Redactable Signatures to Control the Maximum Noise for Differential Privacy in the Smart Grid

Henrich C. Pöhls[1]*, Markus Karwe[2] **

[1] Chair of IT-Security, University of Passau, Germany hp@sec.uni-passau.de,
[2] University Freiburg, Germany markus.karwe@iig.uni-freiburg.de

Abstract. The Smart Grid is currently developed and fundamental security requirements like integrity and origin authentication need to be addressed while minimizing arising privacy issues. This paper balances two opposing goals: On the one hand, we mitigate privacy issues raised by overly precise energy consumption values via data perturbation mechanisms, e.g., add noise. On the other hand we limit the noise's range and keep a verifiable level of integrity of consumption values from the Smart Metering Gateway by using a redactable signature. We propose to use the value obtained by calculating the worst case guarantee of differential privacy as a metric to compare and judge a Smart Grid application's privacy invasiveness.

Keywords: Smart Grid, Differential Privacy, Redactable Signature Schemes

1 Introduction

The transition from nuclear to renewable energy is still in progress and brings stakeholders the burden to improve the overall energy management in order to keep net stability as well as reasonable prices [30, 10]. The Smart Grid (SG) is still in the development phase and can be seen as information overlay network for the traditional energy grid which enables stakeholders to improve the management. While the outlook for SG seems very promising it introduces new challenges like privacy for residential customers.

Corner stones of the SG are the Smart Meter (SM) and the Smart Meter Gateway (SMGW) as depicted in Fig. 3. Note that both devices are trusted and installed by a SG stakeholder, i.e., the power grid provider. A SM sends energy consumption values via the SMGW to a collecting SG stakeholder. Further note, we always assume that the SM produces accurate and timely readings. This allows the stakeholder to get a fine resolution picture of the energy consumption at customer's premises, which can be used for purposes like demand

* The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 609094.
** The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7-SMARTCITIES-2013) under grant agreement n° 608712.

forecasting or creating energy profiles [30]. To counter act malicious tampering, both SM and SMGW protect the integrity and authenticity of the transmitted data. All communication between the SM within a household and the SMGW is secured for wired as well as for wireless connections. Classical digital signatures offer such a protection: they allow detecting any change that occurred after the signature's generation. Cryptographically, a digital signature scheme is said to be *unforgeable*, e.g., RSA-PSS [4]. Hence, data requested by SG stakeholders is encrypted and signed by the SMGW before being sent [12].

Having tampering solved by digital signatures, one problem remains: The fine grained values impose a privacy threat to the residential customer. Several works show that too fine-grained energy values allow detecting appliances within the household [23], detecting the use mode of the appliances [11] as well as deducting the residential customers' behaviour [20]. To mitigate those threats current research and governmental organizations suggest using Privacy Enhancing Technologies (PET). For example, the German "Bundesamt für Sicherheit in der Informationstechnik (BSI)" is using pseudonymization as a privacy protecting mechanism [12]. In [17] it has been shown that de-pseudonymization is feasible in the Smart Grid and pseudonymization is vulnerable to linkage attacks. However, pseudonymization is only one tool from the PET toolbox. PET is rather a holistic concept than one technical solution. One main principle of PET is to reduce the amount of information to a minimum required for a specific application, i.e., data minimization. Another PET tool is the reduction of the data's accuracy or timeliness. However, the application of such a PET would result that in one way or another the data needs to be modified for privacy preserving reasons by a party other than the SM or the SMGW.

1.1 Problem #1: Balancing Data Utility (incl. Integrity and Authenticity) and Privacy

We see one problem in the opposing goals: On the one side the SG stakeholder needs access to integrity protected values gathered by a trusted untampered SM. On the other side consumer requires some trusted privacy component to perform data perturbation to protect the consumer's privacy. The main point we would like to raise is that the entity trusted to generate data is controlled and trusted by the SG stakeholder. With its goals and incentives to gather fine-grained data, this entity is untrusted to maintain the consumer's privacy. Vice versa, the SG stakeholder will not be able to rely on data gathered by an untrusted consumer-controlled device. Figure 1 depicts this situation.

