

4-2016

SECURING THE INTEGRITY OF THE POWER DISTRIBUTION SYSTEM FOR SMART GRID APPLICATIONS

Hosam Yousef Hittini

Follow this and additional works at: https://scholarworks.uaeu.ac.ae/all_theses



Part of the [Systems Architecture Commons](#)

Recommended Citation

Hittini, Hosam Yousef, "SECURING THE INTEGRITY OF THE POWER DISTRIBUTION SYSTEM FOR SMART GRID APPLICATIONS" (2016). *Theses*. 335.

https://scholarworks.uaeu.ac.ae/all_theses/335

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarworks@UAEU. It has been accepted for inclusion in Theses by an authorized administrator of Scholarworks@UAEU. For more information, please contact mariam_aljaberi@uaeu.ac.ae.



جامعة الإمارات العربية المتحدة
United Arab Emirates University

United Arab Emirates University

College of Information Technology

Information Security Track

SECURING THE INTEGRITY OF THE POWER DISTRIBUTION SYSTEM FOR
SMART GRID APPLICATIONS

Hosam Yousef Hittini

This thesis is submitted in partial fulfillment of the requirements for the degree of
Master of Science in Information Security

Under the Supervision of Professor Liren Zhang

April 2016

Declaration of Original Work

I, Hosam Yousef Hittini, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled “*Securing the Integrity of the Power Distribution System for Smart Grid Applications*”, hereby, solemnly declare that this thesis is my own original research work that has been done and prepared by me under the supervision of Professor Liren Zhang, in the College of Information Technology at UAEU. This work has not previously been presented or published, or formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

Student's Signature _____

Date _____

Copyright © 2016 Hosam Yousef Hittini
All Rights Reserved

Advisory Committee

1) Advisor: Liren Zhang

Title: Professor

Networking Track

College of Information Technology

2) Co-advisor: Atef Abdrabou

Title: Associate Professor

Department of Electrical Engineering

College of Engineering

Approval of the Master Thesis

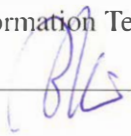
This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Liren Zhang

Title: Professor

Department of Networking

College of Information Technology

Signature 

Date 28/4/2016

2) Member: Atef Abdrabou

Title: Associate Professor

Department Of Electrical Engineering

College of Engineering

Signature 

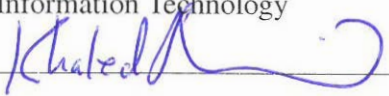
Date 28/4/2016

3) Member: Khaled Shuaib

Title: Professor

Department of Information Security

College of Information Technology

Signature 

Date 28/4/2016

4) Member (External Examiner): Monther Aldwairi

Title: Associate Professor

College of Technological Innovation

Institution: Zayed University

Signature 

Date 28/4/2016


This Master Thesis is accepted by:

Dean of the College of Information Technology: Professor Omar El-Gayar

Signature 

Date May 25, 2016

Dean of the College of Graduate Studies: Professor Nagi T. Wakim

Signature 

Date 25/5/2016

Copy 6 of 9

Abstract

The distribution system is one of the main components in a smart grid, readings are transferred from the distribution substations to the control center. Compromising transferred system data will result in drawing wrong conclusions about current operation status at the control center. Which leads to sending wrong operational commands, that may result in very serious consequences.

Firstly, we propose a scalable communications architecture for future smart grid distribution systems (i.e. Security Aware Distribution System Architecture - SADSA). The architecture is adaptable to use WiFi or other technologies to transfer smart grid information. The architecture is studied from various angles. Both communication and cybersecurity challenges are extracted. In addition, the work provides a detailed discussion on how the proposed architecture meets National Institute of Standards and Technology (NIST) cybersecurity requirements for smart grids.

Secondly, we propose the False Data Injection Prevention Protocol - FDIPP, the protocol prevents packet injection, duplication, alteration and node replication. In other words, it guarantees both system and data integrity. The protocol was analyzed using formal security analysis. Furthermore, Network Simulator 2 is used to evaluate both SADSA and FDIPP. The simulation is used to measure the delay and security overhead introduced from FDIPP and the proposed architecture.

Keywords: Smart Grid Security, Power Distribution System, Security Architecture, Security Analysis, Communication Network Security, Data Integrity, System Integrity, False Data Injection, Node Replication, Scyther, Network Simulator 2.

Title and Abstract (in Arabic)

حماية سلامة البيانات لنظام التوزيع في شبكة الكهرباء الذكية

الملخص

يعدّ نظام التوزيع من أهم أجزاء شبكة الكهرباء الذكية، ويتم نقل القراءات من هذا النظام إلى مركز التحكم، وقد يتسبب اختراق بيانات النظام في إعطاء استنتاجات خاطئة في مركز التحكم. مما يؤدي إلى إرسال تعليمات خاطئة، ومن المحتمل أن يتسبب ذلك في عواقب وخيمة.

في هذه الدراسة نطرح، أولاً، بنية اتصالات لاسلكية قابلة للتوسع والتطوير يمكن استخدامها لنظام التوزيع في شبكة الكهرباء الذكية. لقد أطلقنا اسم [سادسا] على هذه البنية التي تتميز بقدرتها على التكيف مع استخدام تقنية الواي فاي أو تقنيات تواصل لاسلكية أخرى. وقد تمت دراسة هذه البنية من جوانب متعددة، كما تم استخراج تحديات التواصل، والأمن السيبراني الإلكتروني. وبالإضافة إلى ذلك، فإنّ هذا العمل يناقش بالتفصيل مدى توافق البنية مع متطلبات المعهد الوطني للمعايير والتقنية لأمن معلومات شبكات الكهرباء الذكية.

ثانياً، نقدم في هذه الدراسة بروتوكولاً بعنوان إف ديب يحمي شبكات الكهرباء الذكية من لعمليات غير المشروعة كحقن وتكرار وتعديل الرسائل. كما أنه يحمي من التقليد المطابق لنقاط التواصل اللاسلكية. بعبارة أخرى، فإنّ إف ديب يضمن سلامة البيانات المرسلة وسلامة النظام معاً. لقد تم تحليل هذا البروتوكول باستخدام طريقة رسمية لدراسة أمن البروتوكولات. وبالإضافة إلى ذلك، فقد تمت دراسة أداء البنية والبروتوكول المقدمين باستخدام برنامج ان اس ٢ الذي باستطاعته محاكاة شبكات التواصل اللاسلكية، حيث تهدف دراسة الأداء إلى ملاحظة التأخير الناتج عن البنية والبروتوكول.

مفاهيم البحث الرئيسة : أمن شبكة الكهرباء الذكية، نظام توزيع الكهرباء، بنية آمنة، تحليل أمن المعلومات، حماية شبكة الاتصال، سلامة البيانات، سلامة النظام، حقن البيانات غير المشروع، التقليد المطابق لنقاط التواصل، محاكاة.

Acknowledgements

I would like to thank my thesis advisor Prof. Liren Zhang from College of Information Technology at UAE University. His extended years of experience and his unparalleled knowledge were of a great help during the development of this publication. He always steered me in the right direction whenever I approached him for help.

I must also acknowledge the support I received from my thesis co-advisor Dr. Atef Abdrabou from Department of Electrical Engineering at UAE University. His door was always open to answer my questions regarding both research and writing. We had extended discussions and he guided me all the way to produce this work.

My special thanks go to Prof. Khaled Shuaib and Dr. Monther Aldwairi from the examination committee. I'm very grateful for their valuable comments on my thesis. Their thoughtful feedback definitely improved the way this work is presented.

Dedication

To my beloved family, professors, and friends

Table of Contents

Title	i
Declaration of Original Work	ii
Copyright	iii
Advisory Committee	iv
Approval of the Master Thesis	v
Abstract	vii
Title and Abstract (in Arabic)	viii
Acknowledgments	x
Dedication	xi
Table of Contents	xii
List of Tables	xv
List of Figures	xvi
List of Abbreviations, Nomenclatures, and Symbols	xvii
Chapter 1: Introduction	1
Chapter 2: Literature Review	5
2.1 Background	5
2.2 Related Work	10
Chapter 3: Security Aware Distribution System Architecture	12
3.1 Overview	12
3.2 Communications in SADSA	14
3.3 Communication Challenges	16

3.3.1	Low Communication Overhead	16
3.3.2	Scalability	16
3.3.3	Reliability	17
3.3.4	Delay Introduced by Medium Access Control	17
3.3.5	Time Sensitive Protocols	17
3.3.6	Interoperability	18
3.4	Cybersecurity Awareness	18
3.4.1	Concurrent Session Control	19
3.4.2	Remote Session Lock and Termination	19
3.4.3	Permitted Actions without Identification or Authentication . .	20
3.4.4	Remote Access	20
3.4.5	Wireless Access Restrictions	21
3.4.6	User Identification and Authentication	22
3.4.7	Device Identification and Authentication	22
3.4.8	Denial-of-Service Protection	23
3.4.9	Authenticator Feedback	23
3.4.10	Security Function Isolation	24
3.4.11	Boundary Protection	24
3.4.12	Communication Integrity	25
3.4.13	Application Partitioning	25
3.4.14	Software and Information Integrity	26
3.5	Summary	26
Chapter 4:	False Data Injection Prevention Protocol	27
4.1	Background	27
4.1.1	The used EAP	27
4.1.2	Assumptions	28
4.2	FDIPP	28
4.2.1	High Level Protocol Operation	28
4.2.2	Node Authentication	30

4.2.3	Peer Authentication	33
4.2.4	Post Authentication Data Transfer	34
4.3	Key Management	35
4.4	Security Analysis	35
4.4.1	Configuration and Claims	36
4.4.2	Results	36
4.5	Conclusion	38
Chapter 5: Performance Evaluation		39
5.1	Simulation Design	39
5.1.1	Phase 1	39
5.1.2	Phase 2	40
5.2	Simulation Results	41
5.2.1	Phase 1	41
5.2.2	Phase 2	43
5.2.3	End to End Performance	45
5.2.4	Summary	46
Chapter 6: Conclusion		47
Bibliography		48

List of Tables

Tabel 2.1: IEC 61850 Delay Constraints [1]	7
Tabel 5.1: Average Delay Summary	45
Tabel 5.2: End to End Average Delay	46
Tabel 5.3: End to End Average Packet Loss	46

List of Figures

Figure 2.1: Distribution System Components	7
Figure 2.2: Security Evaluation Approaches [2]	9
Figure 3.1: Security Aware Distribution System Archeticture	13
Figure 4.1: False Data Injection Prevention Protocol	30
Figure 4.2: Node Authentication Sequence Diagram	33
Figure 4.3: Peer Authentication Sequence Diagram	34
Figure 4.4: Syther Analysis of FDIPP Peer Authentication Protocol	37
Figure 5.1: Average Delay in Phase 1	42
Figure 5.2: Average Packet Loss in Phase 1	43
Figure 5.3: Average Delay in Phase 2	44
Figure 5.4: Average Packet Loss in Phase 2	45

List of Abbreviations, Nomenclatures, and Symbols

AES	Advanced Encryption Standard
AODV	Ad hoc On-Demand Distance Vector
AP	Access Point
AS	Authentication Server
<i>ASID</i>	Authentication Server ID
CBR	Constant BitRate
CCC	Control Center Cloud
<i>CK</i>	Cloud Key
CSMA	Carrier Sense Multiple Access
DoS	Denial of Service
EAP	Extensible Authentication Protocol
FDIPP	False Data Injection Prevention Protocol
GW	Gateway
HBFW	Host Based Firewall
HMAC	keyed-Hash Message Authentication Code
IEC	International Electrotechnical Commission
IED	Intelligent Electronic Devices
IKE	Internet Key Exchange
IPS	Intrusion Prevention System
k	Long term key
N_{AS}	Authentication server nonce

N_R	Router's nonce
NIST	National Institute of Standards and Technology
NS2	Network Simulator 2
P1	Phase 1
P2	Phase 2
PS	Primary Substation
PSC	Primary Substation Cloud
$PubAS$	Authentication server public key
$PubR$	Router public key
pw	Password
R1	Router 1
RFC	Request For Comments
RID	Router ID
RSA	Rivest, Shamir, and Adelman
RTU	Remote Terminal Units
SADSA	Security Aware Distribution System Architecture
SCADA	Supervisory Control And Data Acquisition
SHA-3	Secure Hash Algorithm 3
SK	Session Key
SS	Secondary Substation
SSBC	SecondarySubstationBackboneCloud
SSC	Secondary Substation Cloud
TK	Temporal key

U	User
UDP	User Datagram Protocol
WIDS	Wireless Intrusion Detection System
WiFi	Wireless Fidelity
y	Random one time key

Chapter 1: Introduction

Smart grids are the next generation power systems, they aim at running the regular power system over a smart communication architecture. Smart grids integrate the usage of solar energy and distributed energy generation. Implementing and using smart grids is envisioned to increase efficiency and reliability of power systems. However, achieving this functionality requires using smart devices, which will introduce many vulnerabilities. Exploiting such vulnerabilities is much more critical than exploiting vulnerabilities in regular networks. Because a successful attack on a smart grid can cause a power outage for large areas and, in turn, results in huge financial losses.

