Master's Thesis

Privacy-preserving Smart Metering Using DC-Nets

Gregor Garten

January 13, 2022

TU Dresden
Faculty of Computer Science
Institute of Systems Architecture
Chair of Privacy and Data Security

Supervisors: Dr. Stefan Köpsell
Dr. Elke Franz
Betreuender Mitarbeiter: Dipl.-Inf. Tim Lackorzynski

Lorem Ipsum

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbstständig erstellt und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Dresden, den **?today?**

Gregor Garten

Abstract

. . . abstract . . .

write abstract

Contents

Todo list

List of Figures

List of Tables

1 Introduction

adopt title page

adopt disclaimer

write introduction

1.1 A Section

Referencing other chapters: 2 3 4 5 6 7

Name	Y	Z
Foo	20,614	23 %
Bar	9,914	11 %
Foo + Bar	30,528	34 %
total	88,215	100 %

Table 1.1: Various very important looking numbers and sums.

More text referencing Table 1.1.

1.2 Another Section

Citing [**bellard2005qfa**] other documents [**bellard2005qfa**; **boileau06**] and Figure 1.1. Something with umlauts and a year/month date: [**becher04:˙feurig˙hacken˙mit˙firew**]. And some online resources: [**green04**], [**patent:4819234**]

1.3 Yet Another Section

add content

1.4 Test commands

DROPS L^{4}LinuxNOVA QEMU `memcpy` A sentence about BASIC. And a correctly formatted one about ECC.

Figure 1.1: A long description of this squirrel figure. Image taken from `http://commons.wikimedia.org/wiki/File:Sciurus-vulgaris_hernandeangelis_stockholm_2008-06-04.jpg`

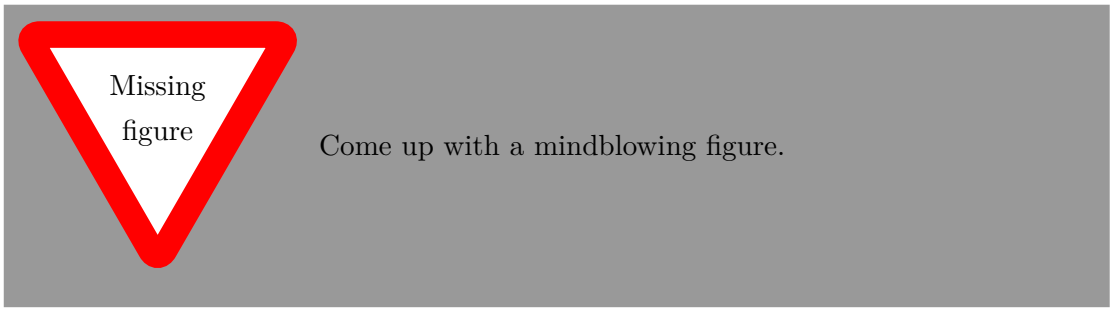

Figure 1.2: A mindblowing figure

2 Background

This section introduces an overview of the basic concepts for this work. Therefore, the key components of the smart grid are explained, what structural changes and what challenges the smart grid will bring. In addition, this chapter discusses the current state of research.

2.1 Smart Grid

The original energy network was mainly considered as a transmission system to send electricity from the generators via a elongated network of cables and transformers to the consumers.Instead of a few electricity producers (e.g. nuclear power plants, coal-fired power plants), which were responsible for a large part of the electricity generation, there are now many smaller producers (e.g. wind turbines). But, renewable power generation is often dependent on external environmental factors. In order for the electricity grid to be stable despite fluctuations in power generation, smart meters have been introduced. Smart Meters enable the electricity provider to receive the electricity consumption of a household every 15 minutes. It offers the possibility to get more easily the current electricity demand from the consumers. Previously, the current electricity demand was simulated from load forecasting models. If the demand should increase spontaneously, peaker plants, mainly consisting of coal-fired power plants, would be turned on to quickly meet this demand. This is costly and environmentally unfriendly. Since then, structural changes have been made to optimize the energy grid and make it more intelligent by exchanging information in near-real-time. This allows the demand to be matched to the available supply. The fundamental component of the smart grid are the smart meters which will be discussed in more detail in the next section.(Quelle:Smart Grid Communications)(Privacy Survey2013)

Smart Meter

Smart meters are the key component in a smart grid. A smart meter is an electricity meter which has an interface to the Internet. The additional functions that a smart meter brings to a regular electricity meter allow a two-way communication between the control center and the smart meter. The interconnection of electricity meters and control center via the Internet is also called Advanced Metering Infrastructure (AMI). The resulting communication between both components improves the quality of the power grid and makes it possible to offer services that would not be feasible without a smart meter. For example it's now practicable to detect power outages for a grid

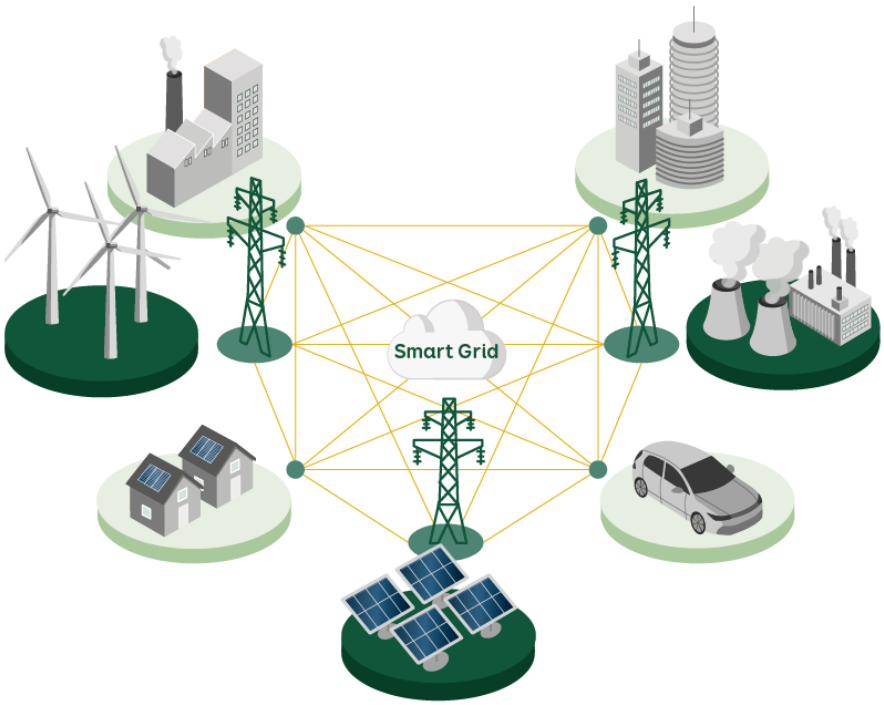

Figure 2.1: An example of a NILM analysis.

operator on its own.Previously, the operator was dependent on customer calls to detect power outages. Another new feature is detailed monitoring of power flows at the smart meter. Beforehand, power flows could only be measured up to substations. Moreover, the advanced functions enables electricity network operators to quickly detect changes in consumption behavior and react to them without having to use peaker plants, which are costly and environmentally unfriendly. Depending on the setting, smart meters can send electricity consumption to the electricity provider at least every 15 minutes. Additionally, in combination with the consumption of all users and the current electricity supply, real-time pricing becomes possible. Not only the customer can be offered a better electricity contract, the smart meters no longer have to be read at home by a technician from the electricity provider. As a result billing becomes easier for customers and electricity providers. Furthermore, customers can also check their current electricity consumption via the interfaces provided by the smart meter in order to analyze their own behavior and to reduce their consumption. (Privacy-Aware Smart Metering)

2.2 Smart Meter Privacy

The main advantage of the smart grid is the advanced communication between the consumers smart meter and the energy suppliers. The 15-minute messages from the electricity meter provide the electricity supplier with a regular update on the status of the electricity grid and there is no longer any need to rely on forecasting models.

However, sending user information in such a short period of time allows for new methods that can be used to create accurate behavioral analyses in one's own home. Sending private electricity consumption data is therefore very sensitive information and has to be protected. This is not an easy task, because on the one hand the electricity consumption must be protected and anonymized, and on the other hand the billing and costs must be clearly assignable to a person. The two problems are referred to as metering for billing and metering for operations. The following paragraph describes how simple behavioral analyses are generated by electricity consumption. Subsequently, solutions to Metering for Billing and Metering for Operations will be presented, which have been discussed in the scientific community so far.(Privacy-Aware Smart Metering)

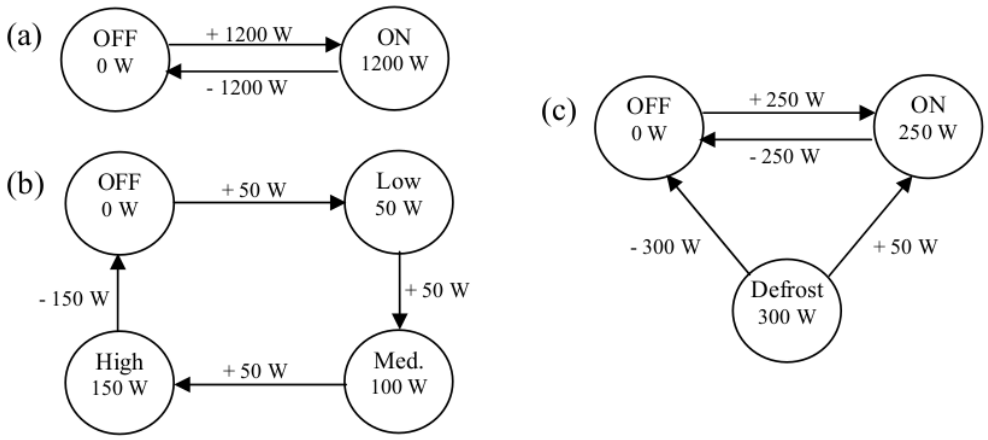

Figure 2.2: An example of a NILM analysis.

2.2.1 Non-intrusive load monitoring

Interpreting power consumption with the intent of identifying devices in the home is called non-intrusive load monitoring (NILM). George Hart and Fred Schweppe were the first to develop non-intrusive load monitors in 1985 and connected them to electricity meters. They were able to record the current power consumption up to every 5 seconds. Then they did the following steps to identify appliances in a household:

1. Edge Detection:
   First, the intercepted electricity consumption was stored. Afterwards, a search for strongly rising or strongly falling edges was performed. These edges indicate that a device may have been switched on or off at that moment.

2. Cluster Analysis:
   The stored events of steeply rising or steeply falling edges are visualized in a graph with the following characteristics. Each event is ordered according to how much power was consumed or how much power was "released" from the device (e.g. when it was switched off). Essentially, a cluster analysis is then applied to the diagram and each cluster represents a household appliance.

3. Appliance Model Construction
   Since different household appliances have been determined by the clusters, appliance models can now be constructed. In this step, different states in which an appliance can be in, are found based on the different power consumption. An example of how the result of a appliance model looks like can be seen in Figure 2.1.

4. Behavior Analysis:
   Once the majority of the household appliances have been identified, the behaviors of people in the household can be analyzed. In real time, it is possible to track the use of devices, since individual signals can be identified as they occur and do not need to be reconstructed anymore. At this point, several approaches can be taken to provide behavioral analysis. A common approach is to track how long a device has been in use and create statistics on how each device has been used. A daily analysis can be viewed in Figure 2.2.

5. Appliance Saving:
   The last approach is to name the household appliances found(washing machine, etc.) and store them in a database. So that in the case of a further household analysis, it is possible to fall back on appliances that have already been found.