1.2 Problem #2: Judging and Comparing Privacy Invasiveness

There is no debate that certain applications of the smart grid will need more data than others. At the moment exact nature of such future smart grid applications is unsure, so is the required data utility. This paper remains open towards future

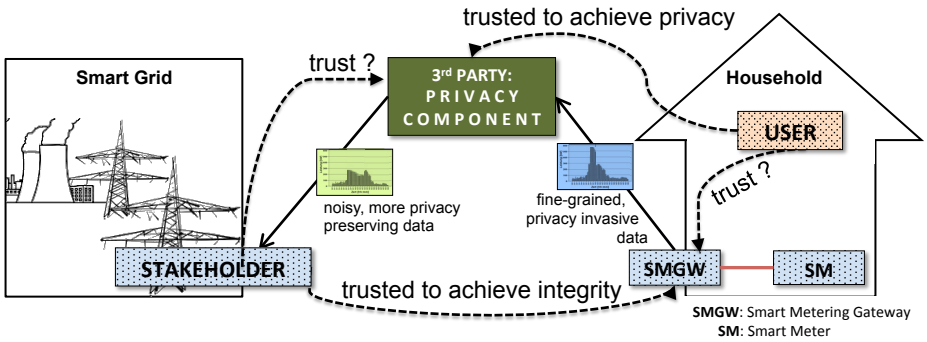

Fig. 1. Trust towards components by SG stakeholders and privacy-aware household

SG applications' need for data utility and future individual consumers' privacy-tolerances. However, we envision the need for a metric to compare and by this also judge the privacy-invasiveness of different applications. We believe that with an informed choice the user's willingness to participate in SG-applications will increase and that SG-applications will hence respect consumer's privacy preferences. Figure 2 shows that participation in applications are possible, if they require a data quality that is below the consumer's privacy preference. Privacy preserving mechanisms or unwillingness to participate limit the maximum data utility.

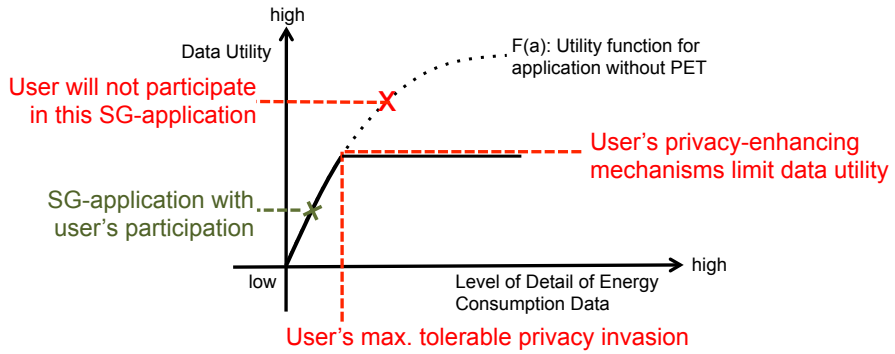

Fig. 2. Data-utility might be hindered by PET, but consumers will participate in applications within consumer's privacy preferences

1.3 Contribution

This paper describes a technology that allows balancing the conflicting interests of privacy and integrity[3]. We follow an approach called data perturbation, which is widely used in the field of privacy preserving data mining and differential privacy [9]. Data perturbation based mechanisms preserve privacy of distinct customers by letting an entity tamper with the data. We will call this entity the *privacy gateway* (PGW). The downsides of data perturbation are twofold: First it obviously must result in a reduced data utility and second the data tampering entity must be trusted. The first is an inherent problem of PET whereas the impact on utility needs to be limited to a level where the application is still executable. We counter the latter by applying a redactable signature instead of a classical digital signature at the SMGW.

The contribution of this paper is to provide a differential privacy guarantee in the BSI Smart Metering Setting (see Fig. 4) and to control the amount of integrity violations needed to achieve the privacy: We achieve control, integrity protection and origin authentication for the SG stakeholder by letting the SMGW sign a *range of values* around actual energy consumption using a redactable signature scheme ($\mathcal{RSS}$). The residential customer's *privacy gateway* (PGW) still has the possibility 'tamper' with the data to increase privacy by choosing *one value* out of the signed range.