The distribution system is one of the major components in legacy power grids. It is responsible for delivering power to end consumers. There are two types of distribution substations, namely, primary substation (PS) and secondary substation or (SS). The latter is connected to consumers from one side and to a primary substation from the other. Every group of secondary substations is connected to one primary substation. Therefore, there is a large number of secondary substations, which are highly distributed by nature.

Acquiring the correct data from secondary substations with appropriate commands is important for future smart grid operation. Receiving wrong information from secondary substations may result in making inappropriate decisions, and sending wrong operational commands, which may cause severe consequences. Researchers have identified many security attacks, which can cause fires, hurt people, or even cause blackout for a whole city. Authors of [3] describe an attack that results in preventing legitimate status messages from being delivered to the control center. The authors in [4] discuss a false data injection attack that will affect the State Estimation System, a system that allows getting an accurate estimate about the power system status, which is a part of Energy Management System. They also identified the minimum number

of compromised nodes needed to cause an unobservable data attack to SCADA system in the control center. In [5], authors show how compromising one control center can result in causing denial of service for state of the art State Estimation system. In addition, using specific technologies such as WiFi or ZigBee for communication in smart grids introduces vulnerabilities related to these technologies.

Before looking at the security service we are targeting in this work, let's define the three core aspects of security, namely, confidentiality, integrity, and availability. Confidentiality is basically ensuring private information is not disclosed to an unauthorized party. Integrity is ensuring that the information sent was not modified on its way in an unauthorized manner and making sure that the system is functioning free from any unauthorized manipulation. Availability is ensuring that users are able to access the service in a timely manner [6].

Confidentiality service is achieved using encryption. There are two types of encryption, symmetric and asymmetric. In symmetric encryption there is a secret key that is used for both encryption and decryption. On the other hand, in asymmetric crypto-systems, every entity has two keys, one is private and one is public. As the names imply, the public key is shared with everyone and it is used for encryption, while the encrypted ciphertext can be only decrypted by using the private key [6].

Altering the content of a transmitted packet is considered a violation of data integrity, whereas, having a node that does not belong to the system injecting data is considered a violation of both data and system integrity. Data integrity can be improved by employing one way hash functions such as SHA-3 [7]. Any change in data will result in changing its hash value, thus, the attempt will be detected. On the other hand, authentication improves system integrity i.e. it denies unauthorized users from accessing the system assets.

To increase security of smart grids, NIST has developed Guidelines for smart grid cybersecurity [8]. The document shows a remarkable effort that can significantly reduce vulnerabilities. It has defined cybersecurity requirement development guide-

lines for communication interfaces between different smart grid domains. The document will be discussed in more details in section 2.1.

On the other hand, communication challenges in smart grids are considered another major concern. Smart grids in their nature use heterogeneous device types, protocols, and communication technologies. Thus, achieving seamless integration between all smart grid components is considered a great challenge. Furthermore, it is known that the power distribution system is made of a huge number of nodes. All these nodes need to communicate with each other in the future smart grid. In addition, the network should be reliable to allow the desired functionality. Thus, scalability and reliability of the underlying communication architecture should be taken into account. Moreover, given the number of nodes in a smart grid communication system, key management should be considered as one of the challenges.

In this work, we focus on the integrity of the distribution system. Although the proposed protocol FDIPP encrypts packets, confidentiality and availability are considered out of the scope of this thesis. Furthermore, security analysis of all the used protocols except FDIPP is considered out of scope as well. That's because the security of all these protocols was verified by their authors. The contributions of this work are as follows:

1. **Security Aware Distribution System Architecture - SADSA:** It's a scalable, adaptable, and security aware communication architecture for smart grid distribution system. Both communication and cybersecurity challenges of the proposed architecture are highlighted. In addition, detailed analysis of the security awareness of SADSA is introduced based on NIST smart grid security requirements.
2. **False Data Injection Prevention Protocol - FDIPP:** FDIPP is an authentication and key management protocol. FDIPP prevents packet injection, duplication, alteration and node replication. Scyther tool is used for the formal security analysis of FDIPP.

3. Performance Evaluation: NS2 software is used to study the delay and security overhead resulting from both SADSA and FDIPP.

Addressing the contributions listed beforehand helps in addressing two active areas in the literature. Firstly, we propose a security aware smart grid communication architecture and analyze the requirements with respect to the proposed architecture. Secondly, we propose a protocol that completely fits the underlying communication architecture. Moreover, the key management algorithm significantly reduces the overhead resulting from key transfers, but still does not use encryption or hashing keys for more than one hour. As a result, the protocol meets forward secrecy property as well. In addition, simulation results help in determining the correct number of nodes per cluster to achieve predefined delay requirements.

The rest of this thesis is organized as follows: Chapter 2 covers the literature review, it covers both background and related work. After that, Chapter 3 covers the Security Aware Distribution System Architecture. It describes the architecture and the communication challenges. It also discusses the security awareness of SADSA and how it aligns with NIST cybersecurity guidelines. Then, Chapter 4 describes the False Data Injection Prevention Protocol. Additionally, it covers the detailed protocol operation, the key management algorithm, and FDIPP security analysis. Chapter 5 addresses the performance evaluation of SADSA and FDIPP. Moreover, it covers both the simulation design and the acquired results. Finally, Chapter 6 concludes this work and highlights future work.

Chapter 2: Literature Review

2.1 Background

The National Institute of Standards and Technology developed a conceptual architecture for smart grids. According to this model, the smart grid is made of 7 domains as follows [9]:

1. Customer: This covers the end users of electricity, but they may generate and store their own electricity using distributed energy resources.
2. Markets: It defines the price of electricity based on supply and demand levels. Dynamic pricing is a component of this domain.
3. Service Provider: The entity that provides electricity to customers.
4. Operations: Managing electricity movement.
5. Generation: Generation and storage of electricity, it may include distributed energy resources as well.
6. Transmission: Carrying bulk electricity over long distances.
7. Distribution: Delivering electricity to and from customers.

Here, we focus on the distribution system, which allows delivering electricity from distribution domain to customers. Electricity passes through substations over its journey to the customers, namely, primary and secondary substations; these substations are used to step down voltage to suit customer needs and they are referred to as PS and SS respectively. Components of distribution domain interact with other domains, such as Operations and Customer domains. These components are called actors in the smart grid arena. Actors are simply software and hardware systems that participate and play a significant role in the smart grid. Actors of the distribution domain are listed below

along with their definitions[8]:

- Distribution Sensors: Devices that measure physical quantities and send them as digital signals to be used by other actors in the system.
- Remote Terminal Units and Intelligent Electronic Devices - RTU and IED: Receive information from various sensors and send commands accordingly.
- Distribution Data Collector: A system that collects data from different sources and modify or transfer these data.
- Distributed Intelligence Capabilities: Autonomous applications that operate separate from centralized control to increase responsiveness and reliability of the system.
- Geographic Information System: A management system that provides asset information and status for other advanced applications.
- Field Crew Tools: Maintenance handheld tools that are used for field engineering.

The described actors are shown in Figure 2.1. Geographic Information System and Field Crew Tools are out of the scope of this work.

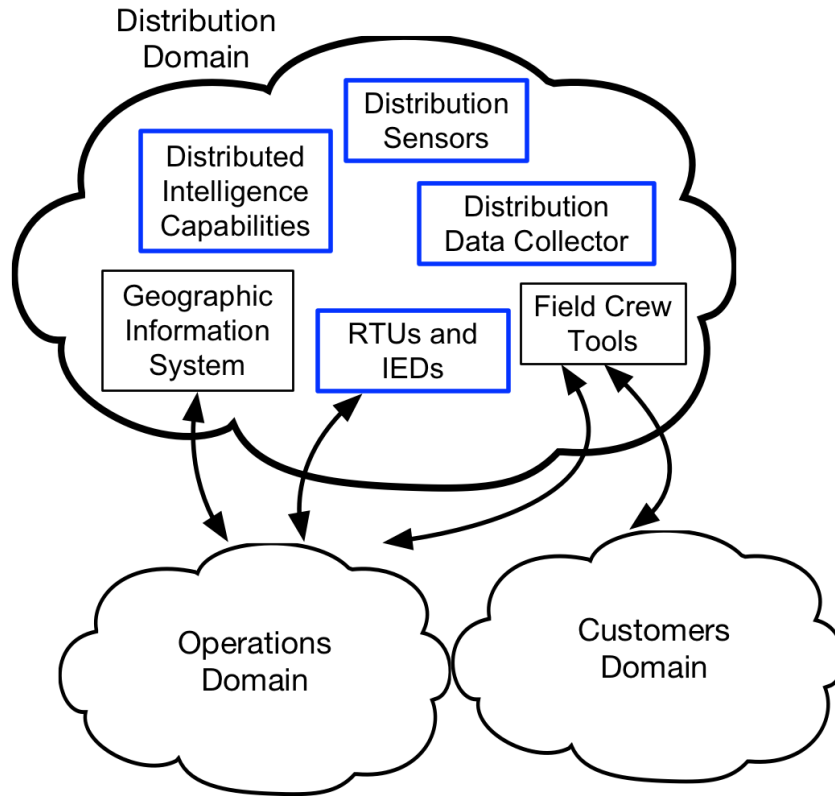


Figure 2.1: Distribution System Components

Communication between substations has specific delay requirements based on the application. These requirements are defined as constraints in IEC 61850 as shown in Table 2.1 [1]. Messages of Type 1A have very strict time requirements because they are used for protection and fault isolation. Type 1B messages are used for traditional communications between different systems. Finally, use of Type 2 and Type 3 messages is restricted to less critical messages such as reading from substations. Because of the importance of such requirements, they should be planned before building any communication network supporting smart grid distribution system.

Message type	Delay constraint(ms)	Usage
Type 1A/P1	3	Fault isolation and protection
Type 1A/P2	10	
Type 1B/P1	100	Routine communication
Type 1B/P2	20	
Type 2	100	Monitoring and readings transfer
Type 3	500	

Table 2.1: IEC 61850 Delay Constraints [1]

A smart grid distribution system is a large scale system because of the number of its components and their high density. Key management includes key generation, distribution, storage, and revocation. Clearly, for a system with this size, key management is needed. Furthermore, maintaining security services is not a simple task because attacks take different shapes. For example, node replication, packet injection, packet duplication, packet alteration, spoofing, and tampering are all considered integrity attacks and every one of them needs to be mitigated using a unique countermeasure.