The founder of NILM G. W. Hart himself said in 1989: "Specifically, I recommend that legal restrictions be enacted or clarified so that electric power usage is considered as private as any phone conversation."(Residential Energy Monitoring) Through Nilm, simple observations can be made without analyzing the household behavior for a longer time. For example, it can be noticed when no one is at home because no lamps are on. It can also be quickly assumed that the house inhabitants are on vacation, if the power consumption is lower than usual over days. For burglars, this information would be particularly useful, as they would have no problem knowing when is a suitable time to break in.

High Resolution Analysis

Since then, research in the field of intrusive monitoring has continued. It was investigated both how much information can be extracted from the household through electricity consumption when electricity consumption was measured particularly frequently. Furthermore, it was investigated whether assumptions can be made about the behavior in the household even at low resolution. For example, in the paper(Multimedia Content), the movie being watched could be determined by the power consumption of an LCD television. TV power consumption is strongly influenced by blacklighting activities and each movie has a unique brightness signature. This unique signature was exploited to make a statement about the film being viewed. In the paper, the power consumption of the TV was measured every 2 seconds and after 5 minutes analysis of the consumption, the content of the program could be determined with high probability. For this purpose, a Power Consumption Prediction Function was trained to estimate the power consumption of the TV based on the brightness of movie sequences. The input of the prediction function is 5 minute sequences of a movie and the output is the possible power consumption of the TV. In order to recognize a movie, the result of the prediction

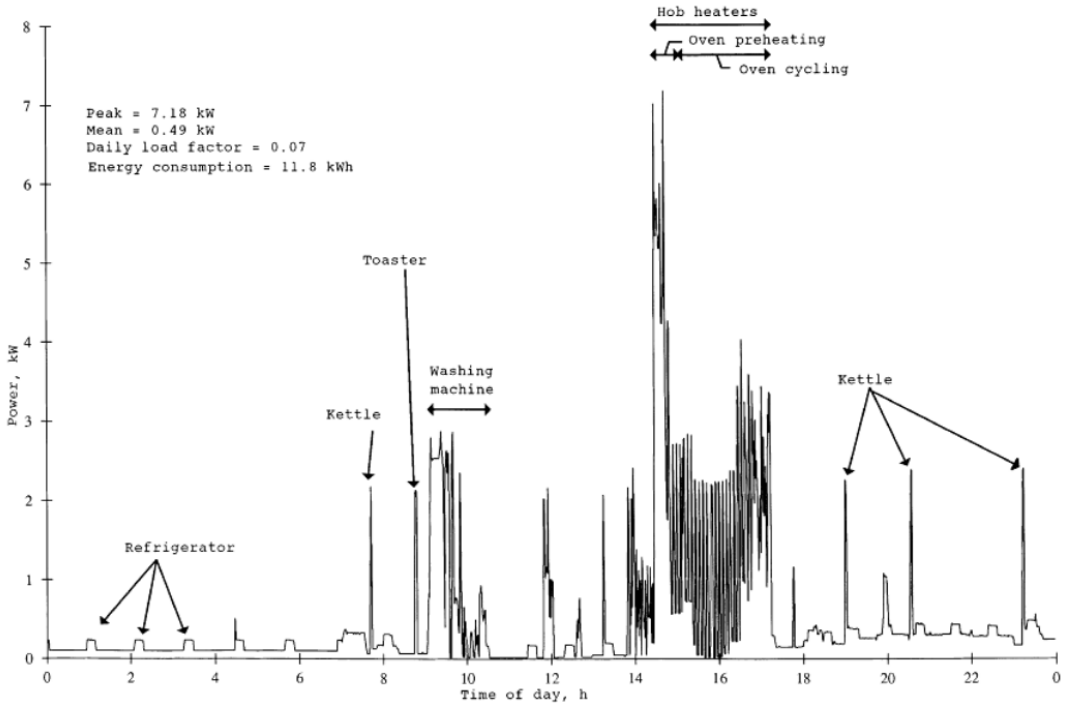

Figure 2.3: An example day of a NILM analysis.

function is stored over the 5 minute movie sequences. If the same 5 minute sequences are running on a TV, the power consumption can be compared with the prediction by a correlation coefficient. For multiple matches with a correlation coefficient higher than 0.85, an additional optimization algorithm is applied to estimate the movie. In the figure the result of the estimation of the prediction function is shown in green and in red you can see the actual power consumption of a LCD TV for the same movie sequence.

Low Resolution Analysis

Another approach is to work backwards from large data sets to obtain detailed information from aggregated electricity consumption. The Nilm method attempted to have household appliances detected from individual households and then behavioral analyses could be generated from the detected household appliances. A common approach with low-resolution electricity data is to identify the numerous factors that influence total electricity consumption and to filter them out. In (A Neuron Nets Based) an artificial neural network was trained to identify household appliances from electricity consumption. The power consumption was measured only every 15 minutes.
In another work(Energy Disaggregation), the power consumption of appliances was measured only every hour. The experiment tried to disaggregate the aggregated power consumption with Discriminative Sparse Coding. The experiment was conducted as follows. First, the disaggregation algorithm was trained with electricity consumption

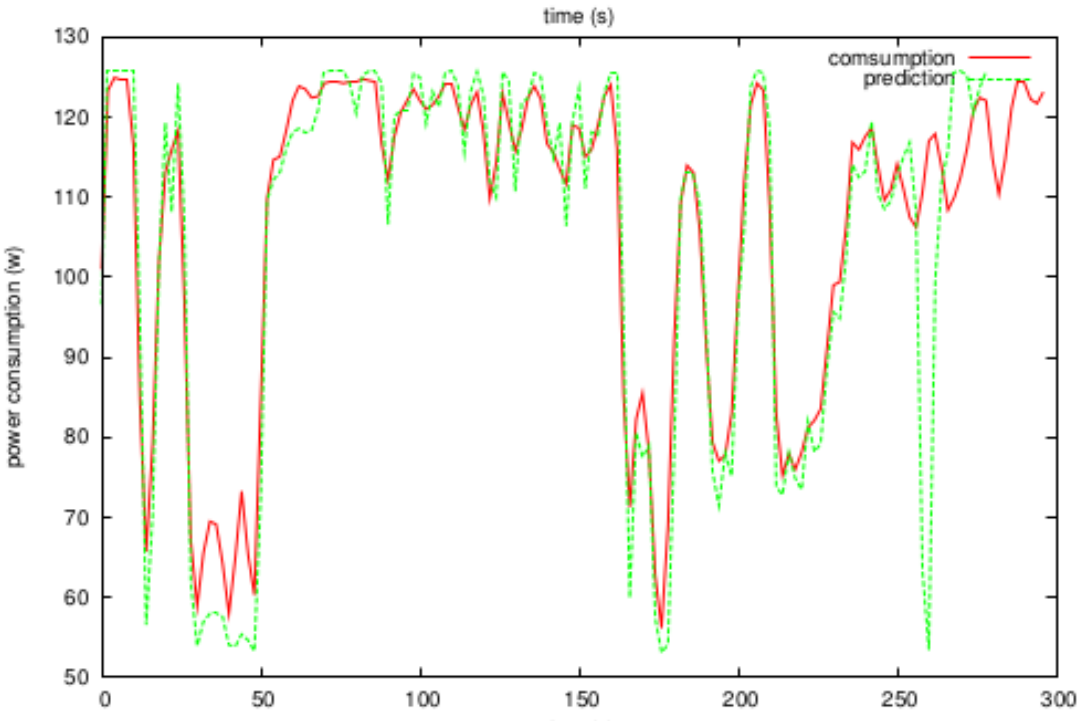

Figure 2.4: An example day of a NILM analysis.

of houses from a data set. Then, the trained algorithm was applied to the aggregated electricity consumption of 2 houses. The algorithm was expected to determine up to 52 individual household appliances from the electricity consumption. The results in the paper showed that the disaggregation algorithm with Discriminative Sparse Coding was up to 55.05 percent correct in its decisions.

2.3 Related Work

2.3.1 Metering for Operations

The paragraph deals with solutions for Metering for Operations, which has been previously discussed in other scientific works. At Metering for Operations, there is currently no established consensus on a solution. Various technical proposals have already been presented in scientific papers, but there is a lack of uniform criteria and often different conditions are set for the power grid. One reason for this could be that smart grids are realized differently in other countries. In the following, the different approaches are divided into categories and presented conceptually.

Anonymization or Pseudonymization Without Aggregation

This approach describes the removal of smart information that allows identification. Identifiable information can also be replaced by pseudonyms. Solutions with

trusted third parties are often used in this case. A trusted third party usually acts as an intermediary between the customer and the power grid provider. The trusted third party must be acknowledged by all participants and take a neutral position. In practice, however, this is difficult to achieve because the trusted third party is often hired as a service provider by the electricity supplier and is therefore also paid by the supplier.

In the paper "A privacy-preserving Concept for Smart Grids" by Petrlic[quelle], a trusted third party is used as an intermediary. In the proposal, a smart meter communicates with a trusted third party. Certificates formed with a public key infrastructure can be used to verify and validate information flows from smart meters at the TTP. As soon as the Trusted Third Party has checked the correctness of the smart meter information, it can pseudonomize/anonymize all the necessary information. Only then is the further processed anonymized information forwarded to the electricity provider by the TTP in encrypted form. This means that the electricity provider cannot assign individual electricity consumption to its customers. With this procedure, smart meters can be anonymized. However, if it is possible for an attacker to record the data traffic between the smart meter and the TTP, then the attacker could forward the time stamps and smart meter identification to the electricity provider. Using these two pieces of information, the electricity provider could at least gain some insight, since it would be possible to match when information is sent to the TTP and when it is received by electricity provider.(Privacy-Aware Smart Metering)

Aggregation with Trusted Third Parties

In the attack just described, the electricity provider tries to link two events. One is the arrival of the message at the TTP and the other is the arrival of the message at the provider itself. One way to prevent this attack is aggregation. In this case, the smart meter sends its electricity consumption to the TTP. Certificates are also sent from the smart meter so that the TTP can check the information for correctness and authenticity. Instead of forwarding the information to the electricity provider, the TTP waits until all smart meters have sent their data for which the TTP is responsible. This data is all added up and a message is sent from the TTP to the electricity provider with the total electricity consumption of all smart meters. From the aggregated value, it is not possible to extract an individual smart meter's electricity consumption, which is why the electricity provider cannot filter out information about individual customers.(A Privacy Model for Smart Metering)
Homomorphic encryption approaches also fall into this category. Homomorphic encryption algorithms allow simple operations such as addition and multiplication to be performed on the encrypted messages. In some homomorphic encryption schemes, only addition OR multiplication are supported. These are then called partial homomorphic encryption. There are also bihomorphic encryption approaches. Here not only the operations on the ciphertexts are homomorphic, but also the operations on the keys. This means that if a plaintext a is encrypted with the key x and a plaintext b is encrypted with the key z, that one can decrypt the ciphertexts enc(a+b) with the keys x+z. A bihomomorphic encryption approach with TTP has been proposed by Vetter et

al.[quelle] In this case, the TTP acts as the key authority. This means that it creates all cryptographic keys and forwards them to the smart meter, which is then used for further communication with a central store. The smart meter encrypts its data and sends it to the central storage. The central storage also stores the incoming data in encrypted form, so that no unencrypted data can be found on the storage. In addition, the central storage has no access to the keys and thus has no way to decrypt the information or access meaningful data. Therefore, the central repository has to be trusted only in terms of functionality. If an electricity provider wants to know the electricity consumption of its customers, it makes a request to the central repository, which sends the aggregated encrypted data to the electricity provider. In order for the electricity provider to decrypt the data, the key authority has to release the correct keys. It is impossible for the electricity provider to query the value of just one smart meter. This is because the key authority can only issue keys that can decrypt aggregated totals. It is guaranteed by the homomorphic encryption method which is used. The advantage of using this approach is that the different functionalities, namely storage of data and key acquisition for confidentiality and authenticity is realised from different participants.(Privacy-Aware Smart Metering)

