Master Thesis

Protection of Consumer Data in the Smart Grid by Aggregation and Homomorphic Encryption

Schutz von Verbraucherdaten im Smart Grid durch Aggregation und homomorphe Verschlüsselung

Anna Biselli
Master Computer Science
Student Number: 3821625

August 26, 2013

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

Contents

1 Introduction

The work at hand will study the applicability of homomorphic encryption mechanisms for the privacy preserving transmission of energy consumption data in smart metering environments.

Many stakeholders from industry, politics and science see the Smart Grid as a great chance to move towards a more sustainable and energy-efficient electricity network in the future. As this is considered the main advantage in Europe, countries like Brazil also focus on a more reliable power distribution.

Fine grained consumption data can help electricity suppliers to exploit the full potential of benefits for load balancing. Together with other factors like the current weather this facilitates the efficient integration of renewable energy sources. Even more benefit could be imaginable if consumers shift their consumption to off-peak times when they can see their consumption footprint. Additional motivation can consist of special tariffs for times of higher energy supply capacity and a low overall demand. At the same time such detailed measurements threaten a customer's privacy and open the possibility of profiling. Informations like absence time, daily routines and devices in use can be derived from his electricity usage in short time intervals [47, 58]. In the worst case, for instance if data is leaked, this personal information can be abused for committing crimes. One can easily imagine that the knowledge when somebody usually is at home or not might also be of interest for burglars and alike – just to visualize one possible scenario.

To render the need for consumption data compatible with customer privacy, an approach to mask the data of an individual and still providing sufficient data for load balancing will be proposed, based on aggregating data from consumer groups. As a starting point, a system published by Mármol et al. is investigated. It is analyzed in terms of its applicability for practical use. Special focus is on the conformity to German legal reality and the technical guideline TR-03109, issued by Germany's Federal Agency for Information Security. It will act as the regulatory framework for smart metering devices in Germany and thus needs to be factored in for systems in productive operation. Therefore necessary adaptions to the proposal of Mármol et al. will be discussed and solutions to achieve conformity to the guideline's postulations are offered.

This work is part of the TruEGrid project, standing for *Trustworthy and Energy-Efficient Smart Grids*. The goal of TruEGrid is to develop solutions that increase the trustworthiness of the Smart Grid targeted at the situation in Brazil. In this context, it will be considered if the approach and the requirements set in the TR-03109 can be suitable and of benefit for the evolvement of a Brazilian Smart Grid solution, too.

A prototype will emulate a framework implementing the most fundamental processing steps which are obligatory, regarding the guideline. The proposal of Mármol et al. will be integrated into this system with developed adaptions to remove incompatibilities

with TR-03109. Based on this, a performance evaluation is conducted, together with an examination of the privacy level, possible security threats and imaginable attack scenarios.

At the beginning, this documentation will provide an introduction into the Smart Grid and smart metering. It will especially outline the current advancement of the smart meter rollout in Europe and Brazil. A further goal of Chapter 2 is to raise awareness for the special chances and challenges accompanying the introduction of this new technology.

As a basis for the classification of the concrete approach, an overview of privacy enhancing techniques will follow in Chapter 3. After this, a synopsis over homomorphic cryptography will be given. Special focus is on related work applying this encryption scheme in smart metering and sensor networks. Those can be seen as a generalization of metering environments with comparable needs. Finally, proposals relying on alternative foundations will also be mentioned.

To broaden the perspective, Chapter 4 offers a review of the most important legal and regulatory factors. It treats their impact on the handling of personal data in Smart Grids in Germany. At this point, also first influences on privacy-enhancing approaches will show up.

The specific design of the prototype and how the technical guideline concretely affects the actual protocol implementation is outlined in Chapter 5. After an analysis of the most essential factors, the prototype, its functionalities and assembly is described. Based on this, Chapter 6 intends to assess the presented approach regarding its performance and security. Besides these functional factors, real-world applicability is judged – bearing in mind conformance to TR-03109 as well as practicability in terms of arising overhead.

Conducting all this, the last chapter gives a résumé of the covered goals and achievements. Future enhancements are suggested to increase the practicability to make the outcome of this work an applicable choice for privacy protection in real-world smart metering systems. Altogether, this will hopefully contribute to a Smart Grid equally serving the needs of industry, society and environment.

2 The Smart Grid

To get a grasp on the chances and challenges of the Smart Grid, it is crucial to provide a definition of the term as such. The main aspect of the Smart Grid propagated by the media and therefore most visible for the end customer of electrical energy is the implementation of smart meters in households. But in fact the Smart Grid covers more than that. According to the *Report to NIST on the Smart Grid Interoperability Standards Roadmap* [29], the Smart Grid aims at modernizing the electricity network to enable it for coping with its various interconnected elements,

> *"from the central and distributed generator through the high-voltage network and distribution system, to industrial users and building automation systems, to energy storage installations and to end-use consumers and their thermostats, electric vehicles, appliances and other household devices."* [29, p. 6].

The electricity network changes from a unidirectional system, providing power from a central power plant to the consumers, towards a bidirectional network where every consumer can also act as an energy producer himself, e.g. by installing photovoltaic devices on his roof and connecting them to the local electrical grid. Induced by that, the Smart Grid is essential towards a flexible and environmentally friendly modernization of power distribution.

The following chapter will deal with the basics of the Smart Grid and smart meters and provide a deeper understanding of chances and challenges, focussing on the conditions in Europe and Brazil.

2.1 Smart Meters

A smart meter is a measuring device able to electronically record the consumption of electricity, gas, heating or water. Those values can then be easily read and processed by the energy supplier.

In contrary to the conventional Ferraris meters, those meters are able to record supplementary data besides the overall consumption value at the time of reading that is needed for billing purposes. They are capable of collecting fine-grained consumption statistics – typically in intervals like 15 minutes – which can be used to create load profiles describing the time course of energy usage and help the supplier with effective load balancing.

Another technical advancement of intelligent meters constitutes the process of remotely reading out the meter's measurements on demand. The values can be accessed

via internet or radio connection without mandatory physical access. This dispenses the need for human readers periodically going from house to house to collect the consumption data. Data can now potentially be transmitted to everyone with an internet connection. This culminates in a variety of possible applications, for example tweeting smart meters[1].

2.1.1 Smart Meter Rollout in Europe

The European Union began to push the implementation of smart meters forward in 2011 and the European Commission presented its communication on the Smart Grid on 12th April 2011 [31]. Key arguments focussing on sustainability were expressed by the Commissioner Günther Oettinger as follows:

> *"When it comes to energy production and consumption, Europe needs more than ever to be on a sustainable path. It is time to invest in higher energy efficiencies and a wider use of renewable sources: this is the best way forward to ensuring safe and competitive energy for us and children. The Smart Grids and the use of smart meters are key for a better use of energy."* [51]

In a press release of 9th March 2012 [32] the Commission set the goal to replace 80% of all conventional electricity meters by smart meters until 2020. On the same day, they also issued a recommendation on preparations for the roll-out of smart metering systems [33]. It contains advice for the creation of a cost-benefit analysis the member states had to conduct, as well as minimum requirements for smart metering systems and security and privacy notions.

A so-called *Smart Meter Coordination Group* was established. It consists of the most important stakeholders from politics, consumer organizations, industry and national committees. Together they are working on a reference model for an advanced metering infrastructure and the connected standards.

Despite the strong engagement of the European Union, the dissemination of smart meters differs very much throughout the member states. A graphical representation of this is given in Figure 2.1. While in Italy, the biggest electricity provider Enel S.p.A. implemented 30 million smart meters, covering almost its entire customer base in 2001 [9], smart meter rollout in the Netherlands has been made voluntary after massive customer protests arising from privacy concerns [50].

2.1.2 Smart Meter Rollout in Brazil

The energy sector in Brazil is growing constantly due to the steadily increasing middle class, the economic growth and the investments made by the government. In 2011, Brazil was already the 7th biggest energy consumer in the world [30] – with upward tendency. Electrical energy mainly originates from hydroelectric power plants, which serve 78.2%

[1] `http://www.spiegel.de/wirtschaft/0,1518,634115,00.html`, `http://www.ladyada.net/make/tweetawatt/index.html`, *Last accessed: 12th July 2013*

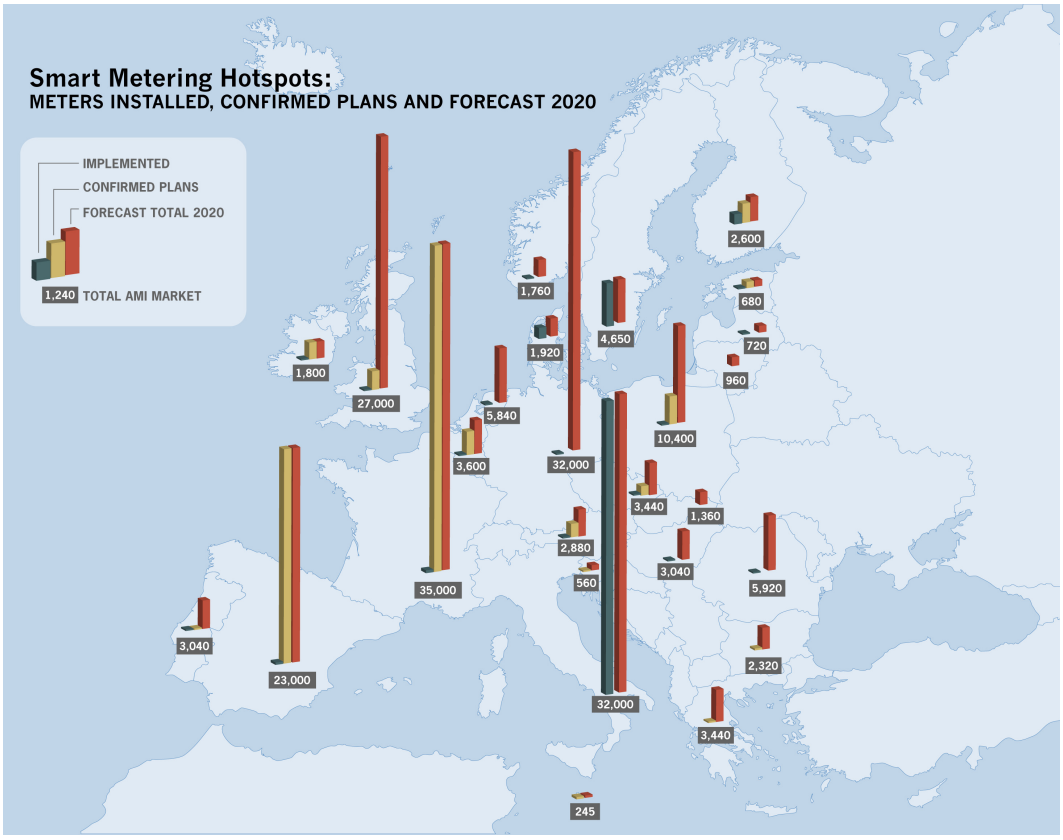

Figure 2.1: Status of European smart meter rollout in 2012 [42]

of the total domestic energy production [44]. Nevertheless, there are highly problematic aspects in the electricity sector in Brazil. Namely, those are losses, peaks in consumption and quality of service.

Losses of electricity are caused by technical issues of distribution and transmission, but also by non-technical ones like cable theft, operational and measurement errors, as well as electricity theft. This last issue alone accounts for up to 24% of the total loss volume [56]. In the matter of load distribution, massive consumption peaks are observed between 6pm and 9pm like illustrated in Figure 2.2. This peak is induced by turning on lights and electric showers. Concerning the quality of service, electricity outage accumulates to 18.77 hours per customer in average [3] which lies above the striven tolerated limit of 16.22 hours. In comparison, the electricity outage in Germany averaged out for 15.31 minutes per head in 2011 [20].

Smart Grid implementation in Brazil mainly aims at bringing these critical points under control. In August 2012, the Brazilian Electricity Regulatory Agency (ANEEL) launched the Resolution #502/2012 [4] for regulating the kick-off of smart meter installation. It is planned to replace 68 millions of conventional meters with smart meters.

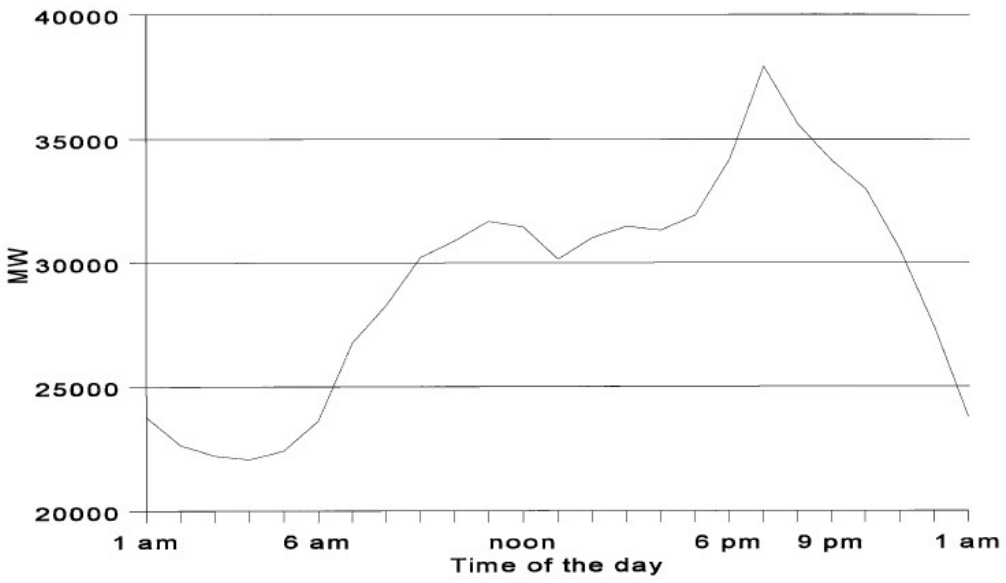

Figure 2.2: Load profile of Brazilian south, south-west and central-west interconnected system [39]

Those meters implement a so-called 'white-tariff' which uses three different price levels correlated to the known peaks (see Figure 2.2). It is hoped for to enable automatic peak reduction by motivating the customers to shift their electricity consumption to cheaper times. But, this effect is naturally limited by the feasibility to alterate consumption habits - while it might be easy to automatically turn on laundry machines during off-peak times, cooking during the night-time is not an acceptable option.

Losses are also expected to be cut back on base of a Smart Grid system because of integrated fault detection and self-healing functionalities. Those also empower to elaborate the quality of service by enabling easier and better maintenance and error localization. In a country as big as Brazil this helps to significantly fasten the patching of faults.

Concerning non-technical losses, fault detection will also help to identify electricity theft. Suspicious consumption behaviour can be monitored and investigated on-site. Furthermore, digital meters will be more difficult to manipulate unobtrusively in comparison to the conventional Ferraris meters, at least without special equipment and sophisticated technical knowledge.

The explanations given above clarify that the motivation of realizing a Smart Grid electricity network in Europe and Brazil are differing. While the first focusses on enabling efficient management of distributed, decentralized power systems and a rise of efficiency and sustainability, the central point of the latter is a better overall electricity quality.

With this challenge overcome, Brazil would have a great potential to boost the integration of renewable energy sources beyond hydroelectricity, especially of solar power. A study conducted by acatech [5] compared ten countries including Brazil, Germany, China and the United States. Among these countries, they rated Brazil as providing the best geographical prerequisites for renewable energy sources in general.

The circumstances and intents specified above have to be taken into account for the analysis and discussion of implementation decisions.

2.2 Chances of the Smart Grid

Like mentioned at the beginning of this chapter, the technology connected to the Smart Grid can help to tackle some challenges of modern electricity networks, which will be described here in short (for detailed information refer to [34]).

Flexible network topology: The existence of various decentralized energy-producing systems from renewable energy sources makes it inevitable to integrate them efficiently into the overall system. This cannot be accomplished by the conventional electricity networks as they were designed for just providing electricity generated by central power plants - not for transporting electricity from the opposite direction like single customers.

Enabling an active consumer: With the smart metering devices, the customer will also get the possibility to gain a detailed insight to his own consumption behaviour. This may lead to a more conscious behaviour and eventual usage reduction. By offering different tariffs in dependence of global peaks, these can eventually be flattened by motivating a transition of electricity utilization towards cheaper time periods.

Efficiency through automation: If together with the Smart Grid, intelligent electric devices are established, those could balance the load even more efficiently than the customers themselves. For instance, it would be imaginable to automatically turn on household-devices with a high need for electricity when the overall load is low. Or to charge the battery of an electric car during cheap tariff times and even use this power for other uses if the car is not needed as such.

Security and reliability: With fine-grained information about the current network status, electricity can be routed in a more flexible way to enable self-healing mechanisms in the system. Also frauds can be detected more easily and pinpointed with an information gain concerning the localization of unusual and suspicious power consumption.

2.3 Risks of the Smart Grid

Compared to the conventional Ferraris meter, the usage of electric smart meters poses a massive increasing of data volume. Potential misapplication of this data is one of the key

concerns of individuals in Europe and especially in Germany. In a survey of 2010, forsa[2] found out that 27% of the interviewed persons expressed doubts about their privacy [35]. These doubts stem from the fear of being profiled and monitored, together with the possibility of criminals to abuse personal data if leakages occured.

Other concerns are mainly related to financial drawbacks, i.e. that the supplier would rise tariffs to the disadvantage of the customer. Also fears of insecure and tampered meters are relevant. The first can be seen as an organizational issue where the electricity suppliers have to establish trust by not acting in conformity with these objections. The latter and the privacy concerns mentioned in the paragraph above have to be addressed by reliable security mechanisms.

This chapter provided the foundations for an understanding of the key issues in Smart Grid implementation. As this work focusses on consumer privacy, a introduction on privacy enhancing techniques in general and related to smart metering will be given in the next sections.

[2] Institute for Social Research and Statistical Analysis

3 Privacy Enhancing Technologies for the Smart Grid

In the *Handbook of Privacy and Privacy-Enhancing Technologies* by Blarkom et al. Privacy Enhancing Technologies (short: PETs) are defined as follows:

> *"Privacy-Enhancing Technologies is a system of ICT measures protecting informational privacy by eliminating or minimising personal data thereby preventing unnecessary or unwanted processing of personal data, without the loss of the functionality of the information system."* [67]

More specific protection measures, leading to concrete implementations can be structured by the following goals:

- Anonymization or pseudonymization of individual users by removing information which identifies a user personally, like e.g. his IP address
- Unlinkability of a user's utilization of different resources and services by identity management approaches like Anonymous Credentials
- Prevention of profiling through data mining and processing, among others by encryption and data aggregation

This chapter focusses on the last point and presents existing approaches to preserve privacy in sensor networks – especially in the Smart Grid.

