Master's Thesis

Privacy-preserving Smart Metering Using DC-Nets

Gregor Garten

March 11, 2022

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

1. Reviewer:	Dr. Stefan Köpsell
2. Reviewer:	Dr. Elke Franz
Supervisor:	Dipl.-Inf. Tim Lackorzynski

AUFGABENSTELLUNG FÜR DIE MASTERARBEIT

Name, Vorname des Studenten:	Garten, Gregor
Immatrikulationsnummer:	4055357
Studiengang:	Master Informatik (2017)

Thema (deutsch): Datenschutzfreundliches Smart Metering Mithilfe von DC-Netzen

Thema (englisch): Privacy-preserving Smart Metering Using DC-Nets

Zielstellung:

In the future, the classic electrical grid will be transformed into the so-called smart grid. This will enable diverse use cases, different power consumption and production patterns as well as bigger and more frequent turnovers. The central component of the smart grid will be an information and communication network, that connects energy producers and consumers.

Endpoints at the end user side will be the so-called smart meters. These electronic devices can send consumption data digitally over an information network, most likely the Internet. Yet, they can also receive data, for example new price schedules, that enable power intensive devices to plan their energy consumption.

The smart meter will be able to emit very fine-grained consumption data. Reading frequency ranges from 30 minutes down to 2 seconds between measures. This makes it possible to infer daily routines, times of absence or even which electrical devices were used by the end users.

To protect the privacy within the smart grid while at the same time allowing for advanced functionalities (dynamic billing, demand side management...), many protocols with vastly differing assumptions have been proposed in the past.

Yet, these assumptions are often not compatible with the real world (homomorphic encryption, availability of trusted third parties, reliance on the energy distribution network).

With the advent of new truly trustworthy embedded systems (as proposed in the Jupiter project), smart metering systems come into reach, which can preserve the end users privacy and at the same time providing the advanced functionalities of the smart grid (real-time metering) while not relying on unrealistic assumptions.

To this end a system shall be proposed, that employs so-called DC-Nets, as proposed by David Chaum, to implement sender anonymity for meter readings, so that an energy utility may construct the proper value of the energy consumption of a whole group of end users for a point in time, while not being able to attribute values to individual users. The leading assumption hereby is, that the smart meters are trustworthy entities, meaning the information they emit can be trusted. The proposed system shall then be prototypically implemented.

Aufgabenstellung für die MASTERARBEIT

The participants in a DC net constitute an anonymity group, where their energy consumption is added up and cannot be attributed individually. Meter readings over time (consumption curves) are very specific markers of an individuals behavior and the quality of the anonymization depends on the number of involved participants.

Therefore, additional research shall be done on the necessary size of those groups. Part of that research is to find a suitable way to measure or judge how well the individuals meter data is masked or anonymized by the aggregation of the DC net. To this end a simulation environment for smart meter networks from a previous project that employs real world data shall be employed.

In der Arbeit sollen schwerpunktmäßig folgende Teilaspekte bearbeitet werden:

- State-of-the-art analysis on related approaches (design and investigations into anonymity group sizes)
- Propose a privacy-preserving smart metering scheme, as described above (including phases for initial setup and establishment of trust between parties)
- Implementation of a prototype
- Explore and measure efficacies of different sizes of anonymity groups in actually anonymizing individual user behaviors
- Evaluation of proposed scheme

Betreuer:	Dipl.-Inf. Tim Lackorzynski
Verantwortlicher Hochschullehrer:	Dr. Stefan Köpsell
2. Betreuer:	Dr. Elke Franz
Beginn am:	12. Juli 2021
Einzureichen am:	12. Dezember

12.07.2021 ....
Datum, Unterschrift der/des Studierenden

....
Unterschrift des betreuenden Hochschullehrers

Selbständigkeitserklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich reiche sie erstmals als Prüfungsleistung ein. Mir ist bekannt, dass ein Betrugsversuch mit der Note "nicht ausreichend" (5,0) geahndet wird und im Wiederholungsfall zum Ausschluss von der Erbringung weiterer Prüfungsleistungen führen kann.

Dresden, den 10. März 2022

Gregor Garten

Contents

List of Figures

List of Tables

1 Introduction

The increased focus on renewable energies in recent years has lead to a redistribution in the traditional power grid. The expansion is in fact a central role of the energy transition in Germany and renewable energies account for 45.3% of gross electricity production in 2020 [Umw21]. The tendency is rising and Germany aims for climate-neutrality by 2035. As a result of the development, the distribution of producers in the power grid is changing from a few large power plants to many smaller renewable electricity producers. The volatile power generation of renewable energies requires efficient communication between all components in the power grid to ensure a stable electricity grid.
This interconnection and communication of components in the power grid is called the smart grid. The core part within the smart grid is the smart meter. An intelligent electricity meter that transmits status information and electricity consumption from consumers to the participating electricity suppliers via the Internet. Additionally, the combined consumption of all users makes real-time pricing possible in the first place. With an integrated overview, customers can check their own electricity consumption. This facilitates the individual user to save electricity and to schedule the electricity consumption for times of the day when electricity is particularly cheap.
Besides the comfortable advantages, increased communication between the components also has remarkable disadvantages. Smart meters on customers-side send electricity consumption data from residential houses at short intervals, that can provide detailed insight into the behavior of individuals in their own private homes. This is a serious invasion of privacy from which the customer must be protected. Moreover, if the information gets into the wrong hands, it could be misused for criminal purposes such as burglary. Since it is very easy to find out whether someone is at home, by tracking their electricity consumption. In addition, with a simple analysis of the electricity consumption, the daily habits or even the religious affiliation of people can be revealed without much effort.
In this master thesis a privacy-preserving solution based on DC networks is presented, which is a privacy-enhancing technology (PET) that is used to achieve anonymity on the communication networks. Therefore, privacy can be guaranteed to the participants and at the same time all advantages of a smart grid can be offered to the energy suppliers. Special attention is paid to the technical guideline-03109 Technical Guideline-03109 (TR-03109) published by the German Federal Office for Information Security Bundesamt für Sicherheit in der Informationstechnik (eng. Federal Office for Information Security) (BSI), which specifies initial requirements for smart meters. The network protocol is based on these requirements and has been designed for easy integration into the smart grid.
The second chapter introduces all the essential terms that are necessary for a basic understanding of this work. All participants in a Smart Grid are introduced and described. Furthermore, other solutions from the scientific community are presented. Finally, the

technical guideline-03109 TR-03109 is discussed in detail and the most important requirements as well as the stakeholder model for the design are described.
The functionality of DC networks and the design follows in the third chapter. The design of the network protocol defines all standard operations and describes what behavior the network protocol should ideally have. Then, it introduces error correction procedures for unpredictable events so that the network remains in a consistent state. This concludes the design and moves on to the implementation phase of the thesis. At the end, the proposed solution is evaluated in terms of performance and security objectives.
In the fifth chapter the experiments, which were accomplished apart from the Design and the implementation, are explained. A minimum size of DC networks is searched for, which prevents a conclusion from aggregated power consumption.
All the results of this work are summarized as a conclusion in the last chapter. Additionally, an outlook on future work is given.

2 Background

The following section explains the key components of the smart grid, the structural changes as well as the occurring challenges that need to be solved. The overview provides the fundamental understanding of this work. Because this thesis requires a basic understanding of the German smart grid, the stakeholders and their interests in the smart grid are expounded. In addition, the technical guideline from the BSI[1] is introduced because it defines a first security-relevant standard required for all smart meters in the German smart grid. Moreover, this chapter discusses the current state of research and the solutions which are discussed in the scientific community.

2.1 Smart Grid

The original energy network was mainly considered as a transmission system to send electricity from the generators via an elongated network of cables and transformers to the consumers.The distribution from the traditional few electricity producers (e.g. nuclear power plants, coal-fired power plants), which were responsible for a large part of the electricity generation, is changing to a rising number of smaller producers (e.g. wind turbines), due to the increased integration of renewable energies. But, renewable power generation is often depending on external environmental factors. In order to maintain a stable grid, smart meters are used to increase the overall grid quality through regular communication of status information and power consumption. Smart meters enable the electricity provider to receive the electricity consumption of a household every 15 minutes. This offers the possibility to access the current electricity demand from the consumers more easily. Previously, the current electricity demand was simulated from load forecasting models. If the demand should increase spontaneously, peaker plants, mainly consisting of gas-fired power plants, would be turned on to quickly meet the demand. This is costly and environmentally unfriendly. Since then, structural changes were made in order to optimize the energy grid and make it more intelligent by exchanging information in near-real-time. This allows the demand to match with the available supply. The fundamental component of the smart grid are the smart meters which will be discussed more detailed in the following paragraph [Fan+12][Zea+13].

Smart Meter

Smart meters are the essential component in a smart grid. It is an electricity meter which has an interface to the Internet. Beside the metering of the power consump-

[1] BSI - Bundesamt für Sicherheit in der Informationstechnik(eng. Federal Office for Information Security)

tion for electricity provider the additional functions of a smart meter are to provide a two-way communication between the electricity grid participants and the smart meter. The participants receive the status data from the smart meter and process it further. The extended transmission of information between smart meters and electricity grid participants via the Internet is also called Advanced Metering Infrastructure Advanced Metering Infrastructure (AMI). The resulting communication between the components improves the quality of the power grid and makes it possible to offer new services. For example, detecting power outages used to be depending on customer calls whereas nowadays the grid operator can do that itself.
Another new feature is a detailed monitoring of power flows at the smart meter, instead of only measuring it up to substations like before. Moreover, the advanced functions enable electricity network operators to quickly detect changes in consumption behavior and react to them without having to use costly and environmentally unfriendly peaker plants. 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 out 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 [FB14].

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 messages every 15 minutes 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 based on data from the past. However, sending user information in such a short period of time allows new methods to be used to create accurate behavioral analyses in an individual 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. At first it is described how simple behavioral analyses are generated by electricity consumption. Subsequently, solutions to Metering for Billing and Metering for Operations will be presented, that have been discussed in the scientific community so far [FB14].

2.2.1 Non-intrusive load monitoring

Interpreting electricity consumption with the intent of identifying devices at home is called non-intrusive load monitoring Non-Intrusive Load Monitoring (NILM). George Hart and Fred Schweppe were the first to develop non-intrusive load monitors in 1985

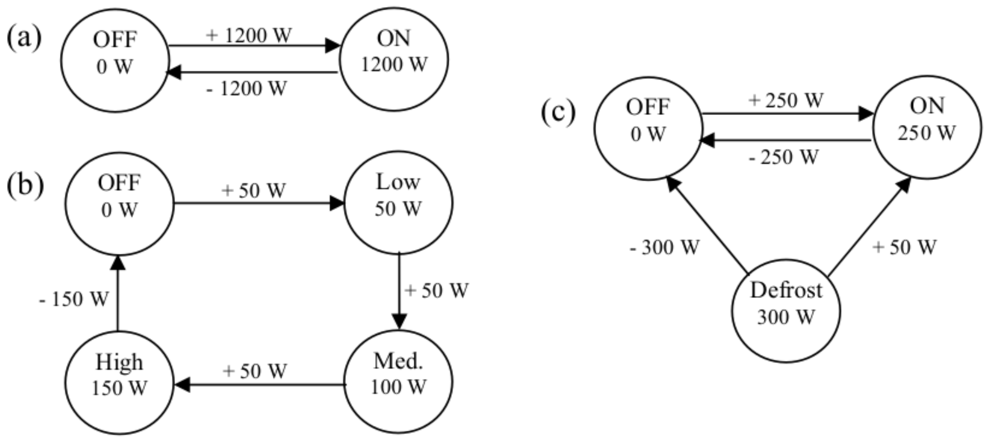


Figure 2.1: Sample Appliances after performing step 3. (a) represents a toaster that has 2 states. (b) is a lamp that has 3 different brightness levels. (c) shows a refrigerator with defrost mode [Qui09].

and connected them to electricity meters [Har89]. They were able to record the current electricity consumption up to every 5 seconds. They developed a five step procedure to detect household appliances that can be explained like this.

1. Edge Detection:
   First, the intercepted electricity consumption is stored. Second, a search in the stored data for strongly rising or strongly falling edges is performed. These edges indicate that a device may have been switched on or off at one specific 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 electricity was consumed or how much electricity was “released” from the device (e.g. when it was switched off). This causes similar events to be recognizable as clusters in the diagram. Essentially, a cluster analysis is then applied to the diagram and each found 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 electricity 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

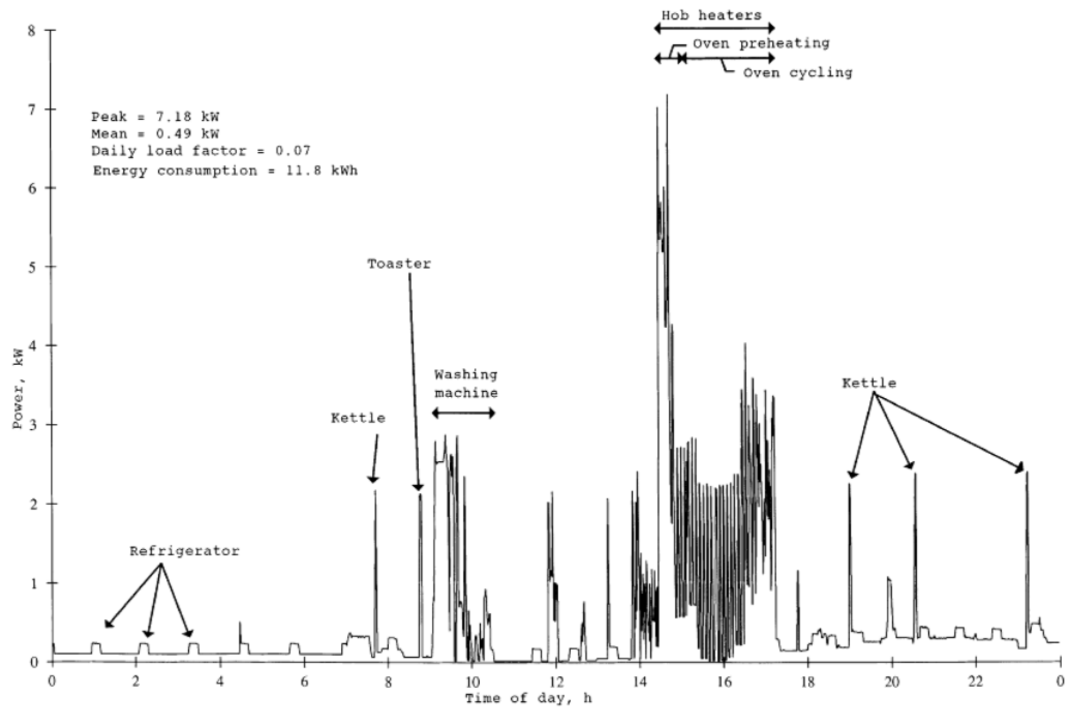


Figure 2.2: The electricity consumption of a household with drawn household appliances detected by NILM [Qui09].

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 all lamps are turned off. It can also be quickly assumed that the house inhabitants are on vacation, if the electricity consumption is lower than usual. 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 [Gre+12], the movie being watched could be determined by the electricity consumption of an LCD television. TV electricity consumption is strongly influenced by blacklighting activities. Each movie has a unique brightness signature, which was exploited to make a statement about the film being viewed.
In the paper, the electricity consumption of the TV was measured and stored every 2 seconds. After 5 minutes of analyzing 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 possible power consumption of the TV by the prediction 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 of the function 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. Later, the paper also showed that the power consumption could be used to identify which TV channel was being watched. In the figure 2.3 the result of the estimation of the prediction function is shown in green and in red the actual power consumption of a LCD TV for the same movie sequence can be seen.

Low Resolution Analysis

Another approach is to work backwards from large data sets to obtain detailed information from aggregated electricity consumption. The NILM procedure attempted to have household appliances detected from individual households and then behavioral analyses could be generated from the 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 [Qui09]. But with low resolution methods, a variety of approaches are being pursued. For example in [Pru02] an artificial neural network was trained to identify household appliances from electricity consumption. The electricity consumption was measured only every 15 minutes. The same time interval is used by Smart Meter Gateway (SMGW)s as well. With the low resolution, only the household appliances that have a high overall impact on electricity consumption, such as refrigerators, were considered. A total of 10 different household appliances were attempted to be correctly identified from the electricity consumption. With the final trained ANN, an accuracy of 90% was found.
In another work [KBN10], the electricity consumption of appliances was measured only every hour. The experiment tried to disaggregate a aggregated electricity consumption with Discriminative Sparse Coding. The experiment was conducted as follows. First, the disaggregation algorithm was trained with electricity consumption of houses from a large data set. Then, the trained algorithm was applied to the aggregated electricity consumption of two houses. The algorithm was expected to determine up to 52 individual

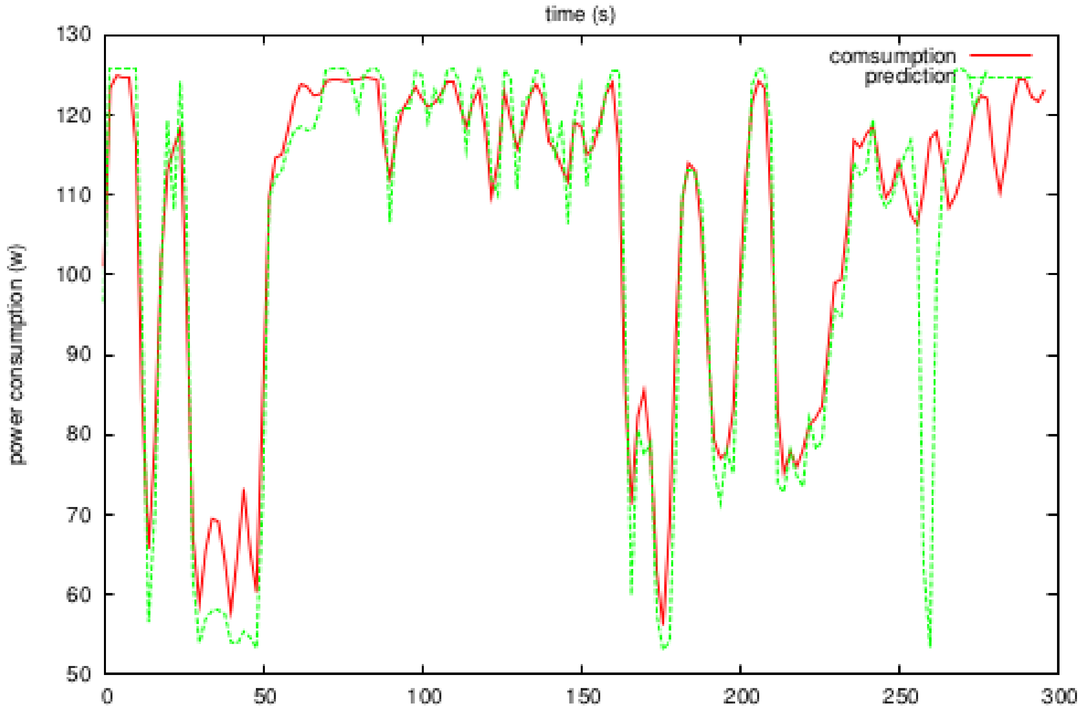


Figure 2.3: The power consumption of a TV in green vs. the estimated power consumption of the TV using the prediction function in red [Gre+12].