NIST has developed a comprehensive report titled Guidelines for smart grid cybersecurity in [8] that covers smart grid cybersecurity guidelines. The report is divided into three volumes as follows:

- **Smart Grid Cybersecurity Strategy, Architecture, and High-Level Requirements:** It covers a systematic strategy for developing security architectures. It also covers the logical architecture developed by NIST and its components. It covers the related cybersecurity requirements as well.
- **Privacy and the Smart Grid:** It defines the privacy terms in smart grids. It determines which personal activities are within the scope of smart grids. After that, it discusses relevant issues, concerns, and risks.
- **Supportive Analyses and References:** It classifies vulnerabilities and highlights security problems in the smart grid. It also gives an overview on how to assess standards against security requirements.

The first volume was used in this work to analyze and verify security awareness of SADS. Although this volume maps cybersecurity requirements to the defined logical architecture, the architecture used is considered high-level because it does not cover the underlying communication network. In addition, the volume only covers definition of requirements. It does not cover how defined requirements are met.

Evaluation of security protocols plays a vital role in determining trustworthi-

ness of developed protocols. It's very hard for humans to identify flaws by inspection. Thus, formal security analysis methods can be used. Authors in [2] have compared multiple security evaluation approaches. As shown in Figure 2.2, evaluation approaches were categorized based on three factors. Namely, model checking or theorem proving, symbolic or cryptographic, and bound or unbound.

	Model checking	Theorem proving
Symbolic	NRL FDR AVISPA	Isabelle/HOL
Cryptographic		BPW(in Isabelle/HOL) Game-based Security Proof (in Coq)

Unbounded

Figure 2.2: Security Evaluation Approaches [2]

In model checking, algorithms are followed to verify security. Whereas proofs should be constructed to verify security in theorem proving. Theorem proving requires significant experience and effort. Dolev-Yao model is usually used for symbolic protocol representation and analysis. On the other hand, probability and complex theories are used for cryptographic analysis. Bounding the analysis limits the number of concurrent protocol sessions an attacker can have. Increasing the number of these sessions may allow an attacker to manipulate replay messages and exploit a weakness in the protocol [2].

Scyther [10] is a formal security analysis tool. It was developed by the authors in [11] based on PhD dissertation. It employs automatic security verification of protocols. It supports bound and unbound parallel sessions. Increasing the number of parallel sessions may exploit new introduced vulnerabilities. It was used to verify IKEv1 (i.e. Internet Key Exchange version 1) and IKEv2, and it found unreported weaknesses as shown in [12]. More details about this tool are covered in Chapter 4.

2.2 Related Work

Authors of [13] have only highlighted the general characteristics and high level requirements for the smart grid, but without a specific proposal. Authors of [14] and [15] have discussed possible communication technologies that can be used. They mainly discussed communication requirements without a focus on security-aware architectures. In [1], the authors surveyed cybersecurity requirements and threats, and evaluated these threats in different scenarios. However, both requirements and threats are not mapped to a specific architecture.

The authors of [16] targeted the data injection attacks on the energy management system, they proposed an algorithm that would find the minimum number of meters if controlled would result in an unobservable attack, but they did not provide any countermeasure. The authors of [17] provided an algorithm to detect integrity attacks if less than 5 meters were compromised, they also propose a countermeasure based on state estimation. However, it allows the attackers to inject data, and they will find their way around it sooner or later, it is better to prevent them from injecting anything at all. The solution doesn't cover the possibility of compromising more than 5 smart meters as well. Authors of [18] use homomorphic signature for aggregated data, the proposed method is computationally inexpensive, but it does not cover system integrity. In addition, it requires data aggregation, which is not always possible in WiFi networks because of data rate limitation. The authors of [19] use historic data to find any inconsistency, a patient attacker can trick this by modifying the data slowly enough to make it undetectable over a long period of time. The authors of [20] propose sending the data over a high speed IP network, and sending a watermark over a low bandwidth secure network, the nature of smart grids makes this solution costly and not scalable.

Authors of [21] have introduced a privacy aware smart grid communication architecture. The proposed architecture is mainly targeting secrecy and anonymity of customers. Furthermore, the authors don't cover performance evaluation. Authors of

[22] have defined information security risks and requirements; they also proposed a framework to support smart grids. Both defined risks and requirements are high level. For example, the authors looked at integrity, confidentiality, and availability requirements without going any deeper. In addition, the proposed framework is not based on a specific architecture. Furthermore, it's defined high level goals in the framework, such data leak prevention, but without describing a process for achieving these goals.

Chapter 3: Security Aware Distribution System Architecture

This chapter describes the proposed architecture, SADSA. At the beginning, an overview of SADSA is given, then communication in SADSA and communication challenges are described. After that, cybersecurity awareness of SADSA is analyzed.

3.1 Overview

The proposed distribution system architecture is shown in Figure 3.1, the scope of this architecture is to define the setup, communication link type and high level communication protocols (i.e. point to point, mesh...etc.) between the Control Center, Primary Substations, and Secondary Substations. Only the components with blue frame in Figure 2.1 are within the scope of this architecture. All nodes are organized to be a part of one or more of the following virtual clouds: Control Center Cloud, Primary Substation Cloud, Secondary Substation Backbone Cloud, or Secondary Substation Cloud. These clouds are described as follows:

- **Control Center Cloud - CCC:** It contains the SCADA system and the authentication server. One of the SCADA system's tasks is to monitor and control remote substations based on received readings. In addition, CCC has managed network interfaces to allow communication to other domains or external networks. Communication through these interfaces goes through a proxy. Furthermore, communication is protected by a firewall and an intrusion prevention system.
- **Primary Substation Cloud - PSC:** It contains all the primary substations; these substations belong to different geographical areas. Every PS is connected to the Secondary Substation Cloud using a minimum of two routers (e.g. R1 and R2) for redundancy.
- **Secondary Substation Backbone Cloud - SSBC:** It connects the gateways (e.g.

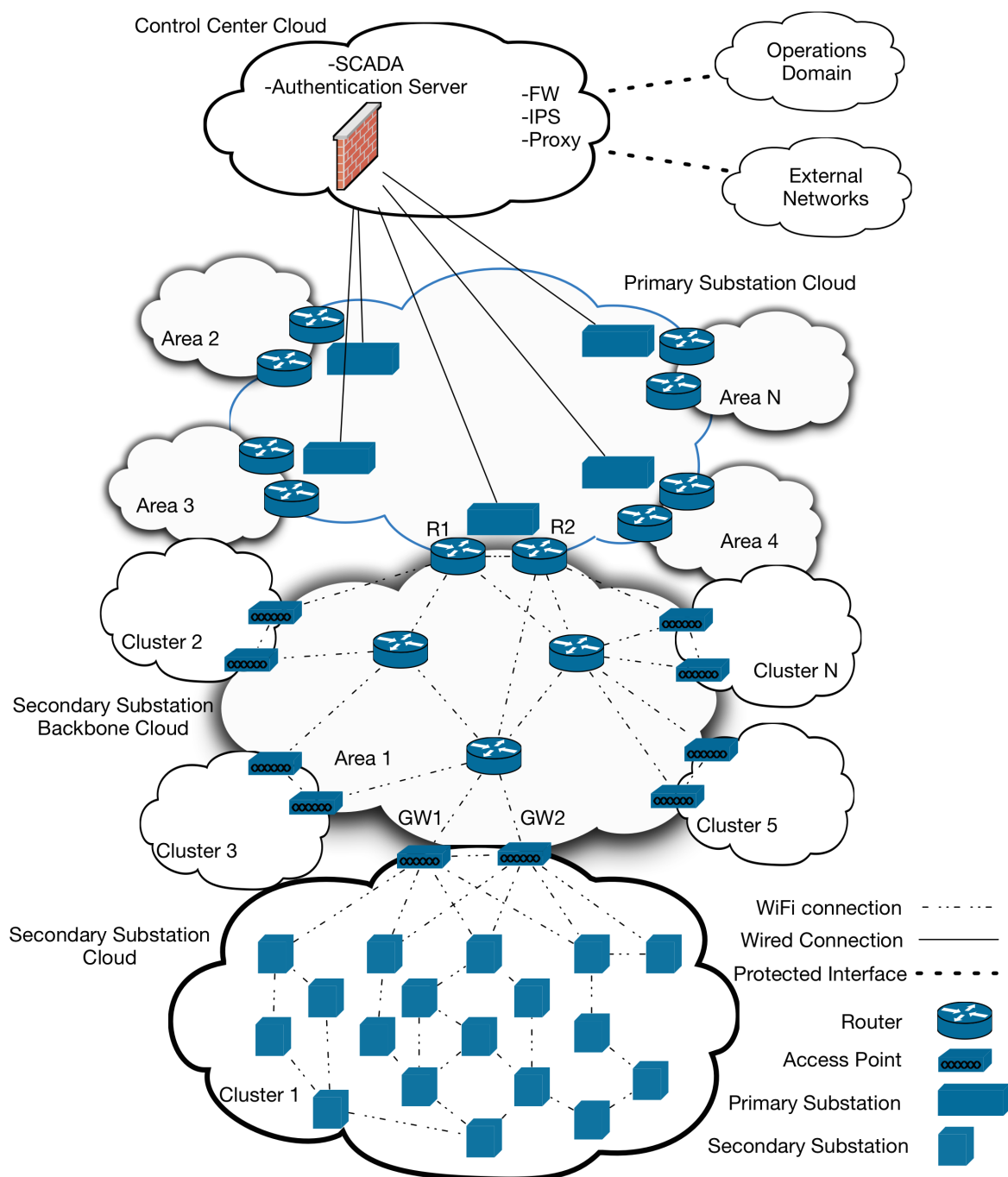


Figure 3.1: Security Aware Distribution System Archeticture

GW1 and GW2) in a Secondary Substation Cloud to a Primary Substation. Gateways are Secondary Substations in reality; they only have an extra responsibility in the network. These gateways should be the least distant ones from the routers in this cloud, which improves the state of wireless communication link between gateways and routers.

- Secondary Substation Cloud - SSC: It connects the Secondary Substations to the Secondary Substation Backbone using gateways, and ultimately, to the Control Center Cloud to exchange operational data and commands. There are two gateways in every Secondary Substations Cloud to provide redundancy.
- Substations and network devices: All substations have smart devices to allow distributed intelligence. They also support in meeting some cybersecurity requirements, which will be described in more details in Section 3.4. All smart devices have customizable software; therefore, they support having a Host Based Firewalls, HBFWs. Furthermore, other network devices such as routers and gateways have HBFWs as well.

Clouds described beforehand are defined in a hierarchical approach to allow better scalability, i.e. CCC collects information and sends commands to all substations, all PSCs are connected to CCC and every one of them covers a different geographical area. SSBC's role is to ease the communication between PSC and SSC.

3.2 Communications in SADS

Different communication technologies are used within the system. They are defined based on the nature and requirement of the cloud as follows:

- Inside CCC: CCC is considered to have a data center structure. Communication inside it is assumed to be using high-speed network channels, such as Fiber or Gigabit Ethernet.
- Between CCC and other domains or external networks: A protected high speed wired connection is used for security reasons, which will be discussed in more

details in Section 3.4.