Aggregation Without a Trusted Third Party

The solution proposed in this thesis is also one of the methods that aggregate without a trusted third party. The advantage of this approach is that no one has to trust a trusted third party. In general, one has to ask the questions who aggregates the data and who generates/uses the keys. In addition, a common problem to consider is how the procedure deals with a few participants/customers.
In the solution of Mármol et al. again a bihomomorphic encryption method is proposed. The approach of Mármol has already been discussed and implemented in a master thesis at this chair.[quelle biselli] As a reminder, bihomorphic encryption algorithms can perform simple operations such as addition on both the ciphertext and the keys. This property is exploited in the presented method of Mármol. Since it aggregates the keys and not the power consumptions as before. Furthermore, it does not matter which bihomomorphic encryption method is used, as long as all smart meters agree on one method. A key is generated from every smart meter in the power grid. Afterwards, the key is used to encrypt the power consumption. The key is then used to encrypt the electricity consumption and the encrypted data is sent to the network operator. The transmission channel to the network operator is chosen in such a way that the identity of the smart meter remains secret. This prevents the smart meter from exposing itself during communication with the operator. Groups are formed among smart meters and a smater meter aggregator is selected in each group. The aggregator is selected randomly and all smart meters send their keys to this aggregator. Subsequently, the keys are summed up at the aggregator and sent to the network operator. The network operator receives a single key and with this key it can only decrypt the messages from one smart meter group. additionally, the operator has to add up all the messages and only then it will be possible to decrypt the messages. There is a possibility that aggregator cooperates with the network operator. The aggregator would then be able to send individual keys

Figure 2.5: The power consumption of a household in a day with battery.

from smart meters to the operator. While the operator would not be able to match the key to any message, by brute force it could decrypt all messages with that key and see which decrypted message has meaningful content. To prevent this attack, an additional measure is taken. All smart meters in a group organize themselves topologically in a ring structure. In this ring structure, all smart meters cooperate with each other and change their keys every round in such a way that the individual key of a smart meter changes, but not the summed key of all smart meters. Even if the aggregator forwards the keys to the network operator, they would no longer be valid in the next round.(nochmal nachlesen, warum nicht eine individuelle Nachricht entschlüsselt werden kann) A disadvantage of this procedure is that if a smart meter leaves the group, then a new aggregated key must be formed.(Privacy-enhanced architecture for smart metering)

Battery Solutions

The battery approach describes a household with a connected battery that is charged, e.g. by grid purchase or photovoltaic panels. The goal of the approach is that the battery feeds energy into the household in such a way that the grid operator can no longer detect private information based on the electricity consumption.

The figure shows the electricity consumption of a private household with a connected battery that is charged via solar panels. It can be seen 3 lines. The red line shows the

electricity consumption of the household. The green line shows when the battery is discharged(when the battery is feeding power to the household). The blue line is the power consumption that the grid operator can see.In the figure you can see that when the battery brings electricity to the household, then grid operator sees that a household does not consume electricity. In other cases, the grid operator sees that electricity is being consumed, but it is much less than the house actually consumes because the battery offsets some of the electricity consumption. In other words, if a household is connected to a battery, the grid operator cannot make correct statements about the behavior of the people in the household.
An algorithm for batteries was proposed in(Protecting consumer privacy from electric load monitoring). This method uses an algorithm that can control the battery to produce a constant characteristic curve in consumption. The algorithm targets a static and fixed current consumption. The target consumption is calculated differently by the algorithm depending on the house consumption and battery capacity. If the power consumption is below the consumption set by the algorithm, then the battery is charged with the difference from the target consumption. If the power consumption is above the target consumption set by the algorithm, the battery is discharged with the difference from the target consumption. If the consumption is significantly higher, so that the battery can no longer absorb the additional consumption, then it is switched to recovery mode. In recovery mode, the target consumption is temporarily increased, so that the battery can charge on the side, even though the house is currently consuming a lot of power. If the recovery mode can be switched off, then a new target consumption is calculated based on the new data. It is important to remember that this method does not anonymize power consumption, so it is even more important to measure how much information can still be extracted from power consumption. There are the following metrics to calculate how much privacy is gained by the algorithm.

1. Relative Entropy:
   Relative entropy is used to compare two sources of information. In this case it would be the power consumption with the algorithm and the power consumption without the algorithm. These two loads form a stochastic process and can then be compared with the relative entropy. (Affordable privacy for home smart meters)

2. Cluster Classification:
   The cluster classification has already been explained for the NILM method and described in this paper at [ref]. The cluster classification has already been explained for the NILM method and described in this paper at [ref]. It can also be used as a metric to evaluate privacy. Here one would perform a cluster analysis with the battery method and once without. Then one looks at the number of clusters in both measurements and if fewer clusters are found with the battery method, then this is considered a privacy gain.

3. Regression Analysis:
   In the regression analysis, first a cross-correlation and afterwards a simple linear regression is performed. More precisely, both power consumptions are ”superimposed” at the point of their maximum cross-correlation. Subsequently, a linear

regression is performed and the privacy is evaluated on the basis of the quality of the predictor.

2.3.2 Metering for Billing

In order to fully protect the privacy of a household, metering for billing procedures must also be applied. Otherwise, conclusions about electricity consumption can be drawn from the billing. A simple solution would be to increase the frequency of the billing period. But at the same time it is also in the interest of the customer to buy electricity as cheaply as possible. The customer can be offered better electricity contracts if the billing period is shorter. In addition, it cannot be guaranteed that the customer's privacy is not violated in more complex electricity contracts by other features, even if a higher billing period is used. This master thesis focuses mainly on the metering for operations problem. By implementing Trusted Platform Modules (TPM) in German smart meters, the problem is considered to be solved. For completeness, frequently proposed solutions in the scientific community are presented.

Billing with a Trusted Third Party

The advantages and disadvantages of a TTP have already been explained in the upper section. The principle is similar to metering for operations. The smart meter sends its measurements to the TTP and the TTP calculates the bill over the time period specified in the electricity contract. The billing is then sent to the electricity provider. On paper, this approach is simple, but important practical questions often remain unanswered. For example, who pays the trusted third party? In this case, the trusted third party provides a service to the electricity provider. However, if the electricity provider pays for the service, then the trusted third party is no longer independent.

Billing with a Trusted Platform Module

Billing can also be implemented on the smart meter with a Trusted Platform Module. A TPM is a chip that is installed within the smart meter and thus additional security features can be used on the smart meter. The TPM contains a cryptographic processor that can generate random numbers, generate RSA keys, generate SHA-1 hashes and it has an encryption-decryption-signature engine. In addition, the TPM can be used to prove that nothing has been tampered with the smart meter afterwards. A secured smart meter can therefore perform correct billings at the customers side and the electricity provider can trust the smart meter. However, the TPM is installed by the electricity provider and it only guarantees the validity of the billing. If the electricity provider decides to send additional sensitive information within the calculations in the TPM, the TPM has no advantage for the end user.(Privacy-Aware Smart Metering)

Billing Secured via Advanced Cryptography

Lastly, there is the cryptographic commitment method. With this approach, no

other participant needs to be trusted. A smart meter can use a cryptographic commitment to prove that each bill was calculated correctly. So with cryptographic commitments, billing can be done on the customer side.
Lastly, there is the cryptographic commitment method. With this approach, no other participant needs to be trusted. A smart meter can use a cryptographic commitment to prove that each bill was calculated correctly. So with cryptographic commitments, billing can be done on the customer side. A commitment is a cryptographic application and works as follows. Both sides agree on the same commitment procedure. Then it is possible that one side can generate a c=(x,r) as an obligation. Here x would be the calculation and r would be a random number. If one wants to check the commitment for correctness, then there is an Open(c,x,r) function that returns True(if correct) or False(if incorrect). Cryptographic commitments are mathematically constructed so that it is easy to compute a c=(x,r), but hard to find an x /= x' with an r' such that open(c,x',r') returns True. In this use case, the following procedure is often used. [pedersen]
Commit(x, r) · Commit(y, s) = Commit(x+y, r+s)
$Commit(x, r)k$ = Commit(x · k, r · k)
The special feature of the method of [pedersen] is that at the same time non-homomorphic properties are satisfied. Without going into exact technical details, cryptographic commitments work as follows in a smart grid. The smart meter generates a cryptographic commitment for each measurement. Via a public key infrastructure, the smart meter and the electricity provider receive cryptographic keys. With these keys, the smart meter can sign its commitments and then send them to the electricity provider. The electricity provider checks the commitments for correctness and if all data is correct, the electricity provider sends back a list with electricity prices and the corresponding time stamps. The meter now knows at each point in time how much electricity was consumed and what the price was. By exploiting the homomorphic properties of the procedure, the smart meter can now calculate the electricity prices. The electricity price in this case would be the variable k.So the smart meter creates new cryptographic obligations with the electricity price calculated on the consumption and sends these new obligations to the electricity provider. The electricity provider can verify the correctness by performing the same calculations as the smart meter. If the results with the Open(c,x,r) method return true, then the calculations were performed correctly by the smart meter. This method is suitable for simple electricity tariffs when only a factor on the electricity consumption needs to be calculated. If an electricity contract is more complex with different conditions such as a higher electricity price if a certain electricity consumption is exceeded, then this approach can no longer be implemented.

2.4 Technical guideline TR-03109

This paragraph will discuss the technical guideline published by the BSI (Federal Office for Information Security). The BSI is the entity of the German federal government that deals with digital security issues and issues recommendations as well as mandatory security guidelines for critical infrastructures. Among other things, technical guidelines are published in which security standards are defined for different IT systems. The

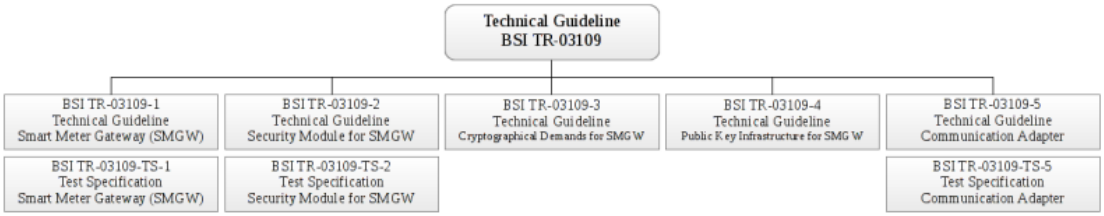

Figure 2.6: An example of a NILM analysis.

technical guideline BSI-TR-03109 defines minimum requirements for the functionality, security and interoperability of smart meters in Germany. The technical guideline BSI-TR-03109 defines minimum requirements for functionality, security and interoperability that individual components of smart meters in Germany must fulfill. The guideline as a whole consists of 6 different documents, which are shown in Figure 3.1. Based on the guidelines, it is possible to have devices certified by test centers. Unless otherwise described, all information are derived from the technical guideline.

Stakeholder in the Smart Grid

Consumer:
The consumer is the person who uses electrical energy, gas, water or heat. In addition, the consumer is the owner of the measurements processed and stored in the SMGW. In order to interact with the SMGW, the consumer uses a communication device. All necessary data can be retrieved and displayed through it.

SMGW administrator:
A Smart Meter Gateway Administrator (GWA) a trusted entity and each SMGW is assigned a GWA. The GWA handles the configuration, monitoring and control of SMGWs and it is even possible to perform updates of SMGWs via the GWA.