3.1 Data Aggregation and Bihomomorphic Encryption

Like already mentioned above, data aggregation is an often used approach when it comes to the protection of an individual against profiling in settings where massive data amounts accrue. The principle can be simply understood as summarizing data from multiple individuals and transmitting the composite value to the entity which requested the information. By performing this aggregation, the single summands are hidden.

3.1.1 Homomorphic Cryptosystems

A drawback of the utilization of plain data aggregation is that an eavesdropper listening on the communication channel that transmits the individual user's value to the aggregator has no difficulties collecting them. Another important issue is the necessity of

a trustworthy aggregator. He comes to know all single values, so if he was malicious, he could simply refuse to carry out the aggregation and forward the individual values directly to the final receiver.

Homomorphic encryption solves these problems by allowing computations on ciphertexts. After decryption, a result corresponding to computations on the plaintexts is retrieved without the need to know the plaintexts at the moment of computation. The operations on plain- and ciphertext need not necessarily be equal. Under these circumstances, trust in the aggregator in terms of eavesdropping is dispensable. He himself does not get to know the plaintexts of the particular transmissions and thus cannot forward them to other entities.

Exemplary homomorphic encryption schemes are the following:

RSA is homomorphic with respect to multiplications:

$$\begin{aligned}\text{enc}(m_1) \cdot \text{enc}(m_2) &= m_1^{k_e} \cdot m_2^{k_e} \mod n \\ &= (m_1 \cdot m_2)^{k_e} \mod n \\ &= \text{enc}(m_1 \cdot m_2)\end{aligned}$$

Paillier cryptoscheme's homomorphic property maps a multiplication of ciphertexts to an addition of plaintexts:

$$\begin{aligned}\text{enc}(m_1) \cdot \text{enc}(m_2) &= (g^{m_1} \cdot r_1^n)(g^{m_2} \cdot r_2^n) \\ &= g^{m_1+m_2} \cdot (r_1 \cdot r_2)^n \\ &= \text{enc}(m_1 + m_2 \mod n)\end{aligned}$$

Modulo Integer Addition is additively homomorphic on both plaintext and key space which removes the necessity of using a system wide key:

$$\begin{aligned}\text{enc}_{k_1}(m_1) + \text{enc}_{k_2}(m_2) &= m_1 + k_1 + m_2 + k_2 \\ &= \text{enc}_{k_1+k_2}(m_1 + m_2)\end{aligned}$$

Schemes like those mentioned above are suitable when a specific operation on the plaintext is desired. This is used in all kinds of data aggregating contexts, like secure voting schemes or privacy preserving methods to evaluate sensor values such as location or consumption data.

Sometimes applications necessitate multiplication as well as addition operations, like in the case of outsourcing the processing of private data to the cloud. This is were fully homomorphic encryption schemes are required. Nowadays these schemes are mainly based on lattices and have the drawback of impracticability due to their high computational effort. As the aggregation of measurement data in the Smart Grid only requires aggregation by addition, those will not be investigated further at this point. More information of this topic can be found in [40, 41].

3.1.2 Related Work

Do Not Snoop My Habits: Preserving Privacy in the Smart Grid

A central point of this work is to analyse an approach published by Mármol, Sorge, Ugus and Perez in May 2012 [48] and discuss adaptions to suit the TR-03109. The publication addresses the problem of privacy loss for customers in a Smart Grid when they transfer their consumption data to the electricity supplier. The paper was followed by a more elaborated and detailed report by the same authors [49] which will also be taken into account here.

In general, aggregation methods try to solve privacy threats by summarizing meter data from multiple customers and sending an already aggregated value to the supplier, so he cannot derive the consumption profiles of a particular household. Hence, a remaining problem is that the aggregation instance has to be trusted because it gets all the measurements and could easily abuse this knowledge.

In the proposed method, the solution to overcome this drawback is a mechanism where the individual customers provide their consumption data directly to the electricity supplier after they encrypted it with a personal key. The key itself is not directly transmitted to the supplier but to another smart meter which acts as a key aggregator for a group of households. This meter sums up the keys from all associated meters and transmits this aggregated sum to the supplier. The role of the key aggregator may change every interval and can be determined by cryptographic election protocols like proposed in another publication of some of the authors [46].

The encryption beforehand is based on a bihomomorphic encryption function developed by Castelluccia et al. [24], described later in this section. It is additively homomorphic in plaintext and key space. This enables the supplier to decrypt the total sum of the values from the contributing meters without gaining knowledge over the individual measurements.

For the authorization of an anonymous smart meter to send data to the electricity supplier, the authors suppose to use either group signatures or anonymous credentials, like the ones proposed in [8, 22, 23, 25, 65]. This allows to authenticate the smart meter for the electricity supplier and detect man-in-the-middle attacks.

The aggregated key has to be sent only once to the supplier when no changes to the meter topology occur and stays the same all the time in general. On the other side, the individual keys of the meters are updated in every interval to enhance security. Here the authors utilize the fact that only the sum of the keys has to stay constant to enable a correct decryption of the aggregation result. This can be achieved if all smart meters choose their own random value to add to their old key. After updating it they transmit this value to another meter. This one can then subtract the received difference from its individual key. As a result the overall sum stays constant. An update of the master key is only inevitable if a meter joins or leaves the group.

In the case a meter fails or acts maliciously, the missing value has to be detected. Otherwise the electricity supplier would not be able to decrypt the aggregated value anymore. To tackle this issue, the paper suggests to introduce a so-called *tokens solution*

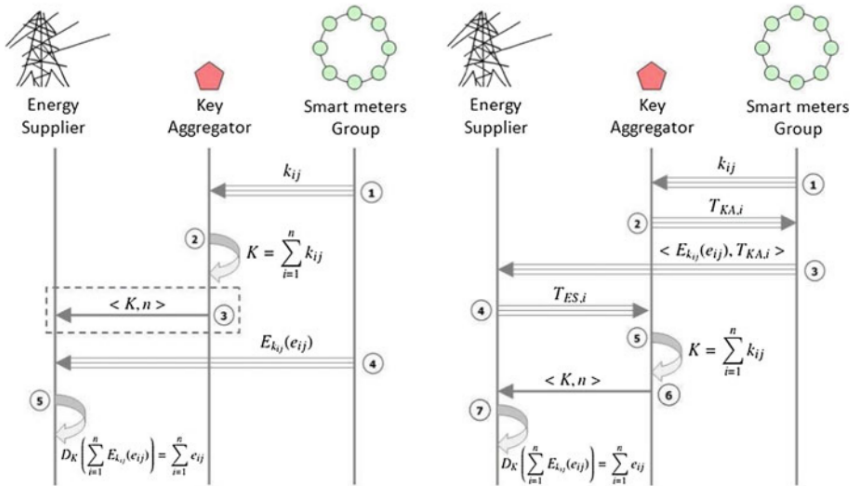

Figure 3.1: Data flow in the approach of Mármol et al. in normal operation (left) and under use of the token solution (right) [49]

which uses an additional token that is generated by the aggregator and sent to the meters after he received their key share. The meters send the token to the supplier together with their measurement. He, as a consequence, forwards the transmitted tokens to the aggregator. This one can now choose the key shares mutually confirmed by a token to aggregate the new master key. The token solution is only intended to be executed in the round after an error was detected. A comparison between the normal operation with the token solution is illustrated in Figure 3.1.

The approach of the authors sounds promising, especially due to the relatively low overhead in data transmission and encryption, which is based on simple additions. Still some problems arise when looking at the suggested scheme:

The authors state that it is not of concern if the aggregating smart meter can be trusted. For instance, a malicious meter could not only transmit the aggregated keys to the supplier but also the individual ones. In the paper it is said that in this case trying out all possible combinations of keys and encrypted measurements to find out the sender's identities would be *"computationally expensive and therefore not worthwhile"* [48]. This statement seems not very plausible because in reality the supplier would have to check $\frac{n(n+1)}{2}$ possibilities to assign each key to the corresponding measurement. As a result, if we assume a group size of for instance 20 households he would have to perform 190 additions and check the results for plausibility. That does not appear to be an excessive computational effort. Therefore in the succeeding paper it is suggested to keep logs of transactions to enable an auditing of entities in order to notice misbehaving.

Another issue addressed by the authors themselves is the possibility to identify a sender on base of other information like his IP address. Here additional protection mechanisms (e.g. onion routing) would be necessary to prevent unwanted identification.

Extra measures to establish a secure channel, like TLS, are also recommended within the latter publication.

Identification of single grid members could also occur if the electricity supplier abused the proposed token solution. There is no measure to prevent the supplier from stating that he only got measurement values from a single (or at least a very small group of) meters. That could reveal the identity of these meters if there is no detection of such an abuse.

Above all this, the most obvious threat is posed by potentially malicious key aggregators. In the second paper a thorough security analysis is conducted. It shows that for more than one malicious meter present, the change in an honest meter's measurements can be leaked. A meter would therefore have to be assumed as a trusted device, which states a similar precondition as a trusted third party.

Besides from problems addressed concerning the scheme itself, it has to be taken into account if it is applicable for a German Smart Grid infrastructure at all because in the BSI protection profile and technical guideline some regulations are made which contradict the suggested solution. Those will be described in detail in Chapter 5.

Secure Data Aggregation with Multiple Encryption

This paper published in 2007 by Melek Önen and Refik Molva [53] focusses on an aggregation mechanism that allows to aggregate values in a rooted tree of sensors without letting the single nodes know the content of the measurements they are aggregating.

An intermediate node – i.e. no leaf node or the sink – aggregates all received values with his own measurement by performing modulo addition on them, utilizing additively homomorphic encryption. It then decrypts the message partly with the keys it shares with its m-th children, where m is the distance to the nodes with which secrets are shared. After that it adds the key shared with the m-th parent node and forwards the message to its direct parent node. This scheme leads to the characteristic that each intermediate node can only strip of a certain part of the secret key but never the whole, unless the node is the sink.

The key distribution in the proposed scheme works as follows: Each intermediate node shares a secret and, to enable an automatic key update, a counter for the children nodes with distance m and its parent node, also in distance m. If the node is a leaf, secrets are shared with all parents up to distance m. Due to the usage of a counter, keys can be updated at each encryption operation.

In case of a non-responding node, an alert is made to all the nodes maximally m away from the fail-node. This allows the affected nodes to remove the keys shared with this node in order to prevent decryption errors in the end.

Goal of the authors was to propose a possibility to ensure data aggregation with an end-to-end encryption scheme without causing a lot of data transmission overhead. Therefore they developed a new additively homomorphic CTR-based aggregation protocol.

Regarding the security level, the encryption scheme ensures confidentiality for a measurement of a node as long as an attacker does not compromise all nodes in a subtree

with the attacked node as root and depth $m - 1$. In that case he could remove all the secret keys to unveil the real value of the node's measurement. So the security level is directly dependent on the size of m. A problem is, that if m on he other hand is very large, at the same time the number of encryption layers increase and the method becomes inefficient.

Another issue is the handling of authentication. In the paper the assumption is made that there would be an additional mechanism to ensure authenticity of nodes outside of the scope of the presented work. At the same time it is stated that the injection of wrong messages from authorized sensors cannot be detected instantly and only due to implausible values afterwards.

Additionally, if taking a look at the case of node failure, it is evidently problematic when a node near to the sink node stays silent. In that case, a large subset of measurements would not be accessible caused by a single point of failure.

Efficient Aggregation of Encrypted Data in Wireless Sensor Networks

Claude Castelluccia, Einar Mykletun and Gene Tsudik propose a method for encrypted data aggregation with a bandwidth-friendly encryption scheme [24]. It is based on modular addition within an additively homomorphic stream cipher.

Like in [53] described before, the network is supposed to have a rooted tree topology. In opposition to the former scheme, single measurements are encrypted by the nodes themselves individually and then added to the aggregation of received encrypted measures from the child nodes before propagating them towards the sink node.

This requires the single nodes to use an individual key to encrypt their data. This key is generated from a long-term key which is shared with the sink node since the beginning. The keystream cipher to produce the current keys can be chosen freely, RC4 is given as a suggestion by the authors, where the id to generate the unique key could be for instance timestamps. The master key from which the sink generates the single node's secrets can then be utilized to decrypt the aggregate.

If a node fails, it is intended that the parent node not receiving an expected value appends the ID of the failing node to the message. So the sink can subtract the key material shared with the defective node from the previous master key.

The proposed scheme appears very lightweight while at the same time providing end-to-end encryption between the single nodes and the sink. Drawbacks in the method can be seen if the sink itself is the attacker. In the case of listening to the values transmitted from the leaves, he would be able to decrypt those without any problems because he knows all the secret keys.

Another issue is the case of failure handling because of the assumption that the sensor IDs have to be known by the neighbours. Thus overhead would be enlarged if many sensors failed. But it is questionable if this really is a relevant drawback for the usage within smart meter networks which are on one hand not too big and on the other hand not expected to fail in such a large degree.

Like in [6, 48, 53], authentication has to be provided by additional measures – hop-by-hop authentication is proposed here – and the correctness of measures cannot be guaranteed.

A lifetime-optimized End-to-end Encryption Scheme for Sensor Networks allowing In-network Processing

The method proposed here [6] has the same goal like the ones mentioned before: Providing an encryption scheme for sensor networks with the aggregation of data and a minimal overhead to respect the limited resources. Special attention is given to silent sensors which will not respond and therefore have to be handled with.

The network topology is supposed to be a tree here, too, with each node being aware of its direct parent and children. Nodes, beginning from the leaves, encrypt their measurement values and send them to their according parent node. This parent node aggregates all received values and adds its own encrypted measurement. This procedure continues until all encrypted data reaches the sink. The sink is then able to decrypt the aggregated value using a master key.

Because the encryption function is bihomomorphic, which means homomorphic in key and plaintext space, this master key is the aggregation of the single nodes' keys. To achieve this, key distribution is organized as follows: During initialization the sink generates a random master key and splits this key, also randomly, into as many parts as it has direct children. These shares are then propagated towards the children which subtract a random key for themselves from this master key and split the rest for their children like the sink did before. This repeats until the keys reach the leaves.

For updating the keys after each aggregation phase, the paper proposes two suggestions: First a coordinated procedure where the sink sends a difference to the old key and this difference is split and spread like the master key before.

The second one is a scheduled variant where no further communication is required. This holds due to the assumption that during the initialization phase, multiple keys have been generated and those are updated following a pre-known function.

In the case of silent, failing nodes there are also two different solutions presented. One is to let the sink generate pre-known dummy values that can be inserted instead of the missing values. These could be recognized and removed by the sink afterwards. The other one is to use the last known value transmitted by the node, with the drawback of leading to slightly inaccurate results.

The proposed method shows good characteristics with respect to the transmission overhead, especially if the autonomous key update is used and the only necessary communication is the transmission of the encrypted, aggregated values. Main drawback of the method would be some security issues. Even with a ciphertext-only attack, information about the plaintexts can be leaked if an attacker can eavesdrop the encrypted messages.

Under this aspect it has to be seriously discussed if such a level of security is still sufficient in a smart metering scenario or if an eventually higher overhead using another method would be more suited with respect to a stronger encryption.

3.2 Alternative PETs in Smart Grid environments

In the following, a brief overview over privacy preserving approaches will be given which utilize differing attempts to hide a user's actual consumption values.

3.2.1 Trusted Computing

Petrlic [55] proposes a method for privacy preservation under the assumption that a tamper-proof metering device, a Trusted Platform Module, is used. Those devices send their readings to a trusted grid operator under a pseudonym. This operator can then aggregate the values and forward them to the electricity supplier, thus masking the IP addresses of the single metering devices. Because the meters can be verified as trusted, bill computation will be accomplished inside the meters themselves and will leak no further information about the consumer profile. While this approach might seem practicable, it implies trust in the grid operator which is not be applicable in all cases. Then a deanonymization using the meter's addresses could be possible.

3.2.2 Adding Noise

The approach of adding noise to a smart meter's sent measurements is described in [59]. Prior to sending, values are distorted following a Gaussian distribution. Therefore a masking of the direct values can be achieved while in average the real value can be read – be it for one meter over a longer time or for a neighbourhood of meters at a single point of time. As this is a statistical approach, accuracy will not reach 100%. Nevertheless the level of accuracy is determinable and controllable by choosing parameters, whereas better accuracy counter-correlates with the quality of masking.

Papadimitriou et al. [54] describe an approach which overcomes the drawback of possibly filtering out white noise. The authors apply perturbation of data's time series. The actual manifestation of perturbation is determined by the properties of the data itself to minimize leakage of information. Adaptability to streaming time series is achieved by a wavelet perturbation scheme.

3.2.3 Masking of Temporal Consumption through Batteries

The last approach mentioned here is to mask the consumption profile of customers by using batteries. A majority of these suggest to meet the customer's electricity demand not directly from the network but from a battery which is recharged and discharged in a way so that it does not leak information about the time of consumption. Kalogridis et al. [37] use a four step model they call *Best Effort Algorithm* that changes the external load whenever the battery is full, empty or there is a rise or fall in demand. A more fine-grained method named *NILL algorithm* [64] utilizes twelve different states for controlling the battery load behaviour. Both methods result in a load profile formed of discrete steps which does not allow a direct derivation of appliances in use, but does not fully hide the presence of electricity usage or times of rising demand.

Similar to the artificial addition of noise, Backes et al. [7] describe another battery based technique where noise is not generated by the smart meter itself but by a battery. It is loaded or discharged randomly and masks the actual consumption. In their paper the authors discuss a noise cascading approach and analyse capacity and throughput constraints.

While the described proposals in the section above aim to pursue the same target like aggregation concepts, they might have some disadvantages in direct comparison. The Trusted Computing approach seems promising but contradicts the target of this work to put as little trust in an external party as possible. Additionally, the possibility to build a perfectly tamper-proof device without back doors for net operators or cheating customers is a very critical question.

In the case of noise-adding approaches, privacy always depends directly on the parameters chosen for the model. If those are picked unfavourably - it does not matter if intentionally or not - information about the real consumption values can be derived. Furthermore, the method would be problematic if the net operators or suppliers were in need of precise values, whereas this still has to be investigated.

