi-Nodooshan, "A novel identity-based key establishment method for advanced metering infrastructure in smart grid," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2834–2842, 2018.

[29] F. Diao, F. Zhang, and X. Cheng, "A privacy-preserving smart metering scheme using linkable anonymous credential," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 461–467, 2015.

[30] (Accessed 11 Jun. 2018) [online]: Qkdsimulator- available at: https://qkdsimulator.com.

[31] M. Mehic, M. Niemiec, and M. Voznak, "Calculation of the key length for quantum key distribution," *Elektronika ir Elektrotechnika*, vol. 21, no. 6, pp. 81–85, 2015.

Chapter IX

Quantitative Analysis of Firewall Security under DDoS Attacks in Smart Grid AMI Networks

Paper H

Quantitative Analysis of Firewall Security under DDoS Attacks in Smart Grid AMI Networks

R.C Diovu and J.T Agee

Published

IEEE Xplore (Presented at the 2017 IEEE NIGERCON Conference)

Abstract

One of the key objectives of distributed denial of service (DDoS) attack on the smart grid advanced metering infrastructure is to threaten the availability of end user's metering data. This will surely disrupt the smooth operations of the grid and third party operators who need this data for billing and other grid control purposes. In previous work, we proposed a cloud-based Openflow firewall for mitigation against DDoS attack in a smart grid AMI. In this paper, PRISM model checker is used to perform a probabilistic best- and worst-case analysis of the firewall with regard to DDoS attack success under different firewall detection probabilities ranging from zero to 1. The results from this quantitative analysis can be useful in determining the extent to which a DDoS attack can undermine the correctness and performance of the firewall. In addition, the study can also be helpful in knowing the extent the firewall can be improved by applying the knowledge derived from the worst-case performance of the firewall.

1 Introduction

The AMI can be defined as the communication hardware, software and its associated data management systems which create a network of networks between the smart meters, utility and authorized third party operators which allow for easy communication of metering data and other control information among all these entities. In many countries, the standard documents for deployment of AMI requires its integration with critical resources which make it a serious target for cyber criminals with different malicious intentions [1, 2]. Thus, cyber security solutions that will protect the entire AMI against numerous attacks targeting its information infrastructures are not only desirable but imperative. A compromise on this important obligation can lead to loss of revenue for grid and third party operators. For example, a highly skilled cybercriminal who wishes to short-change the authorized third party operator in charge of billing in his neighborhood area network (NAN) domain may decide to target the network or his smart meter with a metering data tampering/falsification attack. On the other hand, attacks such as denial of service or distributed denial of service (DDoS/DoS) will threaten the availability of metering data, thus, undermining the smooth operations of AMI back-end systems responsible for billing and other grid control operations. In such circumstance, demand side management will be highly obstructed and grid operations like power quality monitoring would be seriously hampered. The smart grid AMI may be targeted with DDoS attacks such as TCP SYN flooding, UDP flooding attacks or a malicious intension to

disorganize the network routing table in order to affect the packet delivery success of the network [1, 3, 4]. In addition, the cyber attacker targeting the AMI network can saturate the available network bandwidth with volumetric traffic such that little or no resources will be left to service requests from legitimate network users [5]. Baling Fang et al [6] has argued that the smart grid can be incorporated with cloud-based securities in order to reduce the security risks to an acceptable level and to improve the disaster recovery abilities of power systems. Califano et al [7] believes that given the distributed nature of the smart grid, it is almost inevitable for the smart grid and cloud computing to be integrated. In line with this argument, we proposed in our previous work a cloud-based Openflow firewall for mitigation against distributed denial of service (DDoS) attacks in smart grid AMI network [8]. Apart from the mitigation solutions that can be offered against DDoS attack, the Openflow firewall has added advantage of reducing the burden of storage and data computations for the AMI back-end systems. In this paper, a quantitative (probabilistic) analysis of the firewall with an intension of understanding its best- and worst-case performance using the PRISM model checker [9, 10] is carried out. This is achieved by analyzing the DDoS attack success against different firewall detection probabilities and then observing the best- and worst-case performance. Thus, the system is modeled as a markov decision process (MDP). A markov decision process is a well known mathematical formalism employed in the verification of systems with unknown adversaries or scheduling mechanisms or systems with transitions whose next-state probability is not precisely known. In this context, the DDoS attack success and the time it could take the DDoS attack agents to exploit system vulnerabilities that affect the availability of the metering data or the network resources could depend on factors such as network topology and/or the firewall configurations. Therefore, analysis of systems using MDPs generally require the identification of strategies that either minimize or maximize a target function based on the MDP's rewards or costs. An alternative approach is to define formal languages that can express quantitative properties to be analyzed by the MDP which may include those extending PRISM's classical temporal logics with their required probabilistic operators.

This paper provides an extension to the current trend in the analysis of the impact of DDoS attacks on the SG AMI as can be seen from the review of related work presented in section II of this paper. The main focus of such analysis is to provide insights that could be useful in understanding the damage that can be caused by DDoS attacks; thus, providing the knowledge needed in the design and implementation of the necessary countermeasures. The rest of this paper is organized as follows: An overview and quick description on Markov decision modeling in PRISM is presented in section III. In section IV, the PRIM model for advanced metering infrastructure (AMI) network is presented. While section V contains the PRISM experiments carried out in this work, the conclusion of this

research is presented section VI.