Authorized external entities:
External market participants (EMT) are all other authorized participants in the energy network that can establish a communication connection with the SMGW.These include power grid providers and electricity suppliers. The SMGW ignores all other communication requests that do not come from the GWA or EMTs in order to prevent attacks.

There are several other actors such as Controllable Local Systems, service technicians and meters. However, these actors do not play a major role in the protocol that is proposed here.

Stakeholder Motives

It has already been explained in this chapter which participants in the power grid interact

with each other. Now in particular it will be discussed which motives the different participants have and which malicious motives can be pursued by the participants or by attackers.

The smart meter attempts to achieve the 3 security objectives of confidentiality, integrity and availability. The 3 security goals are often summarized as CIA. Another important security goal for this work is anonymity. The 4 definitions are essential for the understanding of this work. Therefore, the terms are explained below.

1. Confidentiality:
   It is not possible for an unauthorized party to gain information about the content of the data sent.

2. Integrity:
   It is not possible for an unauthorized party to modify the content of data without data without this being noticed.

3. Availability:
   It is not possible for an unauthorized party to interfere with the functionality of a service.

4. Anonymity:
   Verbergen der Identität vor dem Kommunikationspartner??? Vllt wer hat die Daten gesendet, wer hat den Stromverbrauch gesendet?

All available attacks will aim at bypassing these security targets to get information about the customer.

Customer's Motives

For the costumer, all the security goals defined above are important. But by far the most important is the security goal of anonymity through the smart meter. Possible attacks on the electricity consumption penetrate deeply into the private sphere of each customer. Therefore, no conclusions may be drawn from the electricity consumption of a customer.
On the other hand, unethical customers may try to steal electricity to save energy costs. The smart meter is located in or on the customer's house. An unethical customer could attempt to tamper directly with the smart meter's hardware or software. The attempts could look like this, a Costumer could try to reduce the recorded electricity consumption at the Smart meter or the Smart meter could be manipulated to measure less electricity when electricity prices are high and more electricity when electricity prices are low.

Electricity Provider's Motives

For the electricity provider, the authenticity of the billing is the most important security objective because From the electricity provider's point of view, the customer is not trustworthy in the calculation of the bill. In addition, the customer has access

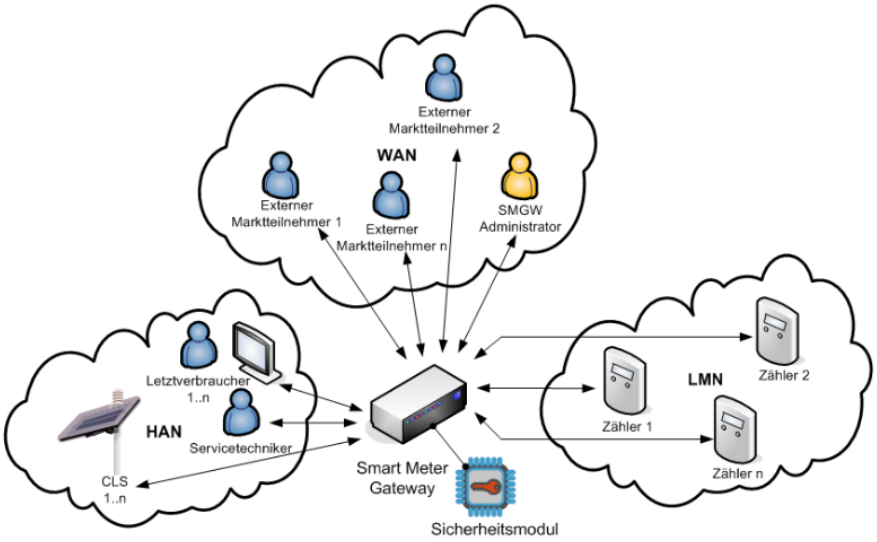

Figure 2.7: An example of a NILM analysis.

to the smart meter at almost any time in an environment trusted by the customer. Unlike analog meters, smart meters cannot be mechanically attacked. But if a customer manages to change the software of the smart meter, the billing can be manipulated at the same time. On the other hand, the electricity provider can also be an overly intrusive electricity provider. In the paragraph [Ref NILM] it was explained how a behavioral analysis can be created from the electricity consumption. This sensitive information could be used to gain an additional source of income. In [SSRN] it was listed which questions could be answered by a Nilm analysis. Quote: "On what days and during what times do you watch TV? How much home time do you spend in front of your computer?" or "Are any of the appliances in your household failing or operating below optimal efficiency? Do you own (and so presumably like) lots of gadgets?". Advertising companies would certainly pay money for this kind of information in order to be able to advertise more accurately. For this reason, it is presumed that the electricity provider is considered an honest-but-curious adversary.

2.4.1 Interfaces and functions of the Smart Meter Gateway

A smart meter or as described in the technical guideline a smart meter gateway(SMGW) must provide 3 different physical interfaces.

1. Local Metrological Network (LMN):
   The LMN is the communication interface in which communication takes place with the connected meters for energy and material quantities (electricity, gas). An SMGW can communicate with one meter from one end user or with several meters from different end users. In practice, however, one SMGW is often responsible for one meter. The measured values are sent from the meters via the LMN to the SMGW and stored there.

2. Wide Area Network (WAN):
   The WAN is the only communication interface with which the SMGW can communicate with EMTs or GWAs over the Internet. If a request is made to the SMGW that was not sent by these authorized participants, then the request is discarded and ignored.

3. Home Area Network (HAN):
   In HAN, an SMGW interacts with Controllable Local Systems (e.g., photovoltaic systems). In addition, users and service technicians can use the HAN interface to display information about power consumption through functions offered by the SMGW.

Functionality of the smart meter gateway

First, the task of SMGW is to store the measurements sent by meters from the LMN. Then, the readings are processed in the SMGW and sent to the authorized EMTs in the WAN after processing. An SMGW must also perform the tasks of a firewall and separate the 3 interfaces. It is therefore impossible for an EMT or GWA to make requests to devices located in the HAN or LMN, even if it is allowed to interact with the SMGW over the WAN. Since the WAN interface is the most important interface for this work, it will be discussed in more detail.

Functions of the SMGW in the WAN

The tasks performed by the WAN have already been explained in the paragraph above. Now the functions and security mechanisms offered by the SMGW to guarantee secure interaction on the WAN will be described.

1. Transmission of measured values based on evaluation and WAN communication profiles:
   Communication profiles of GWAs are stored in SMGW. The communication profiles determine how the data is processed in the SMGW and forwarded to EMTs.

2. Pseudonymization:
   Data that is not relevant for billing must be pseudonomized for data protection reasons. For this purpose, the unique identification number that each SMGW has is replaced by a pseudonym. Subsequently, the information is not sent directly to an EMT, but is forwarded to the EMT via the GWA. This additionally protects the identity of the sending SMGW.Even if peusodonymization does not allow an SMGW to be directly assigned, the described attack in [ref] and the resulting behavioral analysis is still possible. Since no other security mechanisms are available from the SMGW, the question must be asked whether pseudonymization as proposed in the technical guideline is sufficient.

3. Time synchronization:
   In order for the cost electricity consumption to be calculated correctly, it is essential

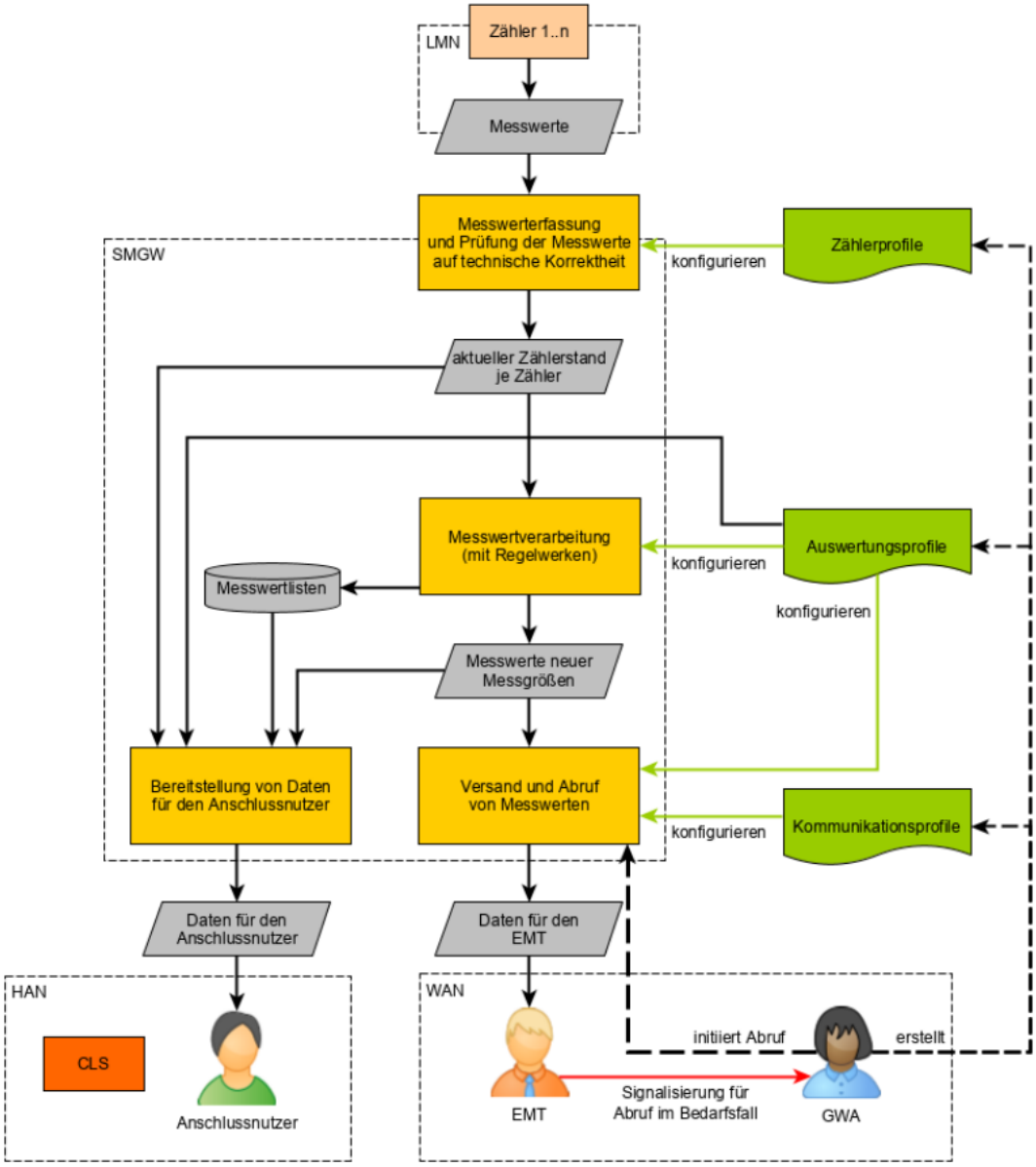

Figure 2.8: An example of a NILM analysis.

that the SMGW have an accurate time. For this purpose, the system time of the SMGW is synchronized with the time server of the GWA at regular intervals.

4. Wake-Up Service:
   A GWA is able to force a communication link with the SMGW. This is done via a data packet signed by the GWA. The SMGW then establishes a fixed preconfigured communication connection to the GWA. This enables the GWA to execute administration commands on the SGMW.