The described set of battery approaches renders critical in two points. First in matters of information leakage which occurs in extremal cases of full or empty batteries. While the last work overcomes this weakness, still the issue of energy efficiency and sustainability remains when using batteries. Particularly the latter is a questionable point. Batteries impose problems due to restricted service life and production as well as disposal impact. Under these circumstances, sustainability as a central goal of the Smart Grid is brought into question.

This chapter provided basic information to be able to understand technical measures for the preservation of an individual's privacy and compare different concepts. Combining this theoretic knowledge and the insights on Smart Grids taken from the previous chapter, only one point to build a real world scheme for privacy enhancement is missing: the legal and regulatory means existing in the country the scheme is used. An analysis of the legal reality in Germany is therefore conducted subsequently.

4 Legal and Technical Regulations for the Smart Grid

The implementation of the Smart Grid is not just a challenge concerning the technical aspects but also concerning regulatory and legal measurements to provide rules. Those are not only necessary to prevent unwanted security and privacy problems. They also help to support interoperability between different manufacturers and components. Therefore this chapter addresses the legislative and regulatory circumstances relevant for this work as they are at the present point of time[1].

4.1 Legal situation

To be a subject to regulatory means in the sense of data protection, data has to be identified as personal data according to the Federal Data Protection Act (BDSG) [26]. Following § 3 I this is defined as follows:

> *" 'Personal data' shall mean any information concerning the personal or material circumstances of an identified or identifiable natural person ('data subject')".*

While power consumption definitely belongs to personal circumstances, especially when profiles can be created, this applies for the subject of smart meters. In the most negative imaginable case those profiles can be abused to spy on individual data subjects and derive their daily routines and habits from the meter readings.

Another critical aspect of smart metering is the loss of control to the instances accessing personal data through the utilization of remote meter readings. In § 4 II the BDSG states that data has to be collected directly from the data subject if there are no additional justifications to act differently. At a first glance, remote meter reading dissents that principle and weakens the customer's ability to control the dissemination of his data. It is hard for him to reproduce who is accessing his measurements without a significantly higher effort than in the conventional procedure. Furthermore, with an actual reading interval of 15 minutes the principles of data reduction and data economy as mentioned in § 3a BDSG can be harmed.

All the crucial points mentioned in the previous paragraphs illustrate that a data subject's privacy can be seriously threatened by smart meters if no proper precautions are taken. A deep and thorough analysis concerning this is made in the *Protection*

[1] Refers to the beginning of this work on 1st April 2013 respectively version 1.0 of TR-03109.

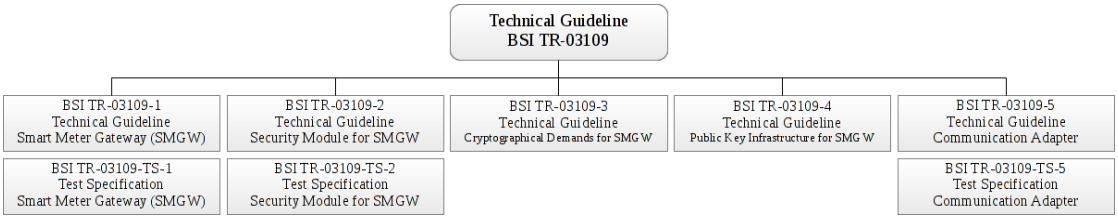

Figure 4.1: Structure of technical guideline TR-03109 [13]

Profile for the Gateway of a Smart Metering System issued by the BSI. It will be briefly described in the following section.

4.2 Technical Guideline TR-03109 and Protection Profiles by the BSI

The BSI (Bundesamt für Sicherheit in der Informationstechnik - in English "Federal Office for Information Security") is responsible for publishing guidelines and regulations describing the requirements regarding security in Smart Grids. On the basis of possible threats for a secure and privacy friendly operation, they define protection measures against those. With the help of the published guidelines, concrete products can be tested and certified.

Currently there exist two complementary protection profiles regarding Smart Grids - one describing the smart meter gateway and another focussing specifically on the security module of this gateway which is responsible for storing the cryptographic material and providing cryptographic methods.

Further technical details on how to implement the requirements in the protection profiles are then stated in the technical guideline TR-03109 – *Requirements for the interoperability of the communication unit of a smart metering system* – and its appendices. This also supports the interoperability between different systems and products by the description and definition of transmission protocols and data structures.

It may be noted that the statements given in the following paragraphs refer to version 1.2 (Final Release) of the protection profile, respectively version 1.0 (Final Release) of the technical guideline TR-03109, both issued on 18th March 2013. While details might be updated or corrected, the author proceeds from the assumption that the general statements will stay valid.

Furthermore, the guidelines cover many aspects of smart metering systems. To keep the focus of this work, especially those who are essential for the implementation of a privacy friendly mechanism to submit detailed measurement data will be discussed.

4.3 Meaning of the Technical Guideline for Privacy-Enhancing Approaches

Protection profile and technical guideline already target at the protection of consumer's data. Nevertheless, there is a need for additional privacy protection. Even pseudonymous data can leak the actual identity of individuals through the combination of different information sources. This especially applies when the data-processing instances are not completely trusted and a malicious cooperation between multiple parties cannot be excluded.

In the preceding chapter, methods for preserving the privacy of a customer by using homomorphic encryption were introduced. All of these methods rely on data aggregation, preventing the electricity supplier from knowing individual consumption values.

To work with these methods, the data to be aggregated has of course to be transmitted to third parties, in most of the described proposals to other sensors. In this case that would imply a direct communication between two smart meter gateways. That this aspect of the proposed methods is problematic will be illustrated in the following paragraphs, together with other relevant aspects for an implementation meeting the regulation's demands.

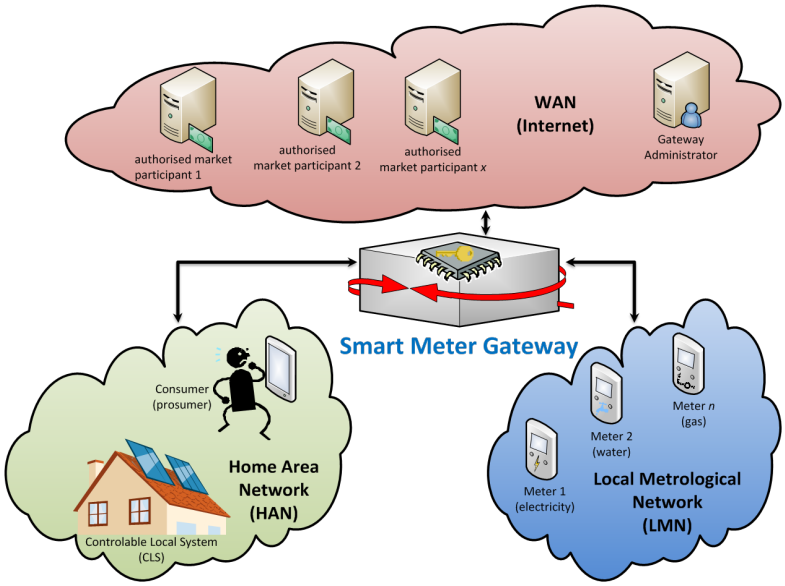

Figure 4.2: Smart Meter Gateway Environment according to TR-03109 [66]

According to TR-03109, the smart metering architecture consists of four main components (see Figure 4.2). Core component hereby is the smart meter gateway, further referred to as SMGW. It serves as a communication unit to store, process and distribute measurement data from the Local Metering Network (LMN) among authorized external market participants:

> *"The TOE[2] is a Communication Gateway. It provides different external communication interfaces and enables the data communication between these interfaces and connected IT systems. It further collects, processes and stores Meter Data."* [12, p. 14]

Moreover, it constitutes the display and monitoring unit for service technicians and customers at the Home Area Network (HAN) interface where it also has proxy server functionalities. It plays an integral part in ensuring privacy and data security because it also acts as a firewall, grants that communication only takes place to trusted entities and separates WAN, HAN, and LMN. Therefore, a direct communication between two gateways is not possible:

> *"The Gateway shall separate devices in the LAN of the consumer from the WAN and shall enforce the following information flow control to control the communication between the networks that the Gateway is attached to: only the Gateway may establish a connection to an external entity in the WAN; specifically connection establishment by an external entity in the WAN or a Meter in the LMN to the WAN is not possible."* [12, p. 19f]

This clearly aims to prevent the SMGW to be attacked from the WAN. But to enable communication from authorized, preconfigured entities to the SMGW a so-called Wake-up Call is introduced. External participants can place a request to communicate with the SMGW to the SMGW administrator. He will then contact the SMGW and tell it who is desiring to establish a connection. The SMGW can then establish the connection by itself if it is confirmed as authorized. Moreover, every connection has to be cryptographically protected, i.e. encrypted, integrity protected and mutually authenticated.

TR-03109-1: Anforderungen an die Interoperabilität der Kommunikationseinheit eines intelligenten Messsystems (Requirements for the interoperability of the communication unit of a smart metering system)

This part of the technical guideline describes how the smart meter gateway can be embedded into a smart metering system. While one of the gateway's main tasks is to handle communication to external entities, the guideline provides rules addressing these communication processes.

[2] Target of Evaluation

Roles of members in the Smart Grid [14, ch. 2.2, p. 13] Besides the service technician there are three kinds of actors defined which should be able to communicate with the SMGW:

- **Consumer:** The consumer owns the measurements in the smart meter gateway and should therefore have all-time access to read them. It should also be mentioned that one smart meter gateway can be used by multiple customers (ch. 2.3.4). Under these circumstances it must be able to handle different users and separate profiles.
- **SMGW administrator:** He is responsible for configuring, supervising and controlling the SMGW. Administration tasks belong to the area of accountability of the measuring point operator.
- **Authorized external entities:** Other stakeholders can request connections to the SMGW as well, if they are known and registered in a list of known communication partners. Authorized external market participants (EMP) are those participants in the WAN (except the SMGW administrator) with whom the SMGW can establish a communication. These are for example the distribution system operators, metering-point operators, measurement services providers, retailers or other authorized service providers.

SMGW in the WAN [14, ch. 2.3.2, p. 16ff] By default, the SMGW ignores communication requests from members of the WAN. On the opposite it can build up a connection on its own behalf at any time towards the SMGW administrator in order to receive his commands. Alternatively, a connection can be kept running constantly or initiated by a wake-up call sent by the SMGW administrator [14, ch. 3.2.5, p. 38ff]. This wake-up call happens if the SMGW administrator sends a specific wake-up packet to the gateway. This packet contains the address of the communication requester and the gateway can build up a connection to this entity if it is configured as trusted.

Application scenarios for WAN communication [14, ch. 3.2.2, p. 20ff] A SMGW is only allowed to establish a WAN connection if it serves one of the purposes mentioned in the guideline :

- Administration and configuration of the SMGW by the SMGW administrator
- Access of the SMGW to services offered by the SMGW administrator (e.g. time synchronization and firmware updates)
- Alarming and noticing the SMGW administrator in case of events, regular or unexpected, in the SMGW
- Transmission of data to the SMGW administrator either meant for himself or for third parties
- Transmission of data to external market participants

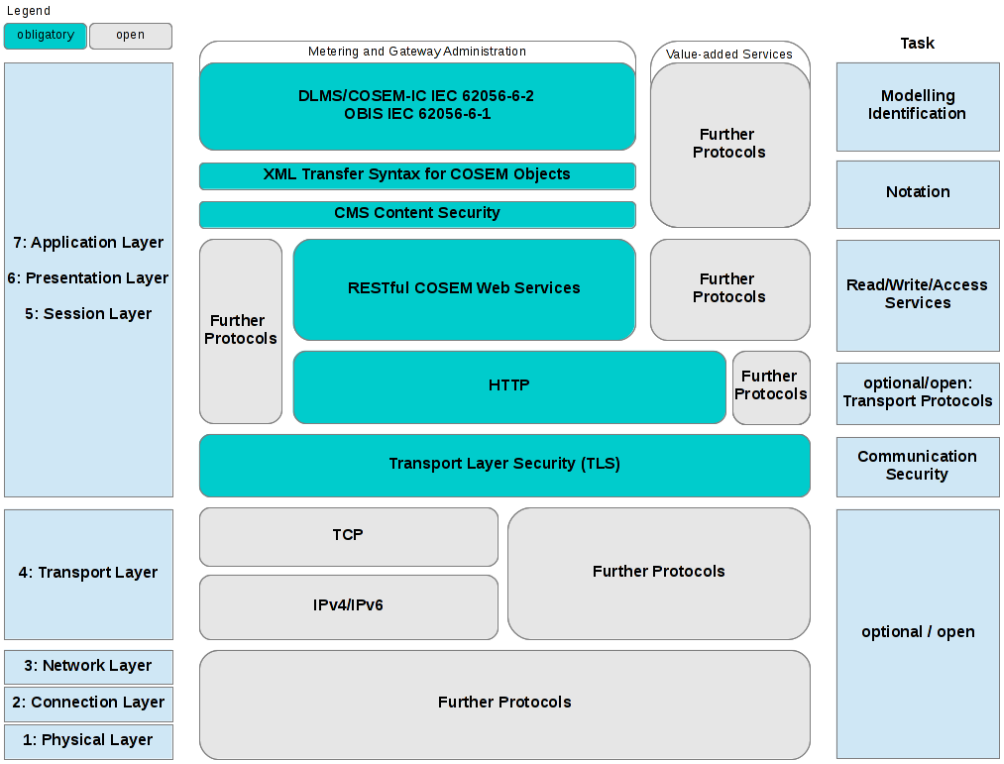

Figure 4.3: Protocol stack for WAN communication of a SMGW [14]

- Communication of external market participants with a controllable local service where the SMGW acts as a proxy
- Wake-up service

Those use cases are divided into the five types MANAGEMENT, ADMIN-SERVICE and INFO-REPORT, NTP-HTTPS and NTP-TLS [14, ch. 3.2.3, p. 24].

The transmission of data in the aggregation context covers the INFO-REPORT scenarios for values designated to the external market participants and ADMIN-SERVICE for the features provided by the administrator, e.g. pseudonymization as it will be described in the following paragraph. In both cases, the SMGW acts as a TLS-Client and has to send its data via a web service request to the communication partners [14, ch. 3.2.3, p. 28f]. The protocol stack demanded for this is shown in Figure 4.3.

Pseudonymization [14, ch. 3.2.4.3, p. 35f] Data not necessary for billing purposes which is transmitted to external entities generally has to be pseudonymized . The final receiver thus gets his data via a third party, the SMGW administrator. The personal ID of the meter is interchanged with a pseudonym and transmitted to the administrator after encryption and authentication for the EMP, additionally signed for the administrator.

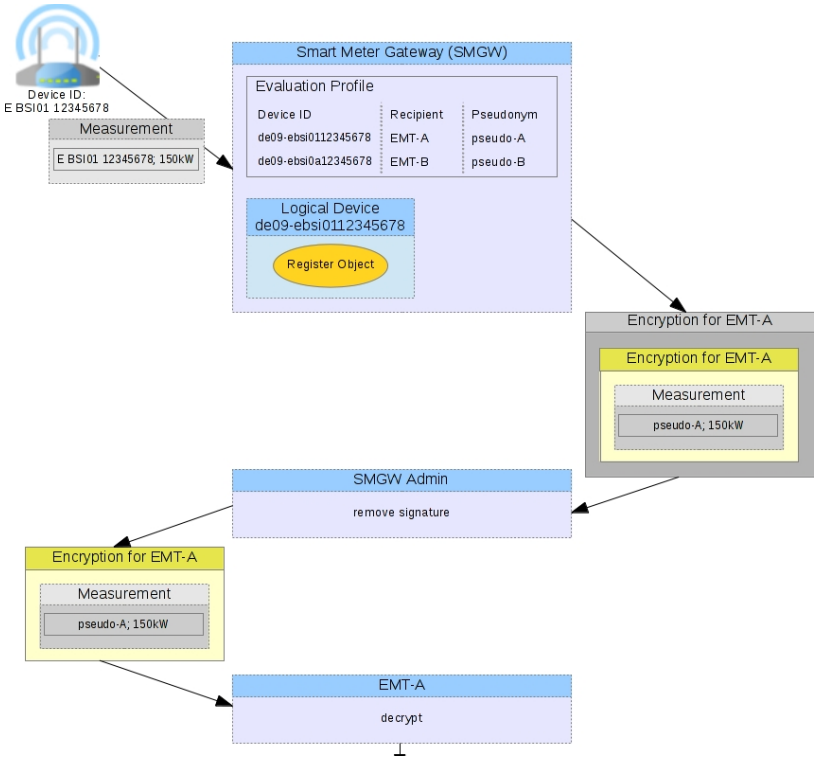

Figure 4.4: Pseudonymization of measurements according to the TR-03109 [14]

He then checks and removes the signature and forwards the measurement value to the final recipient (also see Figure 4.4).

Securing communication with the WAN [14, ch. 3.2.4.4 and 3.2.4.5, 36ff] All content sent via WAN has to be encrypted using Cryptographic Message Syntax. Additionally, the communication with the WAN has to be secured by using the TLS protocol. The gateway has to establish the TLS channel by itself and acts as a TLS-client. Besides that, it may only react to wake-up calls from its administrator and not accept calls from other participants. If those wish to communicate with the SMGW, they have to contact the SMGW administrator first who can send a wake-up packet to make the SMGW initiate a connection if the EMP is known.

Evaluation and communication profiles [14, ch. 4.5.6, p. 119] For every kind of data accruing in the SMGW there must be an evaluation profile in case this data shall be processed further. If additionally a transmission of data is foreseen, those evaluation profiles may contain one to many communication profiles. Only those external market participants are authorized to get information from smart meter gateways that are declared in such a profile. All profiles can only be announced to the SMGW by the administrator and have to contain information about the communication purpose, addresses, keys and certificates. So no spontaneous communication to formerly unknown EMPs will be allowed.

4.4 Regulatory Situation in Brazil

As this study is situated in context of the TruEGrid project, a look at the Brazilian regulatory framework is also important. Major stakeholders influencing regulations around the Brazilian Smart Grid implementation are the following organizations:

- **ANEEL**, the Brazilian electricity regulatory agency. It is an institution linked to the Brazilian Ministry of Mines and Energy. Their task is to regulate the electricity market to keep the relation between agents and customers balanced.
- **ABNT**, the Brazilian Association of Technical Standards. In their responsibility lies the establishment of certification criteria for smart meters.
- **INMETRO**, the National Institute of Metrology, Quality and Technology works together with ABNT in certification. Additional scope of their responsibilities is to provide actual guidelines, specifications and policies to design meters fulfilling the postulations demanded by ABNT.