2 Review of Related Literature

This section provides a quick literature review on the current trend of research related to our study. M.Q Ali et al [11] presented a probabilistic model for checking intrusions in a smart grid AMI. They showed that the behaviour of AMI can be modeled using event logs collected from smart devices and which can be verified using the invariant specifications generated from the AMI devices configurations. In their model, the AMI behaviour was modeled using the 4th order Markov chain and the stochastic model was then probabilistically verified using Linear Temporal Logic specifications. While their work was geared towards modeling malicious intrusions in a smart grid AMI, the focus of this paper is on analyzing the success of DDoS attacks in smart grid AMI under different firewall detection probabilities. Mujahid Mehsin et al [12] proposed the IoT Risk Analyzer framework for quantitative and formal analysis of risks using probabilistic model checking. The authors demonstrated that their framework can be used to realistically model and verify the risk exposure of IoT systems which include smart home profiles. However, their framework may not be suitable for modeling DDoS attacks in a smart grid AMI. Other research works on probabilistic model analysis for DoS related attacks but which were specifically not adapted for environments like the smart grid AMI can be found in [13, 14]. S. Khaled et al [15] presented an analytical queuing model based on the embedded Markov chain employed for the performance analysis of a rule-based firewall. Their model was used to analyze normal flows as well as DoS attack flows targeting different rule positions. Unfortunately, their analytical model was based on complex mathematical formulations and their model was not particularly designed for environment such as smart grid AMI.

Similarly, Yonghe et al [16] conducted a study on the impact of DDoS attack on the SG AMI by simulating DDoS attack scenarios using NS3 simulator. The analysis of their study reveals that compromising about 60 % of the smart meters connected to the AMI network by DDoS agents would lead to a 50% increase in packet loss and an increase in the average and maximum end-to-end delays (13.5 and 5.48 times higher than the normal conditions) respectively. In a similar manner, Asri et al [3] studied the impact of DDoS attack on the SG AMI by simulating the effects of a flood based UDP attack on a server using NS2 simulator. This attack resulted in the grounding of the server at a simulation time of 500 ticks (atomic discrete time unit). Finally, Kallisthenis et al [17] also studied the impact of DoS/DDoS attack on the AMI using OMNET ++ simulator. This study reveals that the above attack had a negative impact on data availability which could hamper the smooth operations of

the AMI back-end systems and utility control centers.

From the above review, some research effort had been given to the analysis of the impact of DoS/DDoS attacks on the SG AMI using both simulation and model checking methodologies. Considering that cyber-attacks such as DoS/DDoS attacks would continue to be a target for the SG AMI in the future, single faced approach in the analysis of the impact (concentration only on either the DDoS impact or firewall performance) on the AMI would not be sufficient in providing necessary insights that could help in mitigating the effects of these attacks. This paper advocates that firewall security would remain one of the most important cyber security measures that can seriously mitigate the effects of DDoS attacks on the SG AMI.

As a result, a double faced approach is applied in this paper where a quantitative analysis of the firewall performance is analyzed under different (probabilistic) DDoS scenarios. The quantitative results obtained from this study will be useful in understanding vividly the dynamics of DDoS attacks, thus, providing reasonable information that can help us to validate the simulation results of our previously designed Openflow firewall. When compared to previous similar studies (as can be seen from review), our study can be differentiated from previous studies in the following ways:

1. The SG AMI was not the focus of some of the previous studies [10], [12], [14, 15].
2. Some of the analysis from previous studies [3, 14–16], have concentrated on either DDoS or firewalls and not both as is the case in this study.
3. Finally, some approaches [3, 16, 17] in which analysis were based on simulation results cannot be compared to the PRISM model checking approach utilized in this study; as results generated from one experiment from a well constructed PRISM model can be equivalent to numerous simulation runs. More importantly, quantitative results generated from this kind of study can be used to validate the simulation results of such approaches as in [3, 16, 17].

3 Modeling using PRISM'S Markov Decision Process (MDP)

Markov decision process is an extension of discrete time Markov chain (DTMC). It has the ability of modeling systems which are stochastic with some non-deterministic behaviour. Each state of an MDP is associated with a set of probability distributions over the states of the MDP. A Markov decision process (MDP) is defined as a tuple which may be given by:

$$Y = (Q, Q_{init}, Steps, rew).$$

where

- Q is a finite set of states.
- Q_{init} is the initial state.
- Steps: $Q \rightarrow 2^{Dist(s)}$ is the probability transition function and
- $rew = Q * Dist(Q) \rightarrow R \geq 0$.

In MDP, a probabilistic transition $\stackrel{\mu}{\rightarrow} \ Q^{'}$ can be made from a state, q by firstly picking non-deterministically a distribution $\mu \in Steps \ (q)$ and then making a probabilistic choice of next targeted state $q^{'}$ depending on the distribution, μ. The function representing reward links the non-negative real values with performing the transition μ from a given state q. In addition, a path of a Markov decision process represents a particular resolution of both probability and non-determinism. Represented formally, a path of an MDP is an infinite or non-empty sequence of probabilistic transitions simply given by:

$$\pi = Q_{init} \stackrel{\mu}{\rightarrow} Q_1 \stackrel{\mu_1}{\rightarrow} Q_2 \stackrel{\mu_2}{\rightarrow} ...$$

such that $\mu_i(Q_i + 1) > 0$ for all values of i. For any infinite path π and set of states k, the total reward accumulated until a state in k is reached along π , is denoted by:

$$r(K, \pi) = \sum_{i=1}^{min|\pi(j) \in k} rew(\pi(i-1), Step(\pi, i-1)) \tag{H.1}$$

provided there exist $j \in N$ such that $\pi(j) \in k$, otherwise, the accumulated reward amounts to infinity.

Due to the presence of non-determinism in a Markov decision process, there is necessarily not a single value corresponding to a given quantitative measurement. On the contrary, best- and worst-case measurements can be considered. Specifically, model checking using the MDP reduces the computations of probabilistic and expected reachabilities properties in the form of minimum and maximum probability of reaching a given set of states and the minimum and maximum expected rewards accumulating from there.

