Security and Privacy methods in Smart Home: Case Study in Smart Meters

Master Thesis Report -
Final results

Author:
Dina Eid Musalam
dinaeid@yahoo.com

Author:
Bueala Paulina Nicodemus
paulina610@gmail.com

Supervisors:
Ahmed Elmesiry

Examiner:
Radu Mihailescu

Final presentation: 03-June-2016
Contact information

Author:
Dina Eid Musalam
E-mail: dinaeid@yahoo.com
Malmo University, Department of Computer Science

Author:
Bueala Paulina Nicodemus
E-mail: paulina610@gmail.com
Malmo University, Department of Computer Science

Supervisors:
Ahmed Elmesiry
E-mail: ahmedmisery@gmail.com
Malmo University, Department of Computer Science
Universidad Tecnica Federico Santa Maria, Department of Electronics Engineering, Chile.

Examiner:
Radu Mihailescu
E-mail: radu.c.mihailescu@mah.se
Malmo University, Department of Computer Science
Abstract

Homes with integrated state of the art technology are considered Smart Homes. A smart home has special systems which enable remote control via remote computer or smart phones. These modern integrated services are supported by electrical power infrastructures called Smart Grids, which provides a stable electrical power environment to support all functionalities, with intelligent power systems that provide increased power quality. Inside a Smart home, a legacy metering system called Smart Meters are installed offering new functionalities such as remote readings of power consumption, and different time usage of tariffs.

Preliminary research has already indicated vulnerability attacks on smart meters which affect the security and privacy in smart meters. Security issues in this system include vulnerabilities and privacy issues includes information leakage in real-time consumption data that is recorded by the smart meters.

Enhancing both security and privacy in the smart meters are the main purpose of this thesis. The principle goal of this research is to provide more understanding about the smart meters from a security and privacy perspective. This thesis investigates issues and problems in smart meters and proposes a secure communication protocol in the application layer, in addition to a proof-of-concept of the final solution.

In this research we introduce the solution by means of two scenarios, we highlight the effects of an attack on the smart meters on levels of simulation and theory. The first scenario is to "take control over the smart meter to access the data consumptions in the smart meters". The proposed solution of using the Smart Phone as a third part to protect the smart meters has been evaluated using a network simulation tool. The result shows that the data captured by attacking node is encrypted and can’t be used for any useful operation. In the second scenario, "monitoring the data consumptions to harm the user’s privacy", in other words, stealing the user’s devices. The proposed solution is using the home electrical power routing to moderate the home’s load signature in order to hide appliance usage information. Data clusters are implemented as a proof of concept to evaluate the data. The results show that there is only 99.5 % correct clustered data with good quality.
Acknowledgement

This thesis was performed at the Malmo University of Science and Technology in Malmo, Sweden. This thesis contributes to the final completion of our studies as Masters in Computer Science. This thesis has been supervised by Dr. Ahmed El-Mesiry and Dr. Johan Holmgren as course coordinators.

We would like to thank Dr. Ahmed K. Elmesiry for the help he offered us and Dr. Johan for his support. We are thankful for Denise Brown and Dr. Ahmad Al-Omari for all the collaboration, motivation and advices throughout the duration of preparing the thesis. We would also like to thank each other for our good team work. The duration of our study has been a mega learning experience and we would like to thank everyone involved in it for making it an incredible journey.

Last but not the least, we want to sincerely thank our families, for there constant support and for all they had to endure from us. Without their love and support, we would not have achieved this.

We are very contented with the entire research and hope that results will be interesting and useful to the community.
## Contents

1 Introduction ............................................. 1
   1.1 Research Objectives ................................. 2
   1.2 Challenges Identification ............................ 2
   1.3 Thesis Statement .................................... 3
   1.4 Research Questions ................................ 4
   1.5 Expected Results .................................. 4
   1.6 Limitations ........................................ 4
   1.7 Formulation of the Paper ............................ 4

2 Research methodology .................................. 6
   2.1 Literature Review .................................. 6
   2.2 Design and Creation ................................ 7
   2.3 Prototyping Model .................................. 8
   2.4 Case Study ......................................... 9
   2.5 Machine Learning .................................. 10
   2.6 Method Description ................................ 11

3 Literature Review ...................................... 12
   3.1 Smart Home ........................................ 12
   3.2 Smart Grid .......................................... 13
   3.3 Smart Meters ....................................... 14
      3.3.1 Smart Meter ................................... 14
      3.3.2 Smart Meters Structure ......................... 15
   3.4 Security and Privacy ............................... 16
      3.4.1 Security ....................................... 17
      3.4.2 Privacy ....................................... 21

4 Terminator .............................................. 25
   4.1 Smart Meter Communication ......................... 25
   4.2 Attack Method ..................................... 27
      4.2.1 Terminator .................................... 28
List of Figures

3.1 Home Network. Smart Home (Lugo et al., 2014). ........................................ 12
3.2 Smart Meters Connection to Power Provider (Wang et al., 2011). .............. 13
3.3 Smart Grid (Wang et al., 2011). ................................................................. 14
3.4 Photograph of a Smart Meter (Molazem, 2012). ....................................... 15
3.5 Smart Meter Structure (Fan et al., 2013). .................................................. 16
3.6 The Consumption Data for Different Devices (Marmol, 2013). ................ 21

4.1 Termineter Framework Interface (Knapp, E. D., & Samani, R. (2013)). .... 28

5.1 Eavesdropping Attack. ................................................................................. 33
5.2 Inject Malicious Codes. ................................................................................ 34
5.3 Code to Hack the Username and Password. .............................................. 34
5.4 Communication Between SP, SM and UP. .................................................. 35
5.5 Security Protocol. ....................................................................................... 37
5.6 Basic Architecture of ns-2 (Issariyakul, T., & Hossain, E. 2011). .......... 38
5.7 Simple Network Topology. ........................................................................... 40
5.8 Code to Implement Malicious Node (“Surajpatilworld Blogspot Com”, 2016). 41
5.9 Malicious Node Dropping the Packets. ....................................................... 41
5.10 Output of Trace File. .................................................................................. 42
5.11 Applied Cryptography Algorithm. .............................................................. 43
5.12 Code for The Security Protocol. .............................................................. 43
5.13 Output from Applying Security Protocol .................................................. 44
5.14 Buffering After Trace File is Loaded. ....................................................... 45
5.15 The Nodes Flows. ...................................................................................... 46
5.16 Protocol Evaluation. .................................................................................. 46
5.17 Protocol Evaluation. .................................................................................. 47
5.18 Malicious Attack. ....................................................................................... 48
5.19 Malicious Attack Code. .............................................................................. 49
5.20 Storage Model. .......................................................................................... 51
5.21 Real Consumption Data for TV, Laptop, Fridge and Lamps. .................. 53
5.22 Clustered Data. ......................................................................................... 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.23</td>
<td>The Result for Clustering the Data.</td>
<td>54</td>
</tr>
<tr>
<td>5.24</td>
<td>One-hour Consumption Data.</td>
<td>55</td>
</tr>
<tr>
<td>5.25</td>
<td>Visualization of Clustered Data.</td>
<td>56</td>
</tr>
<tr>
<td>5.26</td>
<td>The Tree Model for The Clustered Data.</td>
<td>56</td>
</tr>
<tr>
<td>5.27</td>
<td>Clustered Results.</td>
<td>57</td>
</tr>
<tr>
<td>5.28</td>
<td>Clustered Quality Results.</td>
<td>57</td>
</tr>
<tr>
<td>5.29</td>
<td>Result for The Clustered Data Experiments.</td>
<td>58</td>
</tr>
<tr>
<td>5.30</td>
<td>Two-hour Consumption’s Data.</td>
<td>58</td>
</tr>
<tr>
<td>5.31</td>
<td>Visualize the Clustered Data During the Real-time Consumption.</td>
<td>59</td>
</tr>
<tr>
<td>5.32</td>
<td>Privacy Protocol.</td>
<td>59</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Relationship Between Research Questions and Research Methods ............... 6
2.2 Literature Review Keywords ......................................................................... 7
5.1 Parameters Used for ns-2 Simulation .............................................................. 40
6.1 Research Summary ....................................................................................... 62
# List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AODV</td>
<td>Ad hoc On Demand Routing Protocol</td>
</tr>
<tr>
<td>ASCII</td>
<td>American Standard Code for Information Interchange</td>
</tr>
<tr>
<td>AES</td>
<td>Advanced Encrypt Standard</td>
</tr>
<tr>
<td>AFE</td>
<td>Analog Front End</td>
</tr>
<tr>
<td>AMI</td>
<td>Advanced Metering Infrastructure</td>
</tr>
<tr>
<td>BTRAM</td>
<td>Behaviour based Remote Attestation Model</td>
</tr>
<tr>
<td>CBCM</td>
<td>Cipher Block Chaining Message</td>
</tr>
<tr>
<td>D&amp;C</td>
<td>Design and Creation</td>
</tr>
<tr>
<td>DER</td>
<td>Distributed Energy Resources</td>
</tr>
<tr>
<td>DoS</td>
<td>Denial of Service</td>
</tr>
<tr>
<td>DoE</td>
<td>Department of Energy</td>
</tr>
<tr>
<td>DSA</td>
<td>Digital Signature Algorithm</td>
</tr>
<tr>
<td>ECC</td>
<td>Elliptic Curve Cryptography</td>
</tr>
<tr>
<td>ECMQV</td>
<td>Elliptic Curve Menezes Qu Vanstone</td>
</tr>
<tr>
<td>ES</td>
<td>Energy Supplier</td>
</tr>
<tr>
<td>FTP</td>
<td>File Transfer Protocol</td>
</tr>
<tr>
<td>HAN</td>
<td>Home Area Network</td>
</tr>
<tr>
<td>IDS</td>
<td>Intrusion Detection Based</td>
</tr>
<tr>
<td>LS</td>
<td>Load Signature</td>
</tr>
<tr>
<td>LSM</td>
<td>Load Signature Moderation</td>
</tr>
<tr>
<td>MCU</td>
<td>Micro Controller Unit</td>
</tr>
<tr>
<td>MDM</td>
<td>Meter Data Management</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>NAM</td>
<td>Network AniMator</td>
</tr>
<tr>
<td>NIC</td>
<td>Network Interface Control</td>
</tr>
<tr>
<td>NILM</td>
<td>Nonintrusive Load Monitoring</td>
</tr>
<tr>
<td>ns-2</td>
<td>Network Simulator-2</td>
</tr>
<tr>
<td>NSRs</td>
<td>Network System Requirements</td>
</tr>
<tr>
<td>OMAP</td>
<td>One-way Memory Attestation technique and the Pioneer</td>
</tr>
<tr>
<td>OTcl</td>
<td>Object-oriented Tool Command Language</td>
</tr>
<tr>
<td>RSA</td>
<td>Rivest, Shamir and Adleman</td>
</tr>
<tr>
<td>RTC</td>
<td>Real Time Clock</td>
</tr>
<tr>
<td>SG</td>
<td>Smart Grid</td>
</tr>
<tr>
<td>SH</td>
<td>Smart Home</td>
</tr>
<tr>
<td>SKKE</td>
<td>Symmetric Key Key Establishment</td>
</tr>
<tr>
<td>SM</td>
<td>Smart Meter</td>
</tr>
<tr>
<td>SP</td>
<td>Smart Phone</td>
</tr>
<tr>
<td>SUN</td>
<td>Smart Energy Utility Network</td>
</tr>
<tr>
<td>Tcl</td>
<td>Tool command language</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TTP</td>
<td>Trusted third party</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>UP</td>
<td>Utility provider</td>
</tr>
<tr>
<td>WCSS</td>
<td>Within-Cluster Sum of Squares</td>
</tr>
<tr>
<td>WEKA</td>
<td>Waikato Environment for Knowledge Analysis</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Smart home technology has introduced the concept of devices communication in several appliances, which can be remotely controlled from inside your home. Silva et al., (2013), pointed out that the Smart Home (SH) is like an environment that automatically reacts to control home devices depending on the consumer’s behaviour. Smart home intends to cover the entire home’s functions, and it is not just restricted to indoor use, but for outdoor too. This proves to be a gateway for the interaction with other paradigms to improve the levels of quality, and comfort in homes.