2.4.2 Eavesdropping

Eavesdropping may be the weakest type of attack, but successful eavesdropping on the communications of the smart meter could be useful to e.g. intruders. However, curious neighbors might also have an interest in the behavior inside the house. Turning on/off lights implies that someone is at home or leaving the house. Therefore, eavesdropping on electricity consumption could provide information about when is a suitable time to break in. To prevent eavesdropping, smart meter communication is encrypted to maintain confidentiality. Cryptographic algorithms such as AES are widely used today and have been analyzed for weaknesses over the years by a number of researchers. Hence, a successful attack on encrypted data to extract information is extremely unlikely.

2.4.3 Active Attackers

In Germany, the smart grid is one of the critical infrastructures. This means that the failure of the smart grid could lead to a significant compromise of public safety or other serious consequences. Such systems are threatened by active attackers. The objective of active attackers may not necessarily be to analyze a user's electricity consumption. They may want to disable availability through e.g. denial of service attacks. These attacks could leave major damage to the power grid and are definitely a realistic threat [quelle]. But this thesis focuses on smart meters and the anonymization of electricity consumption. That's why it is assumed that the active attacker does not carry out system-wide attacks on the power grid. Rather, it is assumed that the attacker attempts to take control in the proposed DC network. Among other things, it is assumed that the attacker has the theoretical ability to take over one or more SMGW and send messages through the SMGW. In addition, if the attacker has taken over an SMGW, it can perform all operations that are possible through the proposed DC network. In the next section, the conceptual solution of the DC-Net is proposed and how the DC-Net could be implemented in the technical guideline of the BSI. It also describes which attacks on the DC-Net are possible with the defined attacker model.

3 Design

This chapter outlines the conceptual solution of this thesis to achieve privacy-preserving smart meters. The proposed protocol can be categorized as aggregation without a trusted third party. Before discussing the conceptual solution, the technical guideline from the BSI will be explained. The BSI is the cyber-security authority of the German government and is responsible for critical infrastructures such as smart grids in Germany. The technical guideline TR-03109 resolves all security standards and security concepts that must be met by all power grid providers in Germany. Therefore, the technical guideline gives a good overview of the actual structure of the German power grid. After getting an overview of the power grid and its participants, an attacker model will be designed. The attacker model will introduce all necessary participants, what their motives are and what malicious motives they might pursue. Finally, the security protocol will be presented. It will be shown how the protocol can be integrated into the technical policy and how different potentially malicious participants are handled.

3.1 A Privacy-Preserving Aggregation Scheme Using DC-Nets

In [Cha3-85, Cha8-85, Chau-88], David Chaum proposes a protocol which he calls DC network. The DC network offers the possibility to achieve both sender anonymity and receiver anonymity in communication networks. The operation of the DC network is explained in the following.

3.1.1 DC Networks

The DC network uses the property that any finite alphabet can be numerated(e.g a=0, b=1 etc). If an numerated alphabet from 0 is given, then this alphabet forms an abelian group (modulo alphabet size). Because of the abelian group, simple mathematical operations like addition can be performed on the numerated letters in the alphabet. In addition, a DC network assumes that messages are always sent that are of equal length. A participant in a DC network uses one or more keys with which it superposes the messages and one generated key is then communicated to exactly one participant. More precisely, each participant adds locally all key characters it generates. Then, the received keys from other participants are locally subtracted and finally, all meaningful characters (the message) that should be sent are added (modulo alphabet size). The result of the operation is distributed in the communication network and is called local superposition. The distributed superpositions are added together globally and the result is transmitted back to all participants. Thereby only the meaningful messages remain. If a participant does not want to send a meaningful message, the participant sends an empty message. The message consists only of zeros and is superposed with the key. The

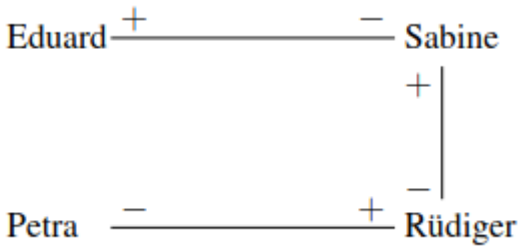

Figure 3.1: An example of a NILM analysis.

empty message reflects the neutral element in this structure. If all participants have sent only empty messages, the global result is a message only containing 0. If one of all participants have sent a meaningful message, the global superposition is the message. If more than one participant sent a meaningful message, then the result is the overlay of all sent messages and a single message from the overlay cannot be recovered. In the last case one speaks also of a collision. In order to solve this problem, the collision resolution algorithm with averaging can be used in a DC network. Exchanging keys to calculate the local sum can be very tedious. In addition, a different key must be exchanged for each message round. Otherwise it would be very easy to calculate the key from previously sent empty messages. Therefore, so-called pseudo-random number generators are commonly used. The participants share the initial values of the pseudo-random number generators with each other when they join the DC network. This can be done in the same way as the exchange of keys (e.g. a cryptographic key exchange procedure). Due to the deterministic property of PRNGs, the same sequence of numbers is always generated from an initial value. This in turn means that the initial value must remain secret and must not be revealed to any other participant, since otherwise the secret keys can be found out from the initial value. The consequence would be the loss of anonymity. The security of the DC network depends largely on how secure the PRNGs are. Therefore, the PRNGs that are used must be cryptographically secure.

The principle of the DC network is illustrated graphically in figure 3.3 using a simple example. Figure 3.3 is also called the key graph of a DC network and shows 4 participants in a DC network that are connected to each other along a communication link. The outer participants have only one partner, the inner participants are connected to 2 partners. Each participant now exchanges keys with its direct partners. The outer partners need to exchange only one key and the inner ones exchange keys with two direct partners. The mathematical operation indicates whether the participant adds or subtracts the exchanged key with the partner. In the example given in the figure, the user Petra would subtract the exchanged key with Rüdiger from the message and would have formed her local superposition. Rudiger would have to add the exchanged key with Petra to his message. He would also have to subtract the key he exchanged with Sabine from the result to calculate his local superposition.

Figure 3.2: An example of a NILM analysis.

3.1.2 DC Network Protocol in a German Smart Grid

The DC network is a scheme that can be used to achieve sender anonymity and receiver anonymity. Considering the use case of the thesis, the receiver anonymity does not have to be implemented. The aim is to anonymize the electricity consumption and send it to the electricity provider. In this case, the electricity provider is a public recipient and known to all participants. Therefore, the identity of the electricity provider does not need to be protected. Unlike in a normal DC network, in the proposed solution the participants do not want to communicate with each other, they only want to send their electricity consumption to the electricity provider. Therefore, the global superposition do not have to be distributed in the network but only calculated at the electricity provider. The only exception is when joining the DC network, there the customers have to perform a key exchange once to configure the initial value for the PRNG as explained in 3.3.2. The electricity supplier is generally regarded as an honest-but-curious adversary. However, in the proposed protocol, the electricity provider partly performs administrative tasks for the dc network. For this reason, an additional aim is to ensure that even stronger attacks can be prevented if the electricity provider takes on the role of the much more dangerous malicious adversary.

Protocol Header

In the network protocol messages are sent which are structured in frames. The structure is shown in the figure 3.4. Each frame consists of a small header and the data part in which usually the local superposition is transfered. The purpose of each field is described further below.
As first field there is a Protocol Identifier field. The purpose of the Protocol Identifier field is to ensure that the application of the DC net protocol is recognized by all participants and that the SMGWs as well as the electricity provider can react correctly to the frames. This is followed by the DC Network Identifier field.The DC network identifier field offers the electricity provider the possibility to operate several DC networks in different regions and to distinguish DC networks. In addition, each SMGW is given a unique identifier so that the SMGWs can be distinguished. If the network needs to perform error correction, SMGWs can be notified by the identification number from the power supplier. Although the SMGWs can be identified by the field, the electricity provider still cannot draw any conclusions about electricity consumption from the local superposition. The transmission bit indicates that an SMGW has sent a message and it is used for error correction procedures. The timestamp field indicates when a message was sent. This allows the electricity provider to classify messages by round and not charge for messages from different rounds. The second to last field is the notification

field. The field is used for correction procedures or notifications for certain operations. An overview of all notification codes is presented in table 3.1. The electricity provider can thus send notifications to the SMGWs to start error correction procedures. In the last field, the Data field, only the local superpositions are sent. The SMGWs do not transmit any information other than the power consumption in the data field. Therefore, it can be assumed that rather small messages are sent with the proposed protocol.

Protocol Initialization

For the Protocol Initialization it is assumed that the electricity provider wants to create a completely new DC network. First, a unique and unchangeable DC Net Identifier is assigned from the electricity provider to the empty DC Net. At least 2 SMGWs have to enter the DC Net. A DC Net with only one participant is not operational and cannot offer anonymity. The SMGWs that enter the network are assigned a subscriber ID by the electricity provider.
To ensure a minimum level of protection for participants in the DC network, the DC network must have a minimum number of user. Even if the electricity provider receives aggregated electricity consumption, individual households may be more noticeable. Different house sizes and number of people in a household leads to a significantly higher electricity consumption, which is visible in the aggregated result for small DC networks. In this master thesis, a stochastic analysis is performed in the experiments chapter to determine a minimum number of participants. Furthermore, it is assumed in the DC network that each participant has at least 3 connections to partners. This reduces the risk of individual SMGWs being disconnected from the DC network or malicious neighbors being able to reconstruct the power consumption from the local total.
According to the Technical guideline TR-03109 from BSI, SMGWs are only allowed to communicate with authorized participants in the smart grid and all foreign requests are ignored. These are EMTs, GWAs and the electricity provider. In order for the DC grid to become operational, two SMGW must exchange an initial value to configure the PRNGs. A start value is exchanged once with which both PRNGs of the clients are initialized. As a result, the same random number sequences are generated independently of each other by the PRNG on both clients. But there is a communication barrier that does not allow SMGWs to communicate with other SMGWs. With the limited communication capabilities, the SMGWs rely on the electricity provider. The SMGWs can use a key exchange protocol like Diffie-Helman to transmit the initial value. Diffie-Hellman is a known key exchange protocol, where 2 users can publicly exchange a secret without a third person being able to figure out the secret.

SMGWs can generate cryptographically secure keys because they have a hardware security module built in. Therefore Key exchange procedures such as Diffie-Hellman can be implemented for the SMGW without any problems. Diffie-Hellman was also only mentioned as an example. There are various attacks on the textbook Diffie-Hellman variant presented.The forwarding of SMGW messages by the electricity provider enables the implementation of other substantially secure key exchange procedures. The advantage of this approach is that SMGWs are anonymous to other SMGWs. When the keys are exchanged, only the partners with whom the key is currently exchanged are aware of it.

Notification Description	Functionality
Notification 1	A SMGW wants to register in a DC Network
Notification 2	A SMGW has succesfully entered a DC Network
Notification 3	A SMGW is exiting a DC Network
Notification 4	Resent local superposition according to the correction procedure
Notification5	Transmitting local superposition

Table 3.1: An overview of all notification messages.

Uninvolved SMGWs do not receive any information about the entry of new users in a DC network. In addition, the participants share their client identifiers during the key exchange. Due to the exchanged communication details, each participant in the DC network knows the identification number of its neighbor. This is later helpful for error correction measures.
The use of a key exchange method also has disadvantages. By forwarding messages, the electricity provider knows which SMGW have exchanged keys with each other. Exchanging keys is equivalent to creating an edge in the key graph.Therefore, the electricity provider can easily replicate the key graph of the DC network. The knowledge about the structure of the key graph alone does not give the electricity provider any further knowledge, but a malicious electricity provider could use the knowledge to launch active attacks on individual SMGW. An example would be that a electricity provider wants to get information about the power consumption of an SMGW. The electricity provider could connect one or more SMGWs it controls to the victim SMGW through a key exchange that the attacker SMGW launches. The electricity provider could now hope that in the future the victim SMGW will only have keys with the attacker SMGWs. Since the electricity provider controls the attacker SMGW and knows the keys of the attacker SMGW, it can reconstruct the electricity consumption of the victim SMGW from the local superposition. In this case, participants must have a minimum number of neighbors to avoid this attack. Furthermore, the electricity provider would have too much power in the DC network if it can control which SMGWs connect to each other upon entry. Therefore, a joining SMGW must be assigned to a random partner in the DC network. Furthermore, in order for error correction measures to be implemented as easily as possible, the resulting key graph in the DC network must be planar.