- Between CCC and PSC: High speed wired network connection is used to allow this communication, such as fiber optical cable. This covers the communication between CCC and every PS.
- Inside PSC: Primary substations do not communicate with each other, thus, there's no communication requirement inside PSC. However, they are connected to the Control Center using high speed wired network connection as described earlier.
- Between PSC and SSBC: PSs have two routers directly connected to them using a wired connection, these routers form a mesh network and they are wirelessly communicating with SSBC. One router is enough for functionality but two are used for redundancy.
- Inside SSBC: Routers on the edge of PSC and routers inside SSBC form a mesh network that uses 802.11s protocol [23]; all of them are communicating wirelessly.
- Between SSBC and SSC: Gateways have two wireless network interfaces, one of them is using 802.11s to be a part of SSBC's mesh network. Whereas the other interface is a part of SSC's mesh network.
- Inside SSC: Every SS in SSC has one wireless network interface, these interfaces form a layer 2 mesh network or a cluster to allow communication between the nodes. Gateways are the cluster heads in these mesh networks. Secondary substations in SSC have hierarchical mesh setup, they are configured to use the cluster head to exit the cluster and reach CCC.

Mesh networks allow multiple nodes to send data at the same time because there's no central access point to collect and forward data. A node can act as a relay to the destination, which results in redundancy and reducing the bottleneck effect. In addition, all devices in the 802.11s mesh are using 802.11n because they naturally transfer more

data than the devices in SSC for example. On the other hand, devices in the Layer 2 mesh network can use 802.11g or 802.11a.

3.3 Communication Challenges

This section addresses the communication network challenges in the proposed architecture.

3.3.1 Low Communication Overhead

To maintain up-to-date status of substations in the distribution system, sensors at substations need to send reading frequently. The size of transferred readings is usually less than 500KB [24], but extra overhead usually comes from the communication protocols. It's important to reduce the communication overhead as much as possible because a small increase in the overhead will have a significant impact on traffic transferred over a system of this scale.

3.3.2 Scalability

Scalability is a big concern for distribution systems. If the architecture does not consider scalability, network congestion will occur because the number of nodes sending data is huge. In addition, the architecture should support future expansion without affecting the performance. For example, integrating substations from City B with the distribution system of City A shouldn't affect the current operation or result in any performance degradation in City A.

The proposed architecture supports scalability because all system nodes are grouped in clouds. Adding a group is simply adding one more cloud. The effect of this is keeping the number of hops as minimal as possible. Furthermore, if adding an SSC resulted in a bottleneck because of bandwidth limitation example, adding one more SSBC will resolve the issue.

3.3.3 Reliability

Reliability is very important in Distribution Systems, because decisions will be made based on received data. Using wireless communications is generally less reliable than using wired ones because wireless channels are susceptible to noise, interference, and attacks.

The proposed architecture increases reliability by adding redundancy to avoid outages resulting from node failures. Redundancy is clearly defined in SADSA, both in redundant devices (i.e. R1, R2 and GW1 GW2) and the communication protocols. Using a mesh network minimizes the dependency on a single node to reach the destination. For example, if one node failed for any reason (or the communication link for that matter), other nodes can be used to transfer information to their desired destination in a timely manner. Automatic recovery is possible but subject to the used mesh protocol.

3.3.4 Delay Introduced by Medium Access Control

Delay caused by CSMA based systems becomes more significant when the number of clients increases. Authors of [25] show how increasing the number of nodes decreases the effective throughput. The architecture minimizes this delay by grouping nodes into horizontal clouds, and controlling the number of nodes per cloud based on the bandwidth requirement. Cluster 1 and Cluster 2 in Figure 3.1 represent horizontal clouds because they belong to the same level in the hierarchy. Note that contiguous clusters and clouds use different channels to minimize interference and achieve better wireless performance.

3.3.5 Time Sensitive Protocols

Applications in Distribution System are sensitive; and they require receiving data in a timely manner based on the message severity as shown in Table 2.1. Ensuring most

packets arrive in a timely manner is considered challenging because of the factors discussed earlier. Medium access control, limited bandwidth, shared channels, interference, and security overhead are examples of such challenges. These factors arise clearly when using wireless communications as in the proposed system. However, maintaining the correct number of nodes in a cloud allows delivering data in a timely manner.

3.3.6 Interoperability

Devices in smart grids need to communicate with each other, but they are of a heterogeneous nature. This implies that these devices may include servers, single board computers, routers, access points, or any other types of sensors. Integrating these devices seamlessly is considered a challenge because they have various operating systems, capabilities, applications, and they support different protocols.

The proposed architecture supports interpretability. For instance, assuming devices in SSC 5 do not support WiFi; they can communicate with each other using ZigBee. The only special requirement lies at the gateways connecting SSC 5 and Area 1. They should support both ZigBee and the communication protocol used in Area 1. Considering that gateways have two network interfaces in their design, these interfaces would allow devices in SSC 5 to communicate seamlessly with the rest of the system.

3.4 Cybersecurity Awareness

In this section, the security awareness features for proposed architecture are identified. Smart grid security guidelines document [8] developed by NIST is used during the process of identifying cybersecurity requirements relevant to SADS. Each one of the following subsections represents one of the NIST guidelines.

3.4.1 Concurrent Session Control

This means that the number of concurrent sessions should be limited. In SADSAs, this requirement should be enforced at gateways, SSBC routers, CCC and nodes themselves. Concurrent Session Control can be achieved by doing the following:

- Limiting the number of sessions at the gateways from SS.
- Limiting the number of sessions at SSBC routers from gateways.
- Limiting the number of sessions at PS from SSBC routers.
- Limiting the number of sessions at CCC from all users and devices in the distribution system.
- Limiting the number of sessions at CCC from Operations domain per actor (e.g. Engineering and Energy Management)
- Limiting the number of sessions at devices using HBFW.

SADSAs make it easier to meet the requirement for two reasons. SADSAs partition the system into multiple clouds and it requires having firewalls. Clouds in the system architecture ease defining and controlling the number of concurrent sessions for nodes that have similar or different roles in the system. Furthermore, both CCC firewalls and HBFW can be configured to limit the number of sessions initiated from a source.

3.4.2 Remote Session Lock and Termination

This requirement highlights the need to define an inactivity period that will result in a session timeout. After this period, users and devices need to re-authenticate to be able to activate the session again. It is important to verify meeting this requirement at the gateways, PSS and CCC in the proposed architecture. Session Lock requirement can be met by:

- Defining an idle inactivity timeout per user or service type or connection type (i.e. wired, wireless..etc)

- Defining the actions should be taken when a session times out. These actions may include session termination and current information hiding.
- Enforcing the Session Lock mechanism in all nodes at CCC, in gateways to monitor the sessions of SSC, and PSC routers to monitor the sessions of PSC.

As can be clearly seen, organization of the proposed architecture makes it easier to define different timeout periods based on the node's role and location in the network. In addition, it helps to monitor the sessions in a hierarchical manner, which better supports scalability to avoid having a single point of failure.

3.4.3 Permitted Actions without Identification or Authentication

The requirement here focuses on defining access levels to users based on their role in the system. SADSAs makes meeting this requirement more practical, because actions can be defined by node role instead of looking at them node by node. In addition, a centralized authorization server, is easier to manage than having authorization implemented at multiple locations.

3.4.4 Remote Access

Remote access is concerned with both wired and wireless access, thus, the requirement breaks down into the following:

- All methods of remote access should be managed, authorized, and monitored.
- Cryptography should be used to protect such communications.
- Wireless access should be protected using both authentication and encryption.
- The spectrum should be monitored to detect any fake access points and take appropriate measures.
- Remote access should be disabled by default, and only enabled when required, approved, and for the required time period only.

In SADS, this requirement applies to WiFi access at PSC, SSBC, SSC. It also applies to all nodes that need to be accessed remotely. This requirement can be met by the following:

- Authenticating and encrypting all WiFi communications.
- Implementing a Wireless Intrusion Detection System -WIDS- that reports to CCC if any attacks were detected.
- Defining if Distribution Engineering and Distributed Generation and Storage Management from Operations domain need to access nodes remotely, if yes, which nodes and in what matter.
- Deny remote access by default, and give permissions for it based on earlier definitions.

SADS makes it easier to meet the requirement because it has the following features:

- All WiFi nodes support state of the art authentication and encryption.
- Substations have smart devices that can support WIDS.
- Remote access is not allowed by default.
- Grouping nodes logically in clusters helps in defining which nodes need to be remotely accessible based on their role in the network.
- HBFW and firewalls at CCC block remote access by default.

3.4.5 Wireless Access Restrictions

This requirement is applicable to PSC, SSBC, and SSC. The way the proposed architecture is organized, the smart devices in all substations, and the fact that all WiFi enabled nodes support state of the art protocols helps in meeting Wireless Access Restrictions requirement. The requirement can be met by doing the following:

- Planning and documenting appropriate WiFi implementation details and use it as guidance.

- Authenticating and encrypting all WiFi communications.
- Implementing a WIDS that reports to CCC if any attacks were detected.

3.4.6 User Identification and Authentication

This requirement highlights that all users should be uniquely identified and authenticated using multi-factor authentication. The following features of the proposed architecture support this requirement:

- Having a centralized authentication server helps in providing multi-factor authentication service for all users.
- The state of the art customizable software at all nodes eases the process of enforcing multi-factor authentication. This is needed because authentication agents need to be installed and configured to communicate with the central authentication server when authentication attempts occur.

3.4.7 Device Identification and Authentication

The requirement can be met by preparing and documenting a list of authentic devices with their details and enforcing authentication for all devices using bi-directional authentication through the authentication server at CCC. Meeting this requirement is considered easier because of the following supporting features in SADSA:

- Organization of the system architecture makes it easier to identify nodes, their types, roles, and location.
- Node information can be used for documentation and authentication server configuration.
- State of the art customizable software on all nodes doesn't restrict authentication methods. Thus, it allows custom protocols or bi-directional authentication.

3.4.8 Denial-of-Service Protection

The aim of this requirement is to ensure that all nodes in the smart grid information system mitigate the effect of DoS attacks. And that the nodes have extra bandwidth and resources to reduce their impact. The following SADS features help in meeting Denial-of-Service Protection requirement:

- If a node fails due to DoS attack, its redundant node can take over and deliver traffic because of planned redundancy in routers and gateways.
- Redundancy in the network in SSBC and SSC helps in mitigating the effect of WiFi DoS attacks. For example, if communication through a network path fails, other paths can be used to reach the destination.
- WIDS will alert CCC when such attacks occur and proper actions can be taken immediately.
- HBFW can be used to limit the number of concurrent sessions per source.
- Defense in depth in CCC can protect from most attacks coming to or through CCC, including DoS and DDoS.

3.4.9 Authenticator Feedback

Obscuring feedback of the authentication server during the authentication process is considered essential, which helps in preventing unauthorized individuals from understanding exploiting the authentication process. This requirement needs to be met at all nodes and it is relatively easy. Information shared and displayed to the user when they try to authenticate should be limited as much as possible. For example, there is no need to tell the user why a login attempt failed.

3.4.10 Security Function Isolation

It's useful to isolate security and non-security functions to avoid unintended information leakage. Which should be done at all nodes in the system. In addition, it is better to employ underlying hardware separation. Smart devices used at substations can be designed or tweaked to support the requirement. The following lists how this requirement can be met.

- Using smart devices whenever a security requirement is involved.
- Using Common Criteria for Information Technology and Security Evaluation [26]. Common Criteria is an international standard that allows organizations to specify their security requirements for computer systems, vendors develop these systems after that, then, the developed systems are tested in certified laboratories to verify their compliance.
- Using routers with custom firmware to isolate security and non-security requirements.

3.4.11 Boundary Protection

This requirement is applicable to the interface between CCC and Operations System, and the interface between CCC and external networks. These interfaces are shown in Figure 3.1. The aim of Boundary Protection is to achieve different goals as follows:

- Defining internal and external boundaries and controlling communication at the defined boundaries.
- Allowing communication to external networks only through protected interfaces, and limiting the number of these interfaces.
- Ensuring the system fails securely if any of boundary protection security controls failed.