For example, Smart Grids are considered the next generation for power grids (Lugo-Cordero et al., 2014). It has included increased incorporation of renewable generation, storage technologies, and user participation through demand response programs and enhanced sensing, control and communication technologies. The Smart Grid system consists of large number of subsystems; one of the core systems is the Advanced Metering Infrastructure (AMI). The AMI via a two-way communication system is responsible for communicating between the users and their devices in order to gather, store, and analyse energy usage data. An AMI system generally consists of three elements: Smart Meters (SM’s) at the customer’s homes, a metering communication infrastructure between the customer’s homes and their devices, and a Meter Data Management system (MDM) (Nilsson et al., 2014).

This research focuses mostly on SM’s. SM’s can manage and monitor home power consumption through a network. Spencer (2016) estimated that by the end of 2017, 280 million SM’s will be installed around the globe because of the benefits from developing SM infrastructures. This rapid deployment of SM raise several security and privacy problems. Security issues in this system are tampering, vulnerabilities and Denial of Service (DoS). Privacy issues include information leakage in real-time consumption data recorded by the SM. This reveals the patterns of customer’s energy usage depending on the current device usage. Previous researches were conducted in this field to improve the current technology, and still more research is needed.
This research aims to give a deeper understanding about the SM from a security and a privacy perspective. In addition, it proposes a framework to test threat models for SM, create new algorithms and architecture to enhance the protection of the software running in SM and the privacy of the user-generated data. The main contributions of this research are presented as follows:

- We exploit the one threat model attack on the SM.
- We propose a protocol with new algorithm to enhance the security in SM’s software, which leads to enhanced privacy of data in SM.
- We consider the constrained resources in SM, and then present a collaborative remote attestation mechanism focusing on the system’s behaviour to detect the malicious attack in SM to enhance the privacy.

This research aims to develop a proof of concept for security and privacy in SM’s.

1.1 Research Objectives

Shepherd and Weldes (2008), define the security as ”the state of being free from danger or threat”. While Acquisti et al., (2013) explained privacy as ”the state in which one is not observed or disturbed by other people”. However, Privacy and security in SM’s software aim to ensure that the communication processes between the SM and provider are executed as expected without any interruption such as the malicious user (Kahmer & Gillott, 2009).

The main goal of this research is to find the possible security threats within the SM’s software and end user sensitive data. Thus, an investigation is needed for the existing threats and protection mechanisms to gain deeper understanding of the current attacks, which could help in achieving the project goal. However, this research focuses on the software threats according to the limitation time. The proposed framework attains security of the SM’s software from tampering, and protects the user’s data from unethical eavesdroppers.

To achieve our goal, machine-learning techniques will be utilizing. The problem will be formulated as a machine-learning problem, where we will utilize a set of algorithms with a set of features and study the importance of each feature to predict the future failures and find the comfort levels of the user.

1.2 Challenges Identification

As mentioned before the operations in SM systems require a huge amount of data transfer between SM and utility server. Automatic and continuous transmission and collection of energy
consumption data are the major parts of SM system to update the user’s measurements and upgrade their services. "This is a tedious and expensive job" (Eldefrawy et al.,2012). During these communications, SM essentially creates privacy and security issues such as the leakage of private data and inability to protect the system from access of malicious activities. McDaniel et al., (2009) presented a scenario where a malicious user attacks an SM to take control over and created havoc in the network.

The SM’s software is similar to any other computer system, where its system can be attacked, and SM system level access could be achieved using any available system vulnerability (Berthier et al., 2011).

1.3 Thesis Statement

SM monitors and measures the electricity consumption of the users appliances and transfers these measurements to the Energy Supplier (ES) in real-time. During these communications, one can figure out various information about the user’s lifestyle, for instance "if the user is using a toaster, it determines he is hungry". This issue regarding privacy needs a new approach to ensure privacy for the user’s livelihood (Berthier et al., 2011). This motivates us to do research in finding an approach to allow privacy of SM users.

Molazem (2012), mentioned that there are many security issues with a SM that need to be solved like vulnerability, network intrusions, etc. In addition, malicious users created several issues in SM security and privacy. The fact that malicious attackers can have partial control over other user’s homes causing the user to not trust the SM. Thus, protection mechanism became a more important issue.

Moreover, security and privacy has a relation in our life. They are correlated to each other, where privacy cannot be achieved without security. Security is the outer layer for privacy.

This research focuses on unethical eavesdroppers as a privacy attack, who gather information for malicious purposes by hacking the SM to collect the data. However, this will create havoc and the user will not feel safe using these electrical devices. Moreover, these attacks can cause damage to the neighbourhood’s electrical devices, for example, the attacker can change the voltage in the network through the attackers control over the SM, which is a security attack (Zonouz et al., 2012). Thus, it is important to protect the consumption data for the users using remote attestation detection algorithms (Finster & Baumgart, 2014). Moreover, a new framework will be recommended for maintaining software security and data privacy of SM.
1.4 Research Questions

The research questions are:
1. What are the potential threats to alter the software deployed in the smart meter?
2. How can we secure the smart meter’s software?
3. How can we enhance the user’s privacy with architecture-based technologies?

1.5 Expected Results

This research will provide information for the reader about thesis topic, by giving a deeper understanding about the background of the topic. In addition, a structure of the framework will be proposed for tamper resistant software security and methods to protect the privacy of consumption data generated by an end user. The framework for this research is a structure of protocols using algorithms. This algorithm model is to quantify how much private information is disclosed. Design a protocol to enhance the security on smart meters by using mobile phone for extra authentication. Storing data in clusters on mobile phone attain privacy for smart meter. Measuring the impact of the proposed methods on the data and the valid results will be involved as well.

1.6 Limitations

The time frame for this thesis to be finished is 20 weeks. Thus, this research is limited for only two attack scenarios which will be defining, one for security and one for privacy. The main focus for this work is on the user privacy perspective. The main goal is to protect the SM and user’s data from the malicious attacks.
This is an extensive work which needs more research such as the authority of user smart phone, secure the communication between smart phone and SM. More details will be providing in the future work section.

1.7 Formulation of the Paper

The rest of the thesis is formulated as follows. Chapter 2 presents the research methodology which will be used for this thesis. Chapter 3 presents the background of the smart meter and its current security and privacy provisions. It also surveys state of the art of ensuring the security and privacy of the smart meter. Chapter 4 describes the Termineter framework and its uses. Chapter 5 discusses the results for the thesis research questions including attack
model, designed solutions and the evaluations. The last chapter concludes this research and discusses future work in this area.
Chapter 2
Research methodology

The main research in this approach is Design and Creation (D&C). A literature review and case study were used as research methods to answer the research questions. In addition, machine-learning techniques were used to determine the utility and privacy level of the proposed algorithms. Table 2.1 briefly describes the relationship between the research questions and the methods used to study and answer them.

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Literature Review</th>
<th>Design and Creation</th>
<th>Case Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the potential threats to alter the software deployed in the smart meter?</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How can we secure the smart meter’s software?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>How can we enhance the user’s privacy with architecture-based technologies?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2.1: Relationship Between Research Questions and Research Methods.

2.1 Literature Review

The research started with Literature review. According to Punch (2013), the literature review is one of the qualitative methods. The purpose of literature review is to understand a particular phenomenon. The author argues that one should try to create a deeper and more complete idea of the phenomenon being studied. A literature review provides the researches and the readers a solid theoretical understanding about the situation evolving SM in regards to privacy
challenges. Thus, scientific papers or books can be used for this purpose.

A literature research approach was used, which provides information for the reader on where the information is collected (Punch, 2013). For this research, all the scientific papers were accumulated from online resources. Digital databases, provided by Malmo University, such as the scientific articles from IEEE database, ERIC and the ACM Digital Library are used for this survey. The chosen literature, as well as the keywords used for search process are shown in Table 2.2.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Home (SH)</td>
<td>SH technology, structure, Advantages of SH and connectivity with SG and SM.</td>
</tr>
<tr>
<td>Smart Grid (SG)</td>
<td>SG technology, structure, difference between conventional power grid and connectivity with SM</td>
</tr>
<tr>
<td>Smart meter (SM)</td>
<td>What is SM, benefits of SM, challenges in SM, security threats in SM software, privacy threats in SM’s software, technical solutions for privacy, security in SM’s software, trusted third party in SM and enhance the security and privacy in SM.</td>
</tr>
<tr>
<td>Machine Learning (ML)</td>
<td>What is ML, evaluation with ML and ML in SM.</td>
</tr>
</tbody>
</table>

Table 2.2: Literature Review Keywords.

2.2 Design and Creation

This research develops a mission generator using D&C. According to Oates (2005), D&C research strategy is to solve a problem focused on developing IT artefacts. The motivation for using D&C is ”The art of designing artefacts to solve intricate problems” (Oates, 2005, p.188). Following D&C research could offer:

- Constructs (The concepts or vocabulary used in the user domain). This research used some technical words such as SM, smart grid, potential threats, trusted third party, security and privacy in SM, case study in SM.

- Model (combinations of constructs that represent a situation, and are used to aid problems, understanding, and solution development). The literature review has been adopted to describe the challenges and the current usage solution to solve those challenges.
• Method (The guidance on the models to be produced). This research proposes a case study to present the attack scenario for security and privacy of SM.

• Process stages (to be followed to solve problems using IT). The research presents a protocol model with algorithms as a solution for security and an enhancement for privacy.

• Instantiation (A working system that demonstrates and constructs, models, methods, ideas, genres or theories). Implements a framework as an instantiation to protect security and enhance privacy in SM. In Addition, design algorithm tests the performance with real-data and then enhances the proposed algorithm.

As a contribution to knowledge and the research, outputs are combinations of these stages mentioned above (Oates, 2005). This makes design and creation a good choice for this research as a building framework for security and privacy in SM that is an instantiation artefact.

According to Oates (2005), an instantiation artefact is a working system that shows the possibility of implementing our framework. For this research, a prototyping approach is planned to use to build a framework as an IT artefact to test our concept. The chosen methodology is influenced by the following reasons. According to (Oates, 2005), D&C research strategy is employed where the IT artefact itself is the main contributor to knowledge, because it followed a strategy to understand and evaluate the IT artefact in use. D&C process based on ”Learning via Making” concept, which uses an iterative process involving five steps:

• Awareness: Understand and identify the problem which is implemented in this research based on literature review.