We gain all the advantages of data perturbation combined with those of redactable signatures:
(1) data perturbation still allowing the stakeholders to address customers individually allowing for applications like providing energy efficiency recommendations;
(2) data perturbation gives an ad omnia privacy guarantee of differential privacy with a small computational overhead;
(3) redactable signatures allow the verifier to gain reassurance that the SMGW actually signed this value. Hence, the signer limits allowed values according to maximum tolerable reduction of data utility;
(4) redactable signatures allow third parties to do the choosing without any interaction with the signer, hence the customer does not need to trust a third party like a Smart Metering Operator (SMO) or the Smart Metering Gateway Administrator to protect her privacy.

2 System Description and Integrity Requirements

The BSI proposed a technical guideline [12] for intelligent metering systems. While this technical guideline is controversial discussed in literature due to its broad as well as expensive security and its slim privacy concept [28], it allows for a controlled data communication between a household and SG stakeholders. The concept is depicted in Fig. 3.

[3] which here includes accuracy

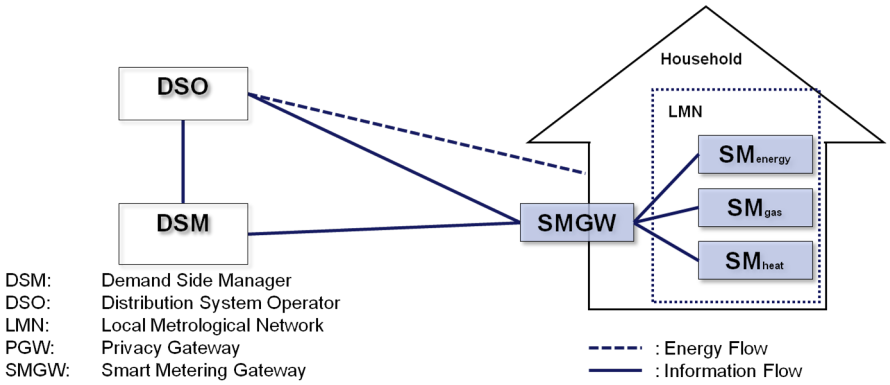

Fig. 3. BSI System Structure

SMGW checks whether a requesting stakeholder like a Distribution System Operator (DSO) or a Demand Side Manager (DSM) are allowed to access values like energy consumption or to send commands to the Controllable Local Systems (CLS). SMGW communicates via the residential Home Area Network (HAN) with CLS. In Addition the SMGW provides over the HAN data for the end consumer as well as the service technician. Within the Local Metrological Network (LMN) SMs for electricity, heat, gas and water are installed. SMs communicate consumption values to SMGW via the LMN.

Stakeholders like the DSO can ask the SMGW to get consumption data. The time interval between the gathering may vary but in the UK a collection rate once every 15 minutes is discussed and considered to be sufficient to guarantee net stability. Even finer grained consumption values are advantageous for forecasting.

3 Privacy Threats

Service providers in the SG like DSO or DSM need to collect data from individual households for their services. This data allows to infer information about households. The general research focus for privacy incursion has been about energy consumption values which are considered the household's output channel. Note that research barely considers the other direction, the input channel to the household. Inferred information of energy consumption values can be structured in the following three categories: First, appliance detection, second, use mode detection, and third, behavior detection. Note that all these attacks are possible for any party that has access to the plain data. Hence, encryption will help to protect the confidentiality during transmission of data, i.e., achieve privacy against third-parties, but will not mitigate privacy attacks by the party finally receiving and decrypting the plain data.

In the first category an analyzer tries to find out which appliances run in a household site. This information can be used for advertising purposes. In the second category an analyzer tries to find out how those devices are used. Experiments with high frequency data shows that even the TV channel can be deduced with a high percentage rate [15]. In the third category data is used to investigate how many people live in a household and what those people do. In [20] wake and sleep cycles as well as presence and absence have been deduced.

The information transmitted over the channel from SG service providers to the household bears privacy risks which depend on the application. Demand Response (DR) application allow to infer incentive sensitivity as well as a customer's preferences. In a simple version of DR the DSM ask the customer to reduce the amount of consumed energy in a certain time frame. In return the customer gets a financial compensation. To measure the compensation amount the DSM needs to know the energy consumption of this time as well as data to compare in order to determine the real reduction. This data can be the consumption from former periods. With this data and to know when the customer accepts and executes DR requests, the DSM can infer incentive sensitivity information of the customer.