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% correct in its decisions.[2]

2.3 Related Work

After NILM was discovered, the scientific community started to look for possible solutions to technically prevent the analysis of electricity consumption. In this chapter, the different approaches in Metering for Operations and Metering for Billing are presented. Although in this thesis only a solution of Metering for Operations is of importance, for completeness the solutions for Metering for Billing are introduced as well.

2.3.1 Metering for Operations

The paragraph deals with solutions for Metering for Operations, which have been discussed in other scientific works so far. At Metering for Operations, there is currently no established consensus on a solution. Various technical proposals have already been

[2] Considering the fact that the time interval is so small and 52 different household appliances were considered, the accuracy of 50% is extraordinarily good. If the algorithm would only guess, the accuracy would be at 1.9%.

published in scientific papers, but there is a lack of standardized criteria and often different conditions are set for the power grid. One reason for this could be the different realization of smart grids in different countries. In the following, the various approaches are divided into categories and propounded conceptually.

Anonymization or Pseudonymization Without Aggregation

This approach describes the removal of smart information that allows identification towards the electricity provider. Identifiable information can also be replaced by pseudonyms. Solutions with trusted third parties Trusted Third Party (TTP) are often used in this case. A TTP usually acts as an intermediary between the customer and the electricity provider. The TTP must be acknowledged by all participants and take a neutral position. In practice, however, this is difficult to achieve because the TTP is often hired as a service provider by the electricity provider and is therefore also paid by the provider.
In the paper "A privacy-preserving Concept for Smart Grids" by Petrlic [Pet10], a TTP is used as an intermediary. In the procedure, a smart meter communicates with a TTP. Certificates formed with a public key infrastructure are used to verify and validate information flows from smart meters at the TTP. As soon as the TTP has checked the correctness of the smart meter information, it can pseudonomize all 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[FB14].

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 additionally 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 [BSU10].
Homomorphic encryption approaches also fall into this category. Homomorphic en-

cryption 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 is supported. These algorithms 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[Vet+12]. In this case, the TTP acts as the key authority. This means that it creates all cryptographic keys and forwards them to the smart meters, which are used for further communication with a central storage. The smart meter encrypts its data and sends it to the central storage. The central storage also saves the incoming data in encrypted form, so that no unencrypted data is stored. 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. Moreover, 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 the approach by Vetter et al. is that the different functionalities, namely storage of data and key acquisition for confidentiality and authenticity are realized from different participants.

Aggregation Without a Trusted Third Party

The solution proposed in this thesis is also one of the methods that aggregates without a TTP. The advantage of this approach is that no one has to trust a TTP. In general, one has to ask the questions who aggregates the data and who generates the keys. In addition, a common problem to consider is, how the procedure deals with a few participants.
In the solution of Mármol et al. [Már+12] again a bihomomorphic encryption method is proposed. The approach of Mármol has already been discussed and implemented in a master thesis at this chair[Bis13]. 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 electricity 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 generated key is used to encrypt the electricity consumption and the encrypted result 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 unknown. This prevents the smart meter from exposing itself during communication with the operator.

Groups are formed among smart meters and a smart meter aggregator is randomly selected in each group. 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 chance that the aggregator cooperates with the network operator. The aggregator would then be able to send individual keys 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 following round. A disadvantage of this procedure is that if a smart meter leaves the group, then a new aggregated key must be formed and the operations are quite computationally expensive.

Battery Solutions

The battery approach describes a household with a connected battery that is charged, e.g. by grid purchase or by 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.

Figure 2.4 shows the electricity consumption of a private household with a connected battery that is charged via solar panels. 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 electricity consumption that the electricity provider can see. As pictured in the figure when the battery is discharged, the electricity provider cannot detect any electricity consumption from the household. In other cases, the electricity provider notices that electricity is being consumed, but it is much less than the actual electricity consumption because the battery of the house offsets some of the electricity consumption. In other words, if a household is connected to a battery, the electricity provider cannot make correct statements about the behavior of the people in the household.

An algorithm for batteries was proposed in [MMA11]. 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 electricity consumption is below the consumption set by the algorithm, then the battery is charged with the difference from the target consumption. If the electricity 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,

Figure 2.4: The electricity consumption of a household during the day (red) with battery (green) and the electricity consumption that the electricity provider can access (blue).

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 electricity. 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 electricity consumption, so it is even more important to measure how much information can still be extracted from electricity consumption. There are the following metrics to calculate how much privacy is gained by the algorithm.

1. Relative Entropy:
   Relative entropy[3] is used to compare two sources of information. In this case it would be the electricity consumption with the algorithm and the electricity consumption without the algorithm. These two loads form a stochastic process and can then be analyzed with the calculated relative entropy [KFB11]. The relative entropy is defined as follows:

$$D(P\|Q) = KL(P,Q) = \sum_{x \in X} P(x) \cdot \log \frac{P(x)}{Q(x)} \tag{2.1}$$

[3] Relative entropy is often referred in literature as Information Gain or Kullback-Leibler divergence.

2. Cluster Classification:
   The cluster classification has already been explained for the NILM method and described in this thesis at 2.2.1. Cluster Classification is well known as a machine learning approach, but 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 [Kal+10].

3. Regression Analysis:
   In the regression analysis, first a cross-correlation and afterwards a simple linear regression is performed. More precisely, both electricity 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 [MMA11].

Drawbacks

Although various approaches to solving Metering for Operations are discussed, the following drawbacks must also be considered.

1. Pseudonymization:
   Pseudonymization does not provide protection against the attacks described in 2.2.1. The only thing pseudonyms protect is the identification of the SMGW. Once it is possible to assign the pseudonym to a customer, the pseudonym becomes invalid and the electricity consumption can be uniquely assigned to a customer.

2. Aggregation with Trusted Third Parties
   The approach of TTPs is difficult to realize in the smart grid scenario. This is because the question of the neutrality of the TTP remains open. The service of the TTP is not free and it remains unclear which entity in the system will pay for the TTP. If the customer were to bear the costs, then the electricity provider could question the neutrality of the TTP. However, it is much more likely that the electricity provider will pay for the TTPs service. In this case, the neutrality of the TTP would be open to doubt by the customer, since the TTP could be dependent on payments from the electricity provider.

3. Battery Approaches:
   Battery methods are a good way to mask actual electricity consumption. However, it is also a physical device that needs to be installed in the home and is costly at the same time. Widespread installation of batteries on a large scale is therefore unlikely in the next few years, as not every household has the funds for this investment.

4. Aggregation without Trusted Third Parties:
   A disadvantage of Aggregation without TTPs procedures is their hire complexity at a conceptual level. In addition, many Aggregation without TTP approaches involve homomorphic encryption schemes that are very computationally intensive. Nevertheless, the author believes that this category is the most appropriate for

protecting client anonymity, because there is no need to rely on TTP or invest a lot of money as in the battery approaches.

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 may 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 and 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 Trusted Platform Module (TPM) in German smart meters, the problem is considered to be solved. But 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 section2.3.1. 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 is paying the TTP? In this case, the TTP provides a service to the electricity provider. However, if the electricity provider pays for the service, then the TTP 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 on 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 after the installation. A secured smart meter can therefore perform correct billings at the customers side and the electricity provider can trust the smart meter, although it is located at the customer's home. 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 [FB14].

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. 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 obligation:

$$c = (x, r) \tag{2.2}$$

Here x would be the measurement and r would be a random number. If one wants to check the commitment for correctness, then there is a validation function:

$$Open(c, x, r) \tag{2.3}$$

2.3 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 $x \neq x'$ with an r' such that $open(c, x', r')$ returns True. In this use case, the following procedure is often used.

$$Commit(x, r) \cdot Commit(y, s) = Commit(x + y, r + s) \tag{2.4}$$

$$Commit(x, r)^k = Commit(x \cdot k, r \cdot k) \tag{2.5}$$

The special feature of the method of [Ped91] 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. Using 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 the exact electricity consumption and the according price. 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 of 2.3 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.

Figure 2.5: All documents from BSI that are subdivided under the Technical Guideline 03109 [Bis13].

2.4 Technical guideline TR-03109

This paragraph will discuss the technical guideline TR-03109 published by the BSI. The BSI is the entity of the German federal government that deals with digital security issues and releases 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 like smart meters. The 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 six different documents, which are shown in Figure 2.5. Based on the guidelines, it is possible to have smart meter devices certified by test centers. Unless otherwise described, all information are derived from the most recent TR-03109 [Sic21].

Stakeholder Model in the German Smart Grid

The TR-03109 describes all stakeholders that participate in a power system. The most important stakeholders for this work are described.

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 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 External market participants (EMT) are all other authorized participants in the energy network that can establish a communication connection with the SMGW. These include electricity providers. The SMGW ignores all other com-

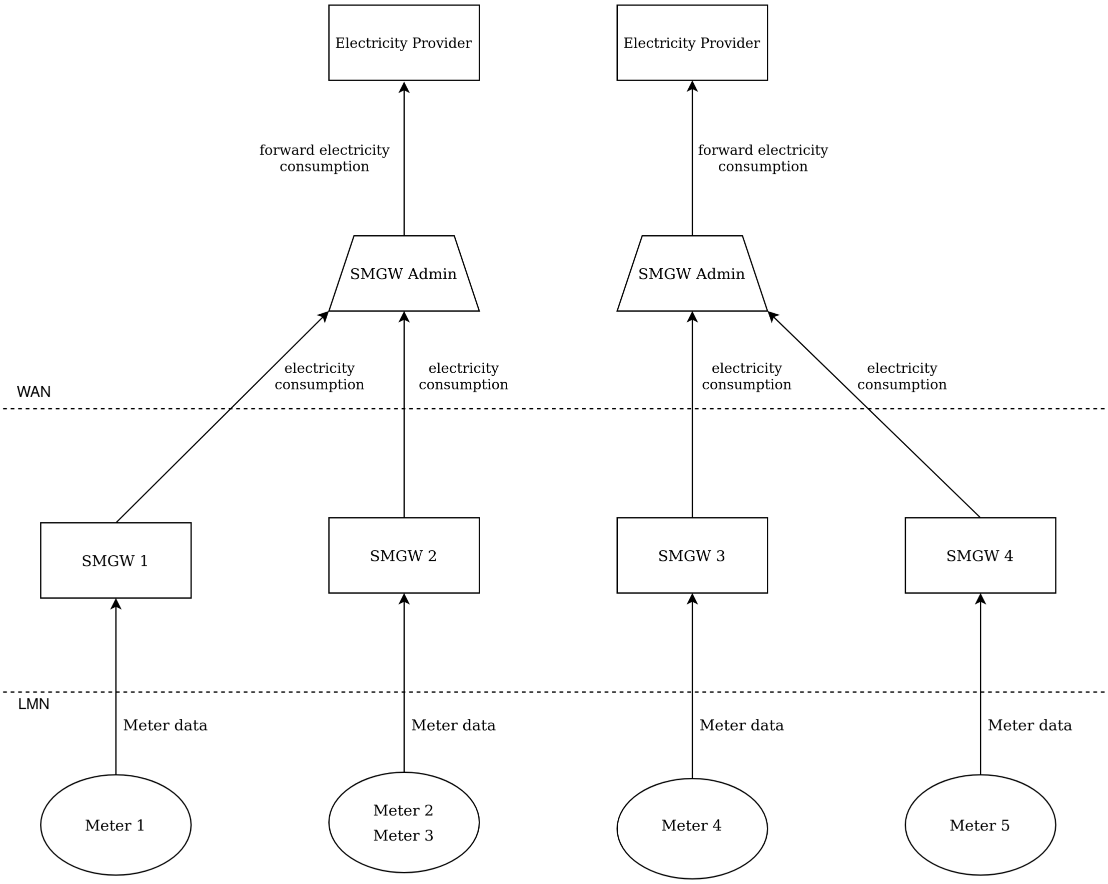


Figure 2.6: An exemplary electricity grid in the TG-03109 with the stakeholders.

munication 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. In Figure 2.6 is a exemplary power grid shown with the Stakeholder in the Smart Grid. Every SMGW is connected to one GWA, which configurates the communication profile of the SMGW. When pseudonymization is activated, the information is not sent directly to the electricity provider, but first to the GWA. Afterwards, the GWA forwards the data to its electricity provider.

2.4.1 Interfaces and functions of the Smart Meter Gateway

A smart meter or as described in the TR-03109 a smart meter gateway SMGW must provide three different physical interfaces as shown in 2.7.

1. Local Metrological Network 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

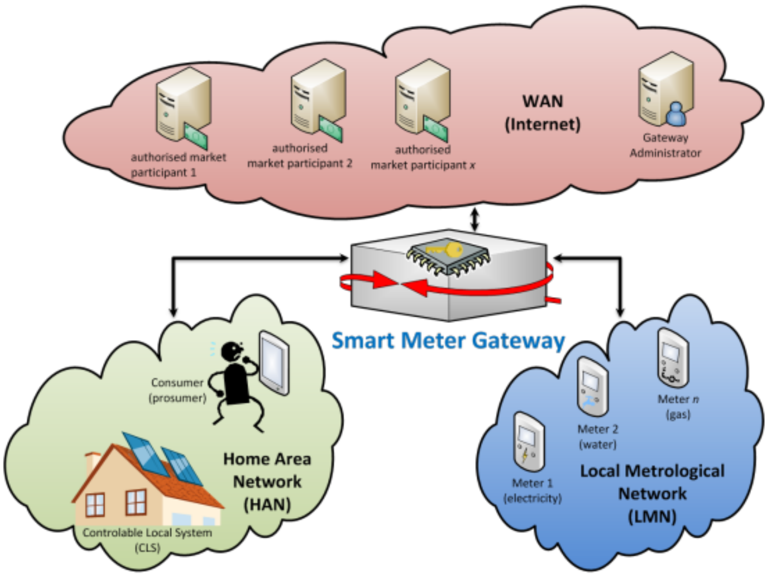


Figure 2.7: An overview of all interfaces and possible stakeholders that can communicate with SMGW [Bis13].

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 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 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 three interfaces. As a result it is 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. The processing of data from the SMGW is shown in Fig 2.8. Since the WAN interface is the most important interface for this work, it will be

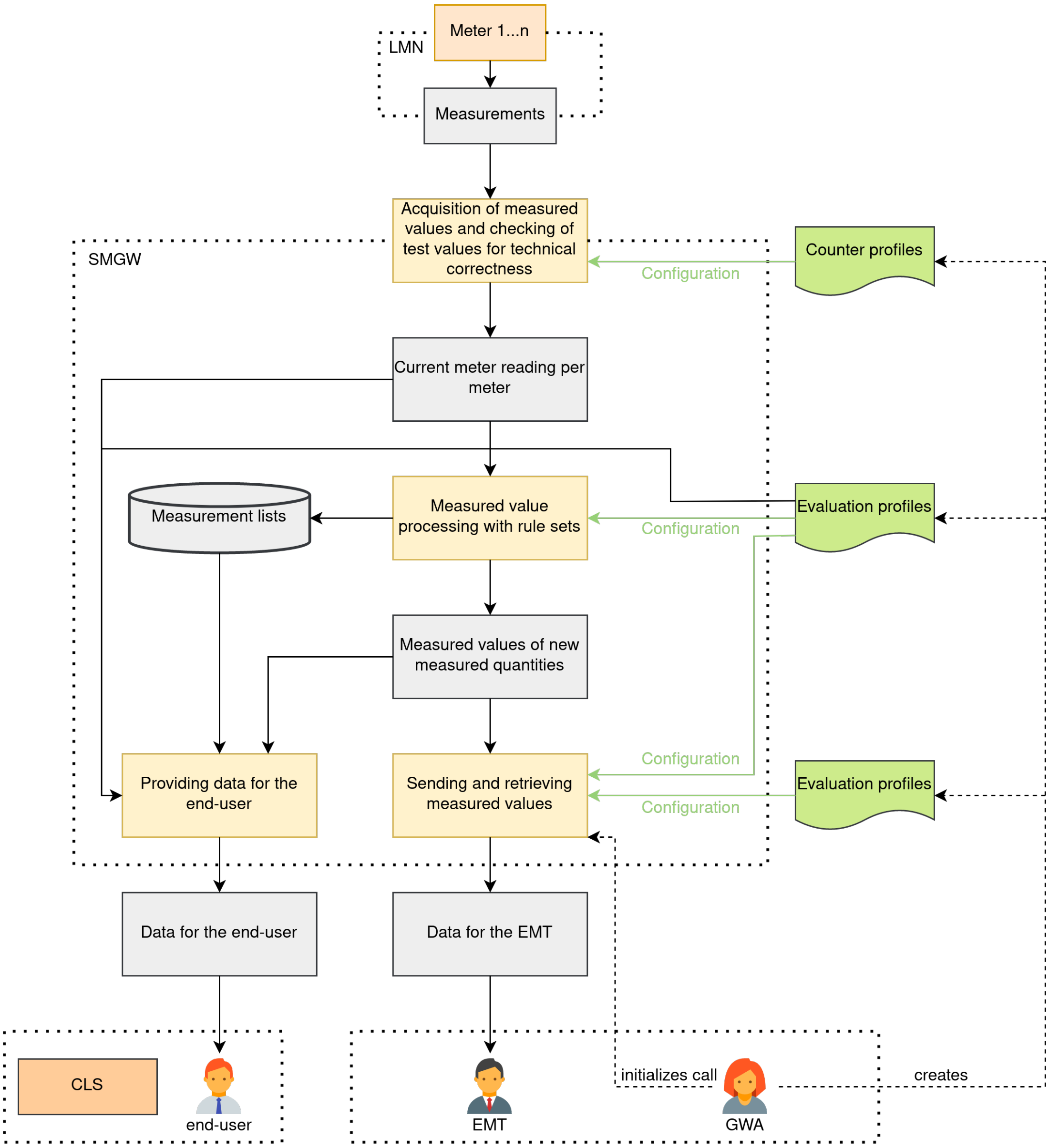


Figure 2.8: An overview of the measured value processing within the SMGW with configuration profiles from the GWA.

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 2.2.1 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 TR-03109 is sufficient.

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

Stakeholder Motives

It has already been explained which participants in the power grid interact with each other. Below the particular motives of the participants are discussed and which malicious intentions can be pursued by the participants.

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 [LeM+07].

Electricity Provider's Motives

For the electricity provider, the authenticity of the billing is the most important security objective because from its point of view, the customer is not trustworthy in the calculation of the bill. In addition, the customer has access 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 2.2.1 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 [Qui09] 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 typical electricity provider is considered an honest-but-curious adversary. But in the proposed network protocol in Chapter 3, the electricity provider gets administrative rights to maintain the network. For this reason, the electricity provider is always assumed to be malicious to be able to protect the network from the potential strongest attacker. But in the proposed protocol in Chapter 3, the electricity provider partly performs administrative tasks for the network protocol. 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 [LeM+07].

2.4.2 Security Objectives

The smart meter attempts to achieve the three security objectives of confidentiality, integrity and availability. The three security goals are often summarized as CIA. Another important security goal for this work is anonymity. The four 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.

In the application field of this thesis, anonymity is considered as follows:

4. Anonymity:
   A user is anonymous if the behavior of the user sent in a message to its recipient remains secret.

In the 3 chapter, the design of the network protocol is presented, which fulfills the mentioned security objectives and thus protects the privacy of the customer from the electricity provider or an attacker. For an attacker his main goal is to try to bypass the security objectives to obtain additional information about the customer.