• Suggestion: create an attack scenario and propose an algorithm to solve the problem

• Development: Develop and design the algorithms that were identified in the previous step. Develop a security and privacy algorithm, which is intended as a construction of a formal proof.

• Evaluation: Examines the developed artefact’s.

• Conclusion: Conclude the results and identify the knowledge.

2.3 Prototyping Model

This model is a development model where a framework is built, tested and revised until an acceptable framework is finally achieved. The prototype in question will be a framework to provide a security and privacy protocol with algorithms for SM. The proof of concept will be
shown via functioning in the prototype to improve the security and privacy of SM. According to Yoshikawa et al., (2015), the prototype approach clarifies the report, including how the implemented solution emerged from repeated cycles of analysis, design and implementation. The repeating cycle in a prototype system involves analysis, design and implementation stages. Then the understanding of these stages is used to modify the analysis and design models, which will be used to create a revised system prototype. These stages are repeated, and the prototype will gradually be modified until a satisfactory implementation is produced. This will help to find a better method to protect the end-user sensitive data. In addition, the prototype’s purpose is to show that the researcher’s designed solution protects the end user data (Yoshikawa et al., 2015).

Our prototype will build a framework to provide security and privacy for SM’s in a run time system. This framework will have two modules, one for security and the other for privacy. A literature research is used to extend on the existing framework in order to build a suitable one that can host our new modules.

2.4 Case Study

For the development of our framework, a case study is the chosen approach to this research. The purpose of the case study is to understand a small number of situations in great depth. The case study research design is also useful for testing whether scientific theories or models actually work in the real world, which in-turn is needed to test the real data from e-on using the proposed framework (Oates, 2005).

Our case study will look at the privacy and security behaviour within its real-life context and focuses on attack issues that constitute the messiness of the real world as observed in studies from literature review. Therefore, plans were made to perform a scenario to cover an attack. A new algorithms or protocols will be invented, or the existing ones will be extended to cover this attack. At the end of the test, steps to measure the level of the security by using this algorithm to check if this technique really prevented this attack will be carried.

Case study included two scenarios, one for a security attack to answer the second research question and another for privacy attack to answer our third research question. These two scenarios will be explained in detail in chapter 4.
2.5 Machine Learning

According to (Rashidi & Cook, 2013) Machine learning, intends to use a building block for privacy protocol and also to evaluate the proposed algorithms. Thus, it is used as a methodology in order to answer our third research question. The reason for self-tuning the system is to maintain its independence and personalization. The system should know the behaviour of the occupants and continuously learn and improve, and in order to achieve that, the system must find clock-work like patterns in the SM’s event data and learn these recurrent actions. (Schweizer et al., 2015).

There are two tasks in which machine learning can be achieved one is supervised learning. The system gets a dataset with different decisions, classification and parameter values, then from which it infers a mathematical function. It automatically draws an input signal to an output signal which then figures out what it is supposed to do (Liang et al., 2013).

Another one is the unsupervised learning, which means that the system acts and monitors the consequences of its actions, without referring to any predefined type other than those previously monitored. This learning is on a trial-and-error basis. Compared to supervised learning, unsupervised methods act badly in the beginning, when they are untrained, but as they regulate themselves, their performance surges. It can be contested that using unsupervised learning; a classifying system should be able to set up hypotheses that no one could understand because of the complexity of unsupervised methods that were used. A machine learning system would have to find a learner stage hypothesis all by itself, which would require more training data than it needs (Rashidi & Cook, 2013). An unsupervised task implementation is intended to be used in this research.

To calculate classifier performance given by a machine-learning program, a specific testing set or a cross validation technique can be used. A test set has pre-classified examples different to those in the training set, and it’s used only for assessing, and not for training. If there is not enough data, the cross validation method could be used so that no data goes unused or waste. This could be useful to improve classifier performance; all data is used for training the classifier, and for testing its performance (Liang et al., 2013).

Machine learning will be used to test the behaviour and evaluate the presented algorithms to protect the user’s data. This research investigates different algorithms that can essentially learn malicious behaviour. Machine learning algorithms specifically unsupervised learning will be used for analysing, and learning new ways, that the user’s data could be protected against attacks.
2.6 Method Description

The previously proposed research questions will be answered through the following phased approaches:

- Related work: Literature research is needed in order to address security, and privacy challenges, and investigate the existing techniques in the literature. An extended version of existing solutions for obtaining security, and privacy will be examined. This approach will help to answer the first question for this research.

- Concept and design of new privacy and security enhancing techniques: Design new methods needed to address the problem of privacy and security protection in SM. These methods cover the main aspects of protecting the user’s consumption data. Apart from these methods, propose the framework that hosts these algorithms to quantify how much private information is disclosed and to measure the impact of the proposed methods on the data and on valid results. To illustrate our approach, we apply it to a case study. This approach will help to answer the second question for this research.

- Implementation and test: Implement and test the proposed algorithms. Then, apply the algorithm to real datasets to test their effectiveness, and to study their performance. This approach will help to answer the second and third questions for this research.

- Evaluation: The evaluation methods using machine learning is to:
  - Compare various privacy, and security enhancing techniques.
  - The impact of specific privacy and security enhancing techniques on a certain task of a SM will be evaluated, by measuring the result of the task with and without applying these algorithms.
  - Evaluate the effectiveness of the algorithm on a real dataset.

The approach that is adopted to evaluate the security by simulating the formal model for security properties of the software running on the SM. To evaluate, privacy needs to employ machine learning techniques focused on the unsupervised techniques, to determine the utility and privacy level of the proposed algorithms. These approaches will help to answer the second and third question for this research.
Chapter 3

Literature Review

The literature review is the first effort made in this research. This section presents the background of the subject with more focus on the current state of the art for the SM domain in order to give the reader a clear understanding of this work.

3.1 Smart Home

![Figure 3.1: Home Network. Smart Home (Lugo et al., 2014).](image)

Literature review begins by explaining the terms and concepts in SH. This will provide the reader a deep understanding about the subject. Smart home technology opened new horizons to use most of the electronic devices in an intelligent way. These functionalities interact the way that people manage their private lives. People play an important part in this technology by adding it, and using it in their home management.

In order to provide a wide understanding of the subject SH, we conduct a literature review
Figure 3.1 illustrate an overview for SH.

SH technology has introduced the concept of device communication in several applications, which can be controlled in your home remotely. Inside the SH, Home Area Network (HAN) connects computers, thermostat and other electrical devices to an energy system. This gives an opportunity to control electrical usage via the web or any other application. This utility provides the usage information that enables better managing of electricity consumption. Inside Smart home there is a SM which communicates with all the devices in SH and supply this data to the energy provider. The smart grid is one of the smart technologies that has been built to enhance the electricity consumption and efficient distribution services (Lugo et al., 2014) for SHs. Figure 3.2 illustrate the connectivity of SM in .

![Figure 3.2: Smart Meters Connection to Power Provider (Wang et al., 2011).](image)

### 3.2 Smart Grid

A Smart Grid (SG) is a modernized electrical grid that utilises digital information and communications technology to gather information, such as the information about the behaviours of suppliers and consumers, in an automated technique to boost the efficiency, reliability, economics, and sustainability of the production and distribution of electricity (Wang et al., 2011). SG technology is being used in all regions of the electronic grid, which include transmission, distribution generation and the markets. The US Department of Energy (DOE) identified several properties needed for future power grids which include storage accommodation, conflict attacks, self-help, and quality of power, user interest and most importantly the enhancement of the markets assets. Many of these improvements enhance the communication, information and processing possibilities (Wang et al., 2011).
Advanced metering infrastructure (AMI) boosts electricity dispersion by positioning SM at the location of the users to reduce costs and reliability of the electricity and support dispersed generation. Generally, an AMI system has three elements: SM at consumer’s home, a metering communication arrangement between the customer’s home and its utilities, and a Meter Data Management system (MDM) (Nilsson et al., 2014). These meters give the user a rough estimate about their energy consumption, making it easier to increase integration of Distributed Energy Resources (DER) through which the users will be benefited from remotely observing blackouts, doing remote meter readings and offering prepaid alternatives (Wang et al., 2011). An overview of the SG is described in the Figure 3.3.

Figure 3.3: Smart Grid (Wang et al., 2011).

### 3.3 Smart Meters

This section provides a detailed overview of a typical SM. First, we describe a SM and its purpose. Second, we describe a typical structure of a SM. Finally, we list its functional components.

#### 3.3.1 Smart Meter

SM is one of the applications for SH along with other applications such as home-automation, multimedia and the key element of the Smart Grid as illustrated in figure 3.4. According to
Molazem, 2012. SM, is an electronic device that records the consumption of electric energy which can manage and monitor home power consumption through a network. Smart metering is a critical part of the ideal smart grid (SG). Its concept is the placement of electric meters that allow a two-way communication between meter and distribution system operator. The two-way communication allows services for the distribution system operator that were near impossible to understand without smart metering (Finster & Baumgart, 2014).

- The "intelligence" of SM is based in three basic functions:
  - Measures the used/generated electricity.
  - Remotely switches the customer off.
  - Remotely controls the maximum electricity consumption.

### 3.3.2 Smart Meters Structure

SM is a networked embedded system that has sensors to receive data of electrical current (Zhang et al., 2012). A brief explanation of the basic components of a SM, and a SM as an essential component of a SG are shown in Figure 3.2.

As shown in Figure 3.4 each SM has a base and a cover that protects the contents of the device. The cover has a seal, which protects the meter from being tampered with. Micro Controller Unit (MCU) is installed inside the meter which transfers data to a flash memory. Through an Analog Front End (AFE) the SM gets information about power and water usage. AFE collects analogue data, and converts it to digital data and sends it to the micro controller. There are memory elements inside the SM that work similar to a flash memory, the micro controller reads data and sends it to the utility server (Berthier et al., 2011).
The meter and the server have a Network Interface Controll (NIC) card so that they can communicate together. It is not practical to (Marmol et al., 2013) connect all meters direct to the utility server because the meters form a LAN and connect to servers through gateways. LANS’s are power line communication networks that send data to servers. Utility servers are connected to collectors through internet cellular networks, which collect all data from them. The SM must be installed with a Real-time Clock (RTC) which is synchronised with the server’s clocks in order to prevent clashes, and it should display the correct time. (Song et al., 2011).

3.4 Security and Privacy

Several approaches have been provided to protect privacy in the SM (Rial, & Danezis, 2011; Shepherd & Weldes, 2008). They could be applied to different solutions to improve privacy; however, they have some gaps. Improving privacy without improving security in the SM device will not make great solution that would protects from privacy issues. Actually, it does not matter how much the solution can protect privacy when an attacker can easily attack the device directly, and retrieve the data. Therefore, the current research will initially improve the security in order to protect privacy.
Firstly, we present the current related state of the art technology in security and privacy. Secondly, the related work has been done in this domain, and the different technologies have been used. Finally, we present open challenges.

3.4.1 Security

As mentioned in chapter 1, Shepherd and Weldes, (2008), define the security as "the state of being free from danger or threat". Security in SM is the state of protecting the user data from unauthorized access. The main focus is to ensure protection of personal information and more importantly to secure the functions of the society, since almost all information systems have become digitalized (Finster & Baumgart, 2014).