4 PRISM Model for the AMI

In this study, the AMI network is modeled as a grid of N by N nodes [18] which can be represented by a square matrix of order, n, and shown by the matrix equation given below:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{H.2}$$

The grid is a special term employed by experts in high-performance computing and networking [19]. The grid network infrastructure can connect network resources such as workstations and supercomputers, storage systems, databases, etc, which may be distributed across different geographical locations and then presented as a unified integrated resource that can be virtualized for the benefit of different organizations. Ever since, grid-based network infrastructures have been employed in different areas like sensor networks, neural networks and even power system networks [20–22]. It can therefore be noted that a grid network of N by N can be used to model networks of desirable node size. If N in equation (H.2) is 3, then an equivalent AMI network is given in Fig.H.2. We assume that each node is connected as a neighbor to four different nodes in the network. However, neighbors outside the border matrix were not considered in this model. With reference to Fig.H.2, the three smart meter nodes represent the smart meters of the users connected to the AMI network whereas the three AMI server nodes represent the servers of AMI back-end systems who may be in charge of authorized third-party operations. Fig.H.1 shows the global variables for the AMI model developed in PRISM. In the model, there is equal probability of DDoS attack affecting all nodes (smart meter nodes, AMI server nodes and openflow firewall nodes) in the network. On the other hand, the probability of the firewall detecting an attack on the AMI servers and smart meter nodes is 0.5. In order to understudy the worst case and best case performance of the firewall, the detection probability of the firewall is varied in the analysis from 0.1 to 0.9. The PRISM model details for the module which represent the first column nodes (AMI seerver node 1, Openflow firewall node 1, and smart meter node 1) have been included in Fig. H.3, H.4, and H.5 respectively for proper understanding. A module normally contains the module name and its commands. The module command describes the module's behaviour which can be understood by the way the module changes its state over time.

In this model, the AMI server nodes and smart meter nodes were initially configured with three probability states which include:

```
// Prism model for N by N Grid AMI Network (N=3)

//Probability of DDoS attack equals 0.5 for all node types
//probability of Openflow firewall detection for AMI servers and Smart Meters equals 0.5
//Probability of detection by the firewall Nodes is varied in the analysis, as such
//is left unspecified in the model.

mdp

// probabilities
const double affect_pro=0.5; // probability a Node is affected by DDoS attack
const double detect_pro1=0.5; // probability DDoS is detected for AMI Servers and SMs
const double detect_pro2; // probability DDoS is detected by Openflow firewall

// AMI Servers

const double detect_pro11=detect1;
const double detect_pro12=detect1;
const double detect_pro13=detect1;

// Openflow Firewall Nodes

const double detect_pro21=detect2;
const double detect_pro22=detect2;
const double detect_pro23=detect2;

// Smart Nodes

const double detect_pro31=detect1;
const double detect_pro32=detect1;
const double detect_pro33=detect1;
```

Fig. H.1: Global Variables for the AMI Prism Model

- Unaffected by DDoS attacks.
- Node Unaffected by DDoS attacks but Openflow firewall's perimeter defense already by-passed.
- Affected by DDoS attacks.

Similarly, the OpenFlow firewall is assigned a DDoS detection probability in the range of 0:0.1:1. The events of the DDoS attack agents (DDoS zombies) bypassing the Openflow firewall perimeter defense in order to attack the AMI server nodes have been modeled as stochastic. In other words, there are chances of success or failure which depends on either the DDoS detection probability of the Openflow firewall. Furthermore, the order or choice of which smart meter nodes to be attacked by the DDoS zombies is non-deterministic as this may depend on the AMI network topology or the level of vulnerabilities present on the nodes.

The number of DDoS attacks targeting the different nodes on the network have been implemented in this study using the PRISM reward structure. In this case, the reward structure was used to implement the expected number of DDoS attacks on the nodes connected to the network. In other words, a reward structure can be used to understand the probability of a model behaving in certain way. More interestingly, reward structure can be used for a wider range of quantitative measurements relating to