2.4.3 Attacks on the Smart Grid

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 all sorts of attackers. The two most common attacks that are relevant to the work are explained below.

Eavesdropping

Eavesdropping may be the weakest type of attack and it is often used by a passive attacker or by an active attacker as a preparation for a larger attack. 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 [LeM+07].

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. 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 network is proposed and how the DC network could be implemented in the TR-03109 of the BSI. It also describes which attacks on the DC network are possible with the defined attacker model.

3 A Privacy-Preserving Aggregation Scheme Using DC-Nets

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 TTP 2.3.1. First, there will be a theoretical introduction to DC networks and an example to underline the theoretical concept. Subsequently, the security protocol is introduced and the functionality of the protocol is described. This includes all security mechanisms and all functions that are important for the correct execution of the protocol. In addition, mechanisms are provided to ensure the stability of the protocol in case of errors. In order to show the correctness of the protocol, the implementation of the protocol in 3.4 is presented and then the protocol is evaluated.

3.1 A Privacy-Preserving Aggregation Scheme Using DC-Nets

In [Cha88], David Chaum proposes a ”round-based” 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 section.

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 regarding the addition (modulo alphabet size). The addition of the inverse element is understood as the subtraction of a character. Because of the abelian group, simple mathematical operations like addition can be performed on the numerated letters in the alphabet. If individual letters can be added up, then words will be also adding up as long as the words have the same length. For this reason, it is assumed that in a DC network the messages have the same length.
Messages in the DC network are encrypted with a secret key to prevent a message from being traced by any participant. The principle is the same for adding up letters. For each key character, a message character is added up so that the content of the message cannot be read from the result. A generated key is shared with exactly one participant in the DC network, so that two participants use the same key to encrypt their message. These two participants are also called neighbors. For new neighbors, it is agreed that one neighbor subtracts the key and the other neighbor adds up the key. It must be considered that if there is only one neighbor in the DC network, then the neighbors can read each other’s messages and no anonymity is achieved. This is because both neighbors

know the key being used to disguise the messages. For this reason, a participant in a DC network exchanges a key with several neighbors. This means that nobody is able to read its neighbors messages in a DC network, because the key from the other neighbor is not known to outsiders. Thus, for each round depending on the arrangement, a message is added or subtracted with the exchanged keys by the partners. 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. The result is called global superposition. The global sum consists of all messages from the participants and all keys that the participants have calculated on their messages. Since each key in the global superposition has been added and subtracted exactly once[4], the keys in the global superposition cancel out and only the meaningful messages remain. After a global superposition is distributed a round is over and a next round starts in which messages can be sent. If a participant does not want to send a meaningful message, the participant will send an empty message. This message consists only of zeros and is superposed with the key. The 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 for the round. If one of all participants has sent a meaningful message, the global superposition is the message for the round. If more than one participant sent a meaningful message, then the result is the total of all sent messages and a single message from the superposition cannot be recovered. The last case is also called a collision. In order to solve this problem, the collision resolution algorithm with averaging can be used in a DC network. But in this proposed protocol it is mandatory that all users send meaningful messages at the same time and that all messages form collisions with each other.

Key Exchange in DC Networks

Exchanging keys to calculate the local superposition 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 Pseudo-Random Number Generator (PRNG) 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. via a cryptographic key exchange procedure. Due to the deterministic property of PRNGs, the same number is always generated from an initial value. The result of the PRNG can be used again as a seed to produce another pseudo-random number and so forth. Therefore, for each message round in a DC network a PRNG can produce a random number after only exchanging one initial value. The random numbers are used as keys for the local superposition. 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

[4] As a reminder, each key is added and subtracted once precisely, because two neighbors share a key and have agreed that one neighbor will add the key and the other neighbor will subtract the key.

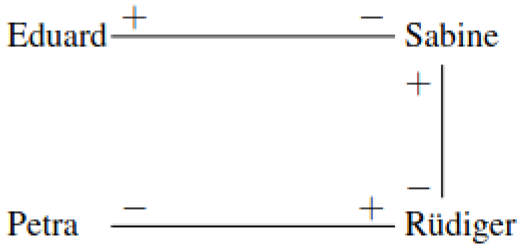


Figure 3.1: A basic example of a DC network key graph[Dat22].

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 3.1 using a simple example. This visualization is also called the key graph of a DC network. A key graph is the underlying graph that is created when two participants in the DC network exchange data with each other. The users form the nodes and when two users have exchanged initial values for the PRNG, they form an edge in the graph. The example DC network shows four participants that are connected to each other along a communication link. The outer participants have only one partner, the inner participants are connected to two partners. Each participant has exchanged keys with its direct partners. The outer partners needed to exchange only one key and the inner ones exchanged 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 3.1, the user Petra would subtract the exchanged key with Rüdiger from the message she wants to send. The result is the superposition of Petra. Rüdiger 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. After everyone in the network has calculated their local superposition, all local superpositions are distributed to every user. Subsequently, every participant in the DC network can calculate the global superposition from all local superpositions [Dat22][Pfi06].

3.2 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. Since the aim is to anonymize the electricity consumption of a customer 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 does not have to be distributed in the network but only has to be calculated at the electricity provider. The only

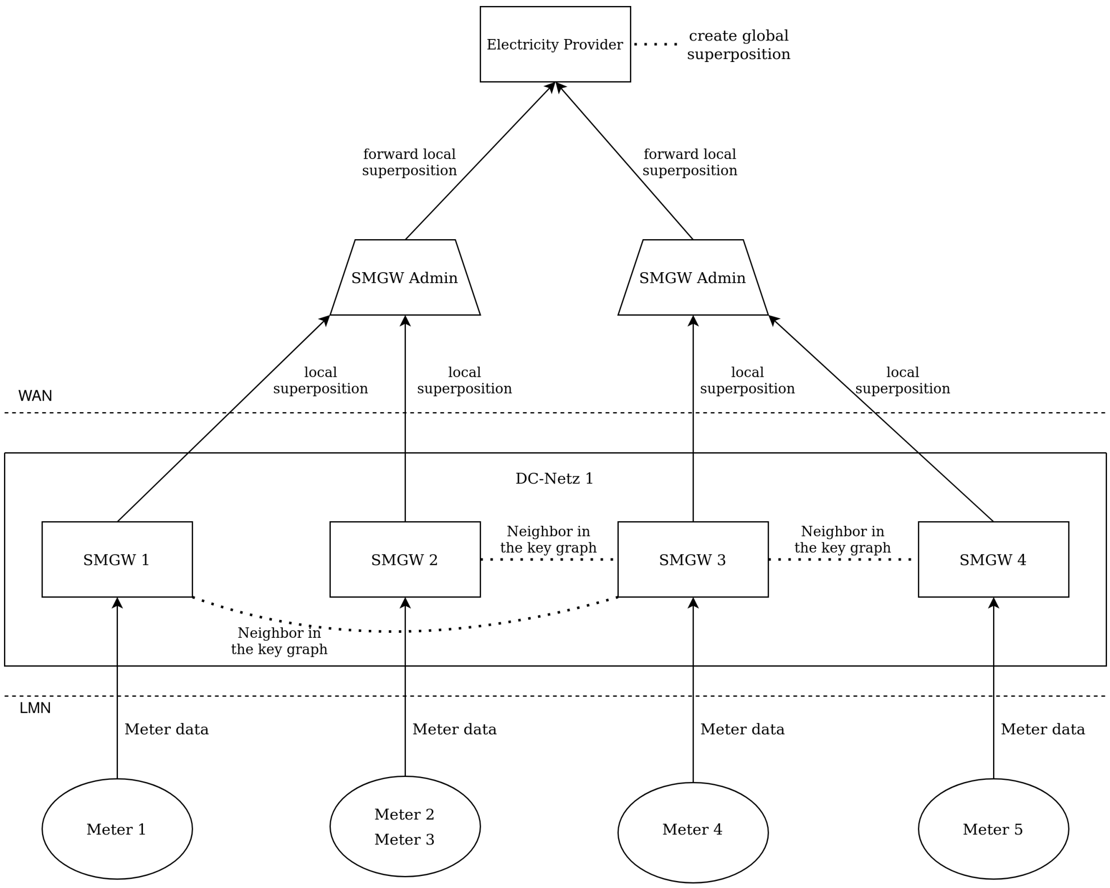


Figure 3.2: In the figure, the seven components of the DC network frame are are shown.

exception is when joining or exiting the DC network. During the registration process the SMGWs have to perform a key exchange to configure the initial value for the PRNG as explained in 3.1.1. Before introducing the core functions of the network protocol, it is described how the protocol can be arranged with the TR-03109 of the BSI [Sic21].

Compatibility with the TR-03109

The DC network must be compatible with the installed infrastructure. Otherwise, the design would not be implementable and impractical for real-world use. Therefore, a large part of the requirements from the TR-03109 were taken into account in the creation phase of the design.
In Figure 3.2, it can be seen how the DC network is integrated into the stakeholder model of the TR-03109. In addition, the interface boundaries of LMN and WAN for the different stakeholders are drawn. The design shows that the DC network protocol needs to be implemented only on the SMGWs. No structural changes need to be made. According to the TR-03109 from the BSI, SMGWs are only allowed to communicate with authorized participants in the smart grid and all foreign requests are ignored. These

Protocol Identifier	DC Network Identifier	Client Identifier	Transmission Bit	Timestamp	Notifications	Data

Figure 3.3: In the figure, the seven components of the DC network data package are are shown.

are EMTs, GWAs and the electricity provider. But it is not possible for two SMGWs to establish a communication link with each other, because SMGWs are not authorized participants in the power grid. Therefore, it is necessary for electricity provider to take over administrative tasks, e.g. to provide necessary communication tunnels for SMGWs to each other. Additionally, the electricity provider is storing the key graph so that possible errors on the DC network can be corrected and assure stability.
A message round in the stakeholder model in combination with the DC network looks as follows. As a comparison sending the electricity consumption without DC network was described in 2.4. The SMGW receives the electricity consumption data from the meters of the different houses. The SMGW implements the communication profiles configured by the GWA. Despite the implementation of the DC network, there is no impact on the processing of the data in the SMGW, since the DC network is performed only after the application of the communication profiles as it can be seen in A.1. The creation of the local superposition is the last processing point before sending the DC network protocol data to the electricity provider. The local superposition is then forwarded to the electricity provider via the GWA. Although the GWA is between the electricity provider and the SMGW during the transmission of the data, the GWA only forwards the data and does not perform any operations or process the protocol data. Therefore, the GWA will not be considered further in the protocol below.
The electricity provider then calculates the global superposition from all local superpositions. In the use case of the thesis the meaningful message contained in the local superposition is the individual electricity consumption of a customer. Therefore, the calculated global superposition is the total electricity consumption of all users in the DC network. In addition, SMGWs send their electricity consumption to the electricity provider every 15 minutes. A round-based network for anonymizing content data like the DC network is particularly suitable for achieving privacy in a smart grid with high transmission intervals.

Protocol Header

In the network protocol messages are sent via data packages. The structure of a package in the DC network protocol is shown in the figure 3.3. Each package 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.
The first field is the Protocol Identifier field. The purpose of the Protocol Identifier field is to ensure that the application of the DC network protocol is recognized by all participants and that the SMGWs as well as the electricity provider can react correctly to the packages. 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 in a network can be distinguished. If the network needs to perform an error correction, SMGWs can be notified by the identification number from the electricity provider. 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 local superposition to the electricity provider in a round. Furthermore, it is used for error correction procedures. The Time Stamp field indicates when a message was sent by a participant. This allows the electricity provider to classify messages by round and not mix up for messages from different rounds. The next 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 electricity 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 new empty DC network. In order to achieve this a few preparations have to be made. First, a unique and unchangeable DC Net Identifier is assigned from the electricity provider to the empty DC network. At least two SMGWs have to enter the DC network. A DC network with only one participant is not operational and cannot offer anonymity. The SMGWs that enter the network are assigned a Client Identifier 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 users. 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 can lead 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 4 to determine a minimum number of participants. It is assumed that in the DC network each participants has at least three connections to neighbors. This reduces the risk of individual SMGWs being disconnected from the DC network or malicious neighbors being able to reconstruct the electricity consumption from the local superposition.

In order for the DC network to become operational, two SMGW must exchange an initial seed value to configure the PRNGs. As a result, the same random number sequences would be generated independently of each other by the PRNG on both clients. But there is a current 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 electricity provider must provide a tunnel for the SMGWs so

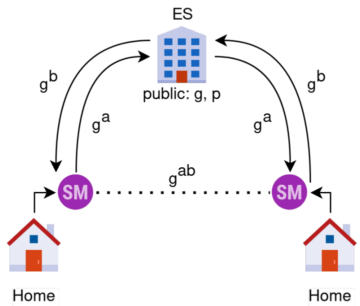


Figure 3.4: The Diffie-Hellman key exchange between two SMGWs via a electricity supplier Electricity Supplier (ES).

that the SMGWs can communicate with each other and exchange the start seed for the PRNG. The tunnel works as follows. An SMGW receives the tunnel information by the electricity provider. The SMGW can then send messages to the tunnel of the electricity provider. When messages arrive at the tunnel, the electricity provider forwards the messages to the receiving SMGW. Of course, the electricity provider has the technical possibility to eavesdrop on the tunnel. For this reason, secret information such as the start value of the PRNG must be exchanged in a cryptographically secured procedure via the tunnel. One method to securely exchange the start value would be the Diffie Hellman key exchange.

Diffie-Hellman is a known key exchange protocol, where two users can publicly exchange a secret over a unsecure channel without a third person being able to figure out the secret. Diffie-Hellman requires a generator g and a prime p, and the two values are public to all users. If two users want to exchange a key via Diffie-Hellman, then both users choose a secret random number a that is between 1 and $p-1$. Each user calculates the public key ka by $ka = g^a$. The public key is distributed to the partner and at the same time the partner's public key $kb = g^b$ is obtained, where b is the random number between 1 and $p-1$ of the partner. After obtaining the public key, both users can calculate $kab = g^{ab}$.Even if an attacker could eavesdrop on the values g^a and g^b, he would not be able to compute g^ab because the computation of the discrete logarithm problem cannot be performed efficiently. How the Diffie-Hellman key exchange is executed in the smart grid use case is shown in Figure 3.4. Both SMGWs receive their public generator and prime number from the electricity provider. Then, g^a and g^b are calculated and sent to the partner via the electricity provider's provisioned tunnel. Afterwards, g^ab can be calculated by both SMGW.

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 executed for the SMGW without any problems. Diffie-Hellman was also only mentioned as an example. The forwarding of SMGW messages by the electricity provider enables the implementation of other secure key exchange procedures as well. 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. Uninvolved SMGWs do not receive any information about the entry of new

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	Electricity provider informs SMGW that its neighbor is leaving the DC network
Notification 4	Resent local superposition according to the correction procedure
Notification 5	Transmitting local superposition
Notification 6	Former defective DC client is rejoining the DC network
Notification 7	An SMGW informs the electricity provider that it is leaving the DC network

Table 3.1: An overview of all notification messages.

users in a DC network. Furthermore, a SMGW can generally obtain little information about other participants in the DC network since it can not communicate directly to other SMGWs. But the participants share their Client Identifiers during the key exchange in the registration process. Due to the exchanged communication details, each participant in the DC network knows the identification number of its neighbor. This is helpful later for error correction measures.

The use of a key exchange method also involves risks. 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, a malicious electricity provider can easily perform an active attack on single SMGWs with the knowledge of the structure of the key graph. An example would be that if a electricity provider wants to get information about the electricity consumption of a 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. To counter the introduced attack, participants must have a minimum number of neighbors. The more neighbors a SMGW has, the smaller the chance that each neighbor is a malicious attacker. Moreover, the electricity provider would have too much electricity 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.

Bootstrap Phase in a DC Network

A SMGW that wants to register in the DC network sends a special defined request to its electricity 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 DC 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 SMGWs to create a communication link via the electricity provider. If an SMGW sends to the tunnel, the message is forwarded to

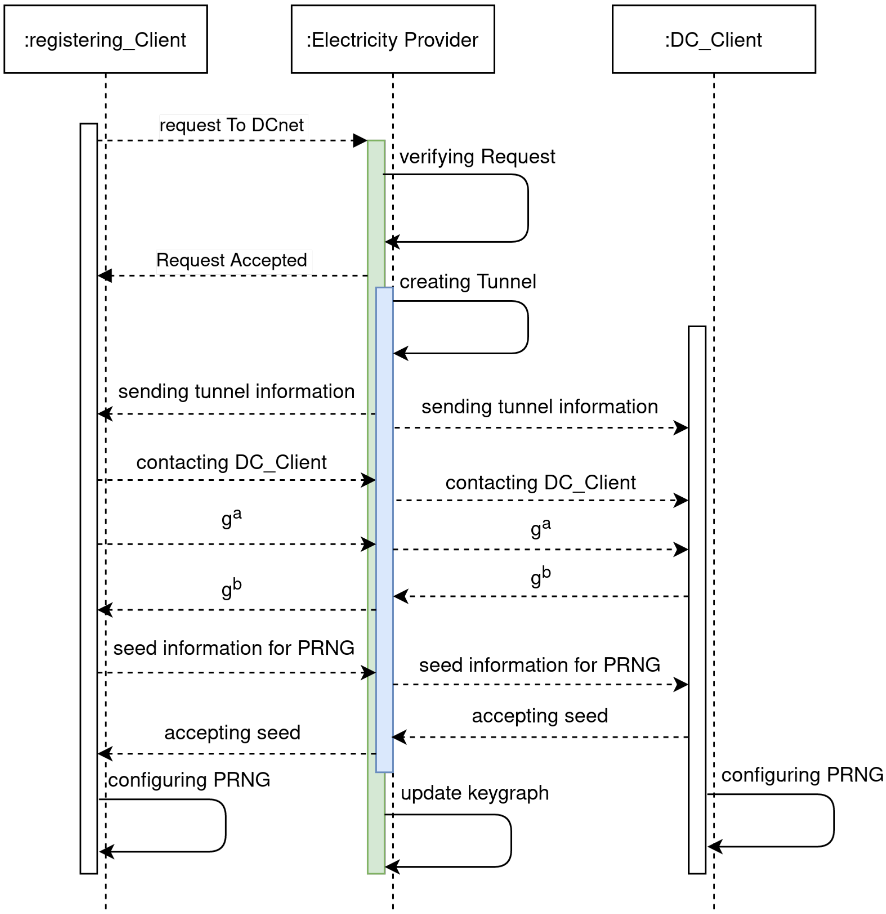


Figure 3.5: A sequence diagram showing the functions for registering an SMGW in a DC network.

its future neighbor. The 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 for the PRNG over the tunnel. To prevent the electricity provider from reading the seeds, the clients swap the seed secured by the Diffie-Helmann key exchange through the tunnels provided.
Once the seeds have been exchanged, the electricity provider is informed by the requesting client that it has successfully entered the DC network by notification 2. The PRNGs are now configured and the generated keys can be added or subtracted to the message in order to create the local superposition. The procedure is explained in more detail in paragraph 3.1.1.The communication exchange between all participants in the registration process of a DC network is graphically illustrated in Fig 3.5. Afterwards all participants in the DC network can send their local superpositions to the electricity provider. The provider forms the global superposition and receives the aggregated electricity consumption of all SMGWs in the DC network. If necessary, the electricity provider must ensure that in the future messages can continue to be

exchanged between registered customers via the same tunnel. In addition, already registered clients can not choose which new communication partners they get. Furthermore, it is assumed that the electricity provider has already been authorized by the GWA. Otherwise, the SMGW would not be able to establish a connection to the provider.