To get a certificate from INMETRO, the following demands on security are stated in *Portaria no. 375* [43]:

> *"A password protection for the access to programmable meters should be available in order to prevent unauthorized access, avoiding unauthorized changes on metrological parameters and the registered information, in case the sealing of the optical port is not available. The meters which do not have physical seals or password must have a counter and data logging of events that always allow the identification and verification of the at least eight last operations occured."*[3]

[3] The above is a free translation of the original section written in Portuguese:

> *"A proteção de acesso via senha, com código de segurança, deve ser disponibilizada para prevenir o acesso não autorizado aos medidores programáveis, evitando mudanças não autorizadas nos parâmetros metrológicos e no arquivo de informações registradas, quando não houver dispositivo de selagem da porta óptica. Os medidores que não possuírem lacre físico ou senha devem possuir contador e registrador de eventos que permita identificar e verificar, a qualquer momento, no mínimo as oito últimas intervenções ocorridas."*

ANEEL also reasons about security requirements in its resolution on smart metering [4] and gives the following recommendations:

> *"In case of the measuring system is provided with a remote communication system, the distributor must adopt procedures and technologies that ensure the security of data traffic, and especially the personal information collected from consumer units.*
> *Unique paragraph: The electrical energy distributor is forbidden to share data collected from consumer units to third parties without the authorization of the owner."*

It can easily be seen that the level of detail given in these regulatory documents reaches the technical guideline by no means. From the aspect of German regulation as it is today, it can therefore be taken for granted that, if the specifications of TR-03109 are satisfied, the Brazilian ones would automatically be fulfilled, too.

This chapter tried to provide information about regulatory circumstances connected to the transmission of personal data, specifically energy consumption values. While it is obvious that those are too extensive to be covered completely, focus is laid on the aspects decisive for the understanding of restrictions and potential difficulties in connection the suggested solution approach which will be outlined in the following part.

5 Development of a Privacy Friendly Aggregation Scheme in Conformity to the BSI Protection Profile

In this chapter the practical implementation of the scheme proposed by Mármol et. al will be looked at. In the previous chapters it has been already appearing in outlines that changes have to be made if conformity to the TR-03109 is striven for. These alterations and the resulting communication flow will be described first. After that, a description of the developed prototype will be given. Connected to this, the additional protocol requirements introduced by the technical regulation are described together with their practical realization.

5.1 Protocol Description

Like described in the chapter before, the first and most obvious restriction if one looks at the technical guideline is, that a direct transmission of data between two gateways is not possible because a SMGW may only accept TLS connections that are initiated by itself or via a wake-up call of the SMGW administrator [14, ch. 3.2, p. 20ff]. Furthermore, the technical guideline [14, ch. 4.4.4, p. 115ff] states that transmissions to external entities may only happen if they are authorized external market participants previously specified by the SMGW administrator in a communication profile. These prerequisites advise the aggregation by an external entity different from the SMGW itself. Another significant difference towards the model described by Mármol et. al. is, that the technical guideline intends a gateway administrator as an additional entity. This can provide some benefits and balance out arising difficulties.

The values which are sent in short time intervals to help the supplier with load balancing and alike are not relevant for billing. A pseudonymous transmission of all sent information, like already mentioned in chapter 4.2, is therefore advised. Here the administrators come into play, acting like a proxy server between the gateway and external market participants, forwarding the end-to-end encrypted messages of the gateway to them without being able to read the actual content. While an external entity can help to overcome the problem of communication restrictions, it introduces a new problem if we aim at putting as little trust in external entities as possible.

In this context it is important to keep in mind that the SMGW administrator must be assumed as trusted, no matter which privacy preserving protocol is implemented. Otherwise he could instantly deanonymize gateways for the electricity supplier ES and the external aggregator, referred to as AGG, which would make a decryption of the

single homomorphically encrypted measurements trivial and render all efforts useless. Therefore ensuring the administrator's reliability and trustworthiness is out of the scope of this work. To achieve an adequate level of organizational security, a detailed description of the demands regarding operations and processes of the SMGW administrator is specified by the BSI in [15]. It may be remarked that - even in case of a misbehaving administrator - the end customer has the possibility to view all of his activities in the customer log. It is created by the SMGW and not accessible by anyone than the customer himself, enabling the detection of potentially malicious SMGW administrators unless the gateway is tampered.

Taking the above factors into account, a system will be described which can minimize the trust needed regarding the external market participant while still utilizing data aggregation to preserve the privacy of single users. First, the standard procedure will be described, assuming that all gateways are available and behaving correctly. An insight into the toleration and detection of errors and special events will be given afterwards.

Prerequisites and Notation

In order to enable the aggregation, it is assumed that a group of n gateways participating in the aggregation is determined before the protocol starts to run. That is likely to happen by contracting with ES and the external market participant performing the aggregation of keys. As a consequence, the goal is to transmit the encrypted consumption values ev_{ij} from metering gateway gw_{ik} belonging to SMGW administrator a_k for the given time intervals j to the supplier ES without letting him know the plaintext values pv_{ij}. Therefore ev_{ij} is encrypted with a key k_{ij} which is unknown to ES.

As long as not explicitly stated otherwise, it is assumed that the data arriving at the administrator has been pseudonomized conform to [14, ch. 3.2.4.3, p. 35f], encrypted and authenticated for the end recipient as well as signed for the gateway administrator respecting the Cryptographic Message Syntax like described in Section 5.2.2. Before the administrator forwards the data he checks and removes the signature. It is pointed out that this process implies that the administrator, due to the end-to-end-encryption between SGMW and EMP, gains no insights into the content of the communication.

In the following it is referred to the SMGWs instead of the single meters. Leaving out the step of data transmission between LMN and SMGW simplifies the description of the solution. Transferring the solution however does not imply alterations to the default communication of the measurements to the gateway, which is assumed as secure and therefore not critical for privacy preservation here.

Initialization phase

To start the protocol, all the gateways and entities have to enable secure communication channels to each other and gain information about the process. For the coordinated exchange of this information, the guideline intends the SMGW administrators a_k to provide evaluation profiles. These determine, according to application case and OBIS code, how the data shall be processed. As each evaluation profile can contain one or

more communication profiles, those specify for whom those data packets are destined and declare certificates and key material for the establishment of a secure communication channel, authentication and key transmission.

As soon as this information is declared to every participant, the actual initialization starts by agreeing on the aggregated master key:

1. Every SMGW gw_{ik} generates a random inital key $k_{i,0}$ and sends it to SMGW administrator a_k.

2. a_k forwards all received keys $k_{i,0}$ to AGG.

3. AGG calculates the master key K by performing aggregation operation $\oplus$,

$$K = \oplus_{i=1}^{n} k_{i,0} \tag{5.1}$$

   which, under the assumption that the aggregation operation equals an modulo integer addition, is specified as:

$$K = \sum_{i=1}^{n} k_{i,0} \mod M \tag{5.2}$$

   Here M denotes a shared modulus which has to be distributed together with the initialization of evaluation profiles to all participants.

4. The determined master key K is sent to ES through a secure channel.

Periodical data transmission

After the initialization process, the gateways will start sending their values in the agreed time intervals. Therefore the SMGW encrypts its measurement pv_{ij} with its round key k_{ij} to obtain $ev_{ij} = enc_{k_{ij}}(pv_{ij})$. Expressed under the modulo integer encryption scheme, the concrete instantiation is defined as:

$$ev_{ij} = pv_{ij} + k_{ij} \mod M \tag{5.3}$$

The encrypted value is transmitted to the corresponding a_k and forwarded to ES.

Decryption by the electricity supplier

Hence the steps described above, the ES receives n encrypted values ev_{ij} every round. To proceed, he aggregates these to a total encryption value EV_j using an aggregation function $\otimes$:

$$EV_j = \otimes_{i=1}^{n} ev_{ij} \tag{5.4}$$

Again expressed with modular addition as:

$$EV_j = \sum_{i=1}^{n} ev_{ij} \mod M = \sum_{i=1}^{n} (pv_{ij} + k_{ij}) \mod M \tag{5.5}$$

$$= \sum_{i=1}^{n} pv_{ij} + \sum_{i=1}^{n} k_{ij} \mod M \tag{5.6}$$

The above finally leads to:

$$EV_j = \sum_{i=1}^{n} pv_{ij} + K \mod M \tag{5.7}$$

Because ES already knows K, he can now easily obtain the decrypted sum of plaintext measurements PV_j by subtracting the master key:

$$PV_j = \sum_{i=1}^{n} pv_{ij} \mod M = (\sum_{i=1}^{n} pv_{ij} + K) - K \mod M \tag{5.8}$$

$$= EV_j - K \mod M \tag{5.9}$$

Key management

In [6] the authors propose two methods for updating and managing keys. One is a coordinated key refreshment and the second is an autonomous key schedule.

Key updates are crucial for the application, because otherwise ES can gain additional information about keys and plaintext. An intuitive setting where this occurs is in the case that measurements stay nearly constant. This might happen for example during the night time or the absence of the customers which would directly correlate with equally constant ciphertexts.

Customizing a coordinated refreshment to the given scenario would result in the following communications:

1. Each gw_{ik} selects a random value $\triangle'_{ij}$. This is used as a key difference which is added to obtain a temporary new round key

$$k^t_{i,j+1} = k_{ij} + \triangle'_{ij} \mod M \tag{5.10}$$

2. All differences $\triangle'_{ij}$ are transmitted to AGG via the administrators.

3. AGG accumulates the key differences to an absolute round difference:

$$\triangle_j = \sum_{i=1}^{n} \triangle'_{ij} \mod M \tag{5.11}$$

4. $\triangle_j$ is randomly split into n parts $\triangle''_{ij}$ so that holds:

$$\triangle_j = \sum_{i=1}^{n} \triangle'_{ij} \mod M = \sum_{i=1}^{n} \triangle''_{ij} \mod M \tag{5.12}$$

5. Via a_k each gateway gw_{ik} requests its second key difference from AGG. AGG randomly allocates the differences $\triangle''_{ij}$.

6. The gateways will now obtain their new round key by subtracting the second key difference:

$$k_{i,j+1} = k^t_{i,j+1} - \triangle''_{ij} \mod M \tag{5.13}$$
$$= k_{ij} + \triangle'_{ij} - \triangle''_{ij} \mod M \tag{5.14}$$

One might be asking if this schedule does not threaten the effectiveness of updated keys. AGG could happen to know the total key distances of the single instances and could maliciously forward them to ES. This drawback only occurs if the same communication profile (i.e. pseudonym) is used for transmitting and retrieving the key differences.

If not, the only thing known to AGG is the originating a_k. In the special case that one administrator would host a single gw_{ik}, this provision would be rendered pointless. Measures to overcome this issue will be proposed later.

A further question eventually coming to one's mind is, why AGG does not simply update the master key K by adding $\triangle_j$ and sending the new key K_{+1} to ES. Here a similar threat to the situation completely without key update would arise. In case ES gets to know the single differences $\triangle'_{ij}$ from a malicious AGG, he can search for correlations with similar differences of ciphertexts for constant values. This needs significantly more effort when he has to perform every possible combination of $\triangle'_{ij}$ and $\triangle''_{ij}$.

For the autonomous key schedule as described in [6] no special provisions have to be taken because an autonomous key schedule implies that no additional communication with the gateways is required to update the keys. The only change in the normal process is, that during setup multiple keys have to be distributed per SMGW to initialize the schedule, e.g. by shift registers. In contrary to the former method, the master key has to be recalculated and transmitted every round. The reason we renounce this method is the intensely higher effort to react to group changes by updating the entire key schedule.

Authentication

Mármol et. al. do not point out a specific measure for authenticating a meter, respectively SMGW, as an entitled member of the aggregation group. Their suggestions are to utilize a group signature or anonymous credential scheme to prevent the identification of single devices and legitimize them nonetheless. In their later publication [49], they recommend revocable group signatures to be capable of excluding failing members from the authorized signing group.

In connection to the protocol stack prescribed by the guideline, additional authentication measures are superfluous. By default, the administrator checks the signature of data received from the gateways, adds his own[1] and the end recipient can additionally check the authenticity as part of the TLS connection.

Anonymity

The authentication mechanism above entails linkability of the periodically sent encrypted measurements ev_{ij}. Therefore ES knows that a subsequent message $ev_{i,j+i}$ originates from the same sender as ev_{ij}. Nevertheless, that does not heavily affect the privacy of customers, given that the group size is sufficiently large and keys are frequently updated.

Due to pseudonymization, ES only gets to know which pseudonyms sent him the values. The authors of [48] mention a deanonymization through OSI layers 1 -3 as a threat. In the present proposal this threat is removed by self-implication of the administrators serving as a proxy. As a consequence, tracing on lower OSI layers ends at the SMGW administrator. Because he is non-anonymous by definition, there is no need to take provisions for the integration of techniques like TOR or the Crowds approach.

To cope with the unlikely and degenerated case that an administrator has only one or very few associated gw_{ik} participating in the protocol, anonymity can be strengthened by the creation of virtual gateways gw'_{ik}. The administrator himself would then generate encrypted values for those gateways so that they sum up to zero when they are decrypted. That does not disturb the aggregation result and at the same time protects the real gateways even when ES and AGG know about the virtual gateway's existence. They are indistinguishable from real gateways and their sending of partly negative values cannot be detected as they are encrypted.

Failure handling

When thinking about failures, two main scenarios can be specified: A gateway might remain silent and not transmit a key share or key update, respectively an encrypted measurement. Or it might send false values. In each case, the aggregation would be disturbed and the resulting decryption wrong and useless for the electricity supplier.

Silent gateways Messages of smart meter gateways can be missing due to different reasons: A gateway might fail or maliciously refuse to send, or the communication channel could be disturbed or disconnected so that messages are lost even when the gateway is behaving correctly.

Mármol et. al. suggest a token solution which is initiated if the number of messages received by ES or AGG does not match the expected group size. That triggers a resending of the round, where additionally a token is attached to both key and measurement value. This token is issued and distributed by the group aggregator. Through comparing the

[1] While not explicitly mentioned in the description of the anonymization process in [14], signing of authenticated and encrypted data is mandatory following [19, p. 19].

tokens received by AGG and ES, the key and measurement aggregation can be corrected by aggregating only messages attached with mutually confirmed tokens.

For the solution presented at hand, this provision is not necessary and an alternative solution approach will be stated. The main difference enabling this approach is the existence of the gateway administrators a_k who are aware of their associated gateways gw_{ik}. Furthermore they are able to link each of them to their pseudonyms, which is mandatory on behalf of the technical guideline [14, ch 3.2.2, p. 21]. This is an advantage, because it simplifies error detection and localization in case of silent meters and thus reduces communication overhead.

Gateway remains silent during key initialization If the administrator notices that a gateway does not send its initial key share $k_{i,0}$, the procedure might just proceed as if the gateway would not be part of the group. Nevertheless, a message has to be sent to AGG and ES so that they can adapt the expected group size to prevent waiting for messages from this gateway.

In case the faulty gateway tries to send an encrypted value afterwards, the administrator will notice by determining that this message stems from a non-initialized gateway and hold it back in order to not disturb the aggregation process. The gateway will gain the opportunity to join the group again in the next round and be handled as in the case that a new gateway would appear (see below).

Gateway does not send encrypted measurement value In case ES receives only a part of the expected values, decryption will fail because key sum and ciphertext sum would not fit.

Because AGG has knowledge about the initial key share of the failing meter, a reconstruction of the current meter's key would be possible if AGG stored all key shares and key updates associated to their pseudonyms over time. If the administrator reveals the pseudonyms of the erroneous meter he could reconstruct a new master key without the key the absent meter contributed.

$$K_{new} = K_{old} - \left(k_{i,0} + \sum_{t=0}^{j} \triangle'_{it} - \sum_{t=0}^{j} \triangle''_{it} \right) \mod M \tag{5.15}$$

$$= K_{old} - k_{ij} \mod M \tag{5.16}$$

This correlates to the case in which a SMGW encrypts the value '0' with its key.

A minimal example is given in Table 5.1, where three gateways correctly update their keys during two rounds and after that the third gateway fails and the key sum can be reconstructed without data from the well-behaving two meters.

X	$k_{i,0}$	$\triangle'_{i,0}$	$\triangle''_{i,0}$	$k_{i,1}$	$\triangle'_{i,1}$	$\triangle''_{i,1}$	$k_{i,2}$
SMGW 1	17	2	1	18	7	8	17
SMGW 2	24	3	5	22	5	8	19
SMGW 3	11	2	1	12	9	5	X
$\sum_{i=1}^{n}$ X	52	7	7	52	21	21	K_{new} =?

Table 5.1: $K_{new} = K_{old} - (k_{3,0} + \triangle'_{3,0} - \triangle''_{3,0} + \triangle'_{3,1} - \triangle''_{3,1}) = 52 - (11 + 2 - 1 + 9 - 5) = 36$

More straightforward, a gateway outage could also trigger a complete key reinitialization like already described above. It would be triggered by AGG after he has been informed about the absent meter by the administrator in charge . This also implies that the encrypted values are resent to ES, because obviously encryption keys have changed. A simple regeneration of a meter's pseudonyms is not helpful in this case because the new pseudonyms can easily be linked among the already existing ones.

While the latter demands a higher additional communication effort, it has two striking advantages: Firstly, it removes the need to keep a key update history reducing eventual linkability between synonyms known to AGG and ES. Additionally, required memory space will be reduced, which might be a significant factor if many gateways are contributing over long time periods.

Secondly, it provides the possibility for the failing SMGW to preserve its anonymity. In the first approach it would be instantly unmasked by the administrator, which means that AGG has the knowledge to combine the sending and requesting pseudonym of the gateway. If he cooperates with ES he would be able to obtain all previous encrypted values sent by the affected SMGW. This is possible because he is aware of all three pseudonyms of the missing gateway – one towards himself and one per sending and receiving key share updates to, respectively from AGG – as well as the key shares and encrypted measurements.

Despite the fact that the consumption profile ES might derive only applies for a pseudonym, even this can threaten the privacy of a customer in some cases. This holds because ES is in possession of additional information like household sizes, coarse grained, but identified consumption values previously used for billing purposes and so on. So at least a significant rise of the probability to unmask a customer will be given.