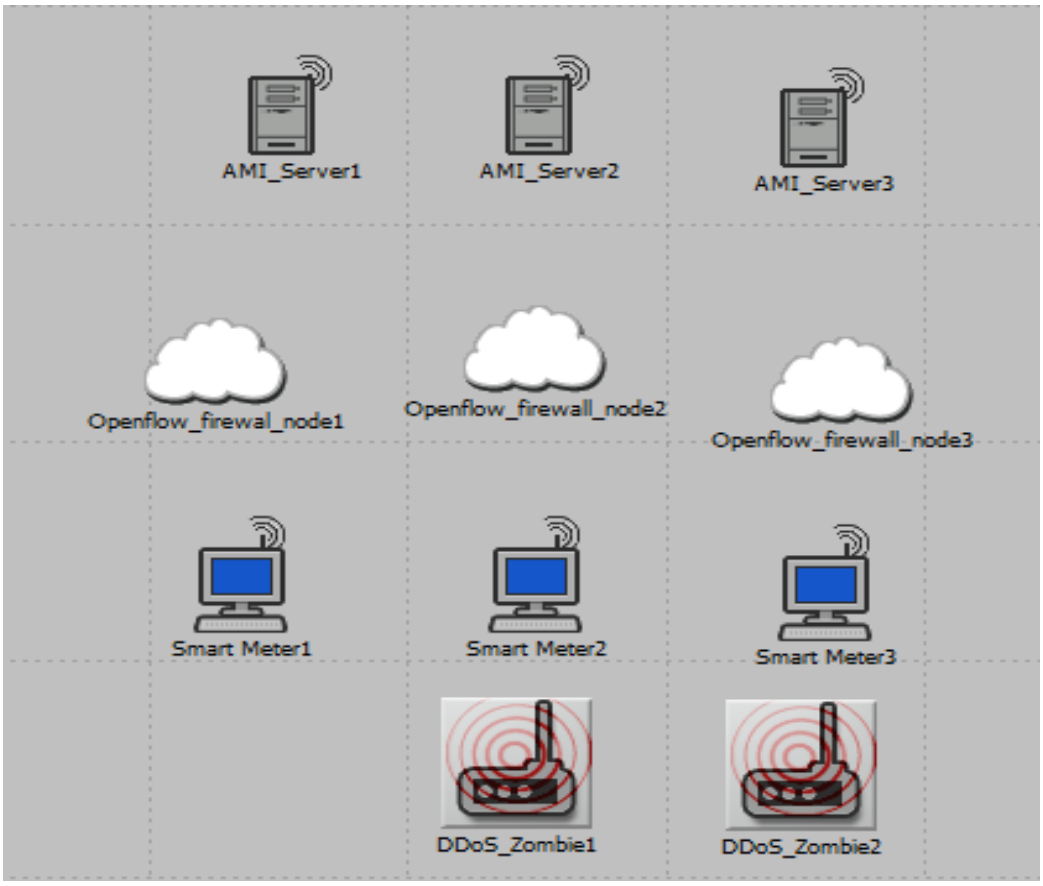

Fig. H.2: Equivalent AMI Network for $N = 3$

```
//module AMI Server Node 1 = n11

module n11
        s11 : [0..2]; // node unaffected
        // 0 - node unaffected
        // 1 - node unffected by DDoS attack but firewall perimeter defense by-passed
        // 2 - node affected by DDoS attack

        // firewall attacked (from an affected neighbour)
        [attack11_21] (s11=0) ->  detect11 : true + (1-detect11) : (s11'=1);
        [attack11_12] (s11=0) ->  detect11 : true + (1-detect11) : (s11'=1);
        // if the firewall perimeter is by-passed attempt is made to affect the node
        [] (s11=1) -> affect : (s11'=2) + (1-affect) : (s11'=0);
        // if the node is affected, then it tries to compromise its neighbouring nodes
        [attack21_11] (s11=2) -> true;
        [attack12_11] (s11=2) -> true;

endmodule
```

Fig. H.3: Module for AMI Server Node 1

```
54 //module Openflow firewall Node 1 = n21
55
56 module n21
57
58         s21 : [0..2]; // node unaffected
59         // 0 - node unaffected
60         // 1 - node unffected by DDoS attack but firewall perimeter defense by-passed
61         // 2 - node affected by DDoS attack
62
63         // firewall attacked (from an affected neighbour)
64         [attack21_31] (s21=0) -> detect21 : true + (1-detect21) : (s21'=1);
65         [attack21_22] (s21=0) -> detect21 : true + (1-detect21) : (s21'=1);
66         [attack21_11] (s21=0) -> detect21 : true + (1-detect21) : (s21'=1);
67         // if the firewall perimeter is by-passed attempt is made to affect the node
68         [] (s21=1) -> affect : (s21'=2) + (1-affect) : (s21'=0);
69         // if the node is affected, then it tries to compromise its neighbouring nodes
70         [attack31_21] (s21=2) -> true;
71         [attack22_21] (s21=2) -> true;
72         [attack11_21] (s21=2) -> true;
73
74 endmodule
```

Fig. H.4: Module for Openflow firewall Node 1

```
76 // //module Smart Meter Node 1 = n31
77
78 module n31=n11[s11=s31,detect11=detect31,attack21_11=attack21_31,
79 attack12_11=attack32_31,attack11_21=attack31_21,attack11_12=attack31_32]
80
81 endmodule
```

Fig. H.5: Module for Smart Meter Node 1

```
// reward structure (number of attacks)
rewards "attacks"
        // Number of attacks for nodes situated at the corner
        [attack11_12] true : 1;
        [attack11_21] true : 1;
        [attack31_21] true : 1;
        [attack31_32] true : 1;
        [attack13_12] true : 1;
        [attack13_23] true : 1;
        [attack33_32] true : 1;
        [attack33_23] true : 1;
        // Number of attacks for nodes at edge
        [attack12_13] true : 1;
        [attack12_11] true : 1;
        [attack12_22] true : 1;
        [attack21_31] true : 1;
        [attack21_11] true : 1;
        [attack21_22] true : 1;
        [attack32_33] true : 1;
        [attack32_31] true : 1;
        [attack32_22] true : 1;
        [attack23_33] true : 1;
        [attack23_13] true : 1;
        [attack23_22] true : 1;
        // Number of attacks for nodes situated at the middle
        [attack22_32] true : 1;
        [attack22_23] true : 1;
        [attack22_12] true : 1;
        [attack22_21] true : 1;
endrewards
```

Fig. H.6: Reward Structure for Attacks

the model behaviour. H.6 shows the multiple reward structure (with a "true" guard and a reward value of 1) implemented in this study.

5 Experimental Results and Analysis

In this section, a brief description of the experiments carried in PRISM is presented. A best- and worst-case analysis has been carried out for two analysis scenarios. In the first scenario, a best- and worst-case analysis for DDoS attack success on AMI server 3 has been carried out. In this analysis, we compute the minimum and maximum probability that the AMI server 3 was attacked within q given number of steps. This was computed by using the following property specifications parameters in PRISM:

$$Constant \;\; q,$$

$$P_{min} =? \;\; [F \Leftarrow qAMI\_Server3 = 2]$$

$$P_{max} =? \ [F \Leftarrow qAMI\_Server3 = 2]$$

Figs. H.7 and H.8 show the graph of minimum probability of DDoS attack success within q steps. Similarly, Figs. H.9 and H.10 show the maximum probabilities of DDoS attack success within q transition states. As can be seen from those graphs, the DDoS attack success decreases as the Openflow firewall detection probabilities increases. It is important to note that for the two pair of graphs, the DDoS attack success increases with increase in q transitions from low to high value states. It can be observed that Figs. H.7 and H.8, Figs. H.9 and H.10 represent the best- and worst-case performance of the Openflow firewall under different firewall detection probabilities. In the second scenario, another best- and worst-case analysis is carried out in PRISM. This was computed by using the following reward property parameter specifications in PRISM:

$$R\{"attacks"\}min =? \ [F \Leftarrow qAMI\_Server3 = 2]$$

$$R\{"attacks"\}max =? \ [F \Leftarrow qAMI\_Server3 = 2]$$

This time, the minimum and maximum number of DDoS attacks carried out by the DDoS zombies until a successful attack on the AMI_server 3 is achieved is computed. Figs.H.11 and H.12 represent respectively the minimum and maximum expected number of DDoS attacks until AMI_Server 3 is attacked. Following the same trend as in the first scenario, these minimum values were affected by increase in the Openflow firewall detection probability.