SMGW Regular 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 transmission interval for the electricity consumption is 15 minutes. Hence, for the DC network, it is most practical if all SMGWs send their local superposition to the electricity provider at the same time or within a short transmission interval (e.g. one minute). This can be done without problems, because according to [Sic21] 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 packet that a SMGW sends to the electricity 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 one bit and is set to one 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 Identifier of the package is stored in a memory structure. The memory structure shows, which client has not sent a local superposition in the round.

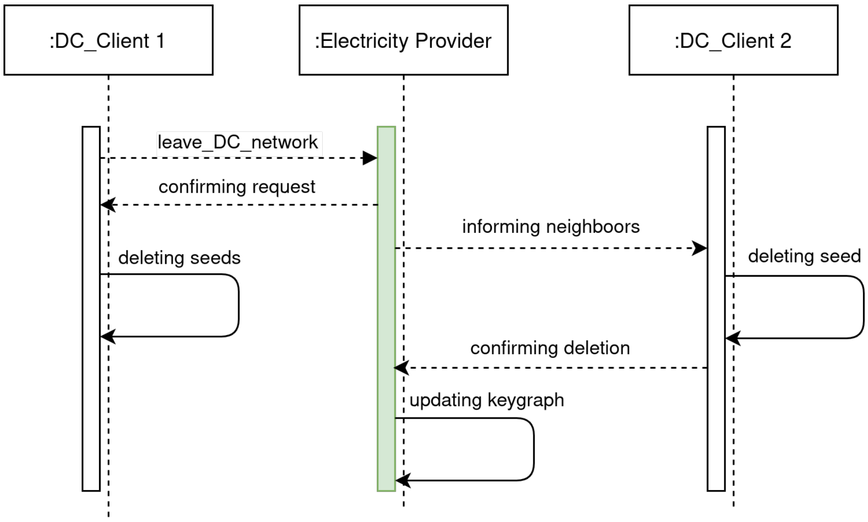


Figure 3.6: A sequence diagram showing the functions for a SMGW, which wants to exit a DC network.

3. Transmission Bit: Each message has a transmission bit set to one. All transmission bits are added up and at the end of the round it can be checked whether all SMGWs have sent their local superposition. If the summed transmission bits do not correspond to the number of participants in the DC network, correction procedures must be applied.

4. Time Stamp: The electricity provider can assign the message to the correct round.

5. Notification: Notification message 5 informs the electricity provider 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 electricity consumption of all SMGWs in the DC network.

Teardown Phase in the DC Network

An exit can be caused, for example, when the customer changes the electricity provider. Then a notification message 7 is sent from a SMGW to the electricity provider, which informs the electricity provider about the exit of the SMGW. The electricity provider instructs the neighbors of the exiting client with a notification message 3. The electricity provider may not abuse Notification 3. Otherwise, a malicious electricity provider would be able to change the structure of the DC network at will. Hence, the misuse must be prevented technically, since the exit from the DC network is one of the standard functions and it is difficult per design to prevent such a misuse. 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 electricity provider that the configuration has 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 superposition.
Furthermore, the key graph must be considered. It could happen that the underlying key graph splits into two DC networks. In this case, two separated DC networks are sending to the same DC network identifier. In the example of Figure 3.1, the scenario could occur 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 electricity 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 three 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. In the sequence diagram in figure 3.6 it is shown which communication exchange is performed so that an SMGW can leave the DC network.

3.3 Error Correction Procedures

It was described how DC networks work and how a DC network works in a normal operation without interference. This section explains which attacks on the network are possible and how the DC network deals with potential disturbances.

SMGW Connection Loss

SMGW can access the Internet by communicating over their WAN interface. 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 superposition 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 electricity

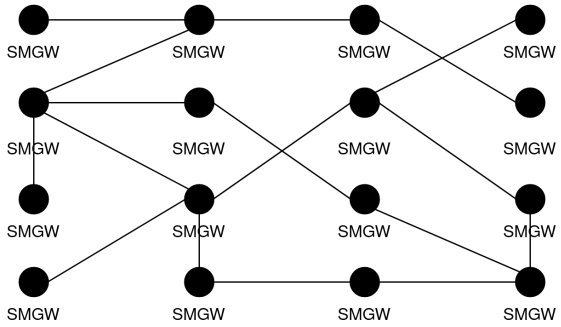


Figure 3.7: An example of a DC network before it is split in the algorithm. The edges between the nodes represent a PRNG configuration.

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 updated 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 rejoins in the DC network by sending a Notification 6 to the electricity provider. Thereupon the electricity provider informs the neighbors that in the next round the key of the former faulty neighbor is to be reused for the calculation of the local superposition. The electricity consumption of the SMGW during the time of Internet loss is not retransmitted. This is because retransmission of the electricity 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. Furthermore, the electricity provider knows how many smart meters are functional at any time and how many are defective through the transmission bit. Therefore, the electricity provider can evaluate how conclusive the data is in the DC network. If not a larger amount of participants of the DC network fails, the electricity provider is still able to ensure a good network stability. In addition, an Internet loss does not have an impact on the billing of a SMGW, because the calculation of the billing is executed via another procedure 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 its 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 electricity provider to calculate a meaningful global superposition. Hence, it is not possible to see the aggregated electricity consumption for the whole DC network. 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 Pfitzmann's error localization and recovery protocol can be applied [Pfi06]. The protocol of Pfitzmann describes two 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 superposition and no SMGW is having Internet issues, this is communicated publicly by the electricity provider to the SMGWs. All SMGWs then save the keys and the electricity consumption from the last round. At the same time the electricity provider saves all receiving local superpositions from the round and enjoys special rights that only prevail in F-mode. The figure 3.7 shows an example DC network to illustrate how the proposed algorithm works.
In the following, an algorithm exploits a property of DC networks that allows a DC network with one meaningful global superposition to separate into two DC networks with two separate meaningful global superpositions. After a round is manipulated by an active attacker, the DC network is halved and the failed round is repeated with the same configurations and electricity consumption. Since the hole round is repeated every participant including the attacker has to resend the same local superposition from the failed round. Since there are now two DC networks, a global superposition will be meaningful and the other global superposition will be unreadable. This allows the electricity provider to narrow down the attacker, since the attacker will always be located in the DC network where the calculation of a meaningful global superposition fails.
The algorithm is executed in four steps:

1. Halve the key graph. The electricity provider has the overview of the key graph and can therefore separate the key graph into two parts.

2. 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 border node by the electricity 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 electricity provider can request an SMGW to throw away a key only in F-mode, because throwing away a key of a neighbor immediately weakens the anonymity of an SMGW.

3. All SMGWs in one half of the key graph now retransmit the local superposition from the last failed round.

The border SMGWs that threw away a key, calculate the new correct local superposition without the key of the neighbor in the other part of the key graph. This procedure results in all nodes sending the same message from the last round except the border nodes and allows two global superposition to be calculated. One global superposition from the first half and one global superposition from the second half of the DC network. So the old DC network 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 superposition if the same local superposition is really resent and can check the correctness in a resent round.

4. If the calculation of a global superposition fails in the first half of the DC network, then the attacker will be located in the first half and the first half is separated again in two halves. If the calculation of a global superposition fails in the second half of the DC network, then the attacker will be located in the second half and the second half is separated again in two halves. If the global superposition fails in both halves or is calculated correctly, then the attacker will be among the border nodes.

The procedure is continued recursively until it is reduced to one SMGW that is eligible to be the attacker. The figure 3.8 shows the first step of the DC network splitting algorithm. There can be the special case that the attacker SMGW is a border node. Since the border nodes send a recalculated local superposition, the electricity provider cannot immediately rule out whether the attacker is among the border nodes. Therefore, for each bisection, an additional subgraph must be formed in which the former border nodes have no neighbors outside the subgraph. In this way it is possible for the electricity provider to control the local superposition when resending the local superposition of the former border nodes.
Since the electricity provider is granted extended rights in F-mode, it must be ensured that the electricity provider does not abuse F-mode by, for example, running the DC network in F-mode all the time. To prevent misuse, all SMGWs in the DC network must be informed at all times as to which mode the DC network is in. If the DC network is conspicuously often in F-mode, this could be an indication that the electricity provider is abusing its rights. In this case, the SMGWs have the option to leave the DC network and terminate the contract with the electricity provider.

DC Network Size

With a small number of participants, conclusions can be drawn about individual participants from the aggregated result. This is the case if the electricity consumption of one user is equal in percentage to the residual consumption of the other users.

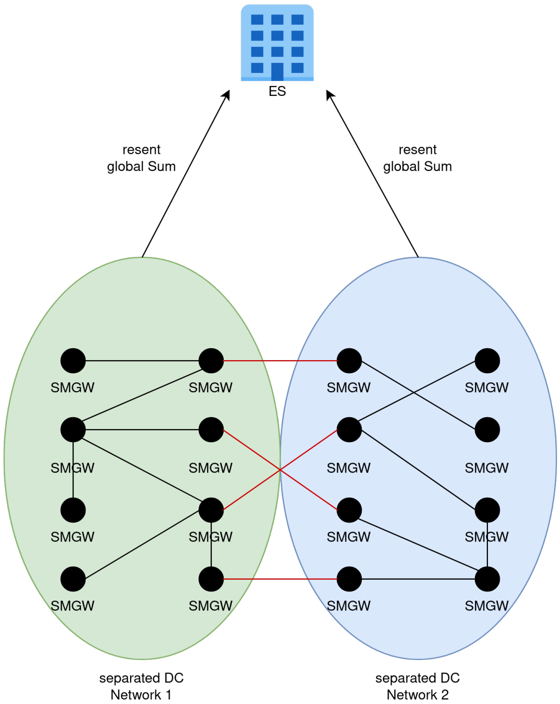


Figure 3.8: The result of the first division of the DC network in F-mode. Red edges are division edges removed from the original network.

Or it is also feasible that a user does not consume any electricity. In a DC network with two users, the electricity 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 [Le 20]. 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 electricity provider 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 participants, the question arises as to how meaningful the results are when hundreds of participants partially fail.

3.4 Implementation

The conceptual approach of the DC network was presented above. 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 section deals with the implementation of the introduced DC network protocol. It is described which technical tools were used and where implementation problems occurred.

3.4.1 Structure of the Testbed

In the practical implementation in this work it was tried to implement a DC network with the same requirements as defined in the section above. But for technical reasons, the exact same structure could not be implemented due to the framework used. If there are any deviations from the defined protocol, then these will be described and explained in this chapter[5]. Four Raspberry Pis are used to realize the testbed of the design, where three Raspberry Pis simulate the SMGW and one Raspberry Pi represents the electricity provider. In the following, the Raspberry Pis that represent the SMGWs are called clients and the Raspberry Pi that represents the electricity provider is simply referred to as the electricity provider. All clients have a communication link via Local Area Network (LAN) to the electricity 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 electricity provider. After the clients join the DC network, the clients build their local superposition and send it periodically to the electricity provider. The electricity provider adds up the local superposition and stores the global superposition 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 for a local superposition is 10 seconds. In addition, the demo was implemented in such a way that after four 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 hole underlying network protocol, gRPC was used as a framework.

gRPC Remote Procedure Calls - gRPC

gRPC [Aut] is an open source remote procedure call 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, the functions are implemented on the client, while the server runs the interface and processes the client requests. On the client-side is a stub that holds the same functions that are on the server. In gRPC server and client can communicate with each other even they were implemented in different programming languages. In this work, both server and client were implemented in Python. A simple application example of gRPC is shown in Figure 3.9.

Protocol Buffers in gRPC

Protocol Buffers are used by default in gRPC and allow structured data to be serialized [Goo22]. The structured data is specified via a message format in the Protocol Buffer file. Afterwards the basic source code of the protocol is automatically generated from the defined message format by executing a Protocol Buffer method. Subsequently, data can be sent from client to server through channels provided by gRPC. Listing 3.1

[5] The source code for the implementation can be found in [Gar22]

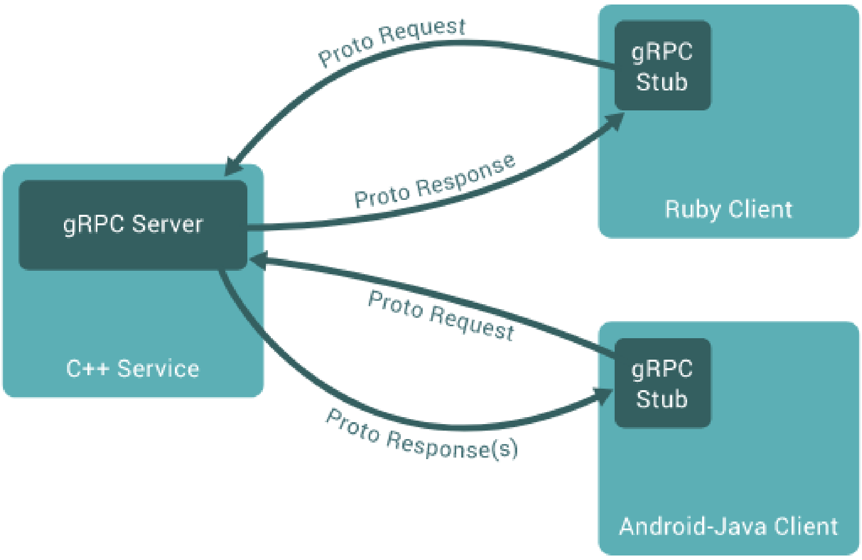


Figure 3.9: An overview of the structure between client and server in gRPC.

shows the implementation of the protocol header by a Protocol Buffer message format. Each header field is assigned a data type and also a unique field number that determines the order of the fields.

```
syntax = "proto2";
package DCnetPackage;

  message DC_net {
    optional int32 dc_net_identifier = 1;
    optional int32 client_identifier = 2;
    optional int32 transmissionBit = 3;
    optional string timestamp = 4;
    optional int32 notification = 5;
    optional int32 localSum = 6;
  }
```

Listing 3.1: The implementation of the DC network protocol header in a Protocol Buffer message format.

3.4.2 Server Implementation

The server implements all functions that are defined as a service in the Protocol Buffer file. 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 3.2 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 accepts a message format. For example, the addClientToDCnet function uses the message format defined in Listing 3.1. Afterwards, the functionality of addClienttoDCnet is implemented using gRPC framework.

```
syntax = "proto2";
package DCnetPackage;

  service DC_round {
    rpc SendLocalSum(DC_net) returns (Acknowlegde)  {}

    rpc addClientToDCnet(DC_net) returns (DC_net) {}

    rpc connectDCClients(DC_net) returns (DC_net) {}

    rpc ExchangeOpenKeysForDH(Secret) returns (Secret) {}

    rpc getDiffieHellman(Empty) returns (DiffieHelman) {}

    rpc ExchangePRNGSeed(Seed) returns (Seed){}

    rpc deleteClient(DC_net) returns (Acknowlegde){}

    rpc sync(TimeStamp) returns (Acknowlegde){}

    rpc updateGlobalSum(DC_net) returns (Acknowlegde){}
    }
```

Listing 3.2: All server functions provided by the DC network service to the client.

3.4.3 Client Implementation

When the in client application starts, a fork function is executed on client-side. The fork function allows a 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 3.10 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 to a random neighbor. In order for the neighbor and the registering client to synchronize the configuration of 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 to the server 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 superposition and send its local superposition to the server for each round.

Generation of the global Superposition

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

Preparations for the next Round

After each sending of the local superposition, 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. Subsequently, 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 three neighbors in the protocol. The flow chart in Figure 3.10 illustrates the structure of the client implementation with the key functions.

3.4.4 Challenges

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

Multiple Client access on the Server

In the experimental environment, three 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 thus particularly difficult in the service functions in which the calculations of the global superposition or the verification of the global superposition were 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 Internet Protocol (IP) address, so the server 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 Raspberry Pi on which a client application is implemented would get an additional server that can communicate with the electricity provider. In addition, a client would have to be implemented on the electricity provider so that requests from the electricity 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 via a stream. In gRPC can be configured depending on the use. For example in a message stream the server can wait for all client requests before responding or it can respond immediately after each message. It 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

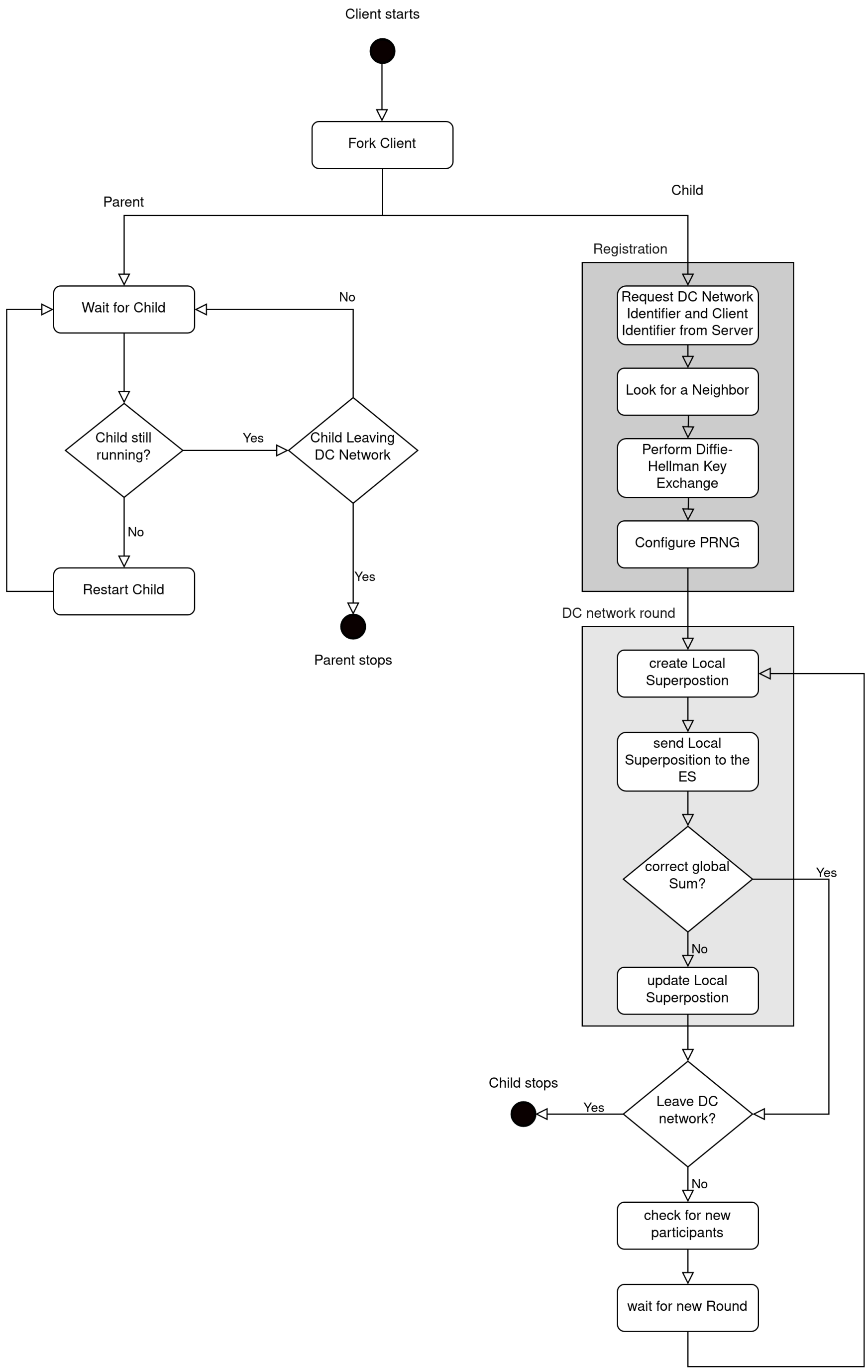


Figure 3.10: The flowchart shows all program operations that are performed in the client.