The complete reinitialization can circumvent this in case that not the gateway itself was down, but victim of temporary transmission errors. In this scenario it can just continue participating by proceeding with the key update as provisioned.

Gateway does not participate correctly during key update In the key update process a gateway can either refuse to send its key difference $\triangle'_{ij}$ or miss requesting $\triangle''_{ij}$. or both, which is handled equivalently to the first case.

When the administrator notices that a gateway remains silent during key update, he will issue a report to announce that a gateway has left the group. Should the gateway

anyhow try to send an encrypted value to ES or request $\triangle''_{ij}$, the administrator holds all messages back until the gateway officially reentrants the group.

As for a missing request, the administrator can inform AGG about the associated $\triangle'_{ij}$ previously sent and also hold back ev_{ij}. That enables him to update the master key accordingly:

$$K_{new} = K_{old} - \triangle'_{ij} + \triangle''_{ij} \mod M \tag{5.17}$$

Shall the correlation between sending and requesting be prevented to give the gateway a chance to behave correctly again, the complete key initialization has to be performed equivalent to the second solution described for coping with missing ev_{ij}.

Incorrect messages Preventing malicious meters or gateways from sending faulty messages while preserving the privacy of honest participants is a nontrivial task. It even seems contradictory to the overall goal - to obviate the publishing of plaintext measurements.

Suggested solutions are for instance described in [38], whereas the authors themselves state that their approach is based on a unilateral trust assumption and has therefore to be refined.

The approach presented here tries to preserve anonymity of the SMGWs as far as possible and makes use of the administrator as an instance taken as trusted.

1. Assuming all expected transmissions occurred, ES will decrypt the aggregations and obtain an aggregated value PV_j. Following data he obtained during the past, be it from aggregations or billing data, he can detect if data is probably highly implausible. It is not in the scope of this work to determine a realistic deviation of consumption values where it is justifiable to assume implausibility. As well, the method does not aim at providing the detection of minor aberrations.

2. If he suspects wrong data, he informs the gateway administrators and thus triggers an error detection mode.

3. Every SMGW administrator requests repeated sending of the measurement values from his gateways. In opposite to the regular transmission, now the usage of another evaluation profile is requested. This profile is exclusively activated during the error detection period. As a difference to before, it defines the administrator as an additional legitimated end recipient of a plaintext measurement. Therefore he will now be able to check for correctness by taking into account a suggested investigation procedure in the TR-03109 [14, ch. 4.3.4.2 p. 111].

   It states that in case of the transmission of status data, additional validity checking can be applied by sending data from the local metering network to authorized entities via the SMGW.

4. If the administrator successfully verifies the results and those appear credible, he aggregates all received plaintext values and directs them to ES. How he decides

about credibility depends on his additional information. For instance it is conceivable that he gets some exemplary average profiles for different household types from ES together with tolerated deviation limits.

5. In case that an administrator detects a certainly incorrect message, he excludes it from aggregation. So even in this case, ES gets a correct aggregation of $n-1$ gateways.

6. After this round, all participants proceed with the standard instructions.

What is done concerning the faulty gateway is up to the agreement entered into by the involved parties. One could imagine excluding the gateway instantly from aggregation or alternatively after a certain limit of misbehavings. For whatever procedure chosen, error toleration methods described above can be applied.

Concerning the failure strategies described here, some concluding remarks have to be made. While the solutions always describe the failure of only one gateway at a time, the same applies for multiple synchronous outages. This holds true because the administrators are able to identify the failing components and can coordinate countermeasures accordingly. The number of faulty participants is consequently known to all entities. In case multiple gateways collaborate to send erroneous data, detecting can be impeded. This happens if the values are not significantly implausible for a single SMGW but the single deviations sum up to a serious aberration. But as this states disproportionate effort, this case can be taken as purely theoretic, in particular because not even financial benefits, for example through the creation of a forged electricity bill, are involved.

Furthermore, it can be assumed that the likelihood of a meter actually failing completely is much lower than an erroneous and distorted communication channel. The latter case can be easily compensated if it occurs only selectively and for short time frames. In every scenario an administrator has the possibility to contact a presumably silent gateway by either requesting a resend using the MANAGEMENT communication scenario or issuing a wake-up call which summons the gateway to establish a communication channel. If any of these measures succeed, a further propagation of the initial error can be dismissed and communication overhead as well as the diminishing of anonymity reduced.

Scalability

In the environment of the Smart Grid it is crucial to ensure adaptability of mechanisms to changing network structures. For the concrete application it is therefore important to enable these gateways, as well as new administrators, to dynamically join, respectively leave an aggregation group. The proposed structure facilitates an easy realization of this. Hence, the only prerequisite for reacting towards changes is the distribution of necessary evaluation profiles and the update of the master key, depending on the added key share as well as the announcement of the new group size.

It is recommended that gateways requesting to join a group which belong to the same administrator are added in bulks. This might be implied as a matter of course by contracting participants for periods, e.g. starting at the beginning of a month or week. Otherwise, correlating newly added synonyms is trivial for AGG and ES. If this is not feasible, anonymity protection can be ensured nonetheless by utilizing the same approach like mentioned for administrators with only one meter - adding virtual meters. Unfortunately this comes at the cost of a rising communication effort.

In the case multiple aggregators are present on the market, the allocation of the master key to the encrypted measurements that have to be decrypted with this key can be made by either an additional piece of information sent by as well administrator and gateway or, even more simple, identified via the RESTful interface. The same applies for locality groups if a supplier has multiple separate gateway groups. As an example, a gateway which is in contract with *'amazingaggregator'* and is situated in the locality *'niceneighbourhood'* can send a PUT request to the following address:

```
https://supplier/niceneighbourhood/amazingaggregator/measurements
```

The aggregator can then send the calculated masterkey to

```
https://supplier/niceneighbourhood/amazingaggregator/masterkey
```

and ES can simply map the encrypted values to the corresponding masterkey. Afterwards he can aggregate all intermediate aggregations and obtain the total sum for *niceneighbourhood* without any additional data to be sent.

5.2 Implementation

The implementation of the protocol has the goal to serve as a proof of concept. To have a testing environment, supplier, aggregators, SMGW administrators and SMGWs are emulated. The reader has to keep in mind that a for a system operating in a real-world environment, the whole implementation would have to be ported to an existing SMGW hardware platform.

5.2.1 Structure

As all participants communicate via WAN, they are modelled as Java Servlet Web Applications – each deployed to an *Apache Tomcat 7.0.42* web server. From the WAN, their functionality is reachable via their web service interfaces, shortly described in the *RESTful/COSEM Webservices* paragraph below. Tomcat assigns a thread from its thread pool for each incoming request. That requires synchronization of data, achieved by a Storage Manager handling the access to the data objects. A schematic view on the server and data storage architecture of one simulated Smart Grid participant can be seen in Figure 5.1. It shows that the COSEM XML data included in the request is

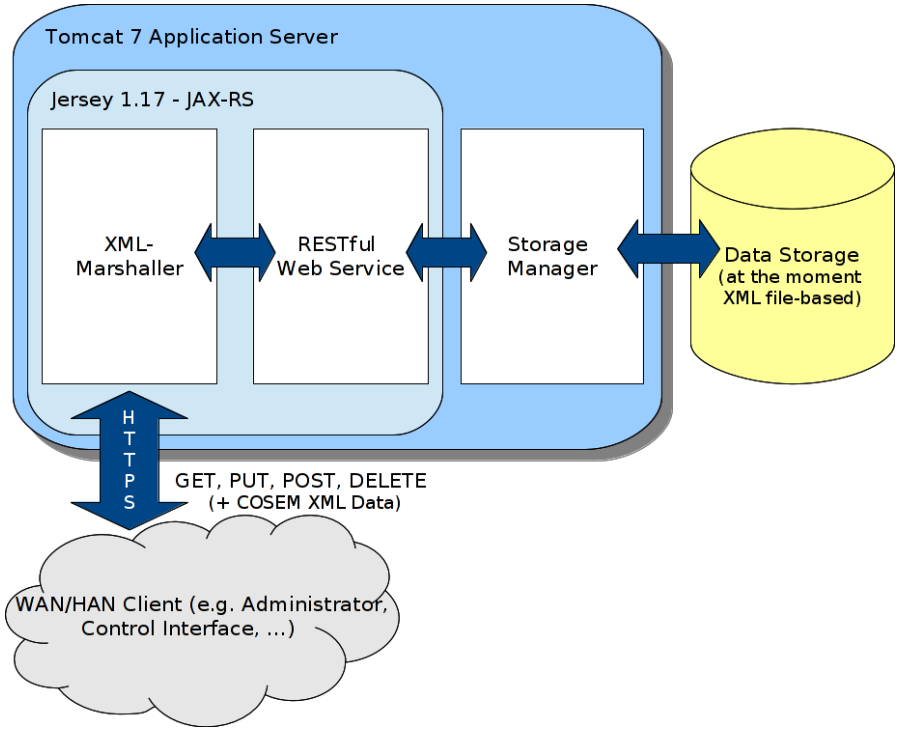

Figure 5.1: Structure of the web application representing a Smart Grid participant

first unmarshalled. Outcome of this processing step are deserialized Java objects which can be handled by the application. These are transmitted to the core web service and processed. To build the response for the requesting entity, the same procedure is run through in reverse. A storage manager was implemented to store and synchronize data objects. At the moment, persistent data is loaded and saved to XML files at server start-up, respectively shut-down. For a more detailed description of structural and technical aspects of the implementation, the reader is referred to the installation notes, the manual and code documentation. They can be found on the attached data volume under `/Implementation/Docs`.

The functionalities the entities have to provide, are hierarchically specified via interfaces as seen in Figure 5.2. Those specify the operations that have to be supported by the different roles a participant can take.

For example, an entity implementing *ISmartGridParticipant* has to realize the most basic functions, like start-up, the exchange of communication keys and alike. More specific roles, like an *IPrivacyProtocolParticipant*, have to provide additional functionality, for example to set the parameters for the homomorphic encryption process. Special members of the protocol need to supply the facilities that correspond to their roles described in Section 5.1 above. For instance, the aggregator has to realize a method to aggregate the master key.

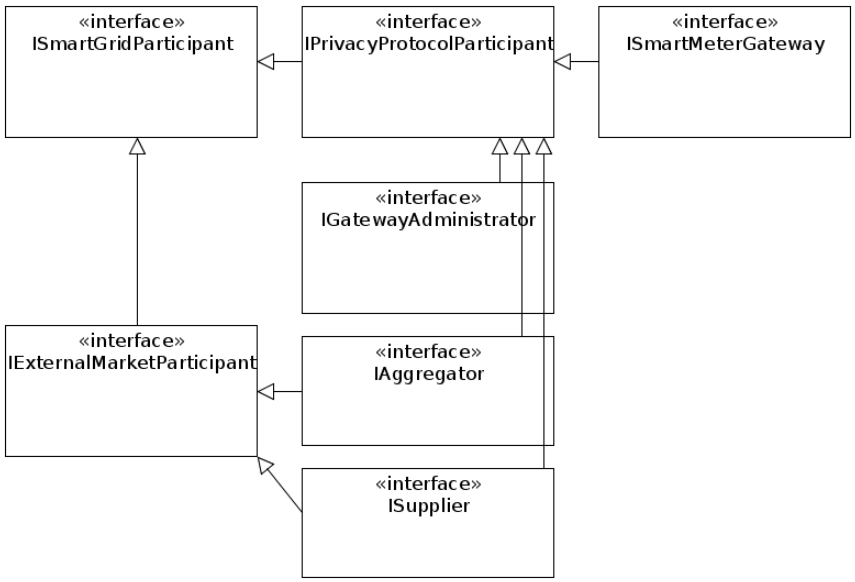

Figure 5.2: Schematic interface inheritance hierarchy

5.2.2 Protocol Stack

As shown in Figure 4.3, a specified stack of protocols has to be respected to fulfil the requirements of TR-03109. The following paragraphs describe how these were realized and which technologies were used to accomplish that task.

HTTP/TLS Like mentioned before, all communication has to be secured via TLS, together with HTTP in the application layer this results in the HTTPS communication protocol with TLS situated in the transport layer between HTTP and TCP. While the syntax matches with HTTP, additional content encryption and authentication is provided between the web server and the client (e.g. a browser or web application).

For the implementation of the TLS communication, the Java Secure Socket Extension (JSSE) is applied. In contrary to the Bouncy Castle libraries [1] consulted for the majority of cryptographic protocols and algorithms in the context of this work, it supports TLS version 1.2 which is demanded as a minimum requirement by the technical guideline. Certificates and keys for the prototype were generated with the Java *keytool*. It allows to manage keypairs and certificates in a `.jks` keystore file and simulate a chain of trust, beginning at a self-signed root certificate authority.

RESTful/COSEM Webservices According to [14, p. 30ff, ch. 324], all communication scenarios between the SMGW, the SMGW administrator and EMPs have to be realized via RESTful web services. Data structures have to correspond to the COSEM interface classes [28] and the OBIS codes [21, 27][2]. For the transport of requests and responses, the usage of HTTP/1.1 syntax is obligatory. The resource tree definition (see Figure 5.3) of the COSEM objects as well as the addressing of objects and attributes is specified in attachment II: *COSEM/HTTP Webservices* [16] of the technical guideline. While the COSEM objects are obliged to be transmitted in XML syntax, a XML schema definition is contained in this attachment, too.

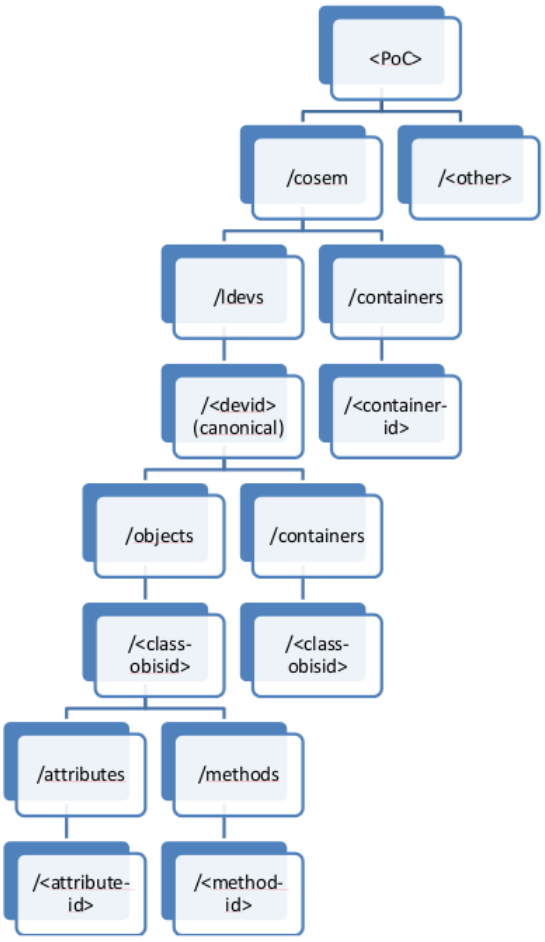

Figure 5.3: URI resource tree of the RESTful web service [16]

It has to be mentioned that a RESTful web service follows a programming paradigm rather than a standard. Therefore no explicit definition of REST can be presumed, despite the fulfilment of the following design principles:

- Services offered by a server have to be addressed with URLs, resources have to be reachable via URIs.
- The provided services might be delivered by the server represented in different formats (such as XML or JSON), according to the demands of the requesting application.
- Operations for the access to services such as GET, POST, PUT and DELETE have to be provided.
- Transitions between states have to be executed by actions identified within hypermedia data types like HTML or XML.

[2] For further explanation see paragraph **DLMS/COSEM OBIS**

In the work presented, *Jersey 1.17* [2], an open source JAX-RS (JSR 311) reference implementation for building RESTful Web services, is used. JAX-RS is the Java API for RESTful Web Services and makes annotations to Java classes to signalize their relevance as resources for the REST service. The transfer of the XML schema definition to Java classes was executed with the help of the *xjc*-Tool (Java Architecture for XML Binding Compiler). It is recalled that the present implementation relies on version 0.3 of the *.xsd*-document provided by the BSI. As this is not the final version and changes have to be expected, future adaptations might be necessary to preserve compliance.

Example 1. An example HTTP GET request to the RESTful interface of a SMGW could look like the following:

```
GET /smgw/cosem/ldevs/edud0000000100.sm/objects/1-00006001ffff
Accept: application/xml
```

This request asks for the device identifier of the logical device *edud0000000100.sm*. It is addressed through the corresponding OBIS code, represented in hexadecimal notation. The defined code for device IDs is *0-0-1-96-255-255* or, in hexadecimal representation, *00006001FFFF* as seen in the URI above. In case of a successful transmission, the gateway's response will be sent to the requester:

```
HTTP/1.1 200 OK
Date: Wed, 14 Aug 2013 14:59:30 GMT
Content-Type: application/xml
Content-Length: 556
```

The response body contains the inquired information in XML representation:

```
<?xml version="1.0" encoding="UTF-8"?>
<object xmlns="urn:k461-dke-de:cor-1"
    xmlns:co="urn:k461-dke-de:cod-1"
        id="3-00006001ffff" version="0">
  <attributes>
    <attribute id="1">
      <co:octet-string>00006001ffff</co:octet-string>
    </attribute>
    <attribute id="2">
      <co:octet-string>
        6564756430303030303030303130302E736D
      </co:octet-string>
    </attribute>
  </attributes>
</object>
```

The XML structure above includes two namespaces: `urn:k461-dke-de:cor-1` for the COSEM resources also shown in the resource tree (see Figure 5.3) and

`urn:k461-dke-de:cod-1` containing supported data types of the transmitted information – based on the ASN.1 specification [45].

For communication of the external entites among each other, a more human-readable URI syntax was chosen within the current prototype implementation. The service interface consists of five different resources:

- **ControlResource** for sending commands to the entities, e.g. *start* and *shutdown* messages.
- **MessageResource** for the exchange of notifications and errors.
- **ValueResource** for the receiving of encrypted or decrypted values originally stemming from the SMGWs.
- **KeyResource** to accept incoming messages related to the encryption keys, like initial key shares, updates and similar.
- **LogResource** for the request of log and status data.