SGMW Registration in a DC Network

An SMGW that wants to register in the DC network sends a special defined request to its power provider. For this purpose the notification field in the header is used and notification 1 is sent. Notification 1 represents a request from the SMGW to register in a DC network. The electricity provider assigns the requesting DC client to a suitable geographical region and suggests a random client (SMGW) which is already registered in the DC network. Afterwards the electricity provider establishes a tunnel and sends the tunnel information to the DC client and the registering client. Via this tunnel it is possible for the two SGMW to create a communication link via the electricity provider. If an SMGW sends to the tunnel, the message is forwarded to the future neighbor. The

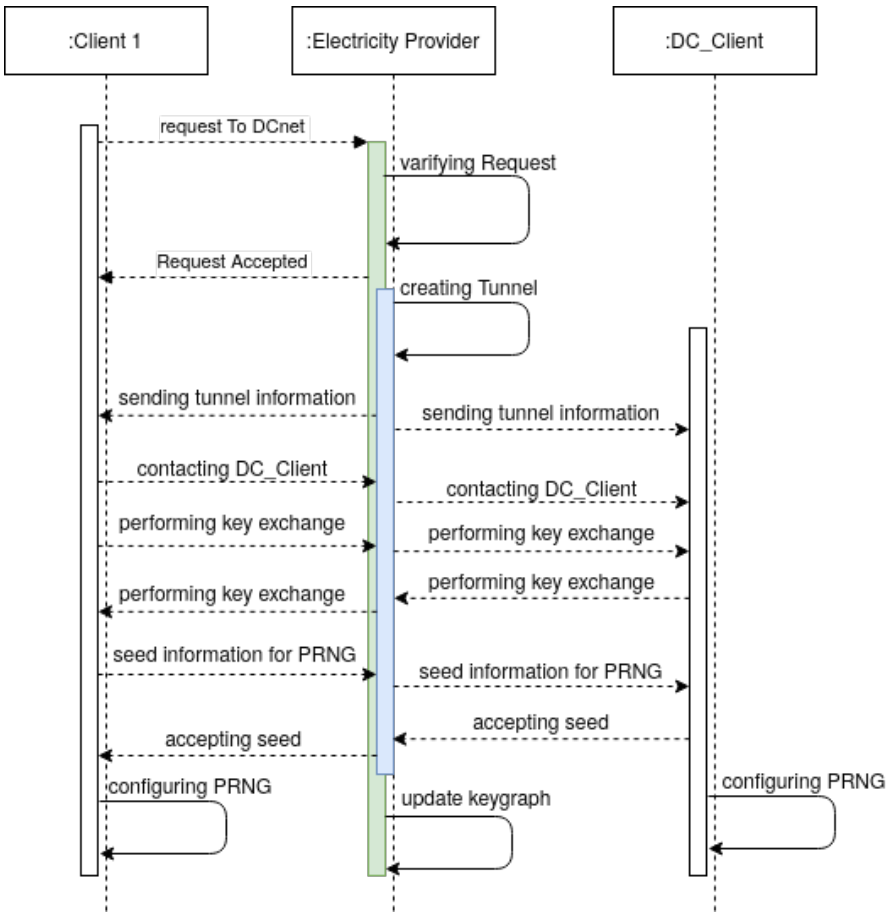

Figure 3.3: An example of a NILM analysis.

DC client which is already present in the DC network is only informed by the electricity provider that it receives a new neighbor and has to exchange contact information. The requesting SMGW needs to send the seeds over the tunnel. To prevent the electricity provider from reading the seeds, the clients exchange Diffie Helmann keys via the tunnels provided.Once the seeds have been exchanged, the power provider is informed by the requesting client that it has successfully entered the DC network by notification 2. The PRNGs generate the keys that are added or subtracted to the message to create the local superposition. The procedure was explained in more detail in ref.Afterwards all participants in the DC network can send their local superpositions to the power provider. The provider forms the global superposition and receives the aggregated power consumption of all SMGWs in the DC network. If necessary, the electricity supplier must ensure that in the future messages can continue to be exchanged between registered customers via the same tunnel. The already registered clients can therefore not choose which new communication partners they get. In addition it is assumed that the electricity provider has already been authorized by the SWA. Otherwise, the SMGW would not be able to establish a connection to the supplier.

SGMW Normal Operation

So far, the steps to initialize a DC grid into the already operating power grid have been explained. Next, a description is given of how the technical process takes place in the DC grid, assuming that no faults occur or corrective measures need to be taken. The SMGW transmits its electricity consumption periodically from the moment it enters the DC network. The time period is defined by the electricity provider. For the DC network, it is most practical if all SMGWs send their local superposition to the power supplier at the same time or within a short transmission interval (e.g. one minute). This can be done without problems, because according to ref 3.1 all SMGW must update their time in regular intervals with NTP servers. If an SMGW has not sent a local superposition within the transmission interval, corrective measures are implemented. The frame that a SGMW sends to the power provider is filled in as follows:

1. DC Net Identifier:
   The DC net in which the SMGW is registered is entered here.
2. Client Identifier:
   The assigned client identifier is sent in this field.
3. Transmission Bit:
   This field is exactly 1 bit and is set to 1 when a local superposition is sent.
4. Time Stamp:
   A time stamp is appended when the frame is generated.
5. Notification:
   Notification message 5 is sent to inform the electricity provider that this message is a local superposition.
6. Data:
   Generated local superposition is entered in the data field.

The electricity provider processes the received frames according to the following procedure:

1. DC Network Identifier:
   DC Network Identifier indicates to which DC net the message is processed.
2. Client Identifier:
   The Client Id of the frame is stored in a memory structure. In the memory structure can be looked up later, which client has not sent a locale superposition in the round.
3. Transmission Bit: Each message has a transmission bit set to 1. All transmission bits are added up and at the end of the round it can be checked whether all SMGWs have sent their local sum. If the summed transmission bits do not correspond to the number of participants in the DC network, correction procedures must be applied.

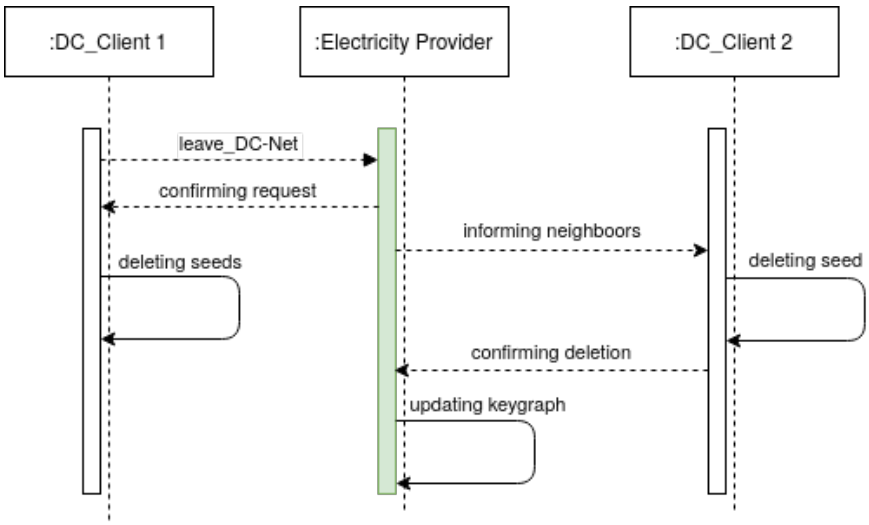

Figure 3.4: An example of a NILM analysis.

4. Time Stamp: The power supplier can assign the message to the correct round.

5. Notfication: Notification message 5 informs the electricity supplier that the local superposition is being transmitted.

6. Data: The local superposition in the field is added up with all other superpositions and the electricity provider gets the global superposition. This is the aggregated power consumption of all SMGWs in the DC network.

SGMW Exit from the DC Network

An exit can be caused, for example, when the customer changes the electricity provider. Then a defined message is sent to the electricity provider. The electricity provider informs the neighbors of the exiting client with a notification message 3. However, to prevent the notification from being misused by the power supplier, notification 3 can only be sent following the exit of an SMGW. Otherwise, a malicious electricity provider would be able to change the structure of the DC network at will. In addition to the notification message 3, the DC Client Identifier of the exiting SMGW is sent as well. This notification signals to the neighbors of the exiting SMGW that they must discard their PRNG configurations to client X and that they must not be used in the calculation of the local superposition in the next round. In order to avoid a synchronization error in the DC network, the ”neighbors” must confirm to the power provider that all seeds have been discarded. Otherwise the case may occur that a SMGW continues to add the old key to its message. This would result in a useless global sum. Furthermore, the key graph must be considered. It could be the case that the underlying key graph splits into two DC networks. If this is the case, two separated DC networks are sending to the same DC network identifier. In the example of Figure 3.3, this could happen if Sabine and Rüdiger throw away their shared key. The result is different depending on the

position where a DC net splits. But at least one DC network experiences a significant loss of anonymity due to the smaller number of participants that can be aggregated. In the case of particularly serious splits, it can even lead to a participant being completely disconnected from the DC network. If a disconnected client notices that it no longer has any neighbors, it sends a special emergency message to the power provider. Then a new registration process is initiated before the next round starts.
To avoid splitting into two DC networks, the exiting SMGW informs its neighbors with which direct partners it was connected. These then initiate a registration process and exchange keys with each other. The fact that all neighbors have exchanged keys with each other guarantees that a DC network does not split when an SMGW leaves. Furthermore, all participants have to have a minimum number of 3 neighbors. This makes the possibility of disconnection from the DC grid much less likely, since several neighbors would have to leave the DC grid at the same time for a participant to be exposed.

SGMW Connection loss

SMGWs have an Internet connection with which they can communicate via the WAN. If the Internet connection is interrupted, this can lead to an SMGW not being able to send its local superposition in time. The result is that the electricity provider cannot calculate a meaningful global sum in the round. The electricity provider notices the error immediately because the global transmission bit does not correspond to the number of participants in the DC network. In this case, the following corrective actions are implemented: The electricity provider detects which SMGW has not sent a local superposition based on the client identifier. Since the SMGW sends a complete header and the client identifier of an SMGW is also sent underneath, an electricity provider only has to check which client identifier was not sent. The missing client identifier is also the client that is defective. Once the defective client is located, notification 4 is sent by the power provider to the neighbors of the defective client. The notification contains the Client Identifier of the defective client and requests the neighbors to recalculate their local superposition, but without using the key of the defective client. Afterwards the local superposition is resent to the electricity provider. Even though the first attempt failed the electricity provider can calculate a meaningful global sum by using the resent local superposition instead of the old local superposition. At the same time, the keys of the defective client are stored by the neighbors in a backup, so that when the defective client re-enters the DC network, the same key graph is restored. The neighbors of the defective client experience no loss of anonymity during the correction process.After this procedure, the defective client is temporarily no longer in the DC network. As soon as the SMGW obtains an Internet connection, it must register again in the DC network. The power consumption of the SMGW during the time of Internet loss is not retransmitted. This is because retransmission the power consumption would lead to a complete loss of anonymity. The electricity provider knows at which time stamp which smart meter was defective and could assign the resent electricity consumption directly. In addition, the electricity provider knows how many smart meters are functional at any time and how many are defective through the transmission bit and the assigned number

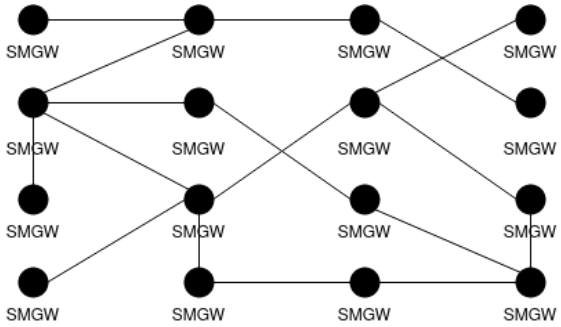

Figure 3.5: An example of a NILM analysis.

of subscribers in the DC network. With this information, the electricity provider is able to ensure good network stability without any problems. Especially since the chance of an Internet outage is unlikely. The billing is also done correctly because the billing is calculated on the SMGW. Therefore, the electricity provider does not have to fear any loss of income, even though the electricity consumption is not sent.