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.

3.5 Evaluation

In this chapter the properties of the DC network are analyzed. Various characteristics are evaluated, including the performance of the process, the security and the necessary minimum size of the network.

3.5.1 Performance

Regarding the performance, it is considered to what extent the proposed network is more efficient or inefficient compared to the TR-03109. In addition, the message and memory volume generated by the protocol is considered.

Computational Performance

The proposed DC network is highly scalable. Due to the fact that on the SMGW side only a simple computation of the electricity consumption with a generated key through a PRNG has to be performed to form the local superposition, hardly any computational electricity will have to be used. Furthermore, a local superposition is sent only every 15 minutes. The computational overhead caused by the sending the local superposition is negligible for the SMGW and the protocol should be able to run without problems even on lightweight systems. It has to be considered that the used PRNG really generates random numbers. The built-in hardware security module in the SMGW offers cryptographically secure PRNG. On the side of the electricity provider, only simple operations need to be performed as well. Every single received local superposition of a SMGW has to be added up to form the global superposition. Addition of many thousands of summands is no problem for ordinary computers and the complexity of the calculation does not deviate from the proposal of the TR-03109, since the transmitted data has to be summed up as well.
In case of the error correction the transmission bit allows the electricity provider to immediately detect which SMGW in the network has failed. Even if several SMGW fail at the same time, it is no issue for the neighboring SMGW to resend the updated local superposition to the electricity provider. Afterwards, the electricity provider efficiently calculates the global superposition 3.3.

Message Overhead

The messages sent in the protocol can be implemented with a small header. The message field in which the local superposition are entered must always be the same size, otherwise the DC network cannot be implemented. Therefore, large data packets are not to be expected, since the local superposition require a sufficiently large message field, which will not require more than several 100 bytes.

Message Volume

When registering a requesting SMGW in a DC network, several messages must be exchanged between the SMGW and the electricity provider as well as between the requesting SMGW and the neighbor SMGW. If a DC network round runs without errors, no additional messages are exchanged. However, an increased message volume is to be expected during error correction. Especially in the case of an active attacker, several phases will be run in F-mode, requiring message exchange between client and server. However, it must be mentioned that most of the time in the DC network no messages are exchanged at all. Messages are only transmitted when a client enters or leaves, or when the local superposition is sent. If the local superposition fails, the system has enough time to correct the error until the next round. Therefore, an increased message volume due to the error correction can be tolerated.

Memory Overhead

The memory overhead of SMGW in a DC network is minimal. Only the PRNG key from the neighbors need to be stored. However, the electricity provider needs to store the key graph. For this reason the electricity provider has a larger storage overhead compared to the TR-03109. Additionally, it must be assumed that the electricity provider will operate multiple DC networks simultaneously. Therefore, a medium storage overhead is expected.

3.5.2 Security

The security evaluation considers the extent to which the security objectives are better or worse protected compared to the TR-03109 from the BSI. Possible attacks on the security objectives have already been explained in the Design chapter 3.3 and in the related work chapter 2.4.3.

Availability

Availability is one of the most important security objectives of the DC network. If an SMGW should fail, e.g. due to a missing Internet connection, then no global superposition can be formed and the functionality of the network is not possible. This is equivalent to an unavailable network for the electricity provider. Measures have been described to deal with this failure. If an attacker succeeds in taking over an SMGW, he can deliberately send incorrect local superposition with the intention of disrupting availability. Against this active attack, an algorithm was proposed that allows the electricity provider to switch to F-mode and search for the attacker and restore availability. According to this, temporary outages may occur in the DC network, but all of them can be solved in a short time by troubleshooting procedures. A degradation of availability is therefore not expected.

Anonymity

By using PRNGs, 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 electricity consumption of other SMGWs. This is because the attacker has no access to the global superposition. 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[6]. Furthermore, the attacker also has no information about the key graph. This further complicates the chances of a successful attack and 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 electricity provider has to take over administrative tasks and therefore possesses a lot of control. To ensure that the electricity provider 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 superposition can be reconstructed and the electricity consumption is visible. This is prevented by forcing the electricity provider to select a random neighbor in the registering phase for a registering SMGW. Additionally, the electricity provider should be technically forced, that it can't remove a random SMGW at will, but only if an SMGW wants to leave the DC network by itself. These measures make it almost impossible for the electricity provider to affect the selection of neighbors in the DC network.

Eavesdropping

In the case of eavesdropping on one or more channels of SMGWs by an attacker, attackers can read but not understand payload data because the local superposition is not a meaningful message. An attacker can therefore only observe when an SMGW sends local superposition. Furthermore, the attacker can assume from an increased message volume that an error correction procedure is being used in the DC network.

Malicious Electricity Provider

The strongest attacker in the DC network is an electricity provider that has malicious intent with the motivation to obtain additional privacy compromising data. Due to the advanced administrative functions available to the electricity provider, the electricity provider could be tricked into exploiting its administrative rights. Therefore, there is a fine line between granting administrative privileges to the electricity provider to maintain the DC network and restricting administrative privileges so that the electricity provider cannot breach security. Careful consideration has been given in the design to what powers are necessary for the electricity provider. However, if a malicious electricity

[6] Assuming the attacker does not have control over a large number of SMGW in the DC network.

provider does tamper with the DC network, the SMGW will have the ability to leave the DC network.

Protocol compatibility with the TR-03109

The protocol has been designed with the structural specifications of the TR-03109. Therefore, after implementation, it could be integrated into the currently specified German smart grid without any major complications. The WAN interface of the SMGW does not need to be extended and existing communication links are not affected by the protocol. But the biggest weakness in the presented protocol is the electricity provider, since it knows the key graph. This design decision was made only because the technical policy does not allow direct communication between two SMGWs via the WAN interface. If this requirement were changed in the TR-03109, then a decentralized DC network could be integrated into the German smart grid. This would make attacks from a malicious electricity provider impractical.

4 Analysis of Aggregation Groups

The DC network from chapter 3 ensures anonymity of user data in a smart grid. It is impossible for an electricity provider to obtain the individual electricity consumption of a user without performing aggressive attacks. But this criterion alone is insufficient. In chapter two, it was described that 52 different household appliances could be correctly identified from the aggregated electricity consumption of two houses with an accuracy of 55% and a measurement interval of one hour (see 2.2.1). Accordingly, it is still possible to draw conclusions about individual electricity consumption or about the usage of household appliances from different households from an aggregated result. In addition, the measurement interval in the German smart grid is much more frequent and amounts 15 minutes. Therefore, the DC network must have a minimum number of participants to make an analysis over the aggregated result impossible. For this purpose, an experimental analysis was performed to achieve the ability of making a statement about the minimum number of participants in a DC network. When researching the information content of aggregated data, only sparse information and no common metric could be found. In particular, the fact of interest is: How much electricity consumption from individual households needs to be aggregated, so that certain characteristics, such as the use of individual household appliances, can no longer be analyzed from the results? In this chapter, the experiment investigates the minimum number of participants needed in a DC network so that it is no longer possible to draw conclusions about individual users from the result with certainty. For this purpose, the assumption and the structure of the experiment and its execution will be described first. Afterwards, the results of the experiments are considered and analyzed.

Assumption

The goal of the experiments is that no inferences can be drawn from the result of the aggregated electricity consumption. This includes that no inviduals can be deanonymized from the result or that no household appliances can be identified from the result. There is no metric that can measure anonymity from electricity consumption, but there are metrics that can measure the information content of electricity consumption. The assumption is therefore that if the information content from the aggregated result of electricity consumption is low, then no conclusions can be drawn about individual households or household appliances.

4.1 Experimental Setup

2.3.1 described which metrics can calculate the information content of aggregated electricity consumptions. For the first experiment, a different procedure was chosen. Electricity consumptions from the London Smart Meter dataset were used [Net22]. In the dataset, the electricity consumption of 5,567 different households between 2011 and 2014 were recorded with measurement intervals of 30 minutes each. In the experiment, N houses were selected and their electricity consumptions were aggregated for each time point. From the aggregated result (ar) the normalized standard deviation ($nominated\, sd$) for each individual time stamp was calculated and compared to the individual electricity consumption of a house at that time. The comparison checks if the electricity consumption of a single house is within the tolerance limit of the aggregated result. The tolerance limit (l) is the following interval:

$$ar - nominated\, sd < l < ar + nominated\, sd \tag{4.1}$$

If the electricity consumption of a house is above or below the tolerance limit at a point in time, then it is assumed that the electricity consumption of the house has a high information content at that point in time and that this information content could be read from the aggregated result. For each time stamp, each house is counted individually for how often the individual electricity consumptions are outside the tolerance limit. In order to perform the experiment, a script was programmed in R, which selects random houses from the London Smart Meter data set. Additionally, the data set got filtered for houses that have readings between 2013 and 2014. For the highest possible precision of the calculation, it was ensured that when a comparison calculation is performed, the time stamp is also correctly identical. In short every 30 minutes measurements in 2013-2014 of a house were compared whether the electricity consumption is above or below the tolerance limit of the aggregated results of N houses. For each calculation, the number of times the values are outside the tolerance limit is counted. This experiment was repeated with different N sizes to analyze how strong the effects of the information content in the aggregated result are measurable.
A second experiment was also executed. This time, not the time span of one year was examined, but the time span of 2 weeks. Due to the shorter time span and the resulting much faster execution of the second experiment, the Kullback Leibler Divergence could additionally be calculated. The Kullback Leibler divergence is often referred to as relative entropy and was introduced in 2.3.1. The second experiment is carried out essentially like the first experiment. It is counted for one week how often the electricity consumption of one house deviates from the standard deviation of the electricity consumption of N houses. In the second week, another house ($N + 1$) is added to the pool of houses already analyzed. It is now analyzed if the added house has a high influence on the overall result and if the added house stands out from the standard deviation of the first week and has especially many deviations. If the added house does not differ from the other houses in terms of deviations, then it can be assumed that new users are added to the DC network without any additional user behavior being noticeable in the result. In addition to the standard deviation as a metric for anonymity, the Kullback Leibler

Divergence was calculated for each house in the second experiment.

Yearlong Results

Figure C.3 - C.10 show the results of the experimental tests where the measured time period is one year. Experiments were conducted with 2, 3, 5, 10, 25, 50, 100 and 150 random chosen houses per group, whose electricity consumptions were aggregated and the experiments were repeated 2 times. The experiments start with a size of 2 houses, because this is the smallest possible DC network. On the X-axis the time in months can be read. On the Y-axis, the average daily electricity consumption within half an hour of a house is shown. It can be seen that in the experiments with little aggregated houses C.3, C.4, C.5, C.6 more significant variations in the amplitude of the characteristic curves are visible. This can be explained by the fact that in the case of less aggregated houses, individual electricity consumptions have a greater impact and can therefore have a more significant effect on the overall result. This could be an indication that the information content of a single house can still be determined. As an example, the diagram with the aggregation group of 2 houses C.3 can be viewed. Between the end of March and April there is a noticeable gap and low electricity consumption in both houses. In London, there were school vacations in 2013 exactly in this period and the assumption is that both households were on vacation in the mentioned period.
From 25 - 150 (C.7 - C.10) houses the characteristic curves approach the same pattern. The seasons can be clearly recognized. The fact that the seasons can be recognized so clearly, can be argued indirectly with the fact that a larger house group is observed. This is because individual household appliances no longer stand out clearly and larger patterns of behavior can be recognized, such as the fact that in the winter, people spend more time inside their homes, while in the summer, they spend more time outside. In addition, it is noticeable that the average base load is different in each graph, but converges briskly for larger aggregation groups.
The average base load can best be viewed in the boxplot diagram by the median C.1. The boxplot diagram summarizes the diagrams C.3 - C.10 with one boxplot each. In conducting the experiments, the preliminary assumption was that with larger aggregations, the result would approach a horizontal constant. The approximation of a constant would be clear evidence that the information content in the result was lost. The experiment cannot prove the prior conjecture. If the results were to converge to a constant at larger aggregations, then the upper quartile and the lower quartile would become narrower at larger aggregation groups. However, this observation cannot be obtained from the boxplots. In addition, the number of occurrences that a house's electricity consumption falls outside the tolerance limit 4.1 for each time stamp in a year was counted. Figure C.2 shows the percentage deviation of the tolerance limit in the aggregation groups. It can be seen that the percentage deviation is very low for 2 houses and increases sharply for three houses and above. The behavior could be observed in both trials of the experiment. One reason for the observation could be that the standard deviation is a measure of the spread of the values of a feature from the average. With so few values, the result of the standard deviation is in many cases larger than the difference of a electricity consumption to the normalized aggregated result of both

houses. Accordingly, almost all values are within the tolerance limit. From three results onwards, this behavior no longer seems to apply and the number of deviations increases sharply. The number of deviations reaches its highest value in the aggregation group of 10 houses and then the percentage deviation drops sharply. The renewed noticeable increase at 50 houses could not be confirmed in the second experiment and is therefore considered an outlier. The experiment shows that the information content has strongly decreased from an aggregation group of 25 houses and the larger the aggregation group becomes, the smaller the number of counted deviations outside of the tolerance limit gets. Therefore, for the yearlong experiments it can be assumed that a minimum DC network size of 25 is sufficient.

Two Weeks Results

The two weeks experiments can be seen in B.1 to B.8. Since the aggregated electricity consumption from the first week is almost identical to the aggregated electricity consumption from the second week, only the electricity consumption from the first week was included in the appendix. The electricity consumption of the first and second week are summarized in the two boxplot diagrams B.9, B.10. In the diagrams of 2 - 5 houses B.1, B.3 strong differences can be recognized concerning the electricity consumption. In diagram B.1 even an unusually high baseload is recognizable, which could not be observed before in any other experiment. The Y-axis had to be adjusted in order to display the results. In diagram B.2, a daily routine can best be read off. Similarly, the daily routine can be recognized in diagram B.3, but with larger fluctuations in amplitude. Starting from an aggregation of 10 houses B.4, the similarity of the electricity consumptions increases and the diagrams hardly differ from each other. This is an indication that already from 10 houses individual electricity consumptions have hardly any influence on the overall result. The results of the Kullback Leibler Divergence confirm an increase in the relative entropy B.11. An increase in the Kullback-Leibler Divergence thus also leads to an increase in anonymity [Kal+10]. In the second experiment, the divergence from the tolerance limit was also measured. The results can be found in B.12. In the weekly view of the aggregated electricity consumption, it can be seen that the deviation is strongly decreasing the larger the aggregation group is. In the aggregation group with 25 houses an outlier is recognizable. This outlier can be explained by the standard deviation, which is exceptionally low for the aggregation group of 25 houses. Thus, the tolerance limit is also particularly small, which leads to many deviations. In summary, the experiment showed, that with a number above 10 houses assumingly no conclusion can be drawn from the results.

5 Conclusion and Future Work

Smart meters offer a great opportunity to automate the existing power grid and increase the quality in the grid 1. At the same time, smart meters can be misused to create accurate behavior profiles of individuals in their own homes 2.2.1. The BSI's TR-03109 is a first standard that establishes uniform requirements for the security of smart meters and provides a minimum level of protection. Nevertheless, analyses of electricity consumption can be carried out by the electricity provider without any major effort, as pseudonymization alone cannot prevent this 2.4. Therefore, a technical solution to the problem must be found and a variety of proposals are already being considered in the scientific community.
This work deals with DC networks, which are classified as aggregation without trusted third party methods. DC networks have not been considered in detail in science as a solution for privacy-preserving smart grids. The proposed design is a dc network based network protocol, which follows the requirements of the BSI TR-03109. Even though not all requirements of the TR-03109 could be implemented in detail, the considered design can be adopted in the German Smart Grid without major structural changes to the existing infrastructure or software 3. The solution therefore deviates from the original DC network as described e.g. in [Cha88], because the restrictions of the TR-03109 would not be compatible with a normal DC network and the classical DC network could not be implemented. Therefore, the electricity provider must be given administrative authority over the DC network as well as over the key graph in order for the DC network to be compliant with the BSI TR-03109. The proposed design changes had to be weighed throughout that, on the one hand, the DC network does not lose in security and anonymity. On the other hand, the proposed DC network must be efficient and operable in the real world with the administrative powers of the power provider. It must be able to dynamically respond to errors and fend off potential attackers. The two requirements of security and operability have always been at odds with each other and have meant that any change to the classic DC network has had to be carefully considered.
In the chapter 3.5, the characteristics of the proposed DC network were considered in terms of security and performance. While there may be an overhead in terms of message volume, the DC network does not lose in security compared to the classical DC network. In the case of an attacker, the power provider has procedures available to protect its network 3.3 and in the case of the strongest attacker, namely a malicious power provider, the clients have options available to protect themselves from the power provider. Furthermore, the design of the network provides many hurdles for the DC network to be structurally protected from a malicious power provider. The increased message volume is negligible since the power consumption is sent only every 15 minutes and likewise error correction actions result in only a slightly higher message volume. Moreover, the proposed DC network is scalable and can handle large numbers of users 3.5.1.

In addition, the proposed DC network was implemented on four Raspberry Pis (three clients and one electricity provider) to simulate real conditions and prove that it is a workable and viable protocol. Furthermore, a stochastic experiment was implemented to study how the information content of electricity consumption is preserved in aggregated results, so that also no conclusions can be drawn from the aggregated result. The findings from the experiments were used to obtain a minimum number of DC network participants. The results show that a high level of anonymity can be achieved with as few as 10 - 25 participants (see 4).

Future Work

In the case of the manipulation of the local superposition 3.3, the algorithm could not be implemented in the demo. Since the algorithm was defined by the author and not implemented, the correctness of the algorithm cannot be answered with complete certainty.
Furthermore, the experiments looked for a minimum number of participants in a DC network. Different metrics were used and the results were analyzed. To get further clarity on the minimum number of DC networks, it would be possible to perform a cluster analysis and linear regression.