Example 2. If the supplier wants to start the 6th round of the privacy preserving protocol, specified with the id *PRIV_PROTOCOL*, he sends a *startround* request to the participants, together with information about the current round counter. A message to an administrator could look like this:

```
GET /control/startround?number=6&operationid=PRIV_PROTOCOL
Accept: application/xml
```

An excerpt of the annotated method processing the request looks like this:

```
@Path("control")
public class ControlResource extends AbstractControlResource {

        @Override
        @GET
        @Path("startround")
        @Produces("application/xml")
        public String startNewRound(@Context HttpServletRequest req,
                        @QueryParam("number") int number,
                        @DefaultValue("") @QueryParam("operationid")
                                String operationid) {

                [...] // Process request

                return "true";
        }

        [...]
}
```

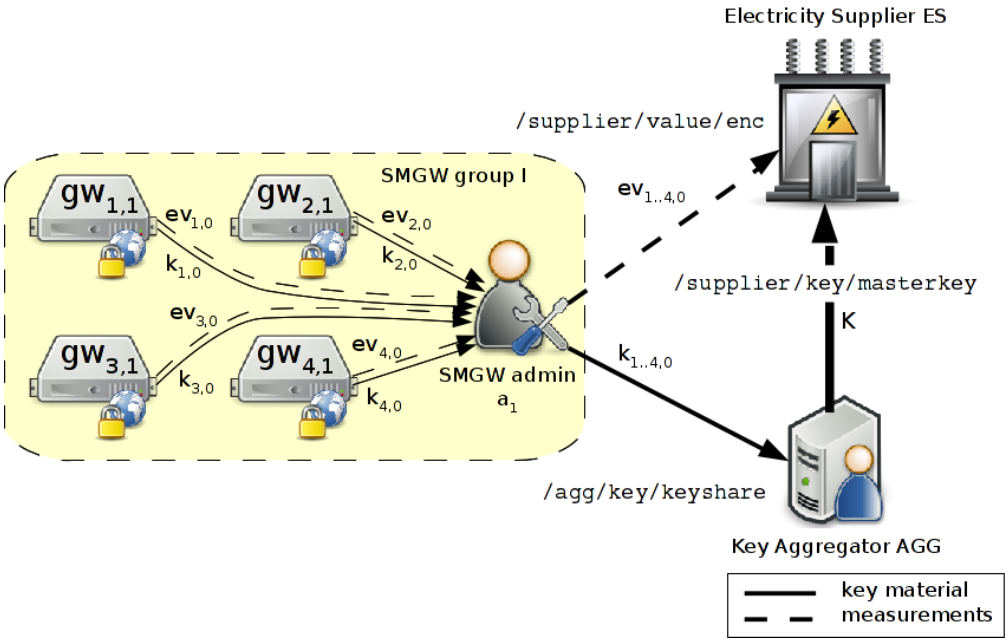

Figure 5.4: Communication flow during and mapping of protocol values to resource URIs during key initialization and first value transmission

Example 3. Another illustrated example is shown in Figure 5.4. Subject is a regular transmission of encrypted values and initial key shares. The SMGW administrator issues PUT requests with the encrypted values $ev_{i,0}$ to the URIs `/supplier/value/enc` while the key shares $k_{i,0}$ are sent to `/agg/key/keyshare`. After aggregation of these, AGG forwards the result K to ES via `/supplier/key/masterkey`.

Cryptographic Message Syntax CMS is a syntax standard for cryptographically protected messages, issued by the Internet Engineering Task Force. It describes the data format of digitally signed, digested, authenticated or encrypted content [61]. Thereby, encapsulation of multiple message types is supported, for example a digitally signed message can be enclosed into an encrypted content type. Furthermore, additional attributes, like the signing time, can be included into a message's signature.

CMS is based on X.509 public key infrastructure [60], describing a key management relying on certificates.

According to the pseudonymization process described in Section 4.3, data needs to be authenticated and encrypted for the external market participant and signed for the SMGW administrator. In attachment I [18] to the technical guideline it is specified that the data format shall be based on CMS. Encrypted, authenticated data has to be encapsulated in the *AuthEnvelopedData* content type which then has to be signed utilizing the *SignedData* content type.

For the realization of the Cryptographic Message Syntax the *CMSAuthEnvelopedData* and *CMSSignedData* classes of the *Bouncy Castle Cryptography API* [1] were used. They transfer the content type definitions of CMS data, defined in ASN.1 notation, into Java classes. An exemplary transformation of the *AuthEnvelopedData* type is given here.

ASN.1 notation

```
AuthEnvelopedData ::= SEQUENCE {
       version CMSVersion,
       originatorInfo [0] IMPLICIT OriginatorInfo OPTIONAL,
       recipientInfos RecipientInfos,
       authEncryptedContentInfo EncryptedContentInfo,
       authAttrs [1] IMPLICIT AuthAttributes OPTIONAL,
       mac MessageAuthenticationCode,
       unauthAttrs [2] IMPLICIT UnauthAttributes OPTIONAL }
```

Corresponding Java class excerpt

```
class CMSAuthEnvelopedData
{
    RecipientInformationStore recipientInfoStore;
    ContentInfo contentInfo;

    private OriginatorInfo originator;
    private AlgorithmIdentifier authEncAlg;
    private ASN1Set authAttrs;
    private byte[] mac;
    private ASN1Set unauthAttrs;

    [...]
}
```

DLMS/COSEM and OBIS As topmost layer of the protocol stack serve the DLMS/COSEM-IC IEC 62056-6-2 and OBIS IEC 62056-6-1 specifications [21, 27]. OBIS stands for *OBject Identification System*, it provides codes to identify the meaning of sent data in DLMS/COSEM compliant metering systems. The structure of this code for measurement data consists of six groups with unambiguous meanings. They are described as follows:

A medium of the item (source of energy); for example water, gas, heating or electricity

B internally, respectively externally used channel of the metering device if more than one channel is used

C type of measured value

D mode of measurement

E tariff rates or classification of time period

F historical values, like most recent measurement

For the application at hand primarily one type of measurement data is transmitted - readings of electricity consumption data during specified time intervals. That leads to the identifier shown in Table 5.2.

group	identifier	comment
A	1	applies for all electricity related objects
B	0	assumed to be irrelevant due to the usage of only one channel
C	1	sum of active power consumption
D	29	time integral 5, stands for the quantity calculated from the beginning of the current recording interval to the instantaneous time point for recording period 1
E	0	total, no tariff rate relevant
F	var.	previous meter reading

Table 5.2: Example of an OBIS code

It was decided not to provide a complete implementation of the DLMS/COSEM-IC IEC 62056-6-2 standard for the reasons that firstly, the standard is not fully available for free. Excerpts of the Blue Book [28], published by the DLMS User Association, describing the COSEM meter object model and the object identification system, are available online. Secondly, parts of the specification are already realized through the RESTful COSEM Webservice.

In this chapter, it was shown how the protocol of Mármol et. al has to be adapted to comply with the TR-03109. At some points different solution alternatives were pointed out, e.g. related to the handling of errors. The actual usage has to take into account the trade-off between communication and computation effort and the degree of privacy preservation. A deeper analysis of this, as well as the discussion of the actual applicability of the approach in a real-world environment, will be conducted in the next chapter.

6 Evaluation

The developed solution has been presented in the preceding chapter. The evaluation will be divided into three parts: The first section looks at the adapted protocol in means of performance and security level, while the second part will focus on the actual functionality of the developed prototype. To conclude everything, a deduction assessing the applicability of the developed protocol for the German and Brazilian situation follows.

6.1 Performance

In terms of performance, it has to be measured how expensive the chosen approach is in comparison to a customary pseudonymous measurement value transmission as anticipated in the TR-03109, Appendix VI for operational processes [17, ch. 4.8, p. 33, no. 32]. As indicators, overhead regarding computations, memory and communication will be taken into account.

Computational Overhead The only difference compared to a direct transmission of measurements is the homomorphic encryption, constituted as a modulo integer addition of the gateway's key share and the update of it by adding a random value and subtracting another one. this sums up to a total of three additional operations per round. As these symmetrical encryption operations are not computationally expensive, computational overhead can be neglected. If other asymmetric homomorphic primitives, like for example the Paillier cryptosystem, are utilized, the effort and complexity grow. But it still almost carries no weight, asymmetric cryptography is already extensively used within the rest of the technical guideline's protocol stack.

Memory Overhead Directly implied additional storage is needed for the homomorphic encryption key of the gateway. Owed to the technical guideline's requirements, the more significant storage effort arises for the additional evaluation and communication profiles. Five evaluation profiles are required, while the one to transmit values to the supplier would be needed in any case:

- Initialization, sending of initial key shares
- Transmission of encrypted measurements to supplier
- Transmission of key share updates to aggregator
- Request of key share updates to aggregator

- Transmission of measurements to administrator in case of implausible aggregation results

In the developed prototype, such an evaluation profile takes up no more than 7 kB. It already contains the communication profile, i.e. also the key material, for the particular market participant – without any optimizing compression mechanism in use. This size, like the computational overhead, hence can also be overlooked.

Communication Overhead The main aspect of overhead the solution carries, comes in terms of necessary communication.

Therefore it will be started with a depiction of the standard communication flow in case of a simple pseudonymous transmission according to [17, ch. 4.9, p. 33], shown in Table 6.1. Communication steps are marked with bold font.

No.	SMGW administrator	SMGW	External market participant
3.2.2		executes pseudonymization	
3.2.3		encrypts data for the recipient and signs it	
3.2.4		**transmits this data to the SMGW administrator**	
3.2.5	checks and removes signature of SMGW		
3.2.6	**transmits data packet to provisioned recipient**		
3.2.7			decrypts the data

Table 6.1: Communication flow for the pseudonymous transmission of measurements [17, ch. 4.9, p. 33]

Following from this, it can be seen that there are basically two message passings necessary. By applying the proposed method this amount is augmented by factor three, as it can be seen in Table 6.2 below.

Additional messages stem from the transmission of key updates to the aggregator and the request for the second update share from the same entity. For the sake of communication minimization it is assumed that the SMGW administrator directly forwards the second update to the SMGW like shown in step 9 (alt.).

For initialization, the five profiles described above have to be passed from the external market participant, respectively the administrator, towards the gateways. Additionally for the initial calculation of the master key, all n gateways have to send their random key shares. This results in a total of $2 \times n$ messages, taking into account the pseudonymization

step. Ideally, this procedure only occurs once, so it can be assumed that the effort does not have to be taken serious. The contrary is the case if errors occur.

In the scenarios where the SMGW administrator simply holds back the gateway's values, i.e. if it remains silent during key initialization, overhead only consists of a message to aggregator and supplier that the provisioned group size has to be decremented. Whereas in all other scenarios, where a reinitialization is suggested as the most privacy-friendly way to proceed, a notable communication effort arises. The round has to be repeated and initial key shares have to be sent again, meaning an overhead of $n \times 2$ messages for key reinitialization and $n \times 6$ for the repetition of value transmission and key update - plus the distribution of the actual error message to the administrators and external market participants.

To come to a conclusion whether this is maintainable, data about the frequency of gateway failures would be necessary. Without that data, an estimation of the real impact of the proposed failure mechanisms on the total overhead has to be purely speculative.

Intuitively, the communication effort needed could be diminished by implementing an autonomous key schedule like proposed by Castellucia et al. [24]. The schedule relies on a number of preconfigured keys from which the subsequent ones are derived. The communication of aggregators and gateways would not be required in normal operation, just a frequent update of the master key and its retransmission from the aggregator to the supplier would come up every round. Thus, $4 \times n$ messages per round could be saved in terms of key update, while one additional arises by a transmission of the new master key from AGG to ES. In cases of a complete reinitialization, more communication volume arises because of the not explicitly specified amount of additional key material.

However, the potential advantages in communication amount come together with the drawback that the obfuscation of the single gateway's keys would not be present anymore. In the implemented solution, this advantage exploits the possibility to send the own key share update using another pseudonym than it is used for receiving the difference distributed by the aggregator. In case of a scheduled key refreshment, the mapping of gateways to initial keys would be known throughout the whole operation. As an additional difference, the initial key shares would be distributed centrally by the aggregator and not randomly generated by the gateways themselves. Therefore here the coordinated key refreshment is favoured, for it significantly impedes the derivation of single gateway's round keys like also described in Section 6.2.4.

No.	SMGW administrator	SMGW	Aggregator	Electricity Supplier
1		creates random key update		
2		executes pseudonymization		
3		encrypts data for the aggregator and signs it		
4		**transmits this data to the SMGW administrator**		
5	checks and removes signature of SMGW			
6	**transmits data packet to aggregator**			
7			decrypts the data	
8			**sends pseudonymous random key update share to SMGW administrator**	
9a		**requests random key update from administrator**		
9b	**transmits random key update to SMGW**			
9 (alt.)	**transmits random key update to SMGW via wake up call**			
10		measures value and encrypts homomorphically		
11		executes pseudonymization		
12		encrypts data for the supplier and signs it		
13		**transmits this data to the SMGW administrator**		
14	checks and removes signature of SMGW			
15	**transmits data packet to supplier**			
16				decrypts the data

Table 6.2: Communication flow of measurement data transmission as proposed in this work

6.2 Security

To define a scale to measure how good the proposed method protects consumer data, it is necessary to clarify the actual protection goals. Those will be examined in the context of data protection and data security. *Data protection* has the goal to protect individuals from consequences arising from their personal data. That means that everyone should have the possibility to self-determine usage and distribution of data related to his person. This includes techniques to provide confidentiality and integrity. *Data security* on the contrary, focusses on every possible kind of data, not necessarily personal. It tries to prevent damages like data loss or unauthorized manipulation.

Derived from the descriptions given above, concrete protection goals can be formulated which help to judge a system's concordance with these. In the following, the most fundamental protection goals regarding data protection and security will be described in brief. A more detailed insight is granted in [10, 57, 62, 63].

Protection Goals of Data Security

Availability demands that data is available and usable within in a timely manner. This includes measures against accidental or malicious destruction of data, but also the protection of the underlying data access process.

Integrity covers completeness, intactness and imputability. This has to be realized with mechanisms to detect and prevent active manipulations. Directed towards personal data, another important aspect are also provisions to correct wrong or outdated data.

Confidentiality means to keep data safe from unauthorized access. It includes that no one shall collect, process or use personal data without consent.

Protection Goals of Data Protection

Transparency claims that an individual who is subject of data processing has to be abilified to gain insight into this process. Which, following the German Data Protection Law, also means that in principle data has to be collected directly and consensual from the data subject. It also implies the right to obtain information about the data which is existing at different instances of data processing entities.

Unlinkability is given when it is not - or with no reasonable effort - possible to connect the data to other information about the data subject in order to gain additional insights. Furthermore this conveys that in case of pseudonymous or anonymous data no subsequent identification through the connection of various data sources shall be accomplishable.

Interveneability ensures that the person affected can exercise his rights. Therefore he must be given the ability to control and see his data in order to correct or lock incorrect information. It also means that at every point of time it has to be possible to withhold consent to further data processing, even if it was given before.

Looking at the outlines above, it becomes apparent that some goals seem to contradict each other. An example are transparency and unlinkability, where it seems implausible that totally transparent processes can exist for unlinkable or even anonymous data processing. Therefore it is crucial to define, regarding whom these goals are formulated, like to ensure transparency for the data subject itself while preserving unlinkability against third parties.

Focussing on the case which shall be discussed here, essential for the preservation of the consumer's privacy are mainly the protection goals of integrity, confidentiality and unlinkability. Integrity aims at the correctness of the exchanged messages, supplementary the authenticity of the communicating parties has to be assured.

The second ensures that the data the user - respectively his SMGW - transmits, cannot be leaked to unauthorized third parties to reveal information about his energy consumption. Unlinkability in this context means that the data he sends does not disclose hints about his real identity and associated electricity consumption. That means, the supplier does not get to know anything besides the consumption sum of a neighbourhood.

6.2.1 Attacker Model

To evaluate how good privacy can be preserved by the described protocol, it needs to be clarified what would mean a break of privacy. The main goal of the approach is the hiding of individual small-grained measurements from third parties – be it unknown outsiders, the electricity supplier or the introduced external market participants. A break of privacy is constituted if these gain knowledge which enables them to derive information about the single values in addition to their total sum.

For a possible attacker, it can be imagined that he can eavesdrop the communication and try to distinguish two self-chosen measurement values encrypted by two gateways with unknown keys. It has to be shown that he has no advantage to successfully determine which encrypted value belongs to which actual measurement. If this states as true, it implies that he cannot distinguish between two arbitrary encrypted values. This would meet the case of the supplier, who gets the individual encrypted measurements from the gateways.

Another scenario imaginable which intuitively appears to be more threatening, is that the key aggregator collaborates with the eavesdropping attacker and hands him the individual key data of the smart meter gateways. The security against such an attack with a misbehaving key aggregator will be treated following the description of the first one in Section 6.2.4.

Here the remark is made that the protocol additionally relies on underlying security measures obligatorily demanded by TR-03109 such as TLS for confidential and authenticated transport of data. Because consequently those measures cannot be influenced by

the proposed protocol, attackers trying to break authentication or confidentiality of the messages sent via the secure channel are not discussed here.

Concluding these assumptions, this means that the attacker can only be represented by the energy supplier. According to the above provisions, he should be the only one able to obtain the homomorphically encrypted measurement values. An attacker therefore is assumed to hold all information a real supplier has at command. This is the strongest attacker to be looked at in this setting.

6.2.2 Domain of Trust

In the original proposal [48], it is stressed that no trusted third party is needed but only an untrusted aggregator. This cannot be fulfilled by the solution presented here. It is caused by the inherent requirements of the technical guideline which sets the gateway's administrator as a trusted third party by definition. Therefore also no malicious administrator is considered as a possible attacker, for he would be present independent from the protocol developed. Hypothetically he would be able to unveil the identity of every gateway. He can map the device IDs to the pseudonyms and therefore thwart any attempted privacy protection.

Up to today it is not yet specified who will be a smart meter administrator in the actual Smart Grid implementation and how many different providers there are about to exist on the market. But nonetheless provisions to ensure a reliable and secure operation of this future administrator are described in [15], an appendix of TR-03109 for the requirements to the administrator's operation. It may be additionally noted that most probably the choice of the concrete administrator will be made via a contract between the meter's owner and the administration service provider, so a non-conform behaviour would lead to legal consequences in any case.

Besides the gateway administrator, trust has also to be put into the SMGW itself. It has to process the measurements in compliance with the evaluation profile and encrypt, authenticate and sign them for the external market participants. Furthermore, it also has to provide keys and certificates for the communication with the original meters, making it responsible for the correct and secured transmissions of the measurements [14, ch. 3.3., p. 45ff].