SADSA has few features that make meeting the requirement easier as follows:

- Boundaries are defined between domains and external networks.
- Communication to other domains is done through protected managed interfaces only.
- Number of managed interfaces is limited to one interface per smart grid domain, and one more interface to external networks.
- Managed interfaces have a two way proxy, IPS and a firewall for protection.

3.4.12 Communication Integrity

The goal of this requirement is to warrant that smart grid information system protects the integrity of data in all communications. This requirement should be met on all communication links, which can be achieved by the following:

- Applying HMAC on all messages for all communications.
- Using a homomorphic signature scheme to sign aggregated data. Such signatures can be applied at the gateways for example after aggregating data from different SS.

3.4.13 Application Partitioning

Application partitioning requirement is made of two components: i.e. separating user and management functionalities, and ensure that management features are not available to all users. This requirement applies to all nodes in the architecture and can be achieved by employing the following two concepts:

- Always assigning the least privilege required by entity to perform its functions.
- Implementing proper access control scheme and enforce using it by all nodes.

3.4.14 Software and Information Integrity

This requirement applies to all nodes in SADSa. It aims at detecting any unauthorized changes to information. It can be achieved by the following:

- Using tamper proof devices, which can be guaranteed by Common Criteria for Information Technology and Security Evaluation [26]. These devices include substation smart devices, routers, HBFWs, firewalls, IPS, and proxies.
- Executing regular tampering checks every organizational defined tampering check period.
- Ensuring physical security.

3.5 Summary

In this chapter, a scalable security aware distribution system architecture is proposed. SADSa has been described along with the communication technologies used to allow functionality between its components and clouds. In addition, multiple communication challenges in SADSa were discussed, namely, Low Communication Overhead, Scalability, Reliability, Medium Access Control Delay, Time Sensitive protocols, and Interoperability.

Furthermore, NIST smart grid cybersecurity guidelines [8] were used to study security awareness of SADSa. It was shown how the requirements can be met and how SADSa helps in achieving that. Various cybersecurity requirements were extracted in that section, namely, Concurrent Session Control, Remote Session Lock and Termination, Permitted Actions without Identification or Authentication, Remote Access, Wireless Access Restrictions, User Identification and Authentication, Device Identification and Authentication, Denial-of-Service Protection, Authenticator Feedback, Security Function Isolation, Boundary Protection, Communication Integrity, Application Partitioning, and Software and Information Integrity.

Chapter 4: False Data Injection Prevention Protocol

In this chapter, False Data Injection Prevention Protocol is introduced. FDIPP protects against packet injection, duplication, alteration and node replication. Integrity is the main security service FDIPP is addressing, which covers both system and data integrity. We start this section by giving a background about the Extensible Authentication Protocol used in FDIPP along with some assumptions. Then, we cover the detailed operational procedure of FDIPP. After that, key management for FDIPP is discussed. Finally, formal security analysis is conducted using Scyther tool.

4.1 Background

4.1.1 The used EAP

We use EAP protocol proposed by the authors in [27] for authenticating entities as a part of our protocol. It is selected for many reasons; it was designed for WiFi networks, it provides forward secrecy, it has low computation and communication cost, and it meets all the security requirements of RFC 4017. The protocol assumes a secure communication channel between the Access Point - AP and the Authentication Server - AS. In addition, it has three main components as follows:

1. Registration phase: This phase is used to exchange credentials between a User - U and the AS. The authors of [27] assume that it was done offline and it is not specified by their protocol. However, they specify that AS and U should negotiate and exchange a long-term key k , a password pw , and a random one time key y in this phase.
2. Normal authentication process: This process is used to authenticate U and assist U and AP in sharing time limited credentials so that they can use the Fast reconnect process.

3. Fast reconnect process: This process is used to accelerate the authentication process, it is used unless the limited time credentials are expired or if this process fails; the normal authentication process is used in these two cases. Fast reconnect process is not used in FDIPP.

4.1.2 Assumptions

- The network cables connecting the Primary Substations to the Control Center are secure.
- The Control Center Cloud is highly secure, and it employs defense in depth techniques to reduce the probability of compromise. Defense in depth techniques include antivirus software, intrusion detections systems, intrusion prevention systems, firewalls, multi-factor authentication, and access control.
- The computers in power substation are tamper proof and all the locations are physically secure.
- There are private/public key pairs generated and stored at all the devices when installing them.
- The authentication server keeps track of the authenticated nodes, their IP addresses, the cloud they belong to and the keys. This helps in revoking access at any time if intrusion attempts was detected.
- Time is synchronized between all nodes.

4.2 FDIPP

4.2.1 High Level Protocol Operation

False Data Injection Prevention Protocol uses the authentication protocol presented in [27] for authentication, RSA [28] for asymmetric cryptography, AES [29] for symmetric cryptography and SHA-3 [30] for hashing. Although the authors of [31] have

presented multiple attacks on RSA, all of these attacks can be mitigated by ensuring physical security. Furthermore, The protocol is made of two major phases as follows:

1. Node Authentication: The goal of this phase is to authenticate nodes with the authentication server.
2. Peer Authentication: This phase is used to authenticate peers in the same network with each other. It is initiated after successfully completing phase 1.

The following describes the operation of FDIPP. In addition, Figure 4.1 reflects this operation.

1. Authentication parameters (i.e. k , pw , and y) are transferred between the nodes and the authentication server. These parameters are encrypted using RSA before being transferred.
2. Nodes attempt to authenticate using the EAP presented in [27].
3. After successful authentication, all communication is encrypted and signed using HMAC and SHA-3.
4. Peers (or nodes in the same cloud) authenticate each other on demand. For example, Node 1 will initiate the mutual authentication process with Node 2, only when Node 1 needs to use Node 2 as its next hop.
5. After peer authentication, nodes start sending data with an HMAC in all messages.

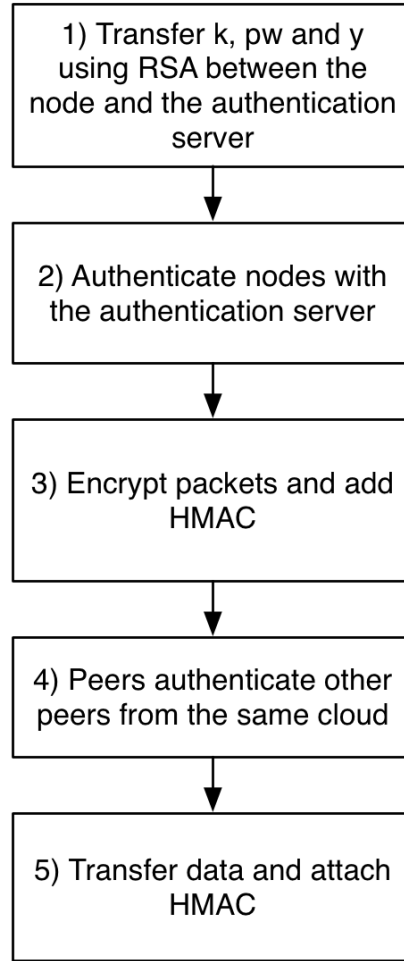


Figure 4.1: False Data Injection Prevention Protocol

4.2.2 Node Authentication

Node authentication is achieved after completing three phases, namely, SSBC Router authentication, Gateway authentication, and Secondary Substation authentication. Figure 4.2 shows the complete sequence diagram for Node Authentication in FDIPP. The three phases are identical and every one of them is made of 11 steps. Furthermore, we describe the steps of SSBC Router authentication phase below. Please note that $||$ symbol is used to represent concatenation throughout this work.

1. SSBC Router \longleftrightarrow Authentication Server: The keys k, y are generated by the authentication server, and the password pw is generated by the router. Then, identities are transferred between both sides. All data transferred is en-

encrypted using RSA. $E_{PubAS}(RID||pw)$ is sent to the Authentication Server and $E_{PubR}(ASID||k||y)$ is sent to SSBC router. Where $PubAS$ and $PubR$ are the public keys of Authentication Serve and SSBC Router respectively, RID is the router ID, $ASID$ is the authentication server ID, k is the long term key, y is a random one time key, and pw is the password.

2. SSBC Router \rightarrow Primary Substation: Authentication start.
3. SSBC Router \leftarrow Primary Substation: Identity request.
4. SSBC Router \rightarrow Primary Substation: The router provides a temporal ID as the identity response. The sent message is $[E_{k \otimes y}(RID), E_{k \otimes y}(RID||N_R)]$ where N_R is the router's nonce.
5. Primary Substation \rightarrow Authentication Server: The primary substation forwards $[E_{k \otimes y}(RID)E_{k \otimes y}(RID||N_R)]$ to the authentication server.
6. Primary Substation \leftarrow Authentication Server: The authentication server sends a challenge made of $H(N_R) \otimes E_{k \otimes y}(ASID||N_{AS}||y_N||y'||TK)$. Where N_{AS} is a nonce generated by the authentication server, y_N is a randomly generated key, y' is another one time key and TK is a temporal key. y' and TK are only used in the fast reconnect process, which is not used by FDIPP. However, they will be still generated and transferred when needed in the normal authentication process to avoid mistakenly reducing the security of the used EAP protocol. For instance, cracking $H(N_R) \otimes E_{k \otimes y}(ASID||N_{AS}||y_N)$ could be easier than cracking $H(N_R) \otimes E_{k \otimes y}(ASID||N_{AS}||y_N||y'||TK)$.
7. SSBC Router \leftarrow Primary Substation: The primary substation forwards $H(N_R) \otimes E_{k \otimes y}(ASID||N_{AS}||y_N||y'||TK)$ to the router. The router will be able to extract N_{AS} and y_N by XORing the received message with $H(N_R)$ again. Then decrypting $(ASID||N_{AS}||y_N||y'||TK)$. After that, it sets $y \leftarrow y_N$.
8. SSBC Router \rightarrow Primary Substation: The router responds with $H(RID||pw||y_N||y'||TK)$.

9. Primary Substation→Authentication Server: The primary substation forwards $H(RID||pw||y_N||y'||TK)$ to the authentication server.
10. Primary Substation←Authentication Server: If the hash was correct, the server sets $y \leftarrow y_N$ and sends Access Accept message. The session key is set to $SK = H(N_{AS} \otimes N_R)$. It also generates and sends the Cloud Key CK in $E_{SK}(CK||H(CK,y))$. CK is kept at the server and is unique per cloud.
11. SSBC Router←Primary Substation: The primary substation sends Authentication Success and forwards CK . The router sets $SK = H(N_{AS} \otimes N_R)$. It also divides CK into two portions (i.e. $CK = CK1||CK2$) to be used for encryption and hashing respectively.

After successful authentication, all transferred traffic is encrypted and signed using HMAC. Thus, the channel between SSBC Router and the Authentication Server is considered secure. The session key is equally divided into two keys (i.e. $SK = SK1||SK2$). $SK1$ is used as the key for AES and $SK2$ is used for HMAC. CK is unique per cloud whereas SK is unique per node.