Appendix A:

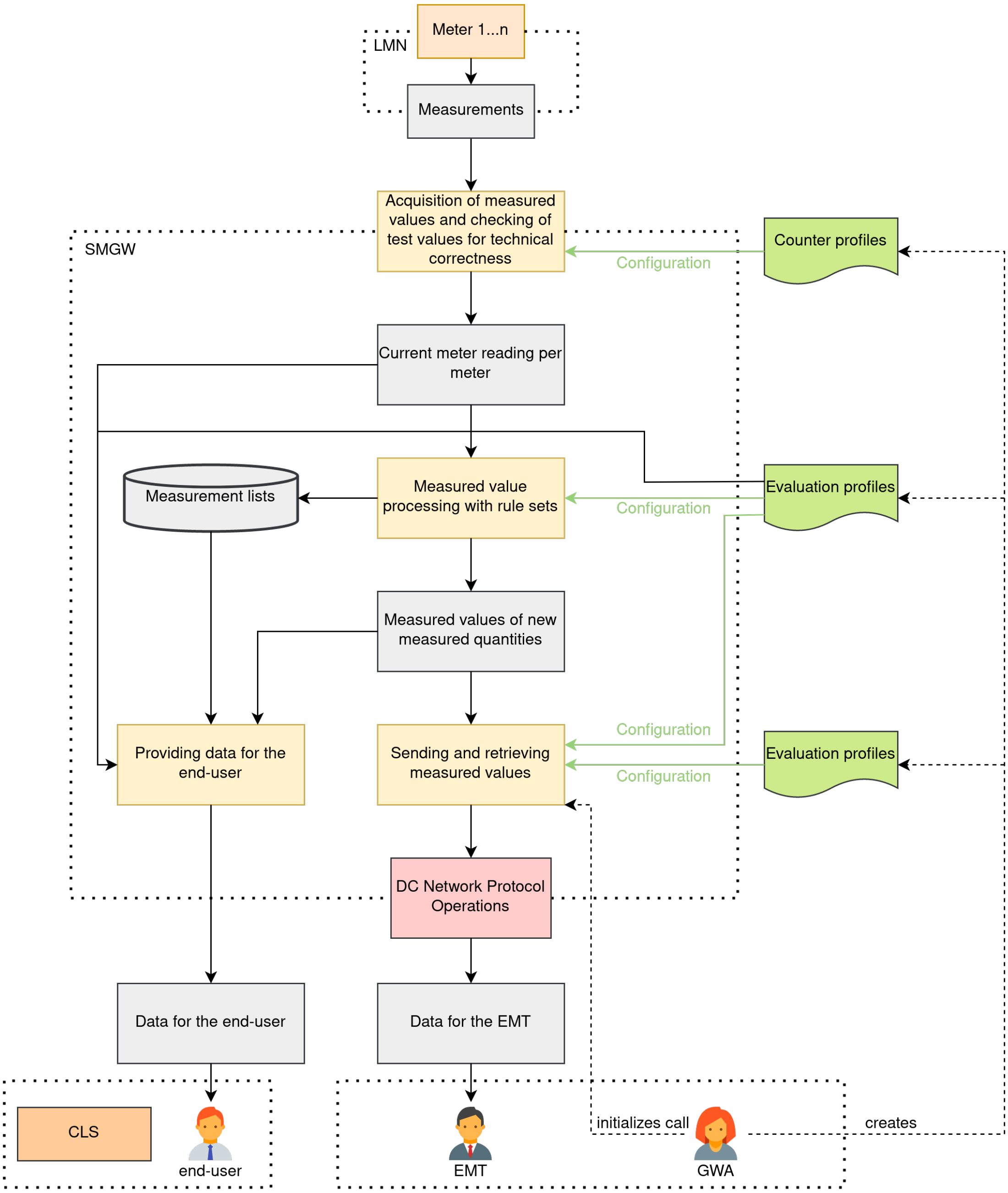


Figure A.1: An overview of the measured value processing within the SMGW with configuration profiles from the GWA and DC Network Protocol Extension.

Appendix B: Two Weeks Experiments

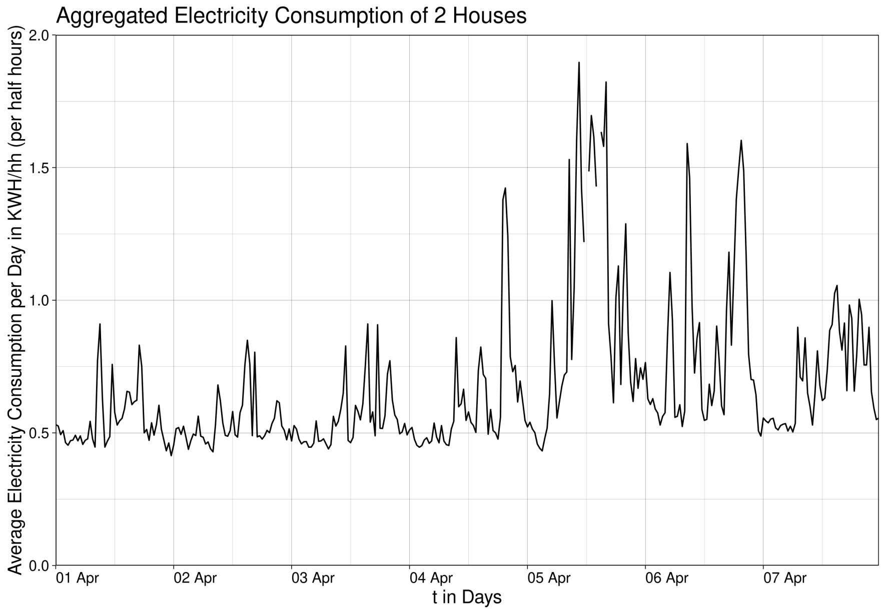


Figure B.1

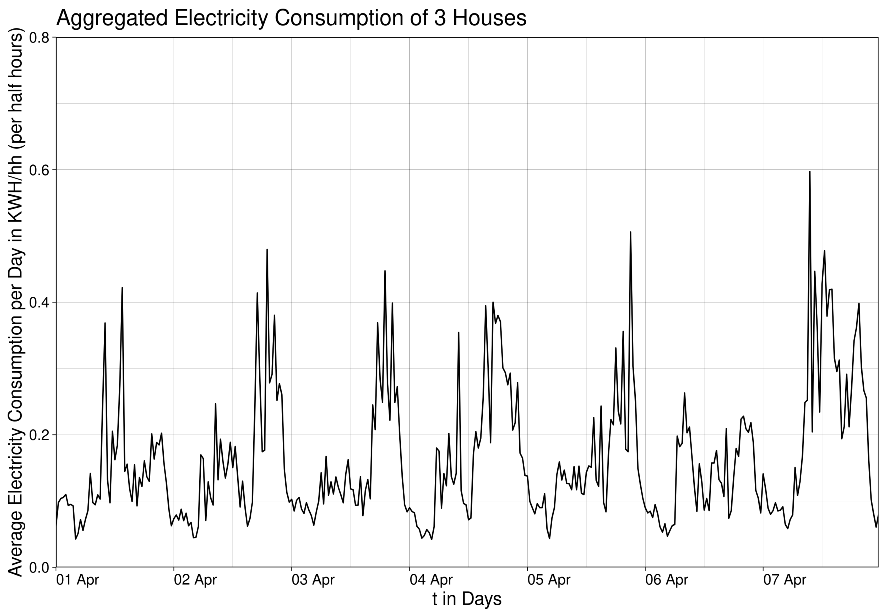


Figure B.2

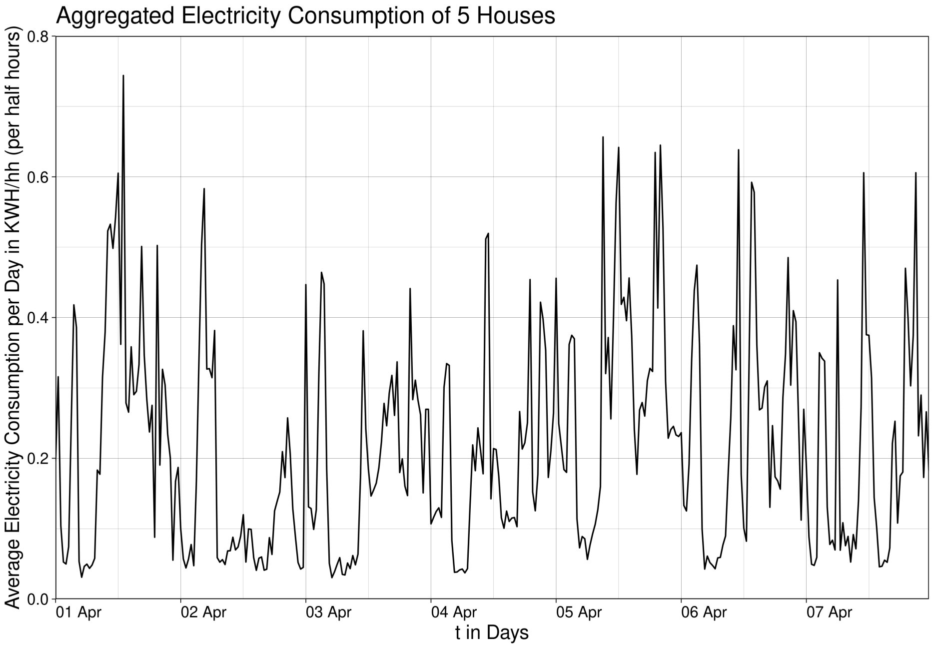


Figure B.3

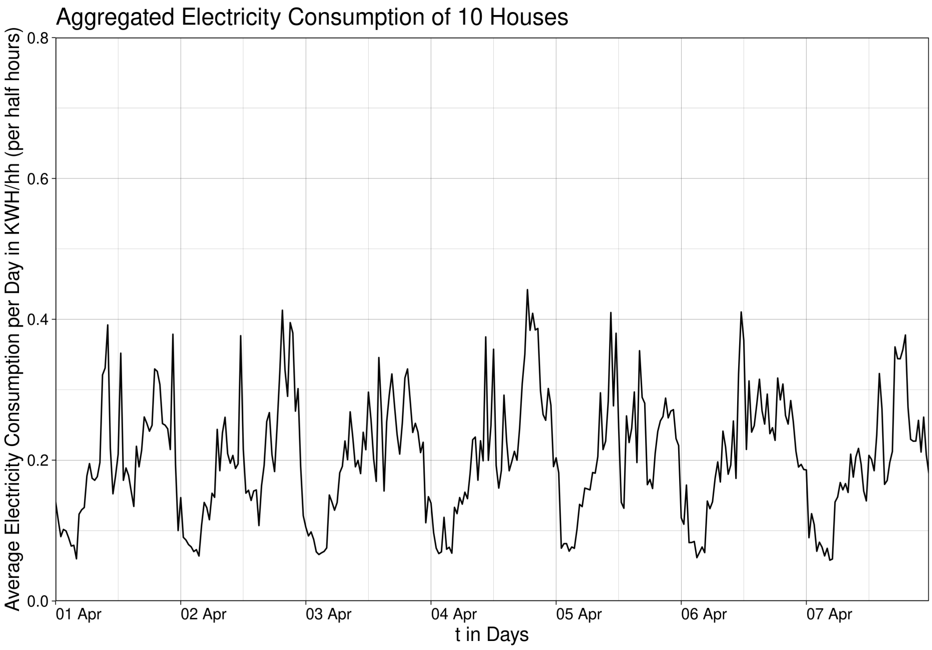


Figure B.4

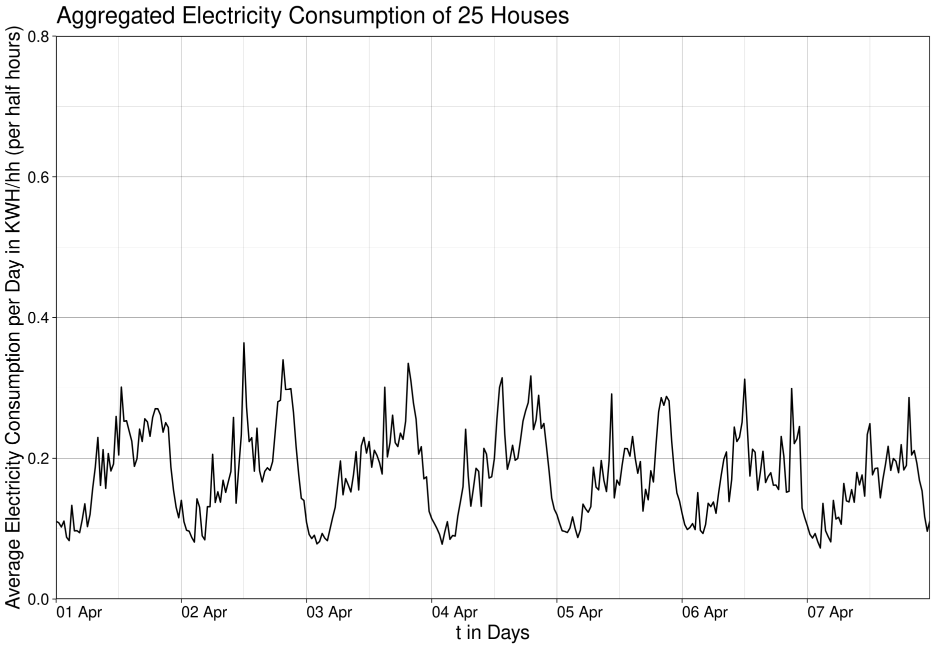


Figure B.5

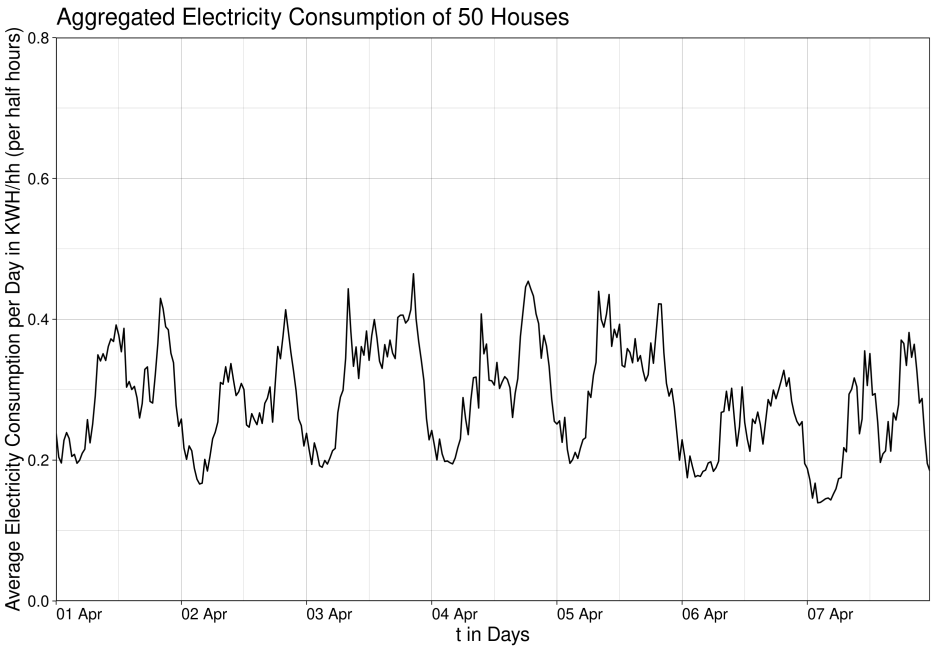


Figure B.6

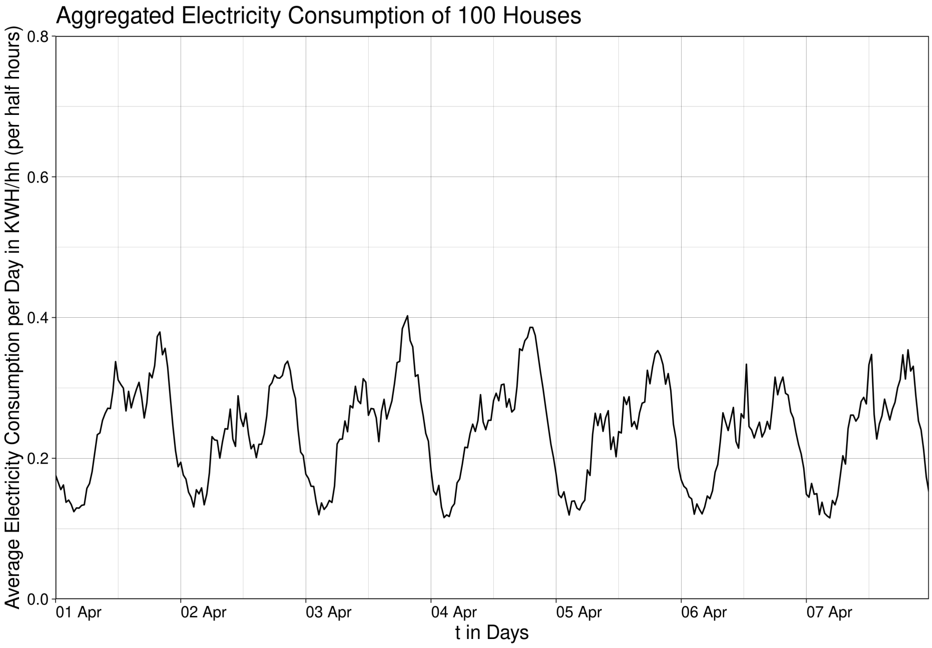


Figure B.7

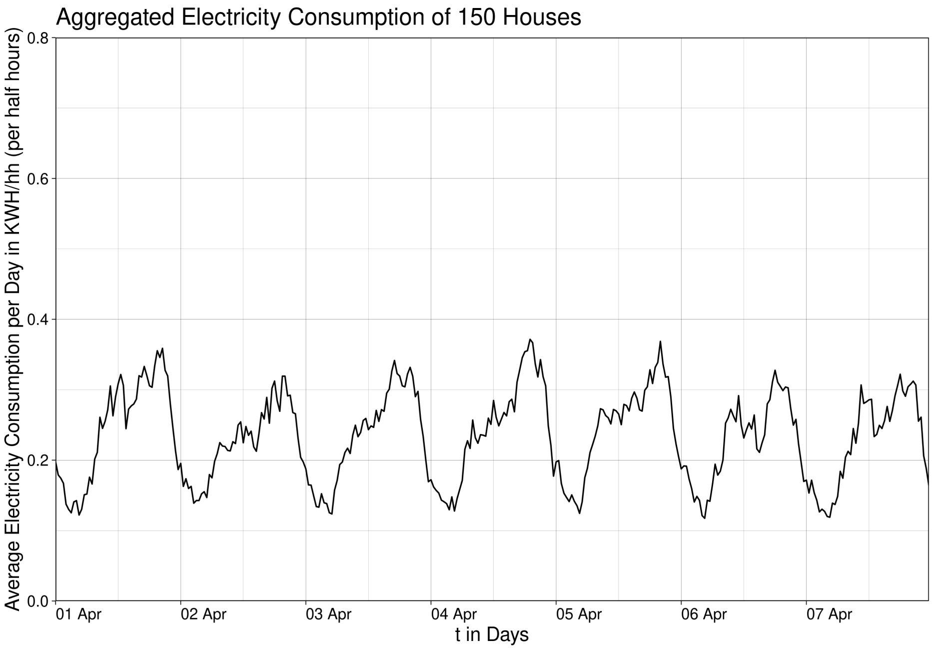


Figure B.8

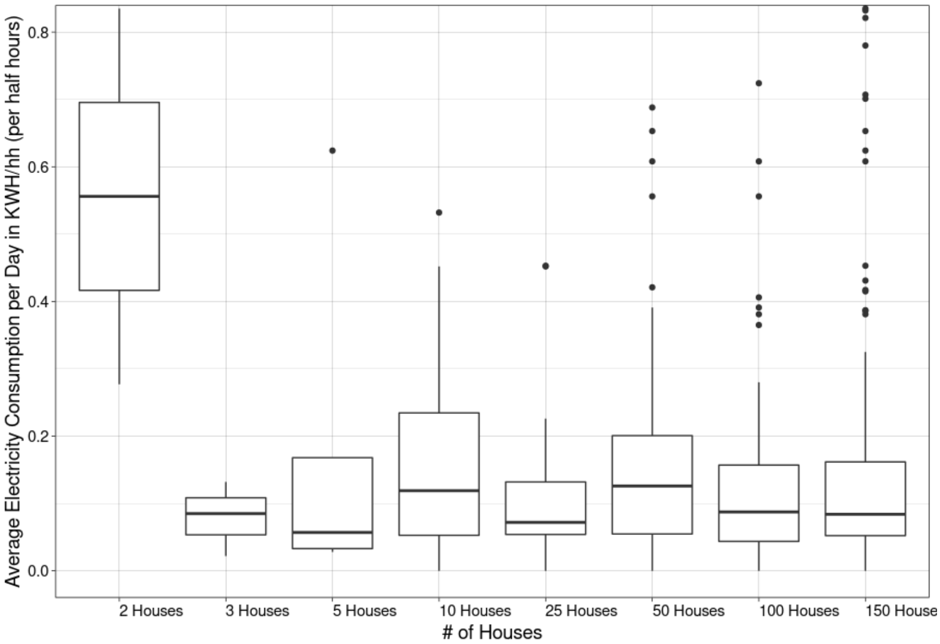


Figure B.9: Boxplot of the average electricity consumption in the first week of the experiment.