The trustworthiness of every entity other than the administrator and SMGW is not taken for granted. A malicious gateway, respectively smart meter could send forged values to disturb the aggregation results and aggregator and supplier could work together to exchange key shares and encrypted values like described in the section above.

6.2.3 Privacy in the Case of Eavesdropping

It was already mentioned in Section 6.2.1 that in the case of an eavesdropping attacker (i.e. the supplier), it has to be proven that no means exist to assign single plaintext measurements to distinct SMGWs. To show this, Mármol et. al have introduced a smart meter privacy break game (SMPB) [49] which builds upon an earlier approach of Bohli, Sorge and Ugus [11]. As details can be found in the denoted papers, only a short

outline will be given here, together with a reasoning why their proof still holds here and is not influenced by the modifications of the present approach.

The base of argumentation is that a customer's privacy is preserved if an attacker A cannot distinguish between two self-chosen energy consumption settings $M_{0,j}$ and $M_{1,j}$ after they ran through the homomorphic encryption scheme. Both of these settings consist of the single measurements pv_{ij}^0 ,respectively pv_{ij}^1, whereas $\sum_{i=0}^{n} pv_{ij}^0 = \sum_{i=0}^{n} pv_{ij}^1$. The single summands pv_{ij}^0 and pv_{ij}^1 have to differ for at least some i.

It then needs to be shown that A cannot distinguish between $M_{0,j}$ and $M_{1,j}$ with a probability better than random guessing, otherwise he would have a SMPB advantage. Here is the game sequence as described in [49][1]:

1. **Setup**: The challenger runs the key-generation algorithm KG for a given set of smart meter gateways $GW = \{gw_i : 1 \leq i \leq n\}$ and generates the keys $\{K, k_{1,0}, \cdots, k_{n,0}\}$ for $n \geq 2$ – groupsize at least has to be two, otherwise privacy cannot exist nor be preserved. The challenger gives the aggregated key K to the adversary A and keeps the gateway encryption keys $\{k_{i,0} : i \leq 1 \leq n\}$.
2. **Challenge**: The adversary decides for two energy consumption scenarios denoted by $M_{0,j} = \left\{pv_{1,j}^0, pv_{2,j}^0, \cdots, pv_{n,j}^0\right\}$ and $M_{1,j} = \left\{pv_{1,j}^1, pv_{2,j}^1, \cdots, pv_{n,j}^1\right\}$ for a period j. The restriction is that $\sum_{i=0}^{n} pv_{ij}^0 = \sum_{i=0}^{n} pv_{ij}^1$ and $pv_{ij}^0 \neq pv_{ij}^1$ for some i. The energy consumption scenarios are given to the challenger.

   The challenger chooses a random bit $b \leftarrow \{0, 1\}$ to select between $M_{0,j}$ and $M_{1,j}$. The challenger then computes the challenge ciphertexts $EV_{b,j} = \left\{enc_{k_{i,j}}\left(pv_{i,j}^b\right) : i \leq 1 \leq n\right\}$ and gives them as a challenge to the adversary.
3. **Guess**: The adversary outputs a bit b' as a guess for b and wins the game if $b' = b$. The SMPB advantage of the adversary is defined as $Adv_A^{SMPB} = |Pr[b' = b] - \frac{1}{2}|$.

The aspects differing in the approach at hand compared to [49] and [11] are: The mode of key generation – compared to the latter one, the key update as well as the additional group members AGG and SMGW administrator. That key generation like done here and by Mármol et al. does not affect the applicability of the proof is already shown in their paper. The explanatory statement is that they, as well as Bohli et al., use a random-based key generation.

Key update in [49] is very similar to the present method. The authors state that a smart meter in the ring generates a random key difference and subtracts this from its own key share. This difference is then forwarded to the succeeding meter in the ring. This implies that every meter also adds the difference received from its predecessor to

[1] Notations have been adapted to the present representation used in Chapter 5.1.

his key share. Those differences correlate with $\triangle'_{ij}$ and $\triangle''_{ij}$ used for key refreshment here. The difference is that at hand no ring topology is assumed, thus $\triangle'_{ij} \neq \triangle''_{i+1,j}$. Nevertheless it applies that $\sum_{i=1}^{n} \triangle'_{ij} = \sum_{i=1}^{n} \triangle''_{ij}$ like also demanded by Mármol et al.. For both methods rely on a pseudo-random function for the generation of key differences, no effect on the comparability occurs.

The presence of AGG and the administrator also does not affect the privacy level. The maximum information level AGG can get is the knowledge of ES, in case they cooperate. The administrator, apart from being a trusted entity, only receives end-to-end encrypted measurements. This transfers the responsibility of privacy preservation to the underlying encryption method. In case he views the plaintext values following the proposed process in case of implausible values, the privacy 'break' is acquiesced. But this failure handling procedure can also be skipped, in case this effect is important to avoid. Concluding this, the proof of semantic security and preservation of one's privacy for the encryption schemes of [11, 49] applies here as well.

6.2.4 Privacy in the Case of a Malicious Key Aggregator

The claim with the presence of a malicious key aggregator is that privacy can also be preserved in case that an attacker holds all keys as well as all encrypted values. In [48], the reasoning for that is based on the expectation that the attacker would not be able to assign keys to encryptions by trying all possible combinations. As already mentioned before in Chapter 3.1.2, this argumentation is taken as too optimistic and an actual, not only theoretical threat is expected by this possibility.

Hence, in the present approach the problem is weakened. The keys used for the actual encryption are only known to the aggregator after an initialization phase. Later, the key update causes a scrambling of the mapping from current round key shares to the single SMGWs. The method exploits the possibility to send the own key difference $\triangle'_{ij}$ under another pseudonym as requesting $\triangle''_{ij}$ because the approach at hand does not rely on a ring structure, but on a completely random key redistribution. Additionally to the combination of all encryptions to all key shares, also a recombination of all possible key differences with all initial keys would have to be made. This states a much higher effort and it can be assumed that, especially in groups with a higher number of participating gateways, a meaningful mapping will hardly be possible in practice. To avoid the ability to maliciously pass the actual keys in the first round, a key update can be made prior to the first encryption. Thereby this threat is instantly removed.

6.3 Prototype Functionality

Focus of this work is to investigate the adaptability of the approach by Mármol et al. to a privacy-preserving protocol which meets the demands of TR-03109. The developed prototype hence is a means to provide a proof-of-concept for the suggested solutions and may serve as a basic framework for further refinements.

To test the prototype's functionality, a small test setup has been assembled, which is described in the following to monitor the behaviour of the protocol members during normal operation and possible exceptional situations and errors.

6.3.1 Test Environment

The testing of the developed prototype was performed in the following experimental setup:

Electricity Supplier The end recipient interested in the total aggregated consumption. It is a single instance receiving values from SMGW administrators and the key aggregator.

Reachable as: `https://supplier/...`

Aggregation EMP Entity for collecting and aggregating the key shares and distributing tokens in case of failure. Serves data from SMGW administrators and the electricity supplier.

Reachable as: `https://aggregator/...`

SMGW Administrators Responsible for the forwarding of data from the gateways to the corresponding EMP. Receives data from the assigned smart meter gateways. In the test setup there exist two administrators, one operating two smart meters and the other four.

Reachable for communication with the gateways of communication scenario **ADMIN-SERVICE** as:

`https://trustworthyadmin:{port}/gwa` resp.
`https://secureadmin:{port}/gwa`

Reachable for external market participants as:

`https://trustworthyadmin/...` resp.
`https://secureadmin/...`

Note that for the SMGWs, every gateway is given its own communication endpoint at the administrator.

Smart Meter Gateways Providing values of the measurement devices and key shares. The prototype does not distinguish between meters and meter gateways because the transmission between meters and gateway does not affect the protocol under test. In case multiple meters are present, it is assumed that their data is already aggregated by the SMGW. For the production of measurement values in the experimental setup, the SMGW provides a generator of random dummy values.

In the present application SMGWs only communicate with their SMGW administrator. They are reachable in the communication scenario **MANAGEMENT** as:

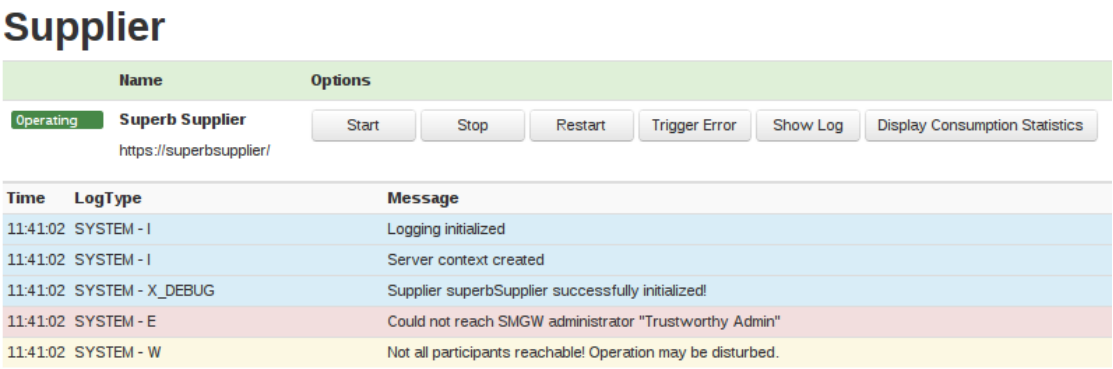

Figure 6.1: Control and logging interface for the electricity supplier

`https://trustworthygateway/{number}/smgw/cosem/...` resp.
`https://securegateway/{number}/smgw/cosem/...`

6.3.2 Test Interface

During actual operation, all participants operate autonomously. In order to simulate and trigger desired events and to be able to monitor the operation of all entities at the same time, a control interface was developed which serves as a testing environment (see Figures 6.1 and 6.2). It consists of a web interface where all entities are listed. Their control and logging functionalities can be addressed via a graphical user interface. It is planned that the interface also provides utility functionality, for example to create communication and evaluation profiles (see Figure 6.3) or change the group setup. The *jQuery 1.10.1* JavaScript API is used to issue Ajax requests to the REST interfaces of the participant server instances and to provide cross-browser compatibility for the control interface.

6.3.3 Experimental Outcome

While running the test setup, the system behaved as expected. It could be determined that an average round of a protocol run requires about five to seven seconds to be completed, starting from a trigger message issued by ES until ES finally calculated the decrypted measurement sum. This value is totally acceptable, in reality a round is expected to last about fifteen minutes. Therefore, no cutbacks have to be expected by the additional appliance of this protocol.

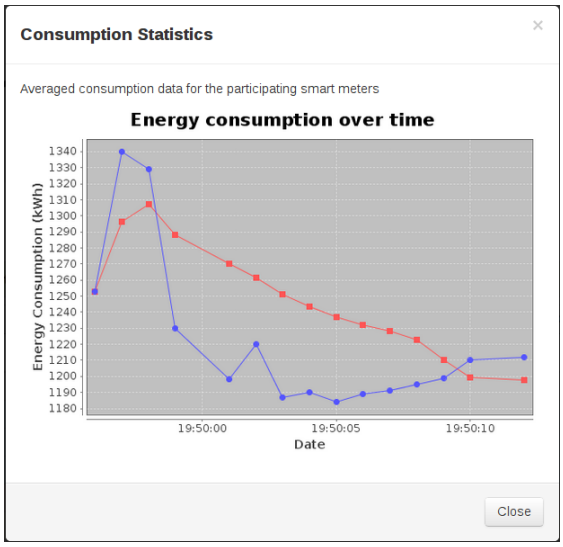

Figure 6.2: Aggregated consumption statistics (blue) generated by the supplier and averaged interval consumption (red)

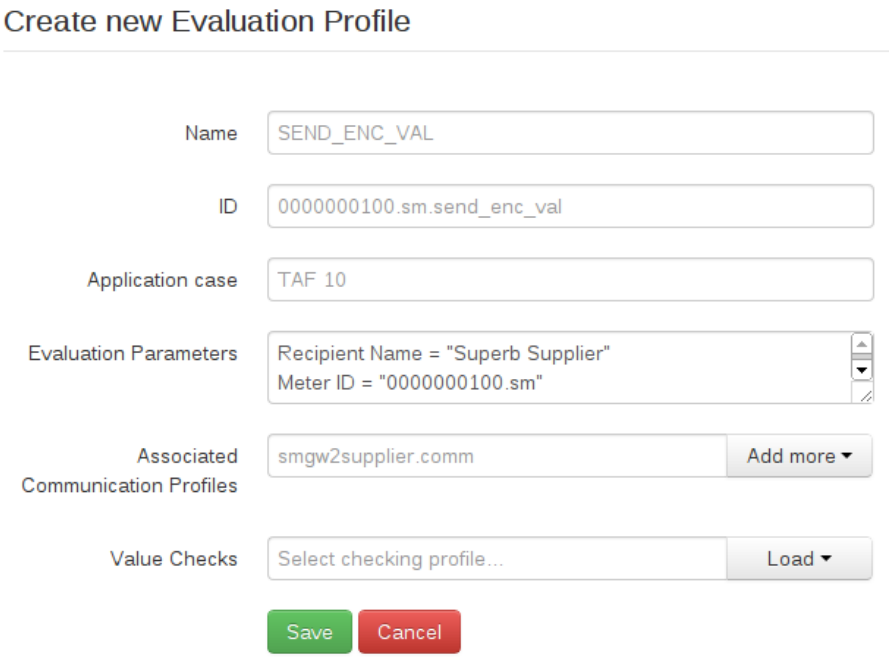

Figure 6.3: Utility interface to create new evaluation profiles

6.4 Protocol Conformity to the Technical Guideline

Due to the prototypical character of the implementation, of course not all specifications described in the technical guideline are realized. In any case this was not the actual goal of this work, which instead was to show that the concept would not contradict the provisions in TR-03109.

By using the administrator as an intermediate instance to pseudonymize a gateway's sendings and forward them either to the supplier or the key aggregator, a direct communication between the gateways is rendered unnecessary. This removes the most obvious contradiction of the underlying approach by Mármol et al. to the guideline's demands.

The conformance of the WAN communication is guaranteed by implementing the protocol stack described in Section 5.2.2. Data to be sent, namely encrypted measurements, keys and key shares as well as informational and error messages are embedded into the postulated COSEM data structure and are assigned the corresponding OBIS codes. A survey of the custom OBIS codes can be found on the attached data volume under `/Implementation/Docs`.. The messages are then packed into the demanded CMS packets and sent via HTTPS using a RESTful interface. Lower layers defining transport, network, data link and physical layer are not addressed here and also left open in the guideline.

The use case for the pseudonymization – ADMIN-SERVICE – is taken into account and the demanded evaluation and communication profiles are configured in conformity to the required fields and declared to the gateways as demanded.

6.5 Applicability of the Approach for the Brazilian Smart Grid

The design and implementation of the protocol developed during this study was focussing on the requirements of TR-03109. But as this work is situated in the context of the TruEGrid project, the applicability of the given proposal for the Brazilian situation also has to be explored. Chapter 4.4 already explained that the Brazilian regulation does not describe detailed prerequisites for a privacy-friendly and secure smart meter operation. Therefore, theoretically nothing intervenes with the application of the work presented here within the Brazilian Smart Grid implementation up to this date.

The level of security formulated by the BSI extensively exceeds the provisions which are postulated throughout Brazilian specifications. An application of those, together with the privacy-preservation protocol elaborated here, would therefore provide a huge security advancement in principle. However, it is questionable if this would be realistic. Main interests of the major stakeholders producing smart meters are certainly of a commercial nature. As additionally security always evokes additional cost, there might be no honest interest to implement those measures, unless there were enforcing regulations. At

least as long as customer awareness does not turn a higher privacy level into a noticeable market advantage for the manufacturer.

In the foregoing chapter the results conducted from prior research and explanation have been evaluated. Firstly, the suggested protocol and its deviations from the underlying publication of Mármol et al. have been discussed. Arising issues of computation, memory and communication overhead have been pointed out and a security evaluation was conducted. Moreover the functionality of the implemented prototype has been demonstrated.

To finalize this study, the next and last chapter will situate these results in the overall context of this work. Concluding remarks will be made and give a glimpse of what can be done to further elaborate the achieved findings.

7 Conclusions and Future Work

In this study it has been pointed out that the implementation of Smart Grids constitutes a great opportunity and at the same time a big challenge in many aspects. One of these aspects is the threat for the privacy of individuals to get profiled and spied on. Fine-grained energy consumption data transmitted by the smart meters in their households can lead to such a scenario. This work is dedicated to investigate methods to dispel this risk - or at least minimize it as far as possible. Privacy enhancing techniques, as they are also used in many other contexts, can help to reach this goal. With a focus on homomorphic encryption and aggregation approaches that are already applied for sensor networks, a selection of related work has been presented in Chapter 3. The work of Mármol et al. which served as a starting point for this thesis has also been introduced there.

But when it comes to real-world applications, predominantly theoretical approaches do not suffice and sometimes cannot be applied without changes. The regulatory situation in Germany, primarily represented by the technical guideline TR-03109, introduces some restrictions and demands on smart meter operation and communication. These, which are described in Chapter 4, inhibit a direct implementation of the considered protocol.

Within this work, solutions were found to overcome those restrictions. Parts of the solution have been adapted where necessary. The initial idea to use a member of the smart metering group as an key aggregator has been changed towards an external market participant. The risks of a malicious cooperation of key aggregator and electricity supplier could even be weakened as an effect of this modification because the traceability of a gateway's key share has been made more complex. Furthermore, the token solution proposed by the authors to handle rounds with failing or misbehaving meters could be eliminated. The guideline provisions a trusted gateway administrator who can accomplish the detection of a faulty gateway, remove it from the group and trigger a correct proceeding.

Additional to the core of the suggested protocol, the framing conditions for the communication of smart metering gateways with entities in the WAN have been investigated. The technical guideline stipulated protocols, data models and other transmission details. A synopsis has been given on those requirements and on this basis the design of the prototype was realized. While it is surely only a skeleton of the guideline's postulations, all necessary protocol steps have been modelled and prototypically implemented.

How and that the evolved solutions and adjustments fit into a potential smart metering system in actual operation has been presented in Chapter 6. It could be shown that the performance of the created concept behaves well in terms of computational and memory overhead, compared to a standard, pseudonymous data transmission like it is envisaged in the guideline. Also the goal to meliorate the preservation of a customer's privacy

could be accomplished. Compared to the solely use of pseudonyms, linkability to an individual's real identity now is way harder. Even in comparison to the underlying approach, the risk imposed by malicious aggregators could be diminished.