Manipulation of the Local Superposition

One of the considerations that absolutely must be made in a DC network is: what happens if an SMGW intentionally manipulates the local superposition? First of all, it must be mentioned that this attack does not help the customer at all to avoid the electricity costs. This is because the electricity costs are calculated by a separate procedure and therefore the billing cannot be affected by the attack. If this problem does occur, it should rather be assumed that it is an external attacker who has taken over an SMGW and wants to sabotage the availability of the DC network. If the local superposition is manipulated, e.g. by deliberately sending a wrong local superposition, it is no longer possible for the power supplier to calculate a meaningful global superposition. Ergo, it is not possible to see the aggregated power consumption for the whole DC grid. In the case of the attack, the electricity provider cannot assume that the situation will resolve itself and must take measures to find the manipulating SMGW. Furthermore, it must be assumed that the attacker is well aware that the electricity provider will be looking for him. Therefore, the procedure must be designed in such a way that the attacker is found even though he tries to conceal his identity. For this purpose, a slightly modified version of Prof. Dr Pfitzmann's error localization and recovery protocol can be applied. The protocol of Prof. Dr Pfitzmann talks about 2 different modes, the anonymity mode (A-Mode), in which the DC net works normally and the fault tolerance mode (F-Mode), in which defective stations are searched for. The F-mode can be extended in this application so that an attacker can also be searched for. If there is an incorrect calculation in the global sum, this is communicated publicly by the power supplier to the SMGW. All SMGWs then save the keys from the last round. The power provider saves all receiving local sums from the round and enjoys special rights that only prevail in F-mode. In the following, the property of DC networks is exploited that allows

a DC network with one meaningful global sum to separate into two DC networks with two separate meaningful global sums. The algorithm is executed as follows: 1. halve the key graph. The power provider has the overview of the key graph and can therefore separate the key graph into two parts. Splitting the graph into two halves should be trivial since the planar separator theorem holds. If a SMGW is exactly on the border of the bisected key graph and has a neighbor in the other half of the key graph, then this SMGW is informed as a special node by the power provider that the neighbor's key is thrown away for this computation. The temporary throwing away of the keys leads to the splitting of the key graph at that point. The power provider can request an SMGW to throw away a key only in F-mode! All SMGWs in one half of the key graph now retransmit the local sum from the last failed round. The adjacent SMGWs that threw away a key calculate the new correct local sum without the neighbor in the other part of the key graph. This results in all nodes sending the same message from the last round except the special nodes. This allows two global sum to be calculated. One global sum from the first half and one global sum from the second half. So the old DC net round is repeated, but in a split net to reduce the number of possible attacker SMGW. The electricity provider can check by the stored local sum if the same local sum would really be resent and can check the correctness. If the formation of a global sum fails in the first half of the dc network, then the attacker is located in the first half and the same procedure is repeated in the first half. If the formation of a global sum fails in the second half of the dc network, then the attacker is located in the second half and the procedure is repeated in the second half. If the global sum fails in both halves or is calculated correctly, then the attacker is among the special nodes. The procedure is continued recursively until it is reduced to one SMGW that is eligible to be the attacker.

There can be the special case that the attacker SMGW is a special node. Since the special nodes send a recalculated local sum, the power provider cannot immediately rule out whether the attacker is among the special nodes. Therefore, for each bisection, an additional subgraph must be formed in which the former special nodes have no neighbors outside the subgraph. In this way it is possible for the electricity provider to control the local sum when resending the local sum of the former special nodes.

DC Network Size

With a small number of participants, conclusions can be drawn about individual participants from the aggregated result. This is the case, for example, if the electricity consumption of one user is equal in percentage to the residual consumption of the other users. Or it is also feasible that a user does not consume any electricity. In a DC network with 2 users, the power consumption would be directly readable even if the individual loads are aggregated. The goal is not to avoid the disclosure of information, but rather to make it hard to draw inferences about an individual user from the aggregated consumption.In this thesis, the experiments chapter determines the minimum size of a DC grid to guarantee that statistical inferences are difficult to realize. On the other hand, it is in the interest of the power supplier not to realize huge DC networks. The more participants a DC network has, the more frequently errors occur that have to be corrected by corrective measures. Although the DC network should scale with many

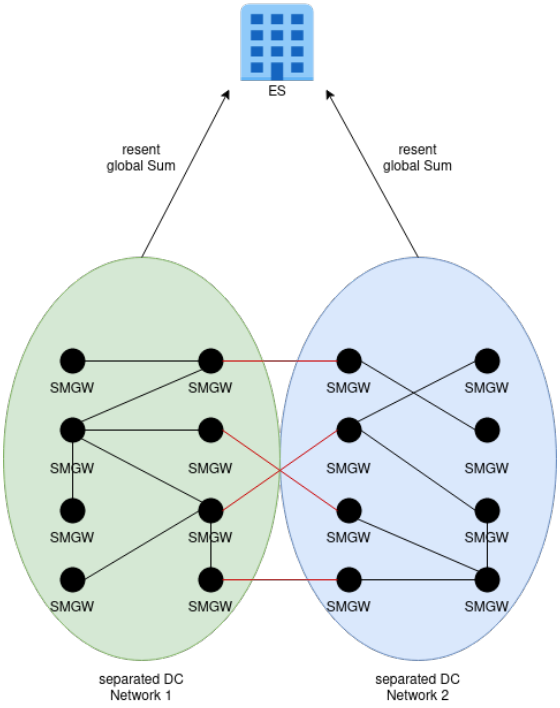

Figure 3.6: An example of a NILM analysis.

participants, the question arises as to how meaningful the results are when hundreds of participants partially fail.

Anonymity

By using PRNG, anonymity decreases from information-theoretically secure to complexity-theoretically secure anonymity. Nevertheless even if an SMGW is controlled by an attacker, it would not be possible for the attacker to read the power consumption of other SMGWs. This is because the attacker has no access to the global total. This can only be calculated by the electricity provider. With the proposed method, the attacker lacks the necessary information to launch a potential attack on the DC network (if the attacker only has control over one SMGW). Furthermore, the attacker also has no information about the key graph. This further complicates the chances of a successful attack to deanonymize the electricity consumption of costumers. The electricity provider is potentially the most dangerous attacker in the protocol, since the SMGWs cannot communicate with each other, they rely on the electricity provider to register in the DC network. As a result, the power provider has to take over administrative tasks and therefore possesses a lot of control. To ensure that the electricity supplier is not too powerful, its competences have been restricted. The best chance of the electricity provider to break the anonymity to its customers is that the administrative powers are abused to connect individual SMGWs to malicious neighbors. Once the electricity provider manages to link an SMGW with only malicious neighbors, the local sum can be

reconstructed and the electricity consumption is visible. This is prevented by forcing the electricity provider to select a random neighbor upon entry and not being able to remove SMGWs from the network on its own. These measures make it almost impossible for the electricity provider to affect the selection of neighbors in the DC network.

4 Implementation

The conceptual approach of the DC network was presented in Chapter 3. Nevertheless, smart grids are a real-world system and it must be shown that the theoretical solution can be implemented in a practical environment. This chapter deals with the implementation of the introduced protocol. It is described which technical tools were used and at which implementation steps problems occurred.

4.1 Structure of the practical implementation

A DC network was implemented with the same requirements as defined in chapter 3. For technical reasons, however, the exact same structure could not be implemented. If there are any deviations from the defined protocol, then these will be described and explained in this chapter. 4 Raspberrypis are used to realise the design, where 3 Raspberrypis simulate the SMGW and 1 Raspberrypi represents the electricity provider. In the following, the Raspberrpis that represent the SMGW are called clients and the Raspberrypi that represents the power provider is simply referred to as the power provider. All clients have a communication link via Lan to the power provider. However, the clients do not have a physical connection to each other. As suggested in the protocol, the only way for the clients to communicate is through the power provider. After the clients join the DC network, the clients build their local sum and send it periodically to the power provider. The electricity provider adds up the local totals and stores the global total in an external text file. In a smart grid, electricity consumption is sent to the electricity provider every 15-60 minutes. Since the implementation is a demo, the sending interval is 10 seconds. In addition, the demo was implemented in such a way that after 4 messages a client fails and a corrective action must be taken. After that the client can re-enter the DC network. To avoid having to implement the application and the related network protocol, gRPC was used as a framework.

gRPC Remote Procedure Calls - GRPC

GRPC is an open source remote procedure call (RPC) system developed by Google since 2015. GRPC relies on a client-server structure and simplifies the construction of linked systems. With GRPC, so-called services can be defined. Each service allows to declare different functions that can communicate via a self-selected message format referred to as Protocol Buffers. Therefore on the client the functions are implemented, while the server runs the interface and processes the client requests. On the client there is a stub that holds the same functions that are on the server. In gRPC server and client can communicate with each other, although they were implemented in different programming languages. In this work, both server and client were implemented in Python. A simple

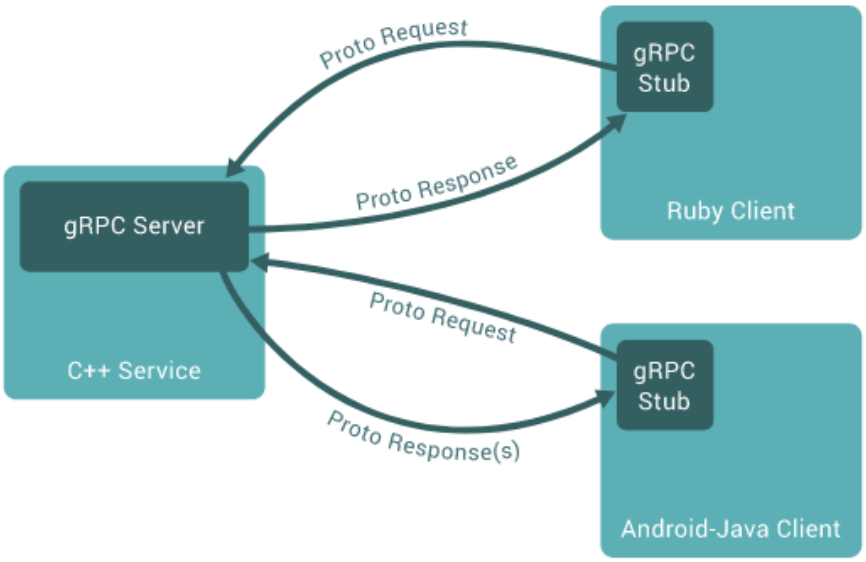

Figure 4.1: An example of a NILM analysis.

application example is shown in Figure 4.1.