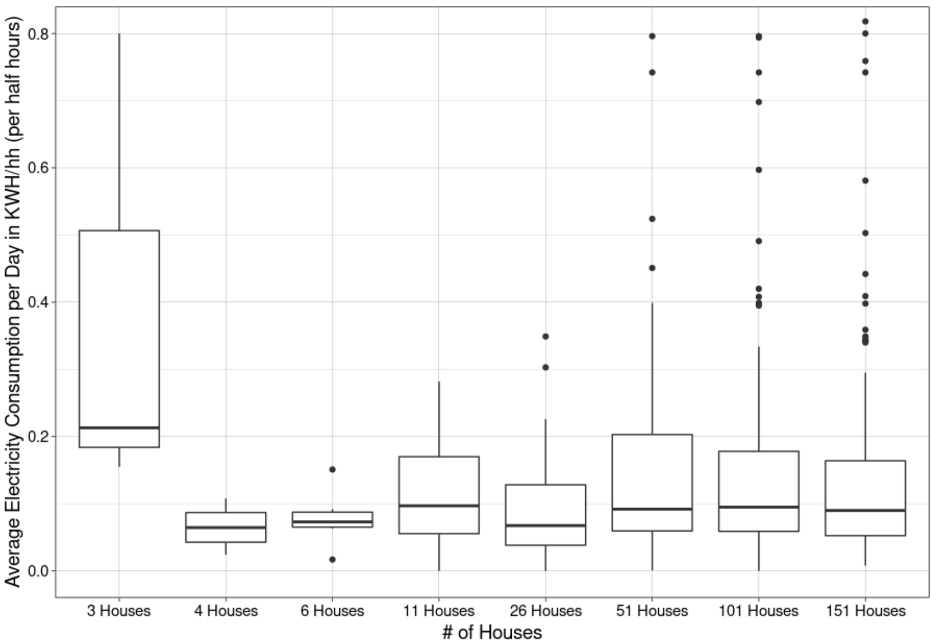


Figure B.10: Boxplot of the average electricity consumption in the second week of the experiment.

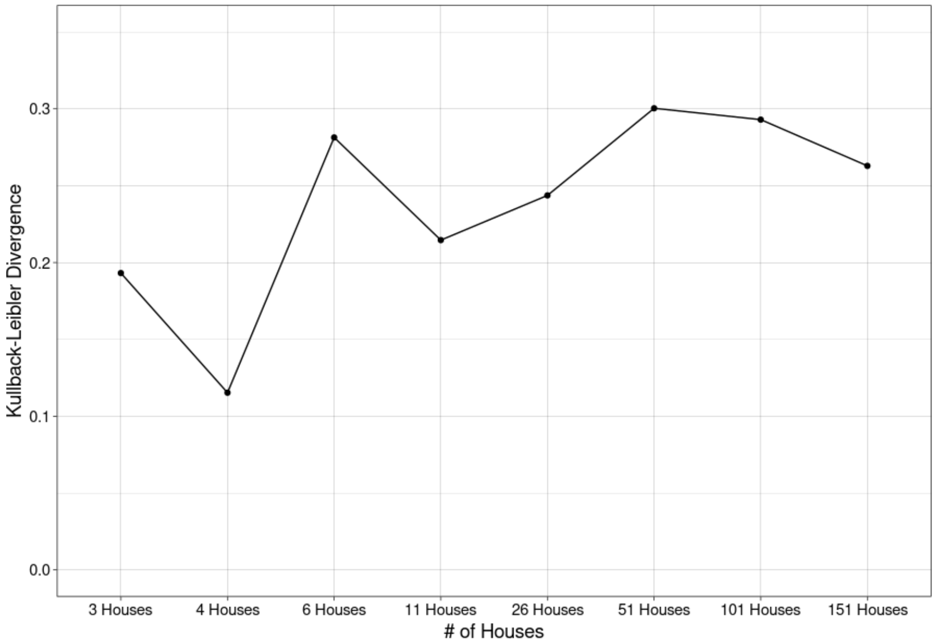


Figure B.11: Kullback Leibler Divergence of the second week.

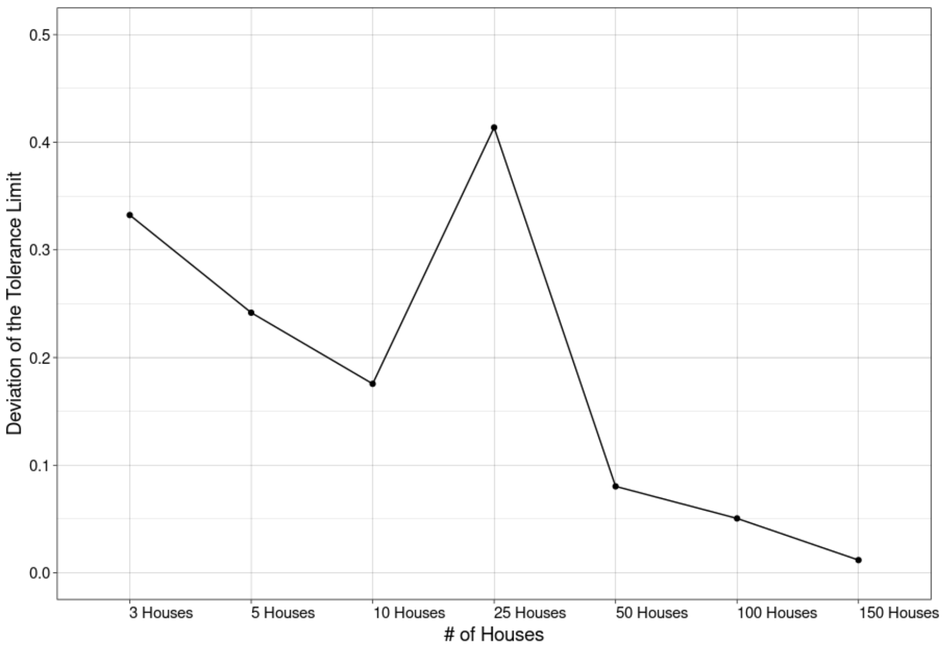


Figure B.12: Deviation from the Tolerance Limit in the first week.

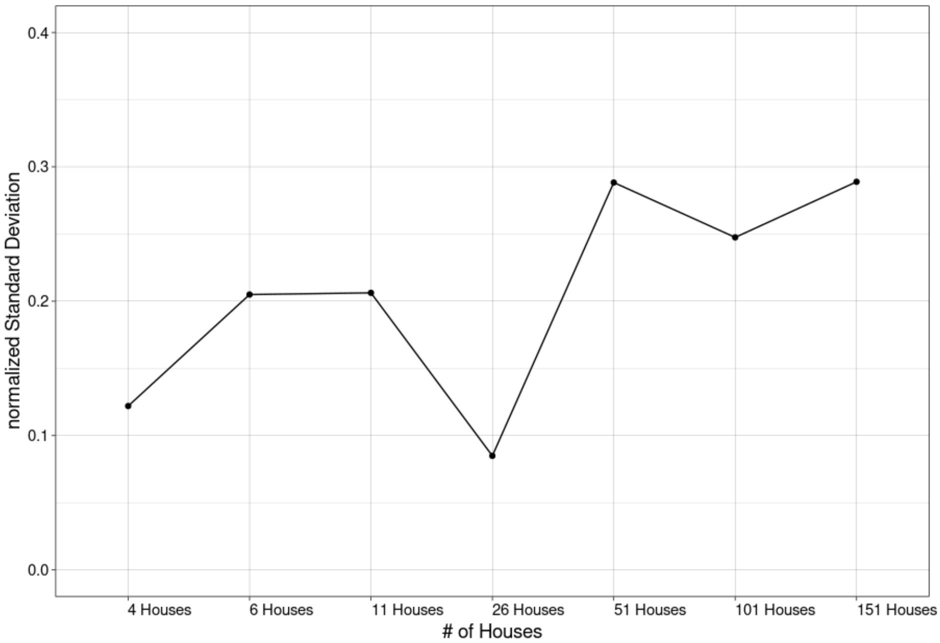


Figure B.13: Normalized Standard Deviation of the second week.

Appendix C: Yearlong Experiments

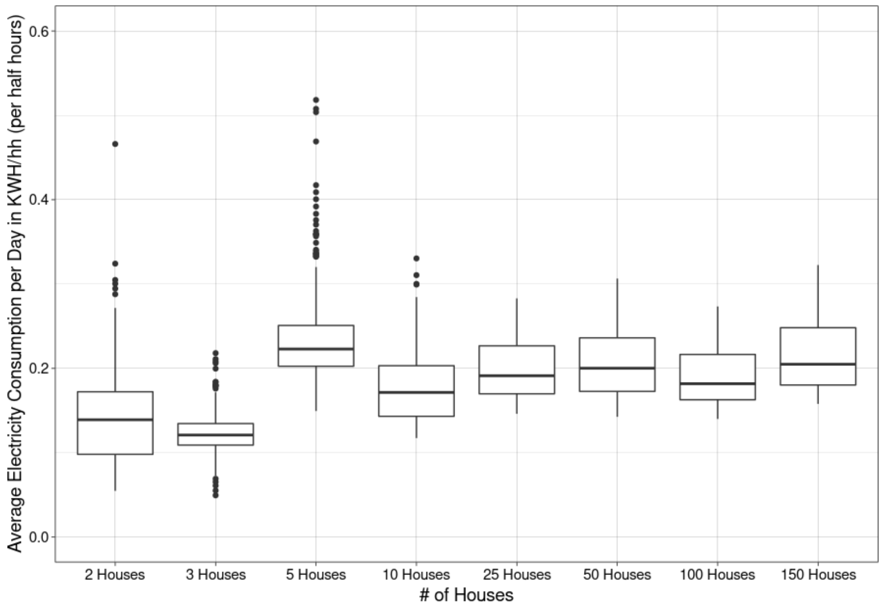


Figure C.1: Boxplots of the average electricity consumption.