Nonintrusive load monitoring (NILM) is the power load signatures that find the appliances that are causing the load. For example, they noticed when appliances were switched on which indicated that the users are at home. They also had knowledge of the usage of an individual's specific appliance; this information was so private that legal measures were considered to safeguard the user’s privacy according to G.W. Hart in 1991. Hart used power loads that were measured every 5 seconds others worked on improving the NILM measuring higher time resolution of 5 to 10 seconds. Efforts were made to lower NILM time resolutions of up to an hour (Gungor et al., 2011).

There are SM that are capable of higher resolutions, which are already installed, but with dubious security features. This data could jeopardise the security of the user if exposed to data thieves who could use this data to burglar the user's home. Fan et al., (2013), explained in his paper that many threats exist in the domain of SM and there are many weaknesses and attacks are present in these systems. The leakage of personal information to malicious users can cause havoc for the user and smart metering systems.

Threats in Security

According to Finster & Baumgart (2014) and Berthier et al., (2011), there are various types of attackers that could be occurring in SM:

- Curious eavesdroppers who are spying on their neighbor's activities.
- Motivated eavesdroppers who gather info for malicious purposes.
- Dishonest users who want to steal electricity without paying for it.
• Intrusive data management agencies who want to use your private information to make user profiles and pass them on to marketing.

• Active attackers who want to make big attacks. Terrorists fall into this category.

• Publicity seekers who want to get famous and do not care about harming the users to gain financial rewards.

• Active attackers who want to make big attacks. Terrorists fall into this category.

• Publicity seekers who want to get famous and do not care about harming the users to gain financial rewards.

The above attacks could be carried out by any opponent and expose the vulnerabilities of the SM. This research will focus on curious eavesdroppers who are spying on their neighbour’s activities. More details about the findings in this domain will be explained in the following section.

Molazem (2012) and Berthier et al., (2011) categorize two software technologies to provide security for SM, Intrusion Detection Based Techniques and the Remote Attestation Based Techniques.

**Intrusion Detection Based Techniques**

Intrusion Detection Based Techniques (IDS) monitors systems and networks for malicious activity such as hacking the password for the SM application in order to divulge user’s data. IDS technique can be host based or network based. Host based are installed on the system to monitor communication between applications and OS. Network based are attached to a network that protects all machines from attacks (Berthier et al., 2011). The attack detection is done by IDS using signature based and anomaly based techniques. The signature based technique cannot detect new attacks, it can predefine sequence of events, and survey network traffic. While the anomaly based has the ability to detect new attacks, its false alarm rates are particularly high (Fan et al., 2013).

Finster and Baumgart, (2014) reported Specification based technique. This method is different from the other two methods because it detects new attacks using the behaviour of the system. On the other hand, this method is very difficult to apply.

However, the IDS is recognised as being very accurate in detecting unfamiliar attacks having low overheads, architecture and resilience of management against severe attacks (Molazem,
In Addition, SM cannot be fully secure with network intrusion detection systems that could have false negatives, which allow attackers to bypass the security mechanism and exploit software weaknesses (Berthier et al., 2011).

Remote Attestation Based Techniques

In remote party authentication, an application is used where both client and server share a secret key to avoid their session from hijacking. Today’s remote attestation techniques will rely mostly on challenge response protocol, a verifier sends a nonce challenge to the target device (Gungor et al., 2011).

The device uses a predetermined verification process to calculate a response to the challenge. A response is then sent to a verifier to be checked to see if it is correct. The reason for performing remote attestation is because malicious codes are being remotely inserted into the target system by attackers (Berthier et al., 2011). As an example, in 2008, a set of unauthorised codes was inserted into the server of a supermarket in US, and these malicious codes managed to gather credit card data of its customers and supplied this data further to a third party server. As a result, Khurana et al., (2008) mentioned that more than ”4200 credit” and debit cards were compromised. Administrators of systems are concerned about their ability to verify the integrity of a system and make sure legitimate codes are being used (Khurana et al., 2008).

However, Gungor et al., (2011) mentioned that the remote software attestation techniques are effective on the condition that it provides low overhead attestation services for SM. They don’t give any security guarantees for the system which means that meters are unmonitored, and an attacker can attack the vulnerable meters without even being detected. Two types of remote attestation techniques were mentioned, software and behaviour attestation (Molazem, 2012).

Software Attestation

Software attestation is used to find genuine software; although it’s a lightweight system, this technique does not guarantee any vulnerability on the meter software. Software attestation is based on a challenge/response protocol, there are two types of software attestation techniques one called OMAP a one-way memory attestation technique, and the Pioneer, a two-way memory attestation technique (Eldefrawy et al., 2012).

The OMAP technique is a one-way memory technique which sends sums of random selected
regions from the memory of the SM to the utility server. The server knows how to calculate the sums and can verify if the memory has been changed or tampered with. The one-way memory will prevent man-in-the-middle attacks against the SM (Brinkhaus et al., 2011). Pioneer a two-way attestation protocol is recommended by Song et al., (2011). It is based on a verification function. The verification function makes a checksum over the code and makes sure the code is not changed.

However, Molazem, (2012) explained in his paper that there are still some challenges with those technologies. For example, the challenger may pre-calculate the checksum, in order to block the attack, the checksum is dependent on the random challenge originally sent by the dispatcher. The challenger may run other code after the checksum has calculated and change the memory values.

**Behaviour Attestation**

In Behaviour based Remote Attestation Model (BTRAM), it monitors the behaviour of the software on the remote device, which will increase the security of the SM. Any vulnerability in the original software running on the SM can be exploited and detect the attack. In this case, software attestation might result in verifying the integrity of the system while the system has actually been compromised. Therefore, the behaviour attestation of the SM is an important approach on which this research is focused on (Ning et al., 2010).

The meter has to record the previously defined events on the system and periodically submit them to the verifier. In semantic based technique, monitoring is done through the lifetime of the application at a specific point in time. Behaviour based technique analyses the behaviour data and decides if it is trusted or not (Zonouz et al., 2012). According to Ning et al., (2010), the behaviour of the application, is defined and classified into two categories, system behaviour and application behaviour. To model the behaviour of the application, several attributes and values are defined. These attributes include auto-transmitting, auto-activating, self-protecting etc.

Gungor et al., (2011), reported that the behaviour attestation techniques are able to detect suspicious activity of SM, their issue remains that they put access overhead on the system. However, the challenges in behaviour attestation are:

1. Behaviour-based attestation techniques are not accurate.
2. Smart metering systems can only afford very low false negative rate due to the scale of the system.
3. Smart metering systems can only afford very low false negative rate due to the overhead of monitoring.

This research is focused on enhancing behaviour attestation techniques for security in SM’s software to be lightweight and effective in keeping the system secure in order to protect SM’s security.

3.4.2 Privacy

Privacy is the state of being free from being observed or disturbed by other people (Rial, & Danezis, 2011). To protect the data, SM will make sure that no unauthorized persons get access to a particular information or data. In some situations, it’s not just securing the content of the information, but to keep it’s existence a secret. Integrity implies that no unauthorized changes are made of the information, neither accidentally nor maliciously (Marmol, 2013).

Finster and Baumgart, (2014), mentioned that privacy is a serious issue with SM. SM identified specific usage data, which can cause a leakage of information about particular devices that are used in homes. This can build profiles of the user’s behaviour. This research will mention solutions for some of these issues.

Goncalves et al., (2011), explained that in many cases appliances can be successfully discriminated considering simple electrical features, as real power and reactive power. The appliances tend to form clusters depending on their critical characteristics. Multiple clusters can be expected for a finite state based machine depending on the differences between states as
shown in figure 3.6.

Gungor et al., (2011), suggested a solution where a trusted third party proxy is involved in meter reading from individual users, and collects data and adds random values to data to protect the privacy of individual users. Another suggestion is the use of homomorphic encryption to prevent electricity service providers from obtaining the usage data of individual households.

Threats in Privacy
The SM privacy concern is to be able to obtain the energy consumption data. SM has valuable information about consumers that could be used to explore the consumer’s lifestyle. This section will review some of the current threats on SM and the current technologies that aim to protect privacy in SM's.

Finster and Baumgart, (2014), recognise some data types from SM’s to check privacy threats against SM.

- Contact details: identifying users from invoices.
- Billing details from direct payments.
- Measurements: periodically collected from meters.
- Payment records: from the user’s payment history.
- High resolution measurements: for real time usage data.
- Smart appliance information.

They categorise the privacy techniques into two classes, architecture based and protocol based techniques to protect privacy.

Architecture Based Technique
The smart grid is modified in architecture based technique to address privacy. This modified architecture with Trusted Third Party (TTP) is needed to validate the identity of the meters and must have access to an endorsement key where each SM is equipped with one and is verified before installation of SM figure 4. The meter’s data are submitted to the server with a different source address or with a different ID, the source of this information is hidden from
the service provider (Petrlic., 2010).

TTP protects privacy, including the usage, by applying the encryption technique that prevents an adversary from observing the identity and the usage. The effort is to build a reliable trustworthy third party service, however, it could be a problem if the third party does not build trust. The whole system in this case would be unsuccessful. Furthermore, it does not provide the technique that prevents a competitor from observing the usage of energy. It only blocks the identities of the consumers, prohibiting the competitor from linking the usage with identified consumers. When a competitor continues to observe the usage of a small group of SM over a long period of time, it is possible to link the data (Efthymiou et al., 2010). However, the challenge is the size of the network, as it is very large for the modification of the smart grid architecture. The third party trust is not a promising approach (Petrlic, 2010).

This research focuses on an architecture-based technique for data privacy on SM that should be lightweight, and effective in keeping the system secure in order to protect the data in SM.

**Protocol Based Technique**

According to Fan et al., (2013) the main focus is on the calculations performed on the smart grid and the modification of the protocols for the privacy of the user. In this approach, it is mentioned that either the communication protocol is modified or data is modified locally to avoid the transfer of unnecessary information.

Gungor et al., (2011) mentioned that Load Signature (LS) moderation is proposed to hide information regarding the usage patterns of the user. LS is identified as a series of time stamped average power loads borrowed from energy values at short intervals. Load signatures can be used to draw information of the user’s activities, for instance, if extra appliances have been turned on it might indicate that the user is home. The main idea is that the users have access to energy storage and energy generator devices, which means that their devices can store free power.

Users do not have to use generated power through the service provider, and when usage is at its lowest the device can recharge through the service provider. This is called Load Signature Moderator (LSM), which smooths out data and also detects possible privacy threats (Lehnert, 2013) and (Peterlic, 2010).
The challenge with protocol based techniques is that they add more overhead to the system by inserting cryptographic calculations on the SM. The expected service time on the meters is around 20 years, and running cryptographic algorithms would be an issue for SM. There is no prototype to analyse attacks to the meter software and a scalability might be needed for the software system in the SM (Finster & Baumgart, 2014).

This chapter covered the background theory for SM starting from SH, SG and more details in SM. In the remainder of this study, we aim to improve SM security and privacy in order to protect consumer privacy.
Chapter 4

Termineter

This chapter presents the Termineter and its process to test the security of SM. This gives deep understanding of the tools that are used in this research. A description of SM communication will be introduced first followed by the attack methods using a Termineter framework.