When it comes to communication, it has to be admitted that the protocol demonstrates a significant rise of effort. The key update scheme requires additional messages to and from the aggregator in every round. In failure scenarios, where a complete reinitialization has to be carried out, non-neglectable overhead arises. Therefore this issue shall be noted as the first part of future work to be done. Starting points for further investigation can be alternative key scheduling solutions, autonomous as well as coordinated, or more sophisticated error correction and tolerance mechanisms. Here the developed prototype serves well as a testing platform for the realization and evaluation of methods independent from specific hardware settings.

Another point demanding future attention is the self-inherent incompleteness of the TR-03109. While it is still a work in progress, especially in the area of WAN communication, updates and changes are potentially implied whenever a new subversion is released. The applicability of the results in this thesis have therefore to be checked frequently to react to changes in the regulatory framework.

Changes also have to be expected on the Brazilian side. Nowadays, the SiBMA (Brazilian System for Advanced Metering) project[1] aims to design a prototype for remotely controlled meters, including the the collection of consumption data from Brazilian households and industrial instances. In the process of the project also open communication protocols are about to be designed, together with a specification, potentially serving as a future national standard for smart metering devices in Brazil. Future work has to keep this development in view and re-evaluate if the approach still fits into the Brazilian regulations.

A last point shall be mentioned here as an important item for the future: A complete reference implementation of the TR-03109 is necessary. Up to the current point in time, no existing solution is known to the author. But to test the procedure wholly, it does not suffice to only look at a single part of the system and show that it does not contradict any specifications. Moreover, it is also of utter importance to have a view on the hardware conditions a SMGW will be potentially running on. This will help to judge the applicability and performance of the protocol not only theoretically but also in a real-world environment.

During the development process of this thesis, the development of frameworks and infrastructure for smart metering systems forged ahead in various institutes, companies and alike. To hint at promising technologies for the implementation of a system suited for practical operation on existing hardware, the *OGEMA* [52] and *OpenMUC* [36] frameworks shall be specifically highlighted. *OGEMA* is an open-source platform to manage communication, applications and electrical devices between a customer and the energy supplying system. *OGEMA 1.1.1* has been released on 10th May 2013. In

[1] For further details, see: `smartgridnews.com.br/blog/2012/01/24/cesar-e-abinee-dao-primeiro-passo-para-smart-grid-no-brasil/`, `www.redeinteligente.com/2011/01/20/abinee-cria-grupo-de-smart-grid-e-trabalha-na-criacao-do-sibma/`

connection to the further development of *OGEMA 2.0* stands *OpenMUC*, a monitoring, logging and controlling system used for Smart Grid projects at the Fraunhofer ISE[2]. It provides support for data logging and access and various communication protocols like MBus and – since 26th July 2013 – DLMS/OBIS. Furthermore, it comes with a web interface and support for different data server types. Both projects are closely connected and actively developed by the *OGEMA Alliance* [3]. They are equally based on Java and OSGi, making them suitable for multiple platforms. *OpenMUC* has been tested to operate on ARM 9 systems and Raspberry Pi boards [4]. A reference implementation on base of these components appears promising and of great value for future development.

As a whole, this study could show that the privacy of a customer and the needs of an electricity supplier for fine-grained consumption data are not mutually exclusive. The proposed solution complies with the guideline and therefore constitutes an opportunity for additional privacy protection in a real-world smart metering environment. With this insight, there is hope for a future Smart Grid which can fully exploit its capabilities to be of a great benefit for all stakeholders, be it producers or consumers of energy. And, last but not least, environment and nature itself.

[2] `http://www.ise.fraunhofer.de`
[3] `http://www.ogemalliance.org/`
[4] `http://www.raspberrypi.org/`

Bibliography

[1] Bouncy Castle Cryptography API for JAVA. 2013. URL http://www.bouncycastle.org. Last access: 2013-04-10. 42, 47

[2] Jersey API for RESTful web services. 2013. URL http://jersey.java.net/. Last access: 2013-04-10. 44

[3] ANEEL. Aneel divulga indicadores de qualidade de energia de 2010, 2011. URL http://www.aneel.gov.br/aplicacoes/noticias/Output_Noticias.cfm?Identidade=3855&id_area=90. Last access: 2013-03-06. 8

[4] ANEEL. Resolução Normativa No 502. 2012. 8, 29

[5] H.-J. Appelrath, H. Kagermann, and C. Mayer. Future Energy Grid, Migration to the Internet of Energy. acatech, National Academy of Science and Engineering, Munich, 2012. 10

[6] F. Armknecht, D. Westhoff, J. Girao, and A. Hessler. A lifetime-optimized end-to-end encryption scheme for sensor networks allowing in-network processing. Computer Communications, 31(4):734–749, 2008. 18, 33, 34

[7] M. Backes and S. Meiser. Differentially Private Smart Metering with Battery Recharging. 2012. 20

[8] M. Bellare, D. Micciancio, and B. Warinschi. Foundations of group signatures: Formal definitions, simplified requirements, and a construction based on general assumptions. In Proceedings of Eurocrypt 2003, pages 614–629. Springer-Verlag, 2003. 14

[9] Bloomberg Businessweek. How italy beat the world to a smarter grid. 2009. URL http://www.businessweek.com/globalbiz/content/nov2009/gb20091116_319929.htm. Last access: 2013-03-08. 7

[10] K. Bock and S. Meissner. Datenschutz-Schutzziele im Recht. Datenschutz und Datensicherheit-DuD, pages 425–431, 2012. URL http://www.springerlink.com/index/Q622V5T10WG18451.pdf. 53

[11] J. M. Bohli, C. Sorge, and O. Ugus. A Privacy Model for Smart Metering. 2010. 55, 56, 57

[12] BSI. Protection Profile for the Gateway of a Smart Metering System. 2012. 24

[13] BSI. Technische Richtlinie BSI-TR-03109. 2012. 22, 72

[14] BSI. Technische Richtlinie BSI-TR-03109-1. 2012. 25, 26, 27, 28, 30, 31, 35, 36, 38, 43, 55, 72

[15] BSI. Technische Richtlinie BSI-TR-03109-1: Anlage V Anforderungen zum Betrieb beim Administrator. 2012. 31, 55

[16] BSI. Technische Richtlinie BSI-TR-03109-1, Anlage II COSEM/HTTP Webservices. 2012. 43, 72

[17] BSI. Technische Richtlinie BSI-TR-03109-1, Anlage VI: Betriebsprozesse. 2013. 49, 50, 71

[18] BSI. Technische Richtlinie BSI-TR-03109-1: Anlage I CMS-Datenformat für die Inhaltsdatenverschlüsselung und -signatur. 2013. 46

[19] BSI. Technische Richtlinie TR-03116-3: eCard-Projekte der Bundesregierung. 2013. 35

[20] Bundesnetzagentur. Versorgungsqualität - Übersicht SAIDI-Werte Strom 2006 - 2011, 2011. URL `https://www.bundesnetzagentur.de/cln_1912/DE/Sachgebiete/ElektrizitaetGas/Sonderthemen/SAIDIWerteStrom/SAIDIWerteStrom_node.html`. Last access: 2013-03-06. 8

[21] Bundesverband der Energie- und Wasserwirtschaft. OBIS-Kennzahlen-System. 2012. 43, 47

[22] J. Camenisch and A. Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In Advances in Cryptology EUROCRYPT 2001, volume 2045, pages 93–118. Springer Berlin Heidelberg, 2001. 14

[23] J. Camenisch and A. Lysyanskaya. Signature Schemes and Anonymous Credentials from Bilinear Maps. 2005. 14

[24] C. Castelluccia, E. Mykletun, and G. Tsudik. Efficient Aggregation of encrypted data in Wireless Sensor Networks. 2005. 14, 17, 51

[25] D. Chaum. Security without identification: transaction systems to make big brother obsolete. Communications ACM, 28(10):1030–1044, 1985. 14

[26] Deutscher Bundestag. Bundesdatenschutzgesetz. 1978. URL `http://www.gesetze-im-internet.de/bundesrecht/bdsg_1990/gesamt.pdf`. 21

[27] DLMS . DLMS UA 1001-6 Ed. 2.3 – List of standardised OBIS codes. 2005. URL `http://dlms.com/documents/members/OBIS_list_v2.3_GK051026.zip`. Last access: 2013-04-10. 43, 47

[28] DLMS User Association. Blue Book. 2010. 43, 48

[29] Electric Power Research Institute. Report to NIST on the Smart Grid Interoperability Standards Roadmap. 2009. 6

[30] Enerdata. Global energy statistical yearbook 2012. URL http://yearbook.enerdata.net. Last access: 2013-03-06. 7

[31] European Commission. Adoption of the communication on smart grids, 2011. URL http://ec.europa.eu/energy/gas_electricity/smartgrids/smartgrids_en.htm. Last access: 2013-03-08. 7

[32] European Commission. Energy: Commission paves the way for massive roll-out of smart metering systems, 2012. URL http://ec.europa.eu/energy/gas_electricity/smartgrids/doc/20120309_smart_grids_press_release.pdf. Last access: 2013-03-08. 7

[33] European Commission. 2012/148/EU: Commission Recommendation of 9 March 2012 on preparations for the roll-out of smart metering systems. 2012. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:073:0009:01:EN:HTML. Last access: 2013-03-08. 7

[34] European Technology Platform SmartGrids. Strategic Deployment Document for Europe's Electricity Networks of the Future. 2010. 10

[35] FORSA. Erfolgsfaktoren von Smart Metering aus Verbrauchersicht. 2010. URL http://onlinelibrary.wiley.com/doi/10.1002/cbdv.200490137/abstract. 11

[36] Fraunhofer ISE. OpenMUC. 2013. URL www.openmuc.org. Last access: 2013-08-18. 64

[37] G. Kalogridis, C. Efthymiou, S. Denic, T. Lewis, and R. Cepeda. Privacy for smart meters: Towards undetectable appliance load signatures. Smart Grid Communications, pages 232 – 237, 2010. 19

[38] F. D. Garcia and B. Jacobs. Privacy-friendly Energy-metering via Homomorphic Encryption. 2010. 38

[39] H. Geller, G. de Martino Jannuzzi, R. Schaeffer, and M. T. Tolmasquim. The efficient use of electricity in brazil: progress and opportunities. Energy Policy, 26 (11):859 – 872, 1998. 9, 72

[40] C. Gentry. A fully homomorphic encryption scheme. PhD thesis, Stanford University, 2009. URL crypto.stanford.edu/craig. 13

[41] C. Gentry. Fully homomorphic encryption using ideal lattices. In Proceedings of the 41st annual ACM symposium on Theory of computing, STOC 09, pages 169–178. ACM, 2009. URL http://doi.acm.org/10.1145/1536414.1536440. 13

[42] GTM Research. The smart grid in europe 2012-2016: Technologies, market forecasts and utility profiles. 2012. URL https://www.greentechmedia.com/research/report/the-smart-grid-in-europe-2012. Last access: 2013-03-08. 8, 72

[43] INMETRO. Portaria Inmetro no. 375. 2011. URL http://www.inmetro.gov.br/legislacao/rtac/pdf/RTAC001738.pdf. Last access: 2013-03-08. 28

[44] International Energy Agency. Key World Energy Statistics. 2012. 8

[45] International Telecommunication Union. ASN.1 Specification. 2013. URL http://www.itu.int/ITU-T/asn1/. Last access: 2013-03-08. 45

[46] S. M, D. Westhof, F. Armknecht, and J. Girao. Non-manipulable aggregator node election protocols for wireless sensor networks. 2007. 14

[47] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin. Private memoirs of a smart meter. In Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, BuildSys '10, pages 61–66. ACM, 2010. 4

[48] F. G. Mármol, C. Sorge, O. Ugus, and G. M. Perez. Do not snoop my habits: Preserving privacy in the smart grid. IEEE Communications Magazine, Special Issue on Communication Protocols and Algorithms for the Smart Grid, 50(5):166 – 172, 2012. 14, 15, 18, 35, 55, 57

[49] F. G. Mármol, C. Sorge, O. Ugus, G. M. Perez, R. Petrlic, and D. Westhof. Privacy-enhanced architecture for smart metering. 2012. 14, 15, 34, 55, 56, 57, 72

[50] NRC. Smart energy meter will not be compulsory. 2009. URL http://vorige.nrc.nl/international/article2207260.ece/Smart_energy_meter_will_not_be_compulsory. Last access: 2013-03-08. 7

[51] G. Oettinger. Eu sustainable energy week: Commission will announce action on smart meters for more savings, 2011. URL http://ec.europa.eu/commission_2010-2014/oettinger/headlines/news/2011/04/20110408_en.htm. Last access: 2013-03-08. 7

[52] OGEMA Alliance. OGEMA 2.0. 2013. URL www.ogema.org. Last access: 2013-08-18. 64

[53] M. Önen and R. Molva. Secure data aggregation with multiple encryption. In Proceedings of the 4th European conference on Wireless sensor networks, pages 117–132. Springer-Verlag, 2007. URL http://dl.acm.org/citation.cfm?id=1758126.1758137. 16, 17, 18

[54] S. Papadimitriou, F. Li, G. Kollios, and P. S. Yu. Time series compressibility and privacy. In Proceedings of the 33rd international conference on very large data bases, pages 459–470. VLDB Endowment, 2007. 19

[55] R. Petrlic. A privacy-preserving concept for smart grids. 2011. 19

[56] Power Technology. Smart grid solutions to latin america's power theft crises. URL http://www.power-technology.com/features/featuresmart-grid-energy-theft-power-brazil-latin-america. Last access: 2013-03-06. 8

[57] T. Probst. Generische Schutzmaßnahmen für Datenschutz-Schutzziele. Datenschutz und Datensicherheit-DuD, pages 439–444, 2012. URL http://www.springerlink.com/index/K4L726783503425L.pdf. 53

[58] E. L. Quinn. Smart Metering and Privacy: Existing Law and Competing Policies. 2013. 4

[59] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor. Smart meter privacy: A utility-privacy framework. In Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on, pages 190–195. IEEE, 2011. 19

[60] RFC 5280. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. (RFC 5280), 2008. 46

[61] RFC 5652. Cryptographic Message Syntax (CMS). (RFC 5652), 2009. 46

[62] M. Rost and K. Bock. Privacy By Design und die Neuen Schutzziele. Datenschutz und Datensicherheit - DuD, 35(1):30–35, 2011. URL http://www.springerlink.com/index/10.1007/s11623-011-0009-y. 53

[63] M. Rost and A. Pfitzmann. Datenschutz-Schutzziele revisited. Datenschutz und Datensicherheit - DuD, 33(6):353–358, 2009. URL http://www.springerlink.com/index/10.1007/s11623-009-0072-9. 53

[64] S. McLaughlin, P. McDaniel, and W. Aiello. Protecting consumer privacy from electric load monitoring. Proceedings of the 18th ACM conference on Computer and communications security, pages 87 – 98, 2011. 19

[65] V. Shoup. Practical threshold signatures. In Proceedings of the 19th international conference on Theory and application of cryptographic techniques, EUROCRYPT'00, pages 207–220, Berlin, Heidelberg, 2000. Springer-Verlag. 14

[66] TruEGrid. Towards a Smart Grid Lab. 2012. URL https://truegrid.eu/images/documents/towardsasmartgridlab.pdf. Last access: 2013-03-08. 23, 72

[67] G. van Blarkom, J. Borking, and J. Olk. Handbook of Privacy and Privacy-Enhancing Technologies. 2003. 12

List of Tables

List of Figures

Glossary

ANEEL Agência Nacional de Energia Eléctrica - in English: Brazilian Electricity Regulatory Agency. Institution linked to the Brazilian Ministry of Mines and Energy. Their task is to regulate the electricity market to keep the relation between agents and customers balanced.

BDSG Bundesdatenschutzgesetz - in English: Federal Data Protection Act. German law regulating data and privacy protection of individuals against companies and governmental institutions.

BSI Bundesamt für Sicherheit in der Informationstechnik - in English: Federal Office for Information Security. Agency of the German government responsible for security of software applications, protection of network infrastructure, cryptographic guidelines, certification and auditing of security products and accreditation of security test laboratories.

CLS Controllable Local System: It is located in the HAN and contains the assessable energy producing and consuming devices of the customer, e.g. combined heat and power or photovoltaic facilities as well as intelligent household devices).

COSEM COmpanion Specification for Energy Metering: provides an object model for metering devices to define a reliable interface for communication and an identification system for all metering data.

DLMS Device Language Message Specification: A general concept for the specification of communication entities.

EMP External Market Participant: entity in the WAN providing services to the customer.

HAN Home Area Network: This entity belongs directly to the end consumer and consists its CLS. Furthermore data for the consumer and service technician are provided and prepared by the SMGW for displaying in the HAN.

IEC International Electrotechnical Commission: international organization that develops international standards for technologies related to the context of electricity and electrical devices.

IEC62056 Standards for electricity metering, describe data exchange for meter reading, tariff and load control. They are the international versions of the DLMS/COSEM specification.

LMN Local Meteorological Network: It describes the meters for electricity, gas, water and heating associated to a household. These meters report their measurements via the LMN to the SMGW.

PET Privacy Enhancing Technology: for explanation refer to chapter 3.

PP Protection Profile: Document issued by the BSI to point out the minimal security requirements for a smart metering system. It will be used for checking SMGW to get a certificate by the BSI.

SMGW Smart Meter Gateway: communication unit of a smart metering system, acting as a physical interface for devices and participants in the WAN, HAN and LMN.

TLS Transport Layer Security: Formerly known as Secure Sockets Layer (SSL) is a protocol for secure data transmission in the internet utilizing hybrid encryption.

TOE Target of Evaluation: System to be evaluated during a Common Criteria security audit, in this work synonymously used for the SMGW.

TR-03109 Technical Guideline 03109: Referring to the PP the technical guideline gives detailed specifications on how to implement a smart metering system with respect to conformity to the protection goals defined by the BSI.

WAN Wide Area Network: communication area of the SMGW with external market participants and especially the SMGW administrator.

Statement of authorship

I confirm that this document has been composed by myself and describes my own work, unless otherwise acknowledged in the text. All verbatim extracts have been distinguished and all sources of information have been specifically acknowledged.
I further declare that this work has not been and will not be submitted in any other application for a degree.

Place, Date Signature