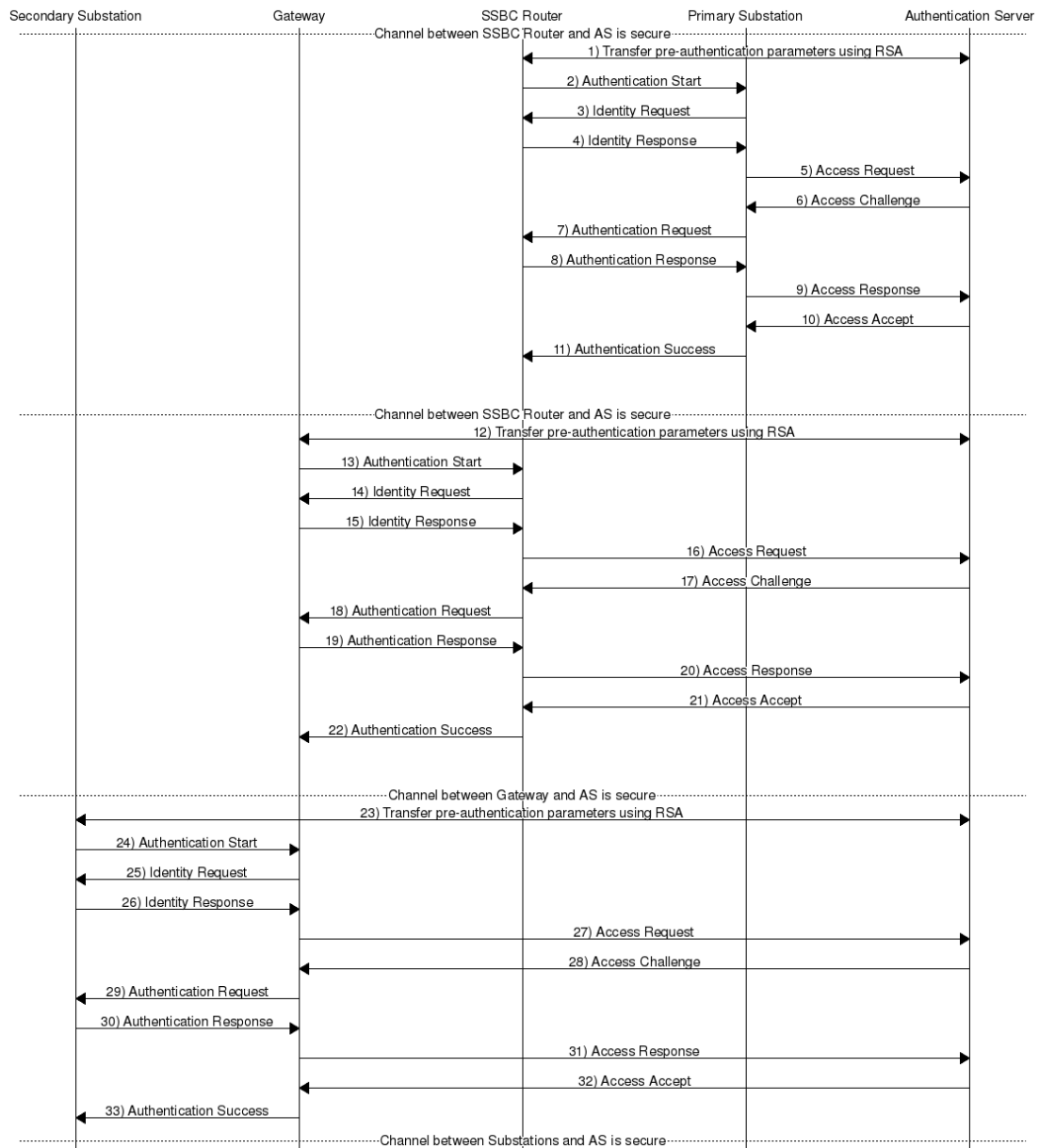


Figure 4.2: Node Authentication Sequence Diagram

4.2.3 Peer Authentication

All nodes from the same cloud will have a layer 2 routing table (e.g AODV routing table). Although nodes are authenticated with the authentication server, they should authenticate each other before node 1 uses node 2 as the next hop for instance. An example for peer authentication method is shown in Figure 4.3. It covers peer authentication between two secondary substations. Furthermore, the method is initiated whenever a node needs to communicate with another node from its routing table, and

it is described in more details as follows:

1. $SS1 \rightarrow SS2$: Secondary substation 1 sends its ID and the nonce N_{SS1} to secondary substation 2 as $E_{CK1}(N_{SS1}||H(SS1||N_{SS1}||CK2))$. SS2 stores the nonce N_{SS1} .
2. $SS1 \leftarrow SS2$: Secondary substation 2 responds with its ID and the nonce N_{SS2} to secondary substation 1 as $E_{CK1}(N_{SS2}||H(SS2||N_{SS2}||CK2))$. SS1 stores the nonce N_{SS2} .
3. $SS1 \rightarrow SS2$: Secondary substation 1 responds with HMAC of both nonces in $E_{CK1}(H(N_{SS1}||N_{SS2}||CK2))$. If the hash was correct, SS2 considers SS1 to be authentic.
4. $SS1 \leftarrow SS2$: Secondary substation 2 responds with HMAC of both nonces in $E_{CK1}(H(N_{SS1}||N_{SS2}||CK2))$. If the hash was correct, SS1 considers SS2 to be authentic.

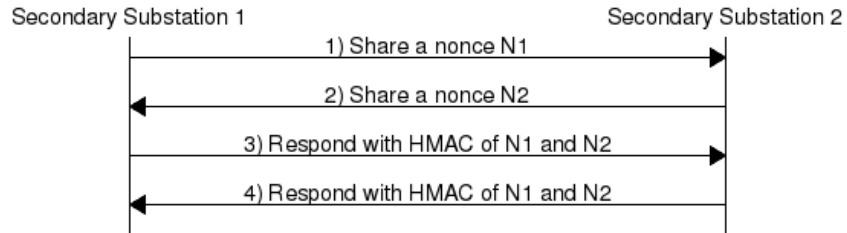


Figure 4.3: Peer Authentication Sequence Diagram

4.2.4 Post Authentication Data Transfer

After both node and peer authentication, nodes are ready to transfer data. All data transfers should be encrypted and signed with HMAC. Thus, when secondary substation 1 sends a data packet it should be in $E_{CK1}(data||H(data||N_1||N_2||CK2))$ form. Where N_1 is the nonce shared with the next hop and N_2 is the nonce shared with the authentication server. $CK1$ and $CK2$ are changed as applicable, i.e. when the packet crosses the boundary of a cloud. In addition, N_1 is changed and the hash is recalculated at every hop in the path.

4.3 Key Management

SADSA has a huge number of nodes, and key management is very important. Key management defines how keys are generated, distributed and revoked. It also defines their expiry. The following principals cover key management in FDIPP.

1. Twelve RSA keys are already stored in all nodes except the authentication server. Every key is one of the public keys of the authentication server. And it may be used for a period of 1 month only.
2. The authentication server has 12 public keys for every node in the system. Each key may be used for a period of 1 month only.
3. Asymmetric RSA keys are changed every month. And they are changed in a predefined order.
4. RSA keys to be physically replaced every year at all nodes in the system.
5. The session key SK is dynamically changing, as it is equal to $H(N_{AS} \otimes N_R)$.
6. The cloud key CK to be changed every hour by finding the hash of the current key, e.g. $CK_{n+1} = MD5(CK_n)$.
7. The cloud key CK to be periodically generated by the authentication server and communicated to all nodes every week.

4.4 Security Analysis

In this section, Peer Authentication portion of FDIPP is analyzed. The formal security analysis of the EAP protocol used in Node Authentication was already covered in [27]. Scyther tool [10] was found to be a perfect fit in our case for two factors. Firstly, it uses symbolic analysis. Which implies that mathematical cryptographic basis in the protocol is not analyzed. In fact, the author in [11] explicitly says that Scyther assumes cryptography is perfect. That property aligns very well with FDIPP because cryptography is beyond the scope of this work. Secondly, it supports both bounded

and unbounded analysis. Thus, it allows setting the maximum number of parallel sessions and identifies attacks within that bound. In the below subsections, configuration, claims, and results are presented.

4.4.1 Configuration and Claims

In Syther, the script was configured to assume $CK1$ and $CK2$ are secret. It was also configured to consider $N1$ and $N2$ as nonces and increment them in every message. After that, the protocol was written and analyzed with the below claims:

1. Secrecy: Secret information are not revealed to an intruder although the communication network is not trusted.
2. Aliveness: Communication partner is alive and he/she initiated an event that was received by the other partner. For example, an intruder replaying messages sent earlier is considered a violation of aliveness claim.
3. Synchronization: Communication parties are in synch, i.e. Agent A sends message 1 to Agent B, then Agent B will respond with message 2. Synchronization covers both ordered and unmodified delivery of messages.
4. Agreement: Communication parties agree on the values of all variables transferred in the protocol.

Both Synchronization and Agreement were claimed to be Non-Injective because Syther does not support any injective synchronization or injective agreement. The authors of Syther tool define injective synchronization to be non-injective Synchronization but it's immune to replay attacks. And the same applies to Agreement.

4.4.2 Results

As shown in Figure 4.4, FDIPP Peer Authentication protocol was proven to be secure and all claims were verified. Although injectivity is not supported in Syther, we know that every transferred message has a unique nonce, which prevents intruders to

replay messages. Thus, if the protocol was verified to have non-injective agreement and synchronization, then, it also matches injective agreement and synchronization.

Scyther results : verify

Claim				Status	Comments	
FDIPPPA	SS1	FDIPPPA,SS11	Secret CK1	Ok	Verified	No attacks.
		FDIPPPA,SS12	Secret CK2	Ok	Verified	No attacks.
		FDIPPPA,SS13	Secret NSS1	Ok	Verified	No attacks.
		FDIPPPA,SS14	Secret NSS2	Ok	Verified	No attacks.
		FDIPPPA,SS15	Alive	Ok	Verified	No attacks.
		FDIPPPA,SS16	Nisynch	Ok	Verified	No attacks.
		FDIPPPA,SS17	Niagree	Ok	Verified	No attacks.
SS2		FDIPPPA,SS21	Secret CK1	Ok	Verified	No attacks.
		FDIPPPA,SS22	Secret CK2	Ok	Verified	No attacks.
		FDIPPPA,SS23	Secret NSS1	Ok	Verified	No attacks.
		FDIPPPA,SS24	Secret NSS2	Ok	Verified	No attacks.
		FDIPPPA,SS25	Alive	Ok	Verified	No attacks.
		FDIPPPA,SS26	Nisynch	Ok	Verified	No attacks.
		FDIPPPA,SS27	Niagree	Ok	Verified	No attacks.

Done.

Figure 4.4: Syther Analysis of FDIPP Peer Authentication Protocol

4.5 Conclusion

False Data Injection Prevention Protocol was introduced, this protocol prevents packet injection, duplication, alteration and node replication. Which guarantees both the system integrity and data integrity in distribution systems. The chapter described Node Authentication, Peer Authentication, and Post Authentication Data Transfer. Key management in FDIPP was also described. In addition, formal security analysis was presented.

Chapter 5: Performance Evaluation

Performance evaluation is covered in this chapter. It's a very important task to identify how well an architecture is performing, and what are the factors that affect. Which will help to know how to customize and improve it. This chapter is divided into three main sections. Section 5.1 covers the simulation design, which is divided into two phases. Section 5.2 presents the simulation results. The results for both phases are separately discussed. Then, results of both phases are integrated and presented.

5.1 Simulation Design

With reference to Figure 3.1, the simulation was divided into two phases. Phase 1 covers communication from secondary substations in SSC to the gateways. Phase 2 covers communication from gateways to the authentication server and SCADA system. Considering that communication links between primary substation routers and both the authentication server and SCADA are completely wired. Routers (e.g. R1 and R2) are assumed to be the destination of traffic destined to CCC. The basis of this assumption is that wired links are using high speed fiber optic cables and will only cause negligible delay and packet loss. Thus, phase 2 of the simulation will cover communication from gateways and primary substation routers. The below subsections cover these phases in more details. In addition, this simulation covers data transfers only. As per FDIPP, authentication messages will not be often transferred, and we assume that it has already completed before the beginning of the simulation. The simulation period was set to be between 58 and 59 seconds in all samples.