Protocol Buffers in gRPC

Protocol buffers are used by default in gRPC and allow structured data to be serialized. With Protocol Buffers, structured data can be specified as message formats. Then the source code is automatically generated from the structured data and it can be sent through channels provided by e.g. gRPC. Listing 4.1 shows the implementation of the protocol header as a protocol buffer. Each header field is assigned a data type and also a unique field number that determines the order of the fields.

```
1  syntax = "proto2";
2  package DCnetPackage;
3
4    message DC_net {
5      optional  int32  dc_net_identifier = 1;
6      optional  int32  client_identifier = 2;
7      optional  int32  transmissionBit = 3;
8      optional  string timestamp = 4;
9      optional  int32  notification = 5;
10     optional  int32  localSum = 6;
11   }
```

Listing 4.1: Protocol Buffers message declaration using `protobuf/lang` and the style in `protobuf/style`.

4.1.1 Server Implementation

The server implements all functions that are defined as a service in the protocol buffer. The client implements most of the logic of the DC network. However, when the client accesses a service, most of the service functionality is implemented on the server. Therefore, the client prepares all necessary data and sends the data over a communication channel to the server. The data is then processed by the server and the result is communicated to the client as a response. Listing 4.2 shows all service functions implemented on the server. The data is then processed by the server and the result is communicated to the client as a response. Listing 4.2 shows all service functions implemented on the server. Each function in the service is defined as an RPC and a name is assigned to the function. In addition, the RPC can accept a message format. For example, the addClientToDCnet function uses the message format defined in Listing 4.1.

```
1  syntax = "proto2";
2  package DCnetPackage;
3
4    service DC_round {
5      rpc SendLocalSum(DC_net) returns (Acknowlegde)  {}
6
7      rpc addClientToDCnet(DC_net) returns (DC_net) {}
8
9      rpc connectDCClients(DC_net) returns (DC_net) {}
10
11     rpc ExchangeOpenKeysForDH(Secret) returns (Secret) {}
12
13     rpc getDiffieHellman(Empty) returns (DiffieHelman) {}
14
15     rpc ExchangePRNGSeed(Seed) returns (Seed){}
16
17     rpc deleteClient(DC_net) returns (Acknowlegde){}
18
19     rpc sync(TimeStamp) returns (Acknowlegde){}
20
21     rpc updateGlobalSum(DC_net) returns (Acknowlegde){}
22     }
```

Listing 4.2: Protocol Buffers message declaration using `protobuf/lang` and the style in `protobuf/style`.

The functions of the service can be divided into 2 categories. First, helper functions for registration or initialization into the DC network. Second, functions in which logic of the DC net is implemented. In the following, the task of the helper functions is described first. Then the more complex functions are described, which also contain logic of the DC network.

4.1.2 Client Implementation

When the in client application starts, a fork function is executed on client side. The fork function allows an process to create a second process by duplicating the address space of the calling process. The calling process is called parent and the duplicated process is called child. In the application, the parent process takes care of the child process. If the child process crashes, the parent process ensures that the child is restarted correctly. The main logic of the DC network is found in the child process of the client. Only in the child process the services provided by the server are called and used. In addition, the communication channel of gRPC is implemented only on the child. The communication channel forms the interface in which the defined protocol buffers can be sent as messages. This means that the parent process has no access to the communication between server and client. Once the communication channel is configured, the interaction between client and server can begin. Figure 4.2 illustrates the structure of the client implementation in a flow chart.

Initialization

First, it is assumed that the server is already started. Otherwise, the client cannot be started in gRPC, since no communication channel can be created. The server waits for requests from the client and processes the requests. Second, the child establishes a connection to the electricity provider. In response, the child receives a DC Net identifier and a client identifier. A registering client is stored with Client Identifier by the server in a list. This gives the server an overview of the number of participants in the DC Net. In addition, the server can identify faster which client could not send its messages in case of error correction measures.

Registration

After a client has received a DC Net identifier and a client identifier, the client requests the server to assign it a random neighbor. In order for the neighbor and the registering client to synchronize their PRNG, both have to perform a Diffie Hellman key exchange. All the necessary information for calculating the public key is provided by the server and is communicated to the clients on request. In the proposed protocol from the previous chapter, it was described that the electricity provider must offer a tunnel so that two SMGWs can communicate and exchange public keys via Diffie Hellman with each other. The tunnel in the protocol represents the communication channel in gRPC. But in gRPC all RPC requests are started by the client. The server only responds to the client's requests. Forwarding messages from a requesting Clients to a non-requesting clients is therefore not possible or very difficult to implement in gRPC. Instead of the public keys being forwarded to the clients by the server, as suggested in the protocol, the public keys are stored by the clients on the server for a short time and the clients make a request for the public keys. Each client only needs to obtain the public key of its neighbor through a request and is able to compute the secret key. After the Diffie Hellman key exchange has been executed and the PRNG has been configured, the client

can create the local sum and send its local sum to the power provider/server for each round.

Generation of the global Superposition

The server receives all the local sums and adds them up to get the global sum. Subsequently the server verifies the correctness of the global sum. If the global sum is incorrect, then a client in a DC network has failed and could not send a local sum. The defective client is located by the server by looking up which client has not sent a local sum to the server. Afterwards the neighboring clients are instructed by the server to recalculate the local sum without using the key of the failed client and resent it to the server. The global sum is then updated by the power provider.

Preparations for the next Round

After each sending of the local sum, all clients check whether a new subscriber wants to register in the DC network. If so, one client is notified by the server that it gets a new neighbor. Then another registration process takes place with the Diffie Hellman key exchange which was described before. Then another registration process takes place with the Diffie Hellman key exchange which was described before. Another deviation from the protocol is that in the implementation each client has no minimum number of neighbors than the required 3 neighbors in the protocol. The flow chart in Figure 4.2 illustrates the structure of the client implementation with the key functions.

4.2 Challenges

Several challenges were encountered during the implementation in this thesis. The most serious and time-consuming 3 errors are described below.

Multiple Client access on the Server

In the experimental environment, 3 clients communicate with a server. A particular challenge was therefore to implement the server cleanly and consistently so that multiple clients could access the same function or even the same line of code at the same time without the server crashing or the program entering an inconsistent state. The implementation was therefore particularly difficult in the service functions in which the calculations of the global sum or the verification of the global sum was carried out.

gRPC Client-to-Client Communication

The protocol defined that the electricity provider must provide a tunnel for SMGW to SMGW communication. It was tried to implement the same structure as in the protocol. However, the clients in the gRPC framework do not receive an IP address, so the server

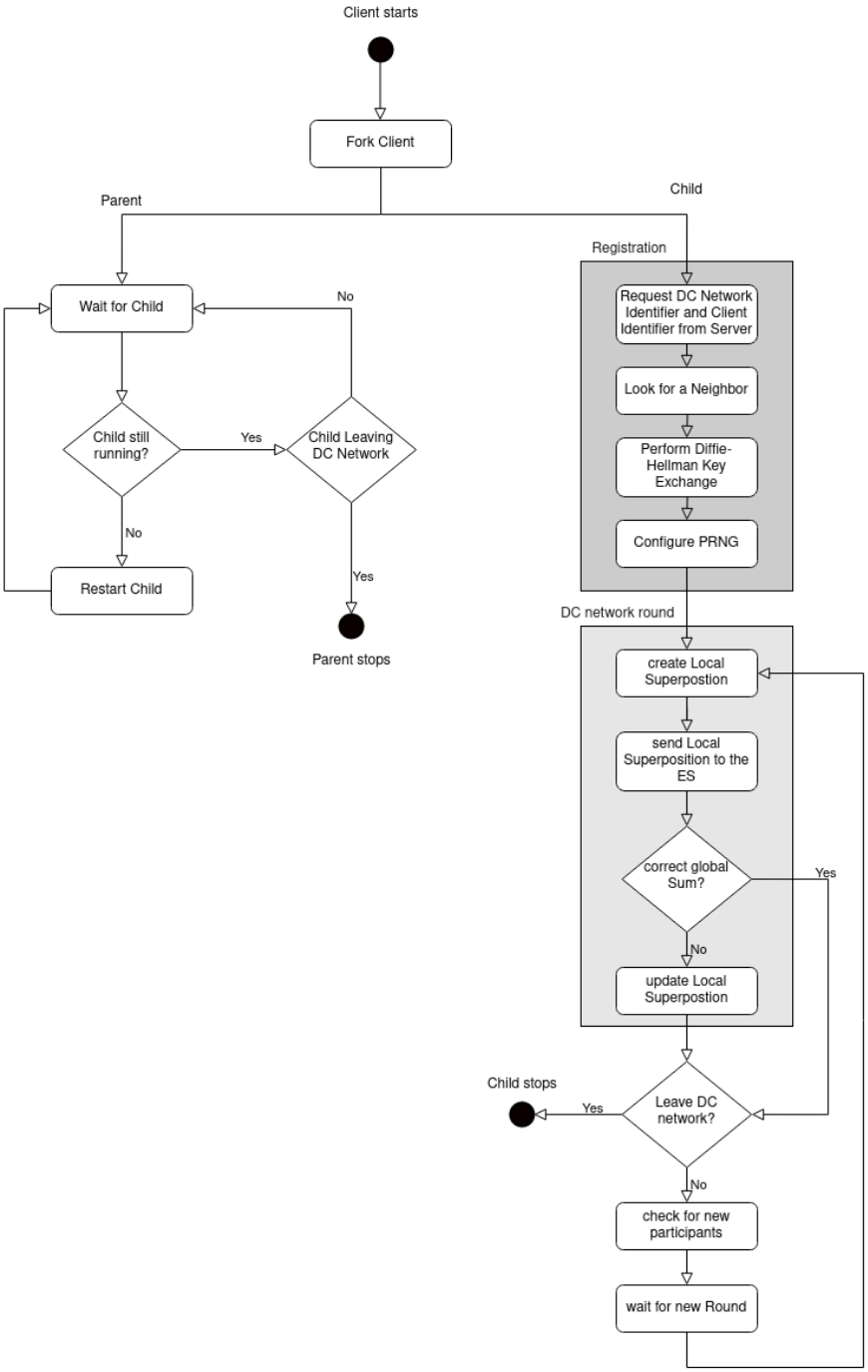

Figure 4.2: An example of a NILM analysis.

cannot distinguish calls from different clients. In search of a solution the following suggestions have been considered.

1. Setup a Server on each Client:
   In this approach, each RaspberryPi on which a client application is implemented would get an additional server that can communicate with the power provider. In addition, a client would have to be implemented on the power provider so that requests from the power provider can be sent to clients over a communication channel. In the work it was decided against it, because it is a considerable additional effort and requires an extra implementation of servers on the clients.

2. Bi-directional Streaming:
   gRPC offers different ways to send messages. One way is a bidirectional streaming RPC where server and client exchange a sequence of messages. In gRPC both streams work independently and can be configured depending on the use.In a message stream the server can wait for all client requests before responding or it can respond immediately after each message. A stream can be defined in the Protocol Buffer syntax with the keyword stream. The stream property can be exploited to implement message forwarding in gRPC. Various solution sketches have been found that promote this approach. However, the proposed solutions were vague and storing the public keys in the short term on the server was much simpler to implement.

Crashing Parent

During the implementation there were indeterministic crashes of the parent in the client. The error could not be traced for a long time, because the crash occurred at different times during the runtime of the application. When the parent crashes, the child continues to work without interruptions. However, the child cannot be restarted if the parent crashes. It has been found by accident that the fork function is not executed correctly after the communication channel has been created.

5 Evaluation

. . . evaluation . . .

write evaluation

6 Future Work

. . . future work . . .

write future work

7 Conclusion And Outlook

. . . conclusion . . .

write conclusion