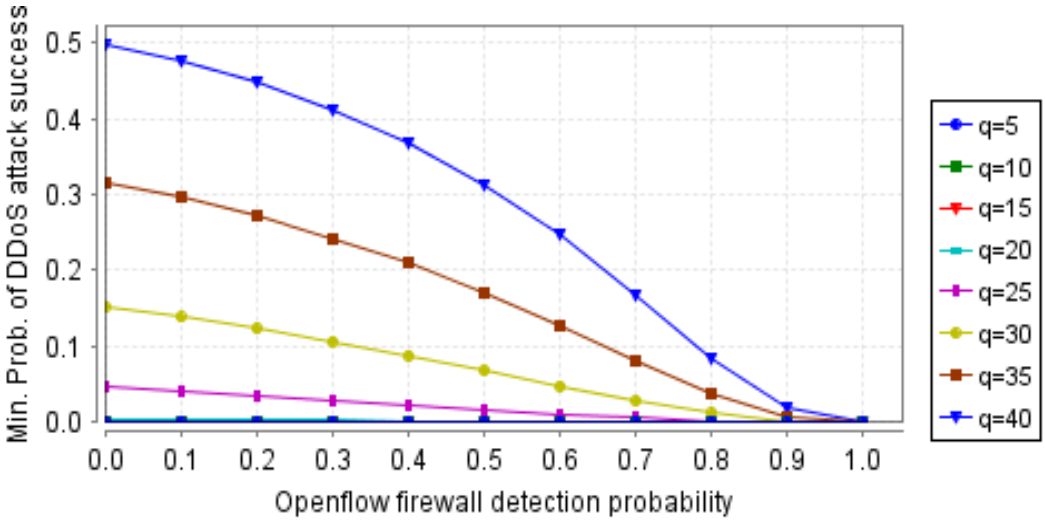

Fig. H.7: Minimum probability of DDoS attack success on AMI_Server3 (low q values)

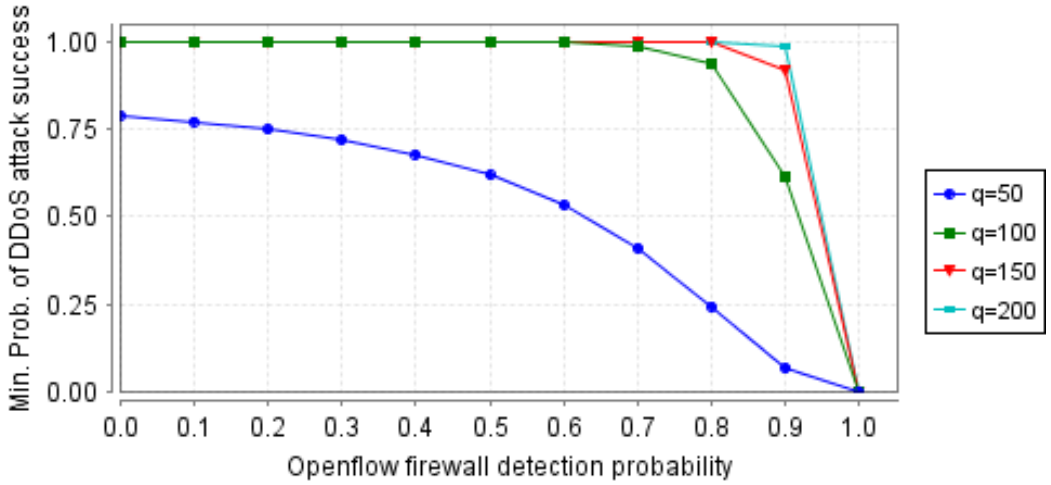

Fig. H.8: Minimum probability of DDoS attack success on AMI_Server3 (high q values)

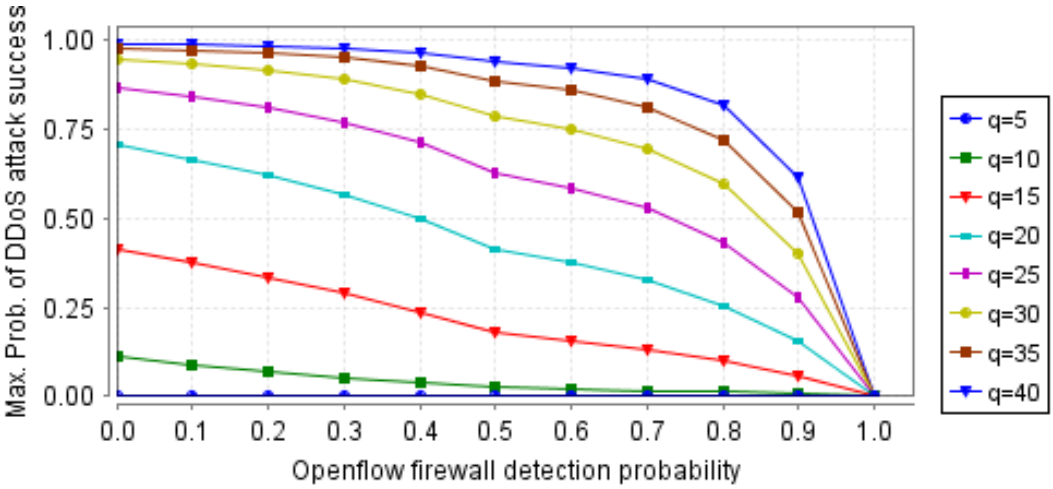

Fig. H.9: Maximum probability of DDoS attack success on AMI_Server3 (low q values)