4.1 Smart Meter Communication

As the Smart Grid grows, it is supported by academic research and industry and is in constant need to define a standard; however little effort has been made. In Canada, the US and in Europe, a number of standards have been in development. An example of these standards is the “Government of Ontario IT Standard Advanced Metering Infrastructure” (Mohassel et al., 2014). Determining each element in the AMI, which protects the user’s privacy, could be out of the scope of the current study, the specific focus is made on the communication between SM with utilities or SM with SM.

One of the standards proposed by Mohassel et al., (2014) is Wireless Personal Area Network (ZigBee) which is preferred for low-rate transmission in SM. The ZigBee standard offers security procedures that protect the network and application layers. The network layer applies the Advanced Encrypt Standard (AES) with the Counter Cipher Block Chaining Message Authentications Code (CBCM) that guarantees authenticity, and privacy which is central to current research. ZigBee meets all needs and polices of Mohassel et al., (2014). Open HAN Network System Requirements (NSRs) Wang, J., et al., (2011), are being developed in the US. Most current utilities use ZigBee, which meet the basic needs.

Batista et al., (2013) made an analysis of power management by using the ZigBee wireless that is based on saving money. ZigBee has several advantages. In fact, the ZigBee is an open
standard, established by the ZigBee Alliance, so it reduces the price of licenses. As shown in Batista et al., (2013), it supports a mesh network; which means that the meters can communicate directly with each other. The implementation of ZigBee in SM has its advantages over Bluetooth. In this research not only the cost efficiency is mentioned, but also protection of user privacy by securing the communication between the meters. The new standards do not address security issues. Joshua Wright showed an attack against ZigBee that sniffed out the data and obtained the Key (Johnny et al., 2010). But, the current research will address the privacy of the users. In effect, the ZigBee could be applied to other techniques of encryption to achieve both cost efficiency and privacy; but current implementation have just achieved cost efficiency.

Cartes et al., (2013) looked at the Security in Wireless Sensor Network based on the ZigBee standard. The ZigBee has several issues that include channel interference, address conflict and weakness in ASE repudiation. A symmetric key encryption is the cause of the weakness in ASE. As a result, the two nodes should swap the key before they communicate; during key swopping any opponent can eavesdrop. The opponent can then easily use the key to compromise the nodes. The National Institute of Standards and Technology considered AES-128 encryption will be secure until 2036; however as on date, it can be broken in just a five-minute attack. Symmetric key encryption has issues with key management when the number of nodes becomes bigger. As Cartes et al., (2013) mentioned, the application of Elliptic Curve Cryptography (ECC) might solve the issues since it’s asymmetric authentication and key swopping Bakhache et al., (2014).

Bakhache et al., (2014), demonstrates that the ZigBee standard is perfect for a low wireless network since it meets all the needs of businesses and manufacturers including a few advantages such as saving power and lowering costs. It also supports security in different layers. But Bakhache et al., (2014) identifies the security issue in ZigBee that is using Symmetric Key Key Establishment (SKKE)swopping . SKKE is fast and has a low-cost implementation; but on the other hand, it has security issues such as key swopping. As Bakhache et al., (2014), suggests applying public key algorithms based on ECC could solve the problem. The advantages of using ECC are scalability and non-repudiation, however ECC uses only one key. Additionally, ECC has advantages over the traditional public key system which have faster computations and less significant key size. For example, Rivest, Shamir and Adleman (RSA), The Digital Signature Algorithm (DSA), and Diffie-Hellman are not good systems because those need large key and massive computing. Bakhache et al., (2014) shows using the ZigBee with ECC will improve the security of communication, and it can fit more than the normal public key system.

In fact, Bakhache et al., (2014) applied Elliptic Curve Menezes Qu Vanstone (ECMQV)
as a key establishment mechanism for ZigBee. But, ECMQV is not approved and has been dropped from the National Security Agency’s cryptographic standards. The reason for that is ECMQV uses signed Diffie-Hellman, which is vulnerable to a man-in-the-middle attack.

Mohassel et al., (2014) presented a solution to safeguard communication between the AMI and Smart Energy Utility Network (SUN). To enhanced security in standard protocol, IEEE 802.15.4 and ZigBee Alliance using SKKE are recommended protocol for Key establishment and management. This idea is based on public key cryptography, and as is shown in [4]; the algorithm which depends on four phases: Key establishment, data encryption/decryption, key and trust data update, and finally orphan node management. As Mohassel et al., (2014) concludes, the proposed algorithm gets better than the SKKE algorithm when the amount of hops is greater than 2 hops. However, this research aims to discover a secure communication algorithm in order to protect the privacy of the user as long as it keeps its cost efficiency. However, Mohassel et al., (2014), does not provide the computing power needed to archive the public key algorithm which is said to be higher than SKKE.

In summary, several approaches have been given to secure a communication channel between SM and utility collectors. But, these approaches have issues either in the security or cost efficiency. Firstly, the issue of security means the approach has been seen as insecure or broken. Secondly, cost efficiency means the approach may be secure, but needs some added functionality that increases the cost of SM significantly. Unlike the above mentioned approaches, the proposed scheme achieves both privacy and cost efficiency.

4.2 Attack Method

Depuru et al., (2011) has shown that SM has a high risk of being compromised, in fact the devices that are installed are in unsecure neighbourhoods. As a result, an adversary can easily launch a physical attack, which can lead to a software attack. In general, privacy cannot be reached without security. Briefly, the SM as a part of AMI has specific characteristics that makes it difficult to secure, however, any attempt to solve the security without considering these characteristics will be unsuccessful.

According to (Schweizer et al., 2015), the SM is a small embedded system that has a modular structure, and three complicated components: the processing unit, the communication module and the electrical meter, which have a big chance of weakness as specified in chapter 3. For example the attacker can use jailbroken firmware, and social engineering can be used to advertise products both instances can reduce electricity usage, and the attacker will gain
not only financially he will also gain access to many smart meters although a remote attack isn’t possible. There are several methods documented on how a rival can take control of SM although this is quite new information, it is to be expected that there are more unknown vulnerabilities (Depuru et al., 2011).

This thesis discusses on one of the framework that has been used to attack the SM called Termineter which will be describe more in the following section.

4.2.1 Termineter

Termineter framework is the proposed framework to implement the attack scenarios in this research. The Termineter framework allows authorised people to test SM for vulnerabilities such as energy consumption fraud, network and hijacking. Many of these vulnerabilities have been brought to attention by the media and opponents have been sent out by law enforcement agencies. This is a big problem for energy companies, as Secure State sees most of these types of vulnerabilities to be having dangerous effects on the security landscape (Knapp, E. D., & Samani, R. (2013).

A public release is needed for this utility to promote security awareness for SM and provide a tool that brings basic testing capabilities to the community and to the meter makers so that security can be improved. Power companies can use the framework to identify internal faults that expose them to fraud and significant vulnerabilities. Users must have good knowledge of the technical works of meters in order to use Termineter (Knapp, E. D., & Samani, R. (2013).

According to Knapp, E. D., & Samani, R. (2013) termineter is a application framework

![Termineter Framework Interface](image)

Figure 4.1: Termineter Framework Interface (Knapp, E. D., & Samani, R. (2013)).

written in python which provides a program application to test SM security.
4.2.2 Basic Steps

Below is a summary of the basic steps to get started with Termineter after the environment configuration ("Termineter Documentation Termineter 0.1.0 documentation", 2016).

- Connect the optical probe to the SM and start Termineter
- Configure the connection options. On Windows, this would be something like COM1 and on Linux is something similar /dev/ttyS0. Check Configuring the Connection for more details.
- Use the connect command, this will check if the meter is responding.

Termineter can be used to modify the software meter, and reduce fee that users pay for electricity. It could also communicate with the accountant to report less consumption. The application will connect to the meter or SM via an interface ("Termineter Documentation Termineter 0.1.0 documentation", 2016). For communication it uses C12.18 and C12.19 protocols, C12.19 with 7-bit character sets supports current SM’s. Using an ANSI type-2 optical probe with a serial interface, termineter communicates SM Termineter communicates with SM through a probe ANSI-2 type optical with an interface series. ("Termineter Documentation Termineter 0.1.0 documentation", 2016).
Chapter 5

Result and Evaluation

This section presents the results of the research. The result is in the same order as the research question for this thesis.

5.1 First Research Question

As mentioned in description method (chapter 2), the literature review was the first activity carried method done for this research to answer the first research question.

Research Question 1:

“What are the potential threats to alter the software deployed in the smart meter?”

Answer:

Several researchers have reported several ways for attacking the SM’s resources. The results for this research question are categorized to security and privacy in order to clarify the result from security and privacy perspective as follows:

- Potential Threats in security

    As mentioned in the literature review, vulnerabilities are taken advantage of by attackers and their reasons are simply to cause havoc or damage to the network, these attackers expertise are paramount they know what they are doing.
The potential threats in security could be as mentioned below, (Finster & Baumgart, 2014), (Berthier et al., 2011), (Gungor et al., 2011), (Molazem, 2012) and (Khurana et al., 2010):

- Malicious attackers are often motivated by their need for challenge and curiosity they challenge themselves to break the operation system and security as if they were playing a game.

- Accidents made by employees could cause errors in the system; Smart grid could be seen as an easy target for terrorists as millions of users would make their cause more noticeable.

- SM’s software takes over control by spreading malicious software in a smart grid with a peer-2-peer topology which would effectively allow an attacker to take control over SM. Controlling a SM’s software would permit attackers to cause huge damage. For instance, forcing thousands of SM to turn off and on simultaneously could cause major trouble due to quickly changing load conditions on the power grid.

- Network Availability: Since smart grid uses IP protocol and TCP/IP stack, it becomes exposed to DoS attacks and to the weaknesses that are a typical feature in the TCP/IP stack. DoS attacks may try to corrupt information, delay, block, or infect SM’s and servers.

**Potential Threats in privacy**

The threats to privacy directed to attack the private data in SM. The following potential threats are summarize based on literature review (Rial, & Danezis, 2011), Gungor et al., 2011), (Petrlic., 2010), (Lehnert, 2013), (Pfitzmann, A., & Hansen, M, 2010), (Aloul et al., 2012), (Skopik et al., 2012)and (Ning et al., 2010):

- Contact details: identifying users through invoices where the attacker uses the invoice information to attack the user’s data.

- Billing details for direct online payments where the attacker uses the payment information to attack the user’s data.

- Measurements: periodically collected from meters where the data will be collected from the SM’s memory.

- Malicious software is used to change or add any function to a device or a system such as sending sensitive information. Consumers bearing a grudge towards other consumers making
them find ways to shut down power in their homes.

- Consumers driven by vengeance and vindictiveness towards other consumers making them figure out ways to shut down their home’s power.

- Employees bearing grudges on the customers or inexperienced employees could cause unintentional errors. Or rivals confronting each other in order to gain financially.

- Competitor’s attacking each other for the sake of financial gain.

- Compromising communication: Attackers can damage parts of the communication equipment, like Multiplexers can directly harm the communication equipment by using it as a backdoor for future attacks.

- Using false information: An attacker can send packets to plant false information into the network, such as fake prices and incorrect meter data.

- Eavesdropping and system behaviour: An opponent can obtain sensitive information by monitoring the system behaviour. Examples of monitoring information include future price information, and control structure of the grid, and its power usage.

Based on “Learning via Making” concept which this research adopted, the above result gives the awareness for this topic and gives suggestions to create the needed attacks scenario in order to find a solution and give answers for the second and third research questions.