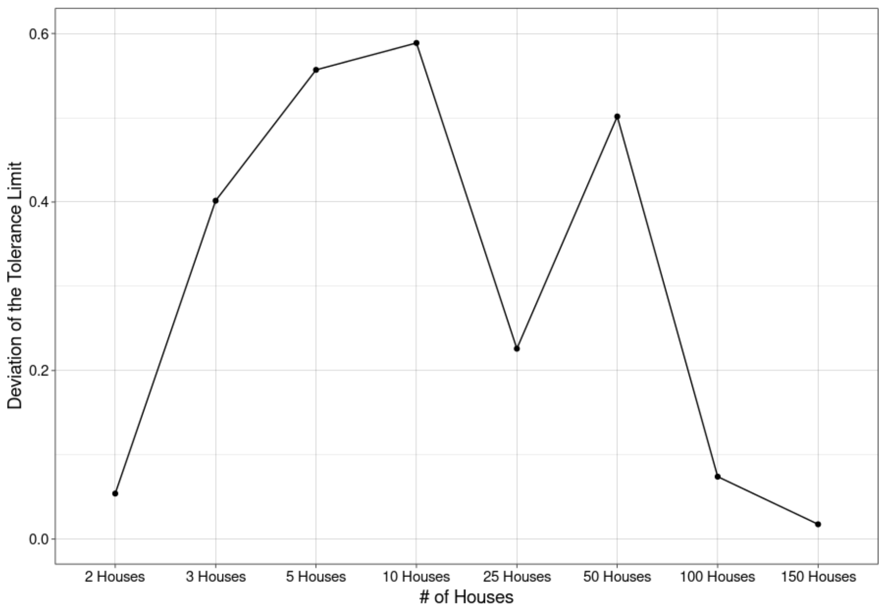


Figure C.2: Deviation from the Tolerance Limit

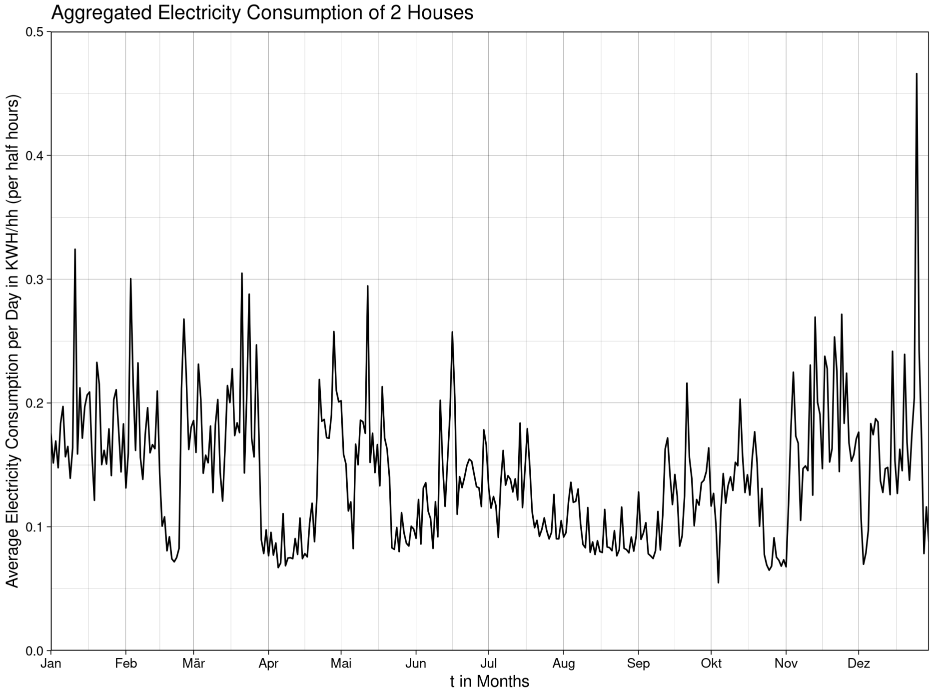


Figure C.3

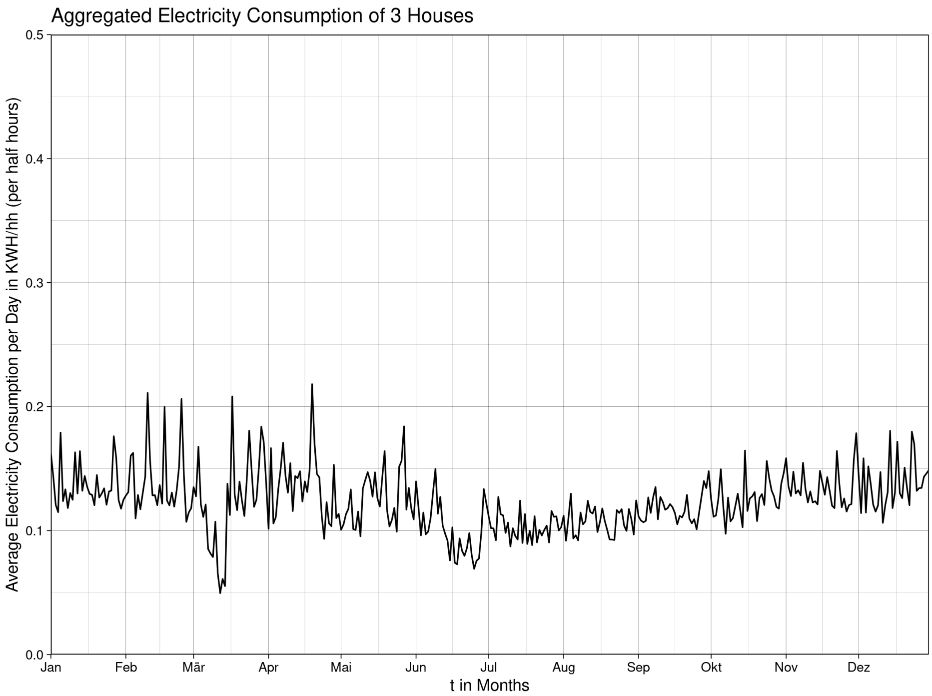


Figure C.4

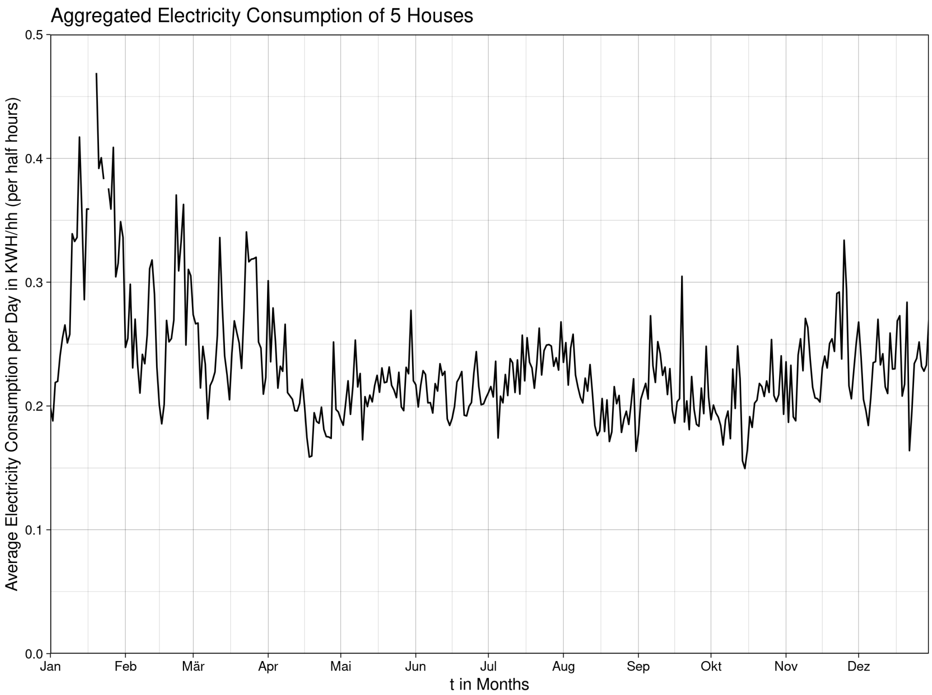


Figure C.5

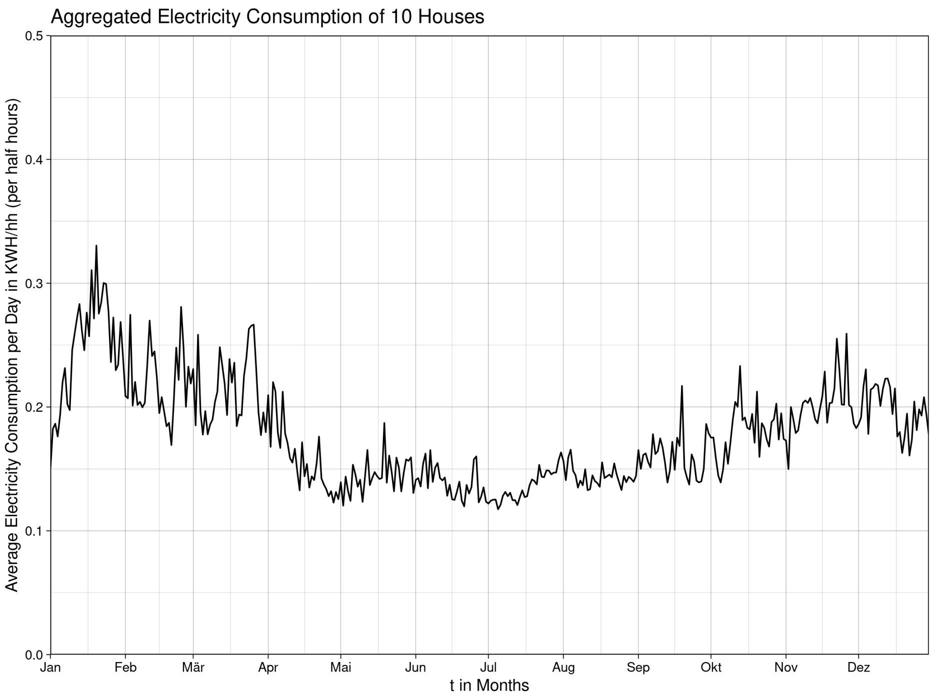


Figure C.6

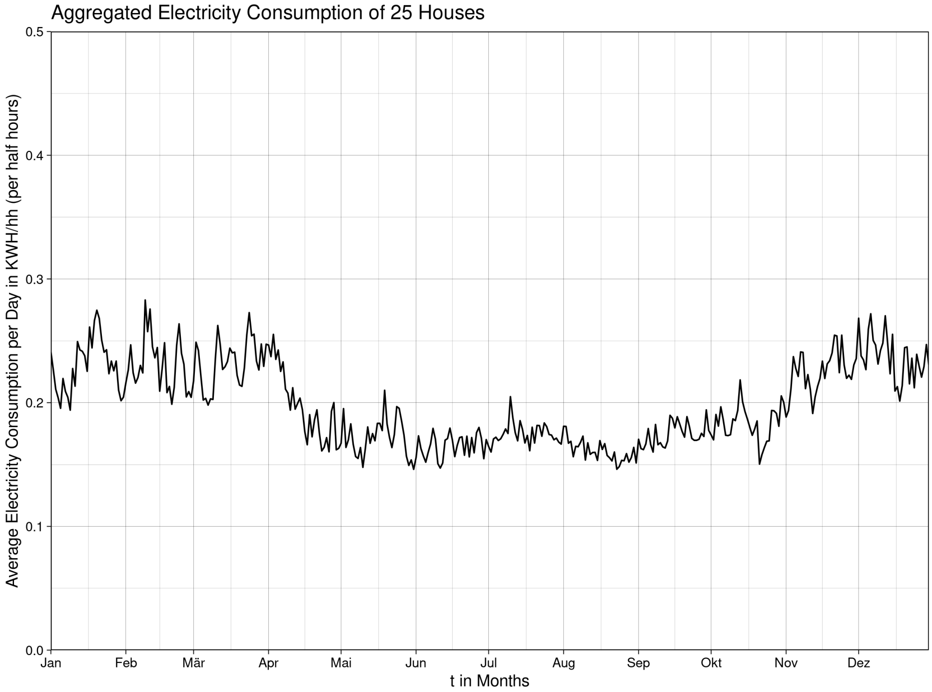


Figure C.7

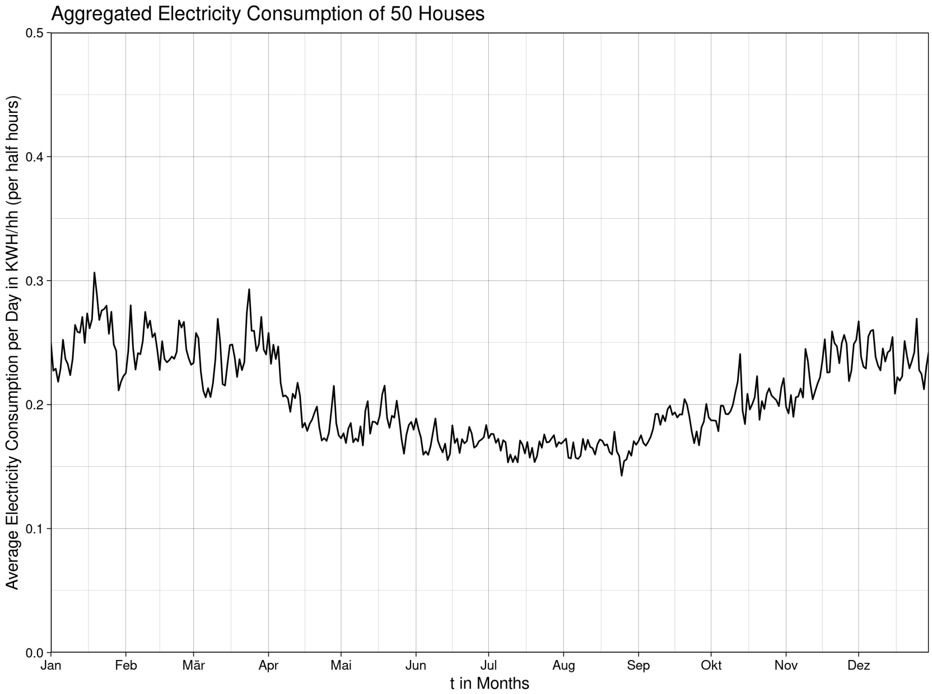


Figure C.8

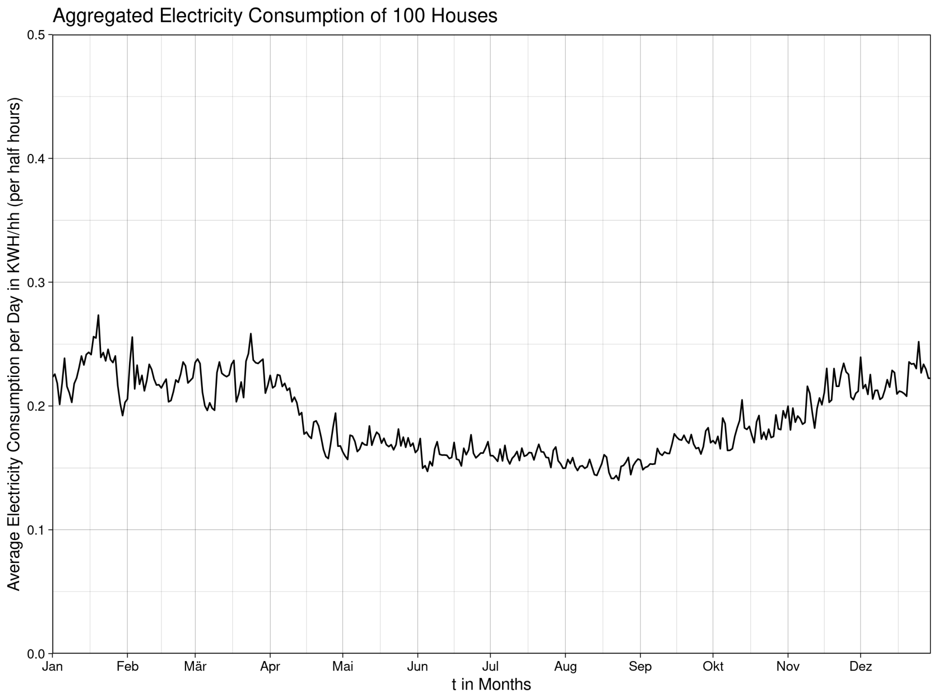


Figure C.9

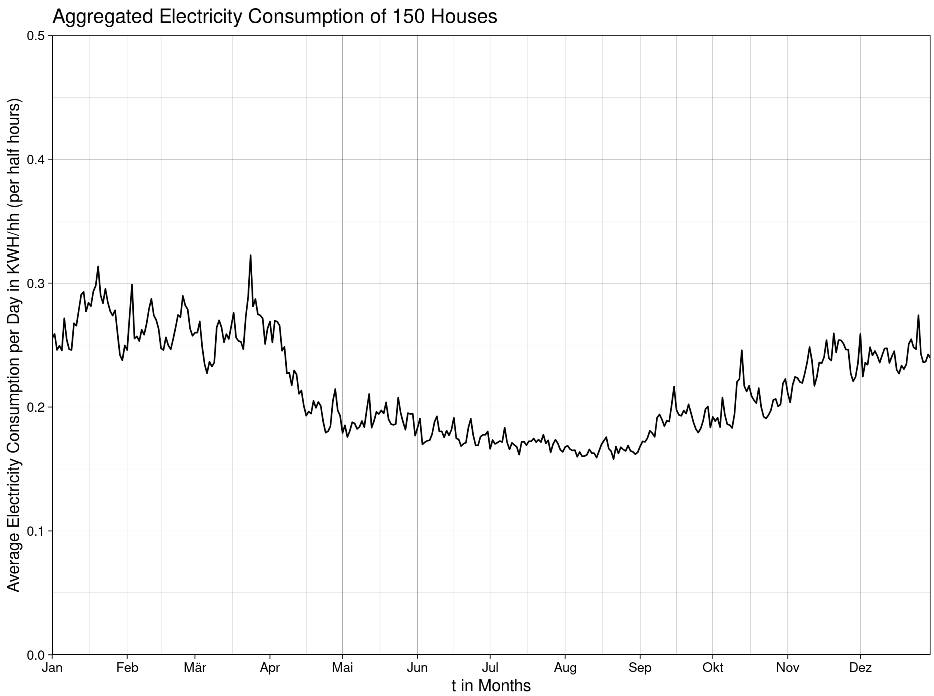
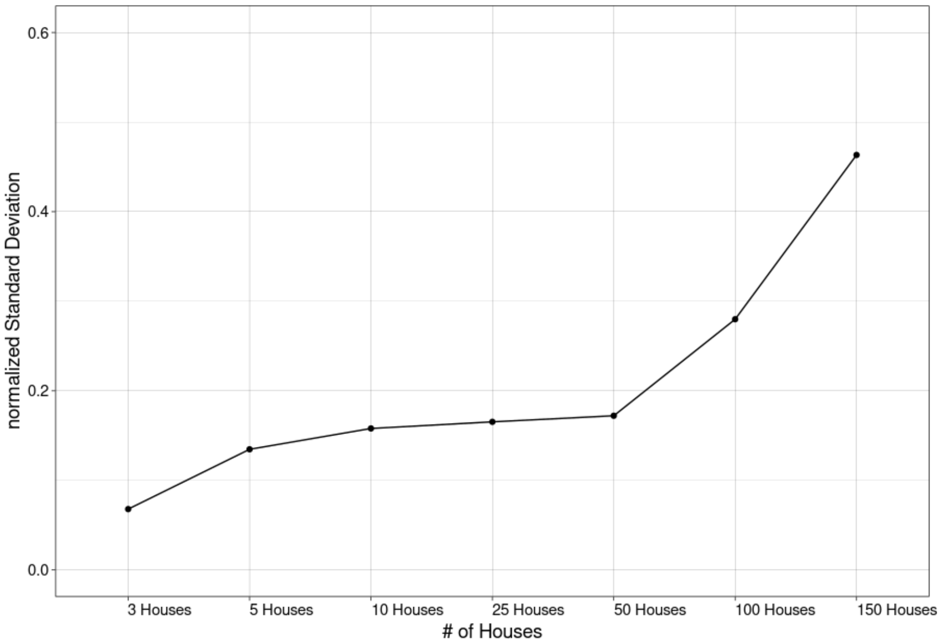


Figure C.10



0.6
0.4
0.2
0.0
normalized Standard Deviation
3 Houses
5 Houses
10 Houses
25 Houses
50 Houses
100 Houses
150 Houses
# of Houses

Figure C.11

Acronyms

AMI	Advanced Metering Infrastructure. 4
BSI	Bundesamt für Sicherheit in der Informationstechnik (eng. Federal Office for Information Security). 1, 3, 16, 23, 27, 47, 54
EMT	External market participants. 16–18, 20, 28
ES	Electricity Supplier. 30
GWA	Smart Meter Gateway Administrator. 16–18, 20, 28, 32
HAN	Home Area Network. 18
IP	Internet Protocol. 44
LAN	Local Area Network. 40
LMN	Local Metrological Network. 17, 18, 27
NILM	Non-Intrusive Load Monitoring. 4, 6–8, 13, 21
PRNG	Pseudo-Random Number Generator. 25–27, 29–32, 34, 36, 43, 46, 47
RPC	Remote Procedure Call. 40, 41, 43, 44
SMGW	Smart Meter Gateway. 7, 13, 16–20, 23, 27–38, 40, 43, 44, 46–49
TPM	Trusted Platform Module. 14
TR-03109	Technical Guideline-03109. 1, 2, 16, 17, 20, 23, 27, 46, 47, 49, 54
TTP	Trusted Third Party. 9, 10, 13, 14, 24
WAN	Wide Area Network. 18, 20, 27, 35, 49

Bibliography

[Aut] gRPC Authors. *gRPC A high performance, open source universal RPC framework*. URL: `https://grpc.io/` (visited on Mar. 5, 2022).

[Bis13] Anna Biselli. “Protection of Consumer Data in the Smart Grid by Aggregation and Homomorphic Encryption.” MA thesis. TU Dresden, Aug. 13, 2013.

[BSU10] Jens-Matthias Bohli, Christoph Sorge, and Osman Ugus. “A privacy model for smart metering.” In: *2010 IEEE International Conference on Communications Workshops*. IEEE. 2010, pp. 1–5.

[Cha88] David Chaum. “The dining cryptographers problem: Unconditional sender and recipient untraceability.” In: *Journal of cryptology* 1.1 (1988), pp. 65–75.

[Dat22] Professur für Datenschutz und Datensicherheit. *Das DC-Netz Anonymes Senden und Empfangen*. 2022. URL: `https://dud.inf.tu-dresden.de/%7esk13/TUD-Web-CMS/LV/Komplexpraktikum/v10_doku.pdf` (visited on Mar. 5, 2022).

[Fan+12] Zhong Fan et al. “Smart grid communications: Overview of research challenges, solutions, and standardization activities.” In: *IEEE Communications Surveys & Tutorials* 15.1 (2012), pp. 21–38.

[FB14] Sören Finster and Ingmar Baumgart. “Privacy-aware smart metering: A survey.” In: *IEEE Communications Surveys & Tutorials* 16.3 (2014), pp. 1732–1745.

[Gar22] Gregor Garten. *Github Repository*. 2022. URL: `https://github.com/GregorUni/masters-thesis` (visited on Mar. 9, 2022).

[Goo22] Google. *Protocol Buffers Framework*. 2022. URL: `https://developers.google.com/protocol-buffers` (visited on Mar. 5, 2022).

[Gre+12] Ulrich Greveler et al. “Multimedia content identification through smart meter power usage profiles.” In: *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*. The Steering Committee of The World Congress in Computer Science, Computer ... 2012, p. 1.

[Har89] George W Hart. “Residential energy monitoring and computerized surveillance via utility power flows.” In: *IEEE Technology and Society Magazine* 8.2 (1989), pp. 12–16.

[Kal+10] Georgios Kalogridis et al. “Privacy for smart meters: Towards undetectable appliance load signatures.” In: *2010 First IEEE International Conference on Smart Grid Communications*. IEEE. 2010, pp. 232–237.

[KBN10] JZNA Kolter, Siddharth Batra, and Andrew Ng. “Energy disaggregation via discriminative sparse coding.” In: *Advances in neural information processing systems* 23 (2010).

[KFB11] Georgios Kalogridis, Zhong Fan, and Sagar Basutkar. “Affordable privacy for home smart meters.” In: *2011 IEEE Ninth International Symposium on Parallel and Distributed Processing with Applications Workshops*. IEEE. 2011, pp. 77–84.

[Le 20] Jerome Le Ny. *Differential Privacy for Dynamic Data*. Springer, 2020.

[LeM+07] Michael LeMay et al. “Unified architecture for large-scale attested metering.” In: *2007 40th Annual Hawaii International Conference on System Sciences (HICSS’07)*. IEEE. 2007, pp. 115–115.

[Már+12] Félix Gómez Mármol et al. “Do not snoop my habits: preserving privacy in the smart grid.” In: *IEEE Communications Magazine* 50.5 (2012), pp. 166–172.

[MMA11] Stephen McLaughlin, Patrick McDaniel, and William Aiello. “Protecting consumer privacy from electric load monitoring.” In: *Proceedings of the 18th ACM conference on Computer and communications security*. 2011, pp. 87–98.

[Net22] UK Power Networks. *SmartMeter Energy Consumption Data in London Households*. 2022. URL: `https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households` (visited on Mar. 5, 2022).

[Ped91] Torben Pryds Pedersen. “Non-interactive and information-theoretic secure verifiable secret sharing.” In: *Annual international cryptology conference*. Springer. 1991, pp. 129–140.

[Pet10] Ronald Petrlic. “A privacy-preserving concept for smart grids.” In: *Sicherheit in vernetzten Systemen* 18 (2010), B1–B14.

[Pfi06] Andreas Pfitzmann. “Security in IT networks: Multilateral security in distributed and by distributed systems.” In: *Lecture Notes, Technische Universitaet Dresden* (2006).

[Pru02] Alberto Prudenzi. “A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel.” In: *2002 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 02CH37309)*. Vol. 2. IEEE. 2002, pp. 941–946.

[Qui09] Elias Leake Quinn. “Privacy and the new energy infrastructure.” In: *Available at SSRN 1370731* (2009).

[Sic21] Bundesamt für Sicherheit in der Informationstechnik. “Technische Richtlinie BSI-TR-03109-1.” In: *Anforderungen an die Interoperabilität der Kommunikationseinheit eines intelligenten Messsystems*. Sept. 17, 2021.

[Umw21] Umweltbundesamt. *Erneuerbare Energien in Zahlen*. 2021. URL: `https://www.umweltbundesamt.de/themen/klima-energie/erneuerbare-energien/erneuerbare-energien-in-zahlen` (visited on Mar. 5, 2022).

[Vet+12] Benjamin Vetter et al. "Homomorphic Primitives for a Privacy-friendly Smart Metering Architecture." In: *SECRYPT*. 2012, pp. 102–112.

[Zea+13] Sherali Zeadally et al. "Towards privacy protection in smart grid." In: *Wireless personal communications* 73.1 (2013), pp. 23–50.