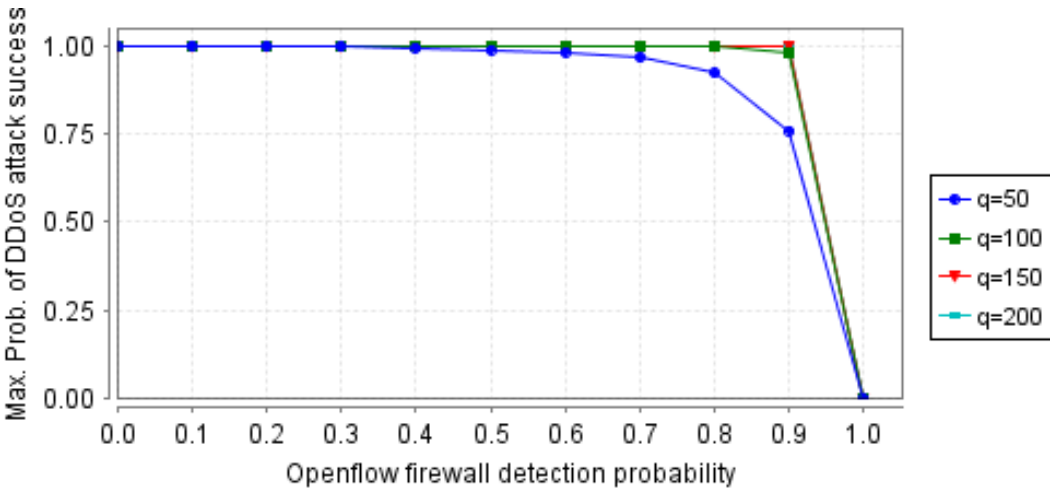

Fig. H.10: Maximum probability of DDoS attack success on AMI_Server3 (high q values)

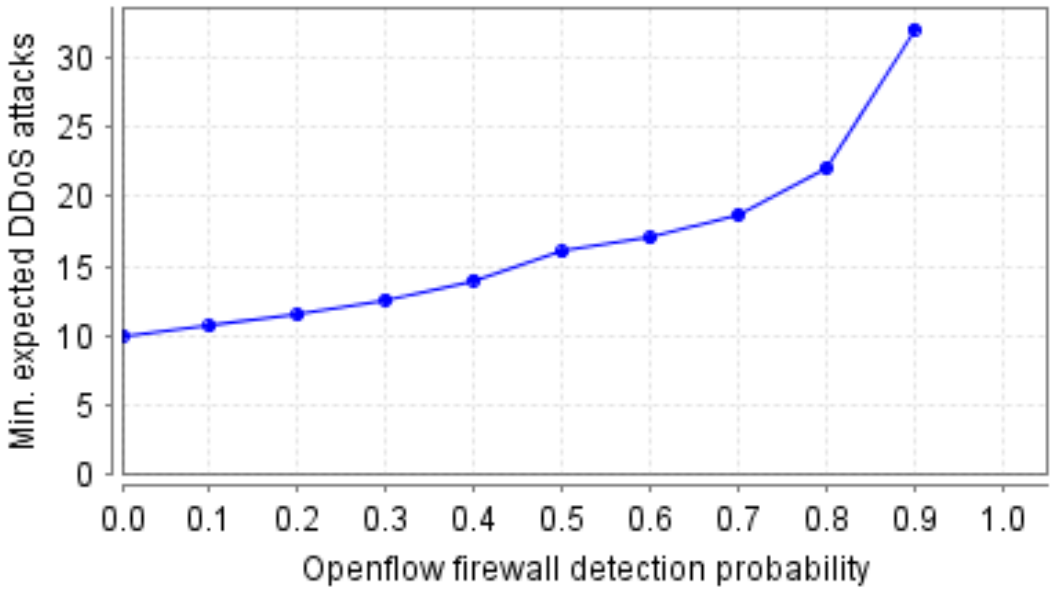

Fig. H.11: Minimum expected number of DDoS attacks until AMI_Server3 is attacked

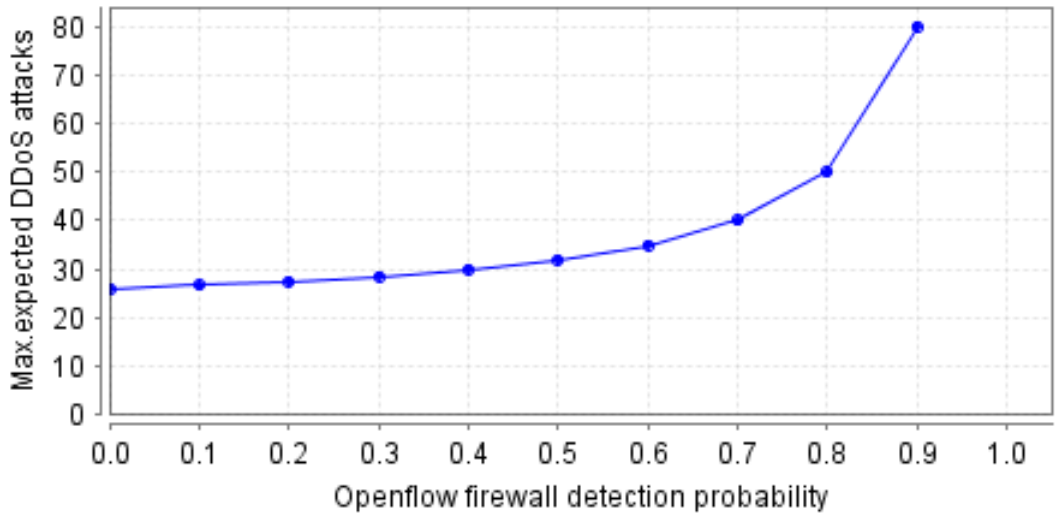

Fig. H.12: Maximum expected number of DDoS attacks until AMI_Server3 is attacked

6 Conclusion

Cyber threats targeting the smart grid AMI will most likely continue to evolve in the nearest future considering the distributed nature of information and communication technology (ICT) resources integrated into it. Firewalls remain one of the most important security defense mechanisms to be implemented in smart grid AMI network. There is therefore an important research need to validate that these firewalls perform according to specifications and design. In this paper, a quantitative analysis of a previously designed Openflow firewall under distributed denial of service (DDoS)

attacks in a smart grid AMI network has been carried using PRISM model checker. The analysis was carried out to ascertain the best- and worst-case performance of the firewall under different firewall's DDoS detection probabilities. From the analysis, the distinction between minimum and maximum (representing best- and worst-case) performance can be attributed to the non-deterministic choice available to DDoS zombies in their quest to attack a target node/s. The results from this study will be helpful in validating the results generated from the simulation of our previously designed cloud-based Openflow firewall for mitigation against DDoS attacks in smart grid AMI.