5.1.1 Phase 1

Given the number of 6.6kV and 11kV substations in Dubai was equal to almost 29,000 substation in 2014. And the area of Dubai is equal to 4,200km². The average number

of secondary substation density is 7 substations per km^2 as shown in equation 5.1.

$$\frac{29,000 \text{ substations}}{4,200 \text{ km}^2} \approx 7 \text{ substations/km}^2 \quad (5.1)$$

In our simulation, we will assume that we cover various secondary substation density. We simulate having 6, 8, 10, 12, and 14 substations per km^2 , which covers a future density increase of up to 200%. All substation nodes send CBR UDP data packets to the gateway. We assume there's only one gateway considering that only one gateway of the High Availability pair is active at a time. Authors of [32] stated that sensing and control traffic size is usually equal to 60 bytes or less. They also state that multi-drop lines are configured at data rates equal to 1200bps.

In our simulation, we will assume that packet payload is equal to 512 bits, and each substation will send 10 messages per second. Security overhead equals to 256 extra bits per message. Where SHA3-256 will result in adding 256 digest bits to all messages [30]. However, AES-128 will only add processing overhead, as the number of cipher text bits for 768 plain text bits is equal to 768 bits. In addition, channel rate is configured to be equal to 54Mbps in this phase to simulate 802.11abg.

5.1.2 Phase 2

This phase was run with two different number of gateways (i.e. 2 and 4). Everyone of these gateways represents a cluster. Which means that all of them send packets during the simulation. Node locations in this phase were not configured to be random. Locations were manually set and were not changed through out the trials to ensure that we have two nodes at the middle which don't send packets. These nodes are similar to the routers inside SSBC (i.e. not on the edges). In addition, the simulation was configured such that all gateways will send to one destination. The destination simulates CCC. In addition, there is one node to at the middle between gateways to

facilitate communication for far nodes.

Packet size was kept same as in Phase 1. However, source rate was increased to 100 messages per second (i.e. 10*10 messages per second). Which was set to simulate 10 secondary substations sending data through the gateway. Furthermore, channel data rate was set to 150Mbps to simulate 802.11n. All other variables were kept the same.

5.2 Simulation Results

In this section, simulation results are presented. Results of Phase 1 and Phase 2 are presented in separate subsections. After that, both results are integrated to calculate end to end performance.

5.2.1 Phase 1

In Phase 1, all delay and packet loss measurements were taken as an average of 48 runs. Furthermore, nodes were configured to have a random location at each run. Figure 5.1 shows the average delay for two packet sizes. As mentioned beforehand, these packet sizes belong to two cases where there is a security overhead resulting from SHA3-256 [30] and where there is no such security overhead. It is clear from the graph that with the defined density of communicating nodes in $1km^2$ introduces a delay of 2ms. It's also observed that increasing the number of substations from 6 to 14 dose not significantly increase the delay.

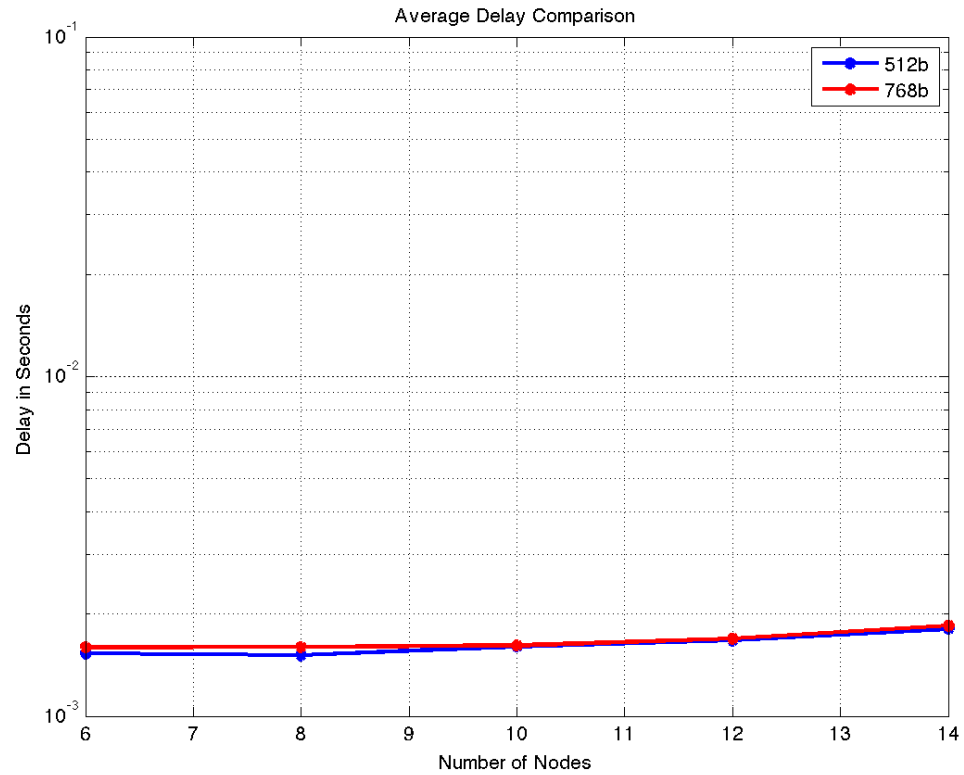


Figure 5.1: Average Delay in Phase 1

In addition, Figure 5.2 shows the average packet loss. The in packet loss matches our expectation based on the delay values acquired earlier. It's very close to 0%. Moreover, increasing the number of substations does not increase the packet loss in an observable manner.

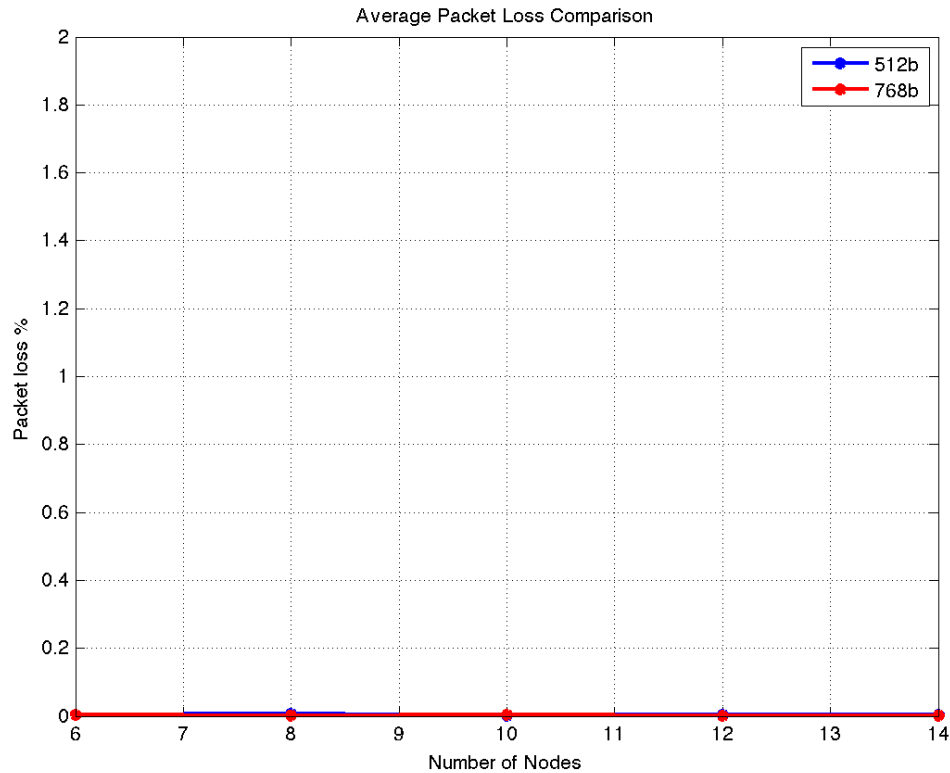


Figure 5.2: Average Packet Loss in Phase 1

Based on simulation results presented in Figures 5.1 and 5.2 we make the following conclusions:

1. Communication in this phase is introducing very low delay and packet loss.
 2. Increasing the number of nodes from 6 to 14 does not significantly affect the communication.
 3. Communication link is still far from saturation and data of larger size can be set.
- The number of nodes can be increased as well.

5.2.2 Phase 2

In this phase, delay values and packet loss measurements represent an average of 5 runs. More runs were taken in Phase 1 because its results required smoothing. Figure 5.3 compares the average delay for a total number of nodes of 4 and 6 (2 nodes are always not sending because they are considered to be at the center). It shows that

the delay goes from 3ms to slightly less than 30ms when increasing the number of gateways from 2 to 4. It also shows that the packet size increase resulting from security overhead addition does not make a major contribution in this phase.

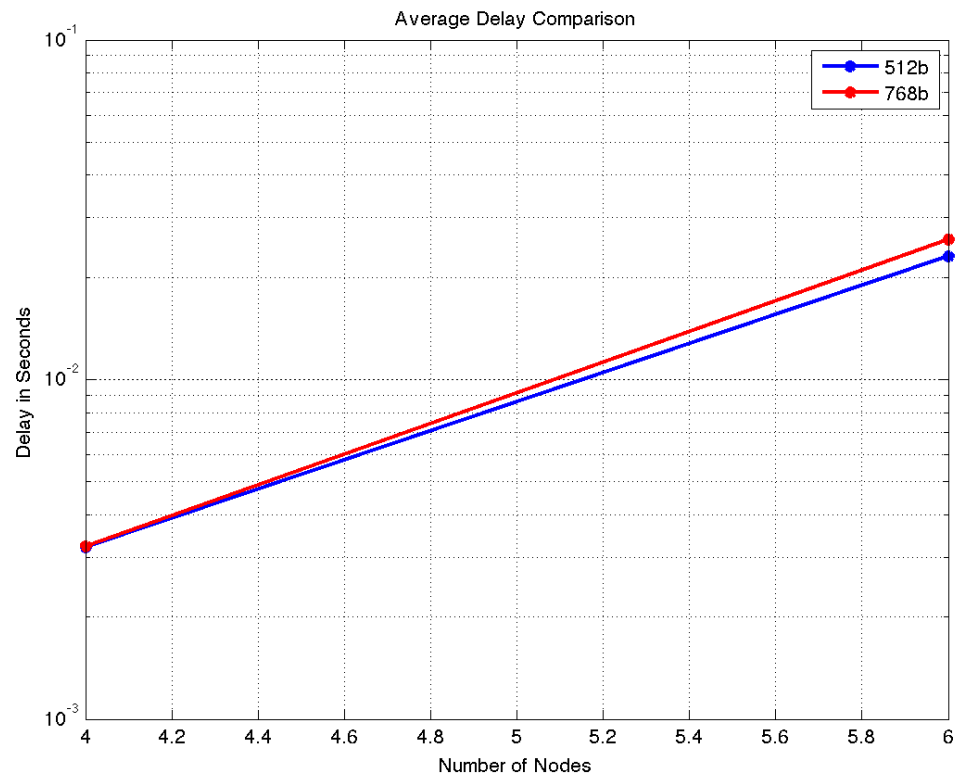


Figure 5.3: Average Delay in Phase 2

In addition, looking at Figure 5.4, we can see that packet loss is very low under all variations. The packet loss is almost 1% with the security overhead. On the other hand, it's very close from 0% without security overhead.