5.2 Second Research Question

To answer the second research question, an attack scenario has been presented based on literature review, then there is a solution to prevent this attack. The results are described as follows.

Research Question 2:

“How can we secure the smart meter’s software?”
Answer:

To secure the software in SM, a remote attestation protocol was chosen. As mentioned in chapter 3, the remote attestation will be used to solve the second research question. Our protocol is the enhancing protocol model for the Ad hoc On Demand Routing Protocol (AODV). The result starting with an attack scenario to take over control of SM, which causes vulnerability for the user’s consumption as follows:

5.2.1 Security Attack

The aim of this attack is to improve the security by developing a new technology. According to the result of this attack, the outcome of the research will apply a new method in order to avoid a SM security attack.

The chosen threats represent circumstances that have the possibilities to cause problems that can put the security features of the system in danger, such as taking over control of SM, while security mechanisms identify possible protection mechanisms of achieving protection objectives.

During the communication between energy provider, and user an attacker can easily break into a network by exploiting the weakness in the system. Then the attacker monitors the transition between SM and energy provider and injects an attack code to obtain valuable information, like serial number for the SM. This will help to run a command to reset the passwords and keys. Once the attacker gets a password, the attacker takes control over the SM and can modify/destroy the device. In other words, the target device has been essentially rooted, and the attacker has absolute control, and the behaviour of an entire system could be altered.

![Figure 5.1: Eavesdropping Attack.](image-url)
**First Scenario steps:**

As illustrated in figure 5.1, take over the control of SM could be done by the following steps:

- **Step 1:** Attacker monitors the network to find the weakness in the system and connecting the SM to the Terminator.

- **Step 2:** Attackers inject malicious codes to identify the SM’s serial number by running the following commands.

  ```
  Use get-info---------- Get basic meter information from reading tables
  Use get-info---------- run
  Use read_table --------- Read Data from A C12.19 Table
  Use run_procedure------ Initiate A Custom Procedure
  Use set_meter_id-------- Set The Meter’s I.D.
  Use set_meter_mode------ Change the Meter’s Operating Mode
  Use reset_table_data (table id, data, offset=None) -------
  reset Data to a C12.19 Table
  
  Figure 5.2: Inject Malicious Codes.
  ```

- **Step 3:** Another code to hack the username and password.

  ```
  Brute-force-login
  Brute-force-login>set USBMax true
  Brute-force-login>run
  set_meter_mode> run ----------- Change the Meter’s Operating Mode
  write_table------------- Write Data to a C12.19 Table
  Exit
  
  Figure 5.3: Code to Hack the Username and Password.
  ```

- **Step 4:** Attacker takes control over the SM to read/write the operations in physical memory. The attacker can easily change the username and password. The attacker has the control over the SM after this step. Then the attacker can stop the SM, run it again,
or even burn the devices at the user’s home.

5.2.2 Solution:

Create a new method to protect the SM. We will consider a protocol to improve the security of SM’s. However, the solution proposed in this research will send an alarm to the user and Utility Provider (UP) at the same time when someone tries to connect to SM. Then the user can shut down the system or reset the application to avoid these attacks. When the SM sends an alarm to the utility, it will provide with performs maintenance.

A smart phone (SP) is used as third party, where the user has complete control over his/her SM even as communication is going on between SM and UP. At the same time authorization is needed for he/she to use SM application from the utility provider.

To perform computations such as energy and efficiency advice and tariff comparisons the UP depends on accurate meter readings. SP platforms could be used to perform such computations on certified readings in a discreet private approach. Figure 5.4 shows the communications between SP, UP and SM.

![Figure 5.4: Communication Between SP, SM and UP.](image)

The answer based on remote attestation is to boost the security in the SM software. The purpose of remote attestation is to allow the utility supplier to clarify the level of trust in the reliability of platform in the customer’s system. A technique known as remote attestation, where a SM/Customer validates their software configuration to a remote utility supplier.

This research uses the remote attestation protocol to boost the security in SM. We put together a protocol to boost the security of the SM and it then assumes that the protocol is secure certain assumptions of trust can be used for remote attestation.
Remote Attestation Protocol

This thesis used AODV. The ultimate goal of the security solutions for AODV protocol is to provide security services, such as authentication, confidentiality, integrity, anonymity and availability to mobile users. In order to achieve these goals, the security solution should provide complete protection spanning the entire protocol stack (Kushwaha, S., & Chaudhary, D. (2015)).

Ad-hoc networks are infrastructure-less. The interconnected nodes coordinate the transmissions with other nodes and they may have to relay messages among several other nodes in order to reach a destination. An ad-hoc network is a self-configuring network of wireless links connecting mobile nodes, where these nodes may be routers and/or hosts (Kushwaha, S., & Chaudhary, D. (2015)).

Packet transmission using multiple channels and Multiple Interfaces. Due to the increasing throughput demand, the idea of exploiting multiple channels is appealing in ad-hoc wireless networks. Having more than one interface on a node allows two different nodes communicating in parallel on different channels. An interface has a capability to dynamically switch to different channels over time. Two adjacent nodes can communicate with each other if they have at least one interface on a common channel (Kushwaha, S., & Chaudhary, D. (2015)).

This research integrates the process of remote attestation security protocol intended to protect data over network connections above the transport layer using cryptography for key exchange, symmetric encryption for privacy, and message authentication codes for message integrity. This process protects data end-to-end by encryption (confidentiality) and hashing (anti-tampering) and it authenticates users typically by means of certificates. This protocol designed to provide communications security in the application layer. This research suggests a modification on the process of the exchange message between SM and UP. Thus, to enhance the protocol to provide the required level of security and trust. To reach this goal, a SP is introduced as a third part in this protocol. In this case, the client is a SM, server is the utility provider, and the third part which is a SP. We assume that the SP application is already installed on the SP and all the settings regarding the communication with the SM has been done as well. The steps for reading the data from the SM will follow the steps below:

- SP sends a message to SM to monitor the electric consumption.
- SM asks permission from the UP along with its identity and the serial number to do the communication between the utility provider and user’s SP.
- Utility provider asks for confirmation to the user’s SP to verify the request. To check
the authentication.

- If the user responds with No or does not respond at all, then the request will be rejected and the process will be terminated.
- If the user reply yes, it should be with a cipher suit or key crypto options.
- Utility provider grants the access together with a digital signature or key to communicate with SM (log in and read the data). This is like a certification for user authorization to use SM.
- SP confirms the key acceptance within the time limit. If the client did not answer during this time, then this certification has to be discarded and a new certification has to be sent to the client. The client has to be informed about the delay, then a new certification will be sent to the client and the client has to answer quick before the certification discard again.
- In case the certification has not received within 3 times the utility provider has to restart the system to prevent any attack to the SM software.
- Utility provider terminates the handshake.

If the UP in the above step trusts the integrity of the client’s platform, then the UP and client continue to exchange messages as in the new protocol to establish a secure session, which is illustrated in figure 5.5.

![Figure 5.5: Security Protocol.](image)

In addition, SP applications are used to send real time alerts to costumer’s SP. In this paper, we suggest using these applications as an alert system as well. The user would receive
a message if any changes in pricing are made or to confirm/refuse to carry out an action which
could be for example, a validation of the operation of appliances.

The adopted detection protocol could protect any SM’s compromise attempts in the early
stage to enhance the security of SM’s software. Different from other protocols, this protocol
involves the user and the provider at the same time to alert them on attacks that are happen-
ing in the SM.

The proposed protocol is evaluated by using network stimulation tools to stimulate the
network environment. This implements the proposed protocol on this environment to observe
the change on the data package.

### 5.2.3 Security Evaluation Approach

The security attack scenario has been simulated using Network Simulator-2 (ns-2). The reason
for the simulation is to present the real system by simplifying assumptions on each possible
detail of the specifications of the system. According to Issariyakul, T., & Hossain, E. (2011),
simulation is defined as "the process of designing a model of a real system and experimenting
with this model for the purpose of understanding the behaviour of the system for evaluating
various strategies for the operation of the system".

Ns-2 is an event-driven simulation tool used to learn the activities of communication net-
works. Ns-2 can be used for both wired and wireless network functions and protocols (e.g.,
routing algorithms, Transmission Control Protocol (TCP), User Datagram Protocol (UDP)
). Figure 5.6 shows the basic architecture of ns-2.Ns-2 is an open source network simulation
application consisting C++ and Object-oriented Tool command language (OTcl) (Issariyakul,
T., & Hossain, E. 2011).

![Basic Architecture of ns-2](Issariyakul, T., & Hossain, E. 2011).
Here the C++ describes the internal structure which is a back-end of the simulation objects, the OTcl sets up simulation by not only "assembling and configuring" the objects but also scheduling events which is a front-end. The C++ and the OTcl are linked together using Tool command language(Tcl). Mapped to a C++ object, variables in the OTcl domains are sometimes referred to as handles. Theoretically, a handle is just a string in the OTcl domain, and does not contain any functionality. Rather, the functionality is defined in the mapped C++ object. In the OTcl domain, a handle is as a front-end which communicates with user and the OTcl objects. It interprets its own procedures and variables to facilitate the communication. Kumar, P.et al., (2015, May) mentioned that most of the protocols has been developed by researchers and adopted into standard version of ns-2 and the current version is ns-2.28 which maintains the standard protocols (Issariyakul & Hossain, 2011).

For specific experiments there are specific requirements in ns-2. The purpose of this thesis is to add security to the user data through encryption. Hence we enchanced existed code from (Kumar, P., Maheshwari, S., & Dubey, H. K. (2015, May)) where they used Caesar cipher to encryption, decryption the algorithms with hashing functions for integrity(Greis, 2005). This might be extended further, We used a basic hashing algorithm to get hash value for the text. This value is added to header of the packet to confirm the data integrity. By the end of communication and after decryption is done, this decrypted text will be hashed again in order to produce new value, which will be checked against the existed value inside the packet header. If they are same, the decrypted text is accepted by ensuring the integrity; otherwise the packet will be discarded. In either cases, an packet should be send to sender to inform the packet status. As of limited time we used a simple hash function algorithm that we obtained from this (Udgata, S. K et al., (2011, June)), where they used Caesar encryption and decryption with "pre-shared key length of 3". These cryptographic function takes string as input and shift the American Standard Code for Information Integration (ASCII) value in three positions to each character in the text (Udgata, S. K et al., (2011, June)).

5.2.4 Simulation Scenario

In order to perform and evaluate the impact of the proposed scenario attack and the solution, we demonstrate a simulation scenario through a simple network topology by using ns-2. It consists of following three steps.

- Step 1: Simulation Design
  
  Figure 5.7 displays the configuration of a network under considerations. This network topology has four nodes n0 to n3, where n0 presents SM, n1 as malicious and n2 as SP and n3 as UP. In this scenario, node n0 sends traffic to node n2, and node n2 transfers
data to node n3 using a File Transfer Protocol (FTP). These carried sources are carried by transport layer protocols through Transmission Control Protocol (TCP). In NS2, the relaying object of these protocols are a TCP agent and a TCP sink agents are the receiver’s. Table 5.1 represents the simulation parameters along with their corresponding values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td>The enhancing AODV protocol</td>
</tr>
<tr>
<td>Simulation time</td>
<td>12s</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>4</td>
</tr>
<tr>
<td>Mobility Model</td>
<td>Constant Position</td>
</tr>
<tr>
<td>Traffic type</td>
<td>TCP</td>
</tr>
<tr>
<td>Rate</td>
<td>54Mbps</td>
</tr>
</tbody>
</table>

Table 5.1: Parameters Used for ns-2 Simulation.