References

[1] S. McLaughlin, D. Podkuiko, S. Miadzvezhanka, A. Delozier, and P. McDaniel, "Multi-vendor penetration testing in the advanced metering infrastructure," in *Proceedings of the 26th Annual Computer Security Applications Conference*. ACM, 2010, pp. 107–116.

[2] E. B. Rice and A. AlMajali, "Mitigating the risk of cyber attack on smart grid systems," *Procedia Computer Science*, vol. 28, pp. 575–582, 2014.

[3] S. Asri and B. Pranggono, "Impact of distributed denial-of-service attack on advanced metering infrastructure," *Wireless Personal Communications*, vol. 83, no. 3, pp. 2211–2223, 2015.

[4] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Detection of invalid routing announcement in the internet," in *Dependable Systems and Networks, 2002. DSN 2002. Proceedings. International Conference on*. IEEE, 2002, pp. 59–68.

[5] D. Dittrich, P. Reiher, and S. Dietrich, *Internet Denial of Service: Attack and Defense Mechanisms*. Pearson Education, 2004.

[6] B. Fang, X. Yin, Y. Tan, C. Li, Y. Gao, Y. Cao, and J. Li, "The contributions of cloud technologies to smart grid," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1326–1331, 2016.

[7] A. Califano, E. Dincelli, and S. Goel, "Using features of cloud computing to defend smart grid against ddos attacks," in *10th Annual symposium on information assurance (Asia 15), ALBANY*, 2015, pp. 44–50.

[8] R. Diovu and J. Agee, "A cloud-based openflow firewall for mitigation against ddos attacks in smart grid ami networks," in *PowerAfrica, 2017 IEEE PES*. IEEE, 2017, pp. 28–33.

[9] (Accessed 11 July, 2017) [online]: Prism model checker- available at: https://www.prismmodelcheker.com.

[10] A. Hinton, M. Kwiatkowska, G. Norman, and D. Parker, "Prism: A tool for automatic verification of probabilistic systems," in *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 2006, pp. 441–444.

[11] M. Q. Ali and E. Al-Shaer, "Probabilistic model checking for ami intrusion detection," in *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*. IEEE, 2013, pp. 468–473.

[12] M. Mohsin, M. U. Sardar, O. Hasan, and Z. Anwar, "Iotriskanalyzer: A probabilistic model checking based framework for formal risk analytics of the internet of things," *IEEE Access*, vol. 5, pp. 5494–5505, 2017.

[13] T. Deshpande, P. Katsaros, S. Basagiannis, and S. A. Smolka, "Formal analysis of the dns bandwidth amplification attack and its countermeasures using probabilistic model checking," in *High-Assurance Systems Engineering (HASE), 2011 IEEE 13th International Symposium on*. IEEE, 2011, pp. 360–367.

[14] S. Basagiannis, P. Katsaros, A. Pombortsis, and N. Alexiou, "A probabilistic attacker model for quantitative verification of dos security threats," in *Computer Software and Applications, 2008. COMPSAC'08. 32nd Annual IEEE International*. IEEE, 2008, pp. 12–19.

[15] K. Salah, K. Elbadawi, and R. Boutaba, "Performance modeling and analysis of network firewalls," *IEEE Transactions on network and service management*, vol. 9, no. 1, pp. 12–21, 2012.

[16] Y. Guo, C.-W. Ten, S. Hu, and W. W. Weaver, "Modeling distributed denial of service attack in advanced metering infrastructure," in *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society*. IEEE, 2015, pp. 1–5.

[17] K. I. Sgouras, A. D. Birda, and D. P. Labridis, "Cyber attack impact on critical smart grid infrastructures," in *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*. IEEE, 2014, pp. 1–5.

[18] M. Kwiatkowska, G. Norman, D. Parker, and M. G. Vigliotti, "Probabilistic mobile ambients," *Theoretical Computer Science*, vol. 410, no. 12-13, pp. 1272–1303, 2009.

[19] I. Foster and C. Kesselman, *Carl Kesselman. The Grid: Blueprint for a Future Computing Infrastructure*. Morgan Kaufmann, 1999.

[20] E. Schikuta and T. Weishaupl, "N2grid: neural networks in the grid," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*, vol. 2. IEEE, 2004, pp. 1409–1414.

[21] L. Ying, Z. Liu, D. Towsley, and C. H. Xia, "Distributed operator placement and data caching in large-scale sensor networks," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 977–985.

[22] R. Moreno and A. Torres, "Network topological analysis to assess the power smart grid," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*. IEEE, 2012, pp. 1–5.

Chapter X

Final Conclusion and Possible Future Works

Final Conclusion

The main focus of this study was to propose novel solutions that can enhance security and privacy in the smart grid advanced metering infrastructure network. Solutions proposed in this study would make the smart grid AMI network to be more resilient against cyber-attacks, threats, and privacy violations so that the full benefits and potentials of the AMI can be fully realized. Consequently, the objectives of the study were set in a way that the main aim of the study can be realized. In chapter two of this study, an extensive review of data aggregation schemes for the SG AMI was carried out. In this chapter presented as Paper A, it was projected that data aggregation of consumption metering data will continue to be a popular research approach for a drastic reduction in privacy violations of the privacy of consumers in an SG AMI network. However, it was discovered that this approach increases the overhead in communication and computations for the network. This increase in overhead in computations and communications can drive the network to a congestive scenario, which can aid a cyber-criminal with the intention of launching data availability attack against the network. In the chapter three presented as Paper B, a systematic review of the state-of-the-art research on ZigBee based SG AMI was carried out with a major focus on security. From the security analysis, it was discovered that the security strength of ZigBee can be adjudged to be strong. However, it was noticed from this part that there are still issues of vulnerabilities, which must be addressed so that the full potentials of ZigBee based AMI network can be realized. In chapter four presented as Paper C, it was discovered that leveraging on cloud computing can impact positively on the smart grid AMI network with regards to an increased capacity for computations and storage. It was also noted that cloud computing can be utilized in enhancing the security of the SG AMI network. It was then concluded

that future security challenges facing the SG AMI can be resolved reasonably by leveraging on cloud computing technology.