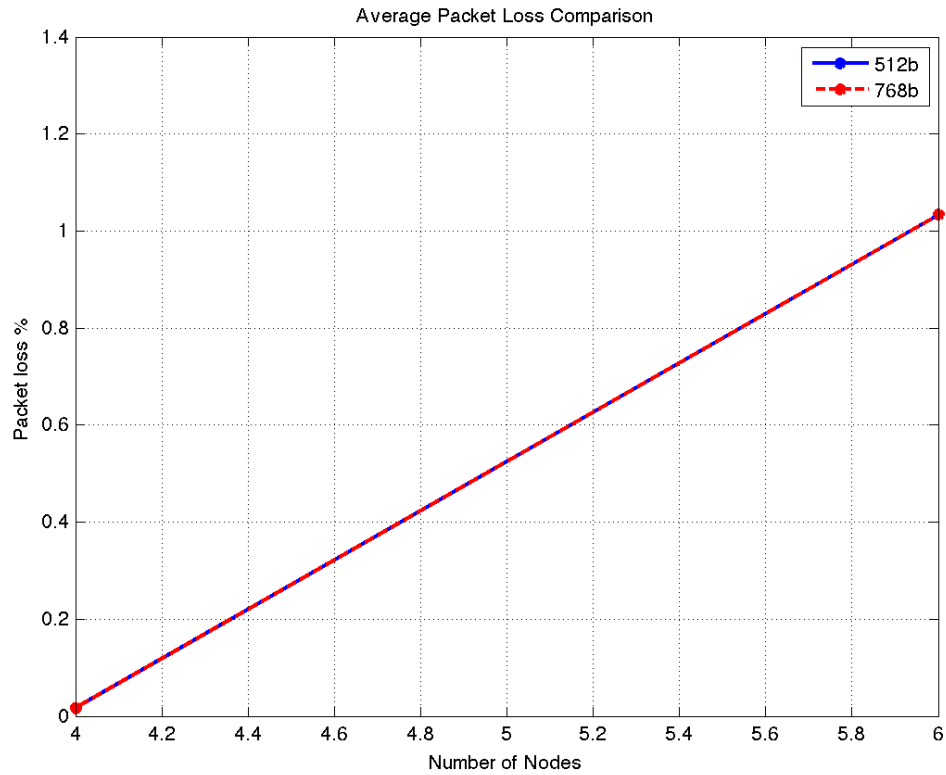


Figure 5.4: Average Packet Loss in Phase 2

5.2.3 End to End Performance

In this subsection, end to end measurements are calculated. Starting with the average delay, Table 5.1 shows a summary of the acquired delay values from simulation. It covers values for both phases (i.e P1 and P2). After that, Table 5.2 shows the end to end average delay.

State	Without Security Overhead (ms)	With Security Overhead (ms)
P1, 6 nodes	1.5	1.6
P1, 8 nodes	1.5	1.6
P1, 10 nodes	1.6	1.6
P1, 12 nodes	1.7	1.7
P1, 14 nodes	1.8	1.8
P2, 2GW	3	3
P2, 4GW	26	27

Table 5.1: Average Delay Summary

State	Without Security Overhead (ms)	With Security Overhead (ms)
6 nodes, 2GW	4.5	4.6
8 nodes, 2GW	4.5	4.6
10 nodes, 2GW	4.6	4.6
12 nodes, 2GW	4.7	4.7
14 nodes, 2GW	4.8	4.8
6 nodes, 4GW	27.5	28.6
8 nodes, 4GW	27.5	28.6
10 nodes, 4GW	27.6	28.6
12 nodes, 4GW	27.7	28.7
14 nodes, 4GW	27.8	28.8

Table 5.2: End to End Average Delay

As packet loss from phase 1 was found to be very close to 0%, we consider it to be negligible when compared to the loss introduced by phase 2. Thus, the packet end to end packet loss is only coming from phase 2 and shown in Table 5.3.

State	2GW	4GW
Without Security Overhead	0%	0%
With Security Overhead	1%	1.2%

Table 5.3: End to End Average Packet Loss

5.2.4 Summary

In this chapter, performance of SADSa and FDIPP were studied. It was found that most packet loss and delay is coming from communication from gateways, which makes sense as they aggregate data from multiple substations. Which can be decreased by increasing the channel rate, or decreasing message frequency. Or maybe, some applications may be able to tolerate such delays. It's all subject to the application requirements. In other words, one size fits all doesn't apply in a network of this nature. Some clusters may need to have higher bandwidth, whereas others may need to share measurements less frequently. However, SADSa is flexible per its design, and it allows tweaking the network to achieve application requirements.

Chapter 6: Conclusion

Smart grids are the future power systems. They are used to increase efficiency and reliability of power systems. Distribution system is one of its major components, and it was the focus of this work. There are three main contributions in this thesis, namely, Security Aware Distribution System Architecture, False Data Injection Prevention Protocol, and Performance Evaluation. SADSAs were described in details, and communication challenges were highlighted. Furthermore, it was shown how SADSAs meet NIST cybersecurity requirements for smart grids. In FDIPP, the protocol was described, key management was covered, and Scyther tool was used for formal security analysis. Moreover, performance of SADSAs and FDIPP was simulated using NS2.

This work can be extended by studying performance of SADSAs and FDIPP more deeply, which can involve testing different protocols and communication technologies. In addition, the concept of SADSAs can be extended to design other smart grid domains, which will result in having security aware domains that are more immune to cyberthreats.

Bibliography

- [1] W. Wang and Z. Lu, “Cyber security in the smart grid: Survey and challenges,” *Computer Networks*, vol. 57, no. 5, pp. 1344–1371, 2013.
- [2] S. Matsuo, K. Miyazaki, A. Otsuka, and D. Basin, “How to Evaluate the Security of Real-Life Cryptographic Protocols?,” *Financial Cryptography and Data Security*, vol. 6054, no. Chapter 16, pp. 182–, 2010.
- [3] D. Jin, D. M. Nicol, and G. Yan, “An event buffer flooding attack in DNP3 controlled SCADA systems,” in *WSC '11: Proceedings of the Winter Simulation Conference*, pp. 2614–2626, Winter Simulation Conference, Dec. 2011.
- [4] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, “Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures,” in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 220–225, IEEE, 2010.
- [5] O. Vukovic and G. Dan, “Security of Fully Distributed Power System State Estimation: Detection and Mitigation of Data Integrity Attacks,” *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 7, pp. 1500–1508, 2014.
- [6] W. Stallings and L. Brown, *Computer Security Principles and Practice*. Prentice Hall, 2008.
- [7] F. PUB, “Secure hash standard (shs),” 2012.
- [8] The Smart Grid Interoperability Panel–Smart Grid Cybersecurity Committee, “Guidelines for smart grid cybersecurity,” tech. rep., National Institute of Standards and Technology, Gaithersburg, MD, Sept. 2014.

- [9] C. Greer, D. A. Wollman, D. E. Prochaska, P. A. Boynton, J. A. Mazer, C. T. Nguyen, G. J. FitzPatrick, T. L. Nelson, G. H. Koepke, A. R. Hefner Jr, V. Y. Pillitteri, T. L. Brewer, N. T. Golmie, D. H. Su, A. C. Eustis, D. G. Holmberg, and S. T. Bushby, “NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 3.0,” tech. rep., National Institute of Standards and Technology, Engineering Laboratory, Gaithersburg, MD, Oct. 2014.
- [10] “Scyther tool,” April 2016. <https://www.cs.ox.ac.uk/people/cas.cremers/scyther/>.
- [11] C. J. F. Cremers, *Scyther: Semantics and verification of security protocols*. Eindhoven University of Technology, 2006.
- [12] C. Cremers, “Key exchange in ipsec revisited: Formal analysis of ikev1 and ikev2,” in *Proceedings of the 16th European Conference on Research in Computer Security*, ESORICS’11, (Berlin, Heidelberg), pp. 315–334, Springer-Verlag, 2011.
- [13] Y. Cunjiang, Z. Huaxun, and Z. Lei, “Architecture Design For Smart Grid,” *Energy Procedia*, vol. 17, pp. 1524–1528, 2012.
- [14] 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.
- [15] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. P. Chen, “A survey of communication/networking in smart grids,” *Future Generation Computer Systems*, vol. 28, no. 2, pp. 391–404, 2012.
- [16] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, “Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures,” in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 220–225, IEEE, 2010.
- [17] A. Giani, E. Bitar, M. J. Garcia, M. McQueen, P. P. Khargonekar, and K. Poolla, “Smart Grid Data Integrity Attacks.,” *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1244–1253, 2013.

- [18] F. Li and B. Luo, "Preserving data integrity for smart grid data aggregation," in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pp. 366–371, IEEE, 2012.
- [19] Y. Guo, C.-W. Ten, and P. Jirutitijaroen, "Data integrity validation framework for distribution system operations," in *CSIIRW '11: Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research*, (New York, New York, USA), p. 1, ACM Request Permissions, Oct. 2011.
- [20] S. Bhattarai, L. Ge, and W. Yu, "A novel architecture against false data injection attacks in smart grid," in *ICC 2012 - 2012 IEEE International Conference on Communications*, pp. 907–911, IEEE.
- [21] C. Callegari, S. De Pietro, S. Giordano, M. Pagano, and G. Procissi, "A Distributed Privacy-Aware Architecture for Communication in Smart Grids," in *2013 IEEE International Conference on High Performance Computing and Communications (HPCC) & 2013 IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 1622–1627, IEEE, Nov. 2013.
- [22] T. Zhang, W. Lin, Y. Wang, S. Deng, C. Shi, and L. Chen, "The design of information security protection framework to support smart grid," in *Power System Technology (POWERCON), 2010 International Conference on*, pp. 1–5, IEEE, 2010.
- [23] G. R. Hiertz, D. Denteneer, S. Max, R. Taori, J. Cardona, L. Berlemann, and B. Walke, "Ieee 802.11 s: the wlan mesh standard," *Wireless Communications, IEEE*, vol. 17, no. 1, pp. 104–111, 2010.
- [24] W. Luan, D. Sharp, and S. Lancashire, "Smart grid communication network capacity planning for power utilities," in *IEEE PES T&D 2010*, pp. 1–4, IEEE, 2010.
- [25] L. Fang, H. Wu, and Z. Huang, "Performance modeling and analysis of ieee 802.11 protocol using poosl," in *Industrial Electronics and Applications*, 2009.

ICIEA 2009. 4th IEEE Conference on, pp. 1330–1335, IEEE, 2009.

- [26] “Publications : New cc portal.” <http://www.commoncriteriaportal.org/cc/>, Jan 2016.
- [27] C.-I. Fan, Y.-H. Lin, and R.-H. Hsu, “Complete EAP Method: User Efficient and Forward Secure Authentication Protocol for IEEE 802.11 Wireless LANs,” *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 4, pp. 672–680, 2013.
- [28] J. Jonsson and B. Kaliski, “Public-Key Cryptography Standards (PKCS) #1: RSA Cryptography Specifications Version 2.1,” tech. rep., Feb. 2003.
- [29] N. F. Pub, “197: Advanced encryption standard (aes),” *Federal Information Processing Standards Publication*, vol. 197, pp. 441–0311, 2001.
- [30] P. FIPS, “Secure hash algorithm-3 (sha-3) standard: Permutation-based hash and extendable-output functions,” *National Institute for Standards and Technology (NIST)*, vol. 202, no. 0, 2014.
- [31] D. Genkin, A. Shamir, and E. Tromer, “Rsa key extraction via low-bandwidth acoustic cryptanalysis,” in *Advances in Cryptology–CRYPTO 2014*, pp. 444–461, Springer, 2014.
- [32] Q. Yang, J. A. Barria, and T. C. Green, “Communication Infrastructures for Distributed Control of Power Distribution Networks,” *IEEE Transactions on Industrial Informatics*, vol. 7, no. 2, pp. 316–327, 2011.