Step 2: Configuring and Running Simulation.

To implement n1 as malicious node, the used commands are shown in figure 5.8. The Tcl simulation script implements the simulation scenario. We run this script by this command: (ns filename.tcl) from the shell environment. After the execution, trace file will be created as shown in figure 5.10 and Network Animator(NAM) will be running to show the output graphically as shows in figure 5.9.

Step 3: Tcl Script Outcome.

Figure 5.7: Simple Network Topology.
Packet tracing records the details of packet flow during a simulation, which is classified into a text-based packets (ex., as.tr files) and a NAM packet tracing. Text-based packet tracing gives details of packets passing through network checkpoints (e.g., nodes and queues). Some part of the trace file obtained by running the above simulation is shown in figure 5.10.

The first column in the figure 5.10 is the trace file shows type identifier and corresponds to possible events. Next column represents the time at which such event occurs. Third column is current node and fourth column is packet name and sixth is size of the packet. Seventh column indicates the flags and abnormalities in the flow. As it shows abnormal-
ity that, the malicious node 1 is dropping the packets, where as ”- - - -” shows no flag (Issariyakul, T., & Hossain, E. (2011)).

NAM trace records simulation detail in a text file as filename.nam, and uses the file to play back the simulation using animation. This NAM file can be initiated directly at the command prompt through the following command ”$ nam filename.nam” (Issariyakul, T., & Hossain, E. (2011)).
The output of this file can be shown in figure 5.11.

![Diagram showing network traffic flow](image)

**Figure 5.11: Applied Cryptography Algorithm.**

As mentioned earlier we applied cryptography algorithm to the above scenario and the output of the script can be seen in figure 5.12. This shows the traffic flow from node n0 to n3 without dropping the packets. A sample of the enhanced code that used from (“Surajpatilworld Blogspot Com”, 2016) is shown in figure 5.12.

```plaintext
#Define a 'recv' function for the class 'Agent/SpecialPacket'
Agent/SpecialPacket instproc recv from rt msg originmsg hash {    $self instvar node_
        puts "node [\$node_id] received packet from \
            from with trip-time \$rt msg - contened: \$mess -
            decrypted \$originmsg -hash: \$hash"
    }
    #Create security agents and attach them to the nodes n0 and n2
set p0 [new Agent/Security_packet]
$n$ attach-agent $n0 $p0
$p0$ set class_1

set p1 [new Agent/Security_packet]
$n$ attach-agent $n1 $p1
$p1$ set class_2
```

**Figure 5.12: Code for The Security Protocol.**
Output can be seen as shown in figure 5.13.

```
paulina@ubuntu:~/Desktop/ns2real/ns-allinone-2.35/ns-.35/taotv$ ns 1.tcl num_nodes is set 4
INITIALIZE THE LIST xListHead Message sent start with hashing 1468
channel.cc:sendUp - Calc highestAntennaZ and distCST
highestAntennaZ_ = 1.5, distCST_ = 550.0
SORTING LISTS ...DONE!
Message sent needpermission with hashing 596190 Message sent permissiongranted with hashing 4378074 data integrity ensured
node 3 received packet from_2 with trip-time 11.3 ms - contend: qhhgshuplvvlrq - decrypted needpermission -hash: 596190
data integrity ensured
node 2 received packet from_3 with trip-time 3.1 ms - contend: shuplvvlrqjudqvhg - decrypted permissiongranted -hash: 4378074
node 2 received packet from_3 with trip-time 14.9 ms - contend: Message Accepted - decrypted _ -hash: 0
Message sent sharedata with hashing 19284 data integrity ensured
node 2 received packet from_3 with trip-time 1.7 ms - contend: vdkuhgdw - decrypted sharedata -hash: 19284
node 3 received packet from_2 with trip-time 8.5 ms - contend: Message Accepted - decrypted _ -hash: 0 node
3 received packet from_2 with trip-time 5.3 ms - contend: Message Accepted - decrypted _ -hash: 0
Message sent sharingdata with hashing 77652.
Message sent receivedterminates with hashing 22027230 data integrity ensured
node 2 received packet from_3 with trip-time 1.6 ms - contend: uhflhybgjwhupldwhv - decrypted receivedterminates
hash: 22027230 data integrity ensured
node 3 received packet from_2 with trip-time 3.3 ms - contend: vdkulqjdwd - decrypted sharingdata -hash: 77652
node 3 received packet from_2 with trip-time 5.0 ms - contend: Message Accepted - decrypted _ -hash: 0 node 2
received packet from_3 with trip-time 7.2 ms - contend: Message Accepted - decrypted _ -hash: 0
```

Figure 5.13: Output from Applying Security Protocol.

This code has been modified from the code listed on ("Surajpatilworld Blogspot Com", 2016).
5.2.5 Security Evaluation Results

In order to visualize the data analysis, we used ns-2 Visual Trace Analyser tool, which is able to easily translate the ns-2 trace files into graphics and to give a summary of what occurred during simulations. This tool is capable of analysing the trace files and making graphics and statistics about the behaviour of the flows in simulation (Issariyakul, T., & Hossain, E. 2011). Loading the simulation files requires two steps:

1. Firstly, load the Tcl file, so that the application can build the topology and get some extra parameters;

2. Second step is to load the trace file, with all packets information.

This Tcl file will help the application to count and place the nodes in the right places; it will also prepare the application to read the trace file properly. After loading the Tcl file, the visual interface will open. It is able to see the nodes position, colour and interact with the view. Loading the trace file is the second stage. After loading the visual interface, it is possible to load the trace file in the simulation tab. Loading the trace file can be an exhaustive task because the trace file can be very large. The process is multi-threaded, buffered and optimized to use the most computer processing power possible. An informative window will appear during the process, showing the current status of the loading process can be seen in the figure 5.14.

![Buffering After Trace File is Loaded.](image)

Figure 5.14: Buffering After Trace File is Loaded.
And also it shows the number of packets that has been transfer and dropped packets is zero which can be shown in figure 5.16.

As a summary, the proposed protocol would be considered as an endeavour to enhance the security requirements of the aodv protocol. Further improvements are needed to make our
solution to be the best by adding digital signing for strong encryption. Thus the protocol proposed, which is implemented, will inherent the added advantages over the security conscious protocols designed for AODV.
5.3 Third Research Question

To answer the third research question, an attack scenario was presented, based on literature review. The second step is to find a solution to prevent this attack. The results are presented as follows.

Research Question 3:

"How can we enhance the user’s privacy with architecture-based technologies?"

Answer:

The chosen attack scenario is to inject false data to the SM and steal the user’s privacy data. The attack scenario explains as follows:

5.3.1 Privacy Attack Scenarios

This attack scenario was designed based on literature review. The chosen threats represent circumstances that have the potential to cause problems, which can put the user’s privacy in danger such as distribution of the user’s sensitive data.

In this scenario, the attack is to tamper with or even delete legitimate consumption data in the physical memory of a SM by injecting malicious codes or false data for the sake of financial benefits. This type of attack is usually called a false data injection attack, which is illustrated in Figure 5.18.

![Figure 5.18: Malicious Attack.](image-url)
Second Scenario steps:

Tempering attack:

- Step 1: the attacker chooses a smart meter to observe.
- Step 2: when SM sends a message to an energy provider, the attacker reports with the time event.
- Step 3: the attacker creates a timetable with event reporting.
- Step 4: After a period of time (e.g. week/month) the attacker can determine the pattern of activities, which is evidence of the householder being absent from the residence.
- Step 5: Then the attacker will modify/change the data in the SM.

According to above attack, this code could help to execute this attack.

```
Step 1: connection.ConnectionBase(device, c1218_settings={}, serial_settings=None, toggle_control=True, **kwargs)
Step 2: get_table_data(tableid, octetcount=None, offset=None)
Step 3: get_info
step 4: run------gives SM basic info
Step 5: set USEHEX true ----gives authorization details
Step 6: run------ give the attacker to modify the username and password
step 7: reset_table_data(table id, data, offset=None)
Step 8: exit
```

Figure 5.19: Malicious Attack Code.

5.3.2 Solution:

This thesis uses architecture-based methods to enhance the privacy of data in SM. The goal of architecture-based methods is to present functionality and benefits comparable with conventional solutions, in which personal data is available to third-parties, without disclosing this data.

The solution is proposed to use an extra storage device on the SP. We introduce a technique to hide information locked in the user’s data by lessening home load signatures, after mixing utility energy with energy supplied and SP storage. We recognise the storage device in SP as an extra memory for SM. The user’s information that was sent to the utility provider from
the SM’s memory will be split between the SM’s memory and the storage on the SP.

So that appliance usage events can’t be detected, the solution is to hide the load signatures in practical situations. Complete privacy state cannot be fulfilled due to a number of physical limitations and cost, which draws attention to an energy mixing elevation problem (Kalogridis, G. et al., 2010).

The idea of "load signature" is to be a series of time-stamped average power loads p(t) copied from (cumulative) energy values e(t) metered at different periods. A "home load signature" is the sum of all home appliance loads. Information locked in a home LS comprises of individual consumption events identified within the home from a privacy point of view. We describe LSM as a load signature transforming method, with which the presence of appliance LSs may be hidden, or made unclear. Changes to the definition of "un-detectability" from (Pfitzmann, A., & Hansen, M., 2010), we discussed that privacy is protected when, a home load signature is given, we cannot really determine whether an appliance load event exists or not. LSM is responsible for shaping LSs via power routing. Note that the LSM and smart meter could be separate devices or it could be integrated into one device, and they could be located near distribution boards. LSM system protects SM data privacy without depending on the security of the SM.

In order to carry out LSM, we suppose that future SHs will have a variety of energy generation and storage devices, and electrical power routing is possible. This would mean the selective control and power mixing of a number of electricity sources to "route" electricity to a number of consumers is achievable. (Kalogridis, G. et al., 2010).

The main aim of LSM is to find and detect privacy threats, and respond by "configuring power routing". The LSM can detect a threat after it identifies a power usage event in the home, this could be a power trigger generated by a user or a supplier, such as a change in power usage like switching a device on or off. Threats can be detected after processing information communicated inside the HAN: smart appliances may notify the LSM of scheduled or predicted future events inside the HAN. For instance, the LSM and the washing machine can plan an operational schedule (e.g. when a private supplier is speculated to sufficiently cover projected energy demands). In this case, the LSM needs to find resources and prepare power routing configuration rules for probable power events. An LSM storage mix, for example, storage in a SP, algorithm should be subject to various limits such as storage efficiency and available SP resources (Kalogridis, G. et al., 2010).