The problem of congestion resulting from an increased burden on computations and communication during data aggregation in an SG AMI was tackled in chapter five (and presented as Paper D). In this regard, a robust communication architecture known as Ring Triangulation Communication Architecture (RTCA) was proposed for a secure transfer of energy user-consumption data during data aggregation in a Wi-Fi-based SG AMI network. It was shown that this proposed scheme can be reliable and efficient in delivering sufficient quality of service (QoS) to the network. Similarly, in chapter six (paper E), a congestion minimizing scheme for enhanced data aggregation in ZigBee based SG AMI was proposed. The proposed scheme utilizes the designed Ring Triangulation Communication Architecture (RTCA) which was reconfigured to be compatible with the ZigBee protocol standards. It was shown that the RTCA congestion minimizing scheme (RTCA_CMS) can be used to relieve transmission congestion and consequently, reduce overheads incurred during data aggregation in the network.

In chapter seven which contains Paper F, a Cloud-Based OpenFlow Firewall for mitigation against distributed denial of service (DDoS) attacks in smart grid AMI networks is presented. It was clearly shown in this paper that the OpenFlow firewall can be utilized to improve the quality of service (QoS) of the SG AMI in an attack scenario. Simulations results showed that with a 250 Gbps voluminous DDoS attack focused on the network, the OpenFlow firewall can provide instant detection and mitigation solutions against vulnerable points on the network. It was concluded that since volumetric DDoS attacks would likely continue to be a target against the smart grid AMI in the near future, the SG AMI would require stronger security control which can be implemented using the cloud-based firewall at a foundational level.

In chapter eight, this study presented another cloud-based system model for the smart grid AMI that can enhance the security of the network. It is expected that the security derivable from this model can be guaranteed at a future period when most security solutions for the SG AMI would have been rendered obsolete by the emergence of quantum computers in commercial quantities. This cloud-based SG AMI network model (presented as Paper G) was designed to leverage on the security features of quantum key distribution (QKD). The proposed model is unique because all the incorporated cryptographic protocols are all based on symmetric cryptosystems unlike most similar protocols in use today which are based on asymmetric cryptosystems. An added advantage of the incorporated protocols is that the overheads incurred as a result of these protocols are far lighter than most similar protocols based on asymmetric cryptosystems. The paper also showed through

simulations that QKD can offer information theoretic security which is guaranteed by some basic laws of Physics such as Heisenberg uncertainty principle and the no-cloning theorem.

Finally, a quantitative analysis of firewall security under a distributed denial of service (DDoS) attacks in smart grid AMI is presented in chapter nine (Paper H). This analysis was carried out using a mathematical tool known as PRISM model checker. The analysis was carried out in this paper to ascertain the best or the worst-case performance of a security firewall designed to mitigate against DDoS attacks under different firewall's detection probabilities. It was concluded that results from experiments conducted in this study, could be used to validate the simulation results of the designed OpenFlow firewall.

Finally, it can be concluded that previous approaches to resolving issues of privacy and security in an SG AMI are not very effective as most research approaches concentrated on the design of cryptographic algorithms. It was discovered from an extensive survey of the related literature that majority of previously proposed solutions cannot be applied to real AMI network while others can bring additional burden in terms of computation, processing, and communications. The study carried out in this thesis have combined different methodologies of information security like cryptographic algorithms, infrastructural security like firewalls and the design of enhanced data aggregation protocols. The study conducted in this thesis also showed that majority of the previously proposed aggregation protocols were not applied to typical communication networks recommended for the SG AMI. This gap was bridged in this thesis by applying the enhanced data aggregation protocols proposed in this study to both Wi-Fi and ZigBee network standards.

Possible Future Works

In the course of this research work, a number of research areas were noted, which are worthy to be taken for a further investigative study.

ZigBee-Based Smart Grid Advanced Metering Infrastructure Network

It can be noted from this study that the security issues resulting from the trust center in a ZigBee-based SG AMI can be considered for a further research. This is because of the so much responsibility on the trust center which can makes it a single point of failure on the network. According to the ZigBee standard, the trust center acts as a trust manager, network manager, and a configuration manager. Optimization models can be developed such that the memory and energy constrained trust center

nodes can be improved especially in the event of network expansion. Appropriate key distributions schemes which can lessen the burden on the trust center can also be designed.

Cloud-Based Smart Grid Advanced Metering Infrastructure Network

The potentials of cloud-based solutions for the SG AMI can be exploited further in a number of ways. First, the openflow firewall designed in this thesis can be improved to tackle cyber-security threats targeting integrity and confidentiality.

It can be very interesting to note that energy theft which has been an age long problem for the traditional electrical power grid would be complicated by the mass deployment of the smart grid advanced metering infrastructure (SG AMI). Therefore, a security solution for the SG AMI can be effectively designed to leverage cloud computing techniques that can be utilized for effective energy theft detection. Therefore, cloud-based techniques for energy theft detection is a clear direction for future research.

Finally, the cloud-based OpenFlow firewall designed in this study can be extended by including features that can be utilized for quick and effective detection of malicious software within the SG AMI network. This can be a good research direction for future research.

Data Aggregation Schemes for the SG AMI

The Ring Triangulation Communication Architecture can be updated for an extended comparative analysis. In this analysis, QoS affected by congestion in the SG AMI network will be analyzed when data aggregation scheme presented in this thesis, some selected data aggregation schemes and TCP congestion control schemes are compared.