For distributing the consumptions data, Cluster algorithm is recommended (Pfitzmann,
A., & Hansen, M., 2010). This algorithm is modified for its use with SP to carry some of the output load when the user is at home using their devices. The storage on the SP sends this data to SM’s memory gradually when the user is not at home. This method manages to show the data in SM at all times, it shows the user is at home and active. At the end of the time interval, the complete usage data is sent and saved in the SM’s memory. This method does not have an effect on the billing information for each month. The only difference in this case is that, the user will be paying for the same amount of electricity for each hour instead of paying a different amount depending on the usage of electricity at different times (Pfitzmann, A., & Hansen, M., 2010). For example, the only difference in this case is that, the electricity rate per hour will stay the same even though at certain times through the day less electricity is consumed. For example, if the user uses 20 Peaks of electricity between 7 am and 8 am, then the user leaves the home for 8 hours, when the user comes home he/she uses 30 Peaks between the hours 8 pm and 9 pm. The SM will send to the UP the consumption data for 20 Peaks between 8 am and 9 am, nothing during the 8 hours the user isn’t home and 30 Peaks during 8 pm and 9 pm. According to our algorithms the consumption data will be distributed for 24 hours, and the SM will send 2.1 Peaks/hourly during the 24-hour period. This means the SM will send the same consumption data for the user all the time. This hides the user’s availability to protect his privacy. Notice that at the end of each interval the whole data will be transferred to SM and the user will get a bill for the whole consumption period as usual. The storage device on the SP will be empty and ready for the next interval to save the usage data.

The Storage Device Model

![Storage Model Diagram](image)

Figure 5.20: Storage Model.

In figure 5.20, X is the accumulated load of the appliances. The output load Y is the combined load of the appliances and the SP, as described by the SM to the SP. Power, flows right-to-left from the SP through the storage to the household’s appliances. At any time, the storage device in SM may perform a mixture of the following actions, transmission of power
directly from the utility to the appliances; store energy from the utility company for future use; deliver previously stored energy to the appliances. In this way, charging and discharging the battery can manipulate the output load $Y$, hiding some of the information inside the input load $X$.

**Algorithms Model**

We propose a cluster algorithm to cluster the data in two clusters. One of the clusters must be saved in the SM and the other in the SP. We used the cluster as a tool to evaluate the solution.

Kidane, P., & Bonds, A. (2015) explained that clustering aims to partition $n$ data into $x$ clusters in which each observation belongs to the cluster with the nearest means. Each cluster serving as a prototype. The cluster based on distance, presented the nearest cluster by using the smallest Euclidean distance. Those authors explained that, the clustering clustering algorithm aims to partition the given set of $n$ observations $(n_1, n_2, \ldots, n_m)$, into $x$ ($n$) sets $S = \{S_1, S_2, \ldots, S_k\}$ so as to decrease the internal-cluster sum of squares (WCSS) (sum of distance functions of each point in the cluster to the $X$ centre) in order to find the smallest Euclidean distance. In other words, its objective is to find:

$$\arg \min_S \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2$$

Where $i$ is the mean of points in $S_i$.

In addition these algorithm proceeds by alternating between two steps, the first step being assignment of each observation to the cluster. Mathematically, this means partitioning the observations based on distance generated by the means.

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\|^2 \leq \|x_p - m_j^{(t)}\|^2 \ \forall j, 1 \leq j \leq k\},$$

Where each $x_p$ is assigned to exactly one $S^{(t)}$, even if it could be assigned to two or more.

The second step is to calculate the new means to be the centroids (where the mean position of all the data points in all of the coordinate directions) of the observations in the new clusters.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x \in S_i^{(t)}} x_j$$

The algorithm has converged when the assignments no longer change, which we will improve in the next section. Figure 5.21 represents an example of the real-data for electrical
consumption data during one hour. We measure the consumption data for one hour for each device TV, Laptop, Fridge and Lamps.

Data Cluster

We use Weka to cluster the data consumption. "Waikato Environment for Knowledge Analysis (WEKA)” is an open source software written in JAVA, that performs data mining tasks by the collection of machine learning algorithms. It contains tools for data pre-processing, classification, regression, clustering and visualisation. (Hall et al., 2009). Through WEKA, any raw data can be changed into meaningful information.

We used Explorer, which is an environment for exploring data with WEKA. Once Explorer is started, it is required to open a file (data set) which is prior to explore the data. This file should be in ARFF format or CSV format. Thus the data can be clustered through CLUSTER option. The result can be visualising by clicking the VISUALISE option.

First step was to filter the data using unsupervised methods in order to add the cluster number as a nominal attribute to the data processed. This makes the post-processing or analysing of the cluster assignments easier. Then we apply the algorithms to cluster the data to two clusters, which we proposed in figure 5.22.

Figure 5.21: Real Consumption Data for TV, Laptop, Fridge and Lamps.
Figure 5.22: Clustered Data.

The clustered data is shown in Figure 5.23.

The output represents that each cluster comes together, with a "1" meaning everyone in that cluster shares the same value of one, and a "0" meaning everyone in that cluster has a value of zero for that attribute. Numbers are the average value of everyone in the cluster. Each cluster shows a type of behaviour in the consumption’s data, which can help to draw some conclusions:

- **Cluster 0** This group has 20
• Cluster 1 This group has 80

We assume that the user uses 5w each hour when the user is at home. When the user is not at home there is non-consumptions data. To avoid that, we need to re-send the 20% of the data, which was saved on the SP and send it back to SM in the same interval as the user’s normal usage, which is 5W in our case. Figure 5.24, shows how the consumption data (in Volts) during one hour.

![Figure 5.24: One-hour Consumption Data.](image)

Figure 5.25 visualises the data after applying the clustered algorithms.

### 5.3.3 Privacy Evaluation Approach

We utilise WEKA for evaluation as well as to build the privacy boosting protocol. We used the result through applying the cluster algorithms and utilise J48 algorithm to observe the data and create the decision tree. These generated model will be evaluating each cluster to observe the level of the quality for our solution. Figure 5.26 below shows the tree view of the model.

The generated model has been evaluated by using Weka evaluation tools. The results are shown on figure 5.27.

The result shows that the correct classified instances are 99.9 % indicating our model to be a very good model. Moreover, the confusion matrix shows that 356 is true positive, 1437 true negative, 1 false negative and 0 false Positive 0.06 % incorrect classified data. Pares test percent was used to evaluate the clustered data quality, and the result shows that only 0.3 %
incorrect clustered instances, as shown in figure 5.28.

5.3.4 Privacy Evaluation Results

To evaluate the data model an experiment on the generating data model was performed to measure the quality level of the model.

Three different experiments were made to compare the results for each cluster. The results are shown in the figure 5.29.
The results show that the clustered data have very good quality, indicating this model can be used successfully to enhance the user’s privacy.

The consumption’s data has to be clustered into two clusters 20% of the data will be saved
on the SP, and 80% of the data will be saved on the SM. When the user is not at home, we need to return this data to the SM in the same interval.

The SP has to send the consumption data to the SM whenever the user is not at home, which will show that the user is at home the whole day. Figure 5.30 displays the consumption’s data for those devices during one hour before the clustered data.

This consumption’s data is supposed to be sent during the day, at the same interval. A visual for the clustered data during the time is shown in figure 5.31, where the clusters are
distributed during a time that gives a high privacy for the user.


5.3.5 Enhancing protocol for privacy

After evaluating the security protocol, it was enhance in order to improve the communication between SM, SP and UP during the data transformation. The protocol has been boosted to prevent any malicious attacks against data in SM. Moreover, the remote attestation protocol is used for privacy as well, during the communication between SP and supplier. Figure 5.32 describes the changed protocol where clustered 0 data in red and clustered 1 data in blue.

In the previous protocol the UP allowed the access together with a digital signature or key to communicate with SM (log in and read the data). Then, the SP confirms the key acceptance within the time limit. If the client did not answer during this time, then this certification must be discarded and a new certification must be sent to the client on request. The same concept is used for this protocol as well during the communication between the SP and the SM.

Figure 5.32: Privacy Protocol.
As a summary, using clustered algorithms on the consumption data will provide a good model for protecting the user’s data. This indicating this model can be used successfully to enhance the user’s privacy.

The consumption’s data has to be clustered into two clusters 20% of the data will be saved on the SP, and 80% of the data will be saved on the SM. When the user is not at home, we need to return this data to the SM in the same interval. The SP has to send the consumption data to the SM whenever the user is not at home, which will show that the user is at home the whole day.
Chapter 6

Conclusion and Future Work

This research focuses on SM’s from a security and privacy perspective. Security issues in this system are tampering, vulnerabilities and DoS. Privacy issues include information leakage in real-time consumption data recorded by the SM. This research proposed a new concept for solving the above mentioned problems. Using the SP as a third part to prevent any malicious operations on the SM was the proposed solution for the security issue. Clustering the consumption data by using an extra storage device on the SP to hide information locked in the user’s data by lessening home load signatures, after mixing utility energy with energy supplied and SP storage, was the proposed solution for the privacy issue.

For security evaluation, we designed an attack scenario and proposed a protocol which is implemented on ns-2. The results show that connecting SP to SM, provides a higher security for SM and a none malicious operation could be done on the data package during transferring the data between UP, SP and SM.

For privacy evaluation, we employed machine learning techniques focused on the unsupervised techniques, to determine the quality of the clustered data and privacy level of the proposed algorithms. The results show that clustering the consumption’s data gives higher levels of privacy. Any kind of monitoring the SM will give a reading that the user is at home during a period of 24 hours. If the user is using 5 peaks of data hourly, then 3 peaks will be saved in the SM and 2 peaks will be saved on the SP. When the user isn’t at home, the data which has been stored, will be gradually sent to the SM in order to hide the user’s privacy. Table 4 shows a summary of the research questions and answers after evaluating the results for this research.

Based on this thesis result, these are the steps to build the framework:

1. Storage data between the SM and SP.
2. Secure the channel communication between the SM, SP and UP.

3. Run authentication with security protocol collaboratively between communication channels.

The framework needs real time experiment however we simulated and evaluated these steps. Our future work will include implementation of the whole framework.

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Evaluation</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the potential threats to alter the software deployed in the smart meter?</td>
<td>Literature Research</td>
<td>Based on literature review, preliminary research has already indicated several potential threats and vulnerabilities attacks on SMs which effecting the security and privacy in smart meters.</td>
</tr>
<tr>
<td>How can we secure the smart meter's software?</td>
<td>Stimulate the environment using ns-2</td>
<td>Using the SP as a third party provides a higher security for SM and operations could be done with SP along encryption.</td>
</tr>
<tr>
<td>How can we enhance the user’s privacy with architecture-based technologies?</td>
<td>Cluster the real-data consumptions using ML</td>
<td>Clustering the consumption’s data gives a high levels of privacy</td>
</tr>
</tbody>
</table>

Table 6.1: Research Summary.

Further improvement is needed to enhance the cryptography. In Addition, this research is focused on only two attack scenarios and there are many more to test. It will be interesting to evaluate the usability of our approach in other types of scenarios and we consider this as one of our future works for this research.

Furthermore, a system has to be developed to identify and verify if the user phone is the actual phone to send the data and an application to secure the communication between the SP and the SM. In addition, improvement in the communication way between the SP and SM in distance, when the user is outboard, for instance, is needed. Also, a system has to be built to secure the data storage and distributed system between SP and SM, and how the digital
signature authenticates the data between the communication between SM and SP.
6.1 Reference


Knapp, E. D., & Samani, R. (2013). Applied cyber security and the smart grid: implementing security controls into the modern power infrastructure. (pp. 73-74)


In Information Networking (ICOIN), 2011 International Conference on (pp. 114-119). IEEE.