To mitigate privacy threats appliance and use mode detection as well as behaviour deduction, several privacy enhancing technologies have been introduced. PET are based upon the principle of data minimization and concealing. The main drawback of those techniques are that either customers can not be addressed individually or that fine granular data is not available.

4 Differential Privacy: Perturbation to protect Privacy

A different approach than data minimization and concealing is the addition of noise to consumption data. While the outlook from the standpoint of privacy protection is very promising, the effect of the introduced error to data utility in SG is still in research. Data perturbation done in a right way, allows to reach the differential privacy ad omnia guarantee. The data perturbation is in general defined as the function $k()$. The following definitions are from [9].

Definition 1 (ϵ-differential privacy). *Be k a sanitizing algorithm, D_1 and D_2 two Databases which differ in at most one element, ϵ a privacy parameter which can be chosen and $S \subseteq Range(k)$.*

$$\frac{Pr(k(D_1) \in S)}{Pr(k(D_2) \in S)} \leq e^{\epsilon}$$

We use differential privacy as the basis for our metric. Especially, we use the calculation for the guarantee that if a single data record joins a dataset, the worst information leakage is e^{ϵ}. This rigid notion can be reached with limited computational overhead.

As an instantiation of using k to achieve privacy consider a DSO asking SMGW for current consumption data. The SMGW is retrieving this information and uses a function k, that adds noise taken from a Laplace distribution.

Definition 2 (Sanitizing Mechanism k). *The Sanitizing Mechanism k is* : $k(D) = f(D) + \mathfrak{L}(\frac{\Delta(f)}{\epsilon})$. *The mechanism is ϵ-differential private for all functions $f : D \to R^x$, where $\mathfrak{L}(\frac{\Delta(f)}{\epsilon})$ denotes the noise which is taken from the Laplace distribution, $\Delta f = max||f(D_1) - f(D_2)||$ and where D_1, D_2 differ in exactly one single dataset.*

Addition of noise as well as function f performed over the data base are done by a trusted entity, known as curator. In the SM case, the database needs to hold stored consumption values for specific points in time.

5 Redactable Signatures ($\mathcal{RSS}$): Fine control of Integrity

Assume the message to be signed is a set which contains ℓ values as elements: $\mathcal{M} = \{m_1, \ldots, m_\ell\}$. This paper uses a set-like notation without loss of generality.[4] The fundamental difference to classic signatures is that a $\mathcal{RSS}$ allows anyone to *redact* an element from the signed list, such that the signature still verifies. Basically, a redacted list no longer contains all elements from $\mathcal{M}$. Assume $\mathcal{R} \subseteq \mathcal{M}$, than removing elements in $\mathcal{R}$ from $\mathcal{M}$ leaves a subset $\mathcal{M}' = \mathcal{M} \setminus \mathcal{R}$. The most important differentiator between a classical signature is that a redactable signature scheme allows deriving an adapted signature σ', which still verifies. This action is called *redaction* and can be performed by anyone; the secret signing key is not required. Hence the original signer is not involved. However, a secure $\mathcal{RSS}$ is unforgivable comparable to classic digital signature schemes; this ensures that each element $m_i \in \mathcal{M}$ is protected against modifications other than complete removal. To continue the example, assume you redact all the other $\ell - 1$ elements, leaving only one value m_i in the signed set: $\mathcal{M}' = \{m_i\}$. Due to the $\mathcal{RSS}$ you can adjust the signature to σ'. A positive consecutive verification of the signature σ' over $\mathcal{M}'$ means that all elements in $\mathcal{M}'$ are authentic. In other words without use of the secret signing key you can produce a valid signature for remaining unchanged elements. Hence m_i that remained in $\mathcal{M}'$ can be verified to having not been altered and originating from the original signer, which remains identifiable via its public key.

Algortihmic Description of $\mathcal{RSS}$ The following notation is derived from [25], which is based of Brzuska et al. [5].

Definition 3 (Redactable Signature Schemes). *An $\mathcal{RSS}$ consists of four efficient algorithms $\mathcal{RSS} := (\mathsf{KeyGen}, \mathsf{Sign}, \mathsf{Verify}, \mathsf{Redact})$:*

KeyGen. *The algorithm* KeyGen *outputs the public key* pk *and private key* sk *of the signer, where λ denotes the security parameter:*

$$(\mathrm{pk}, \mathrm{sk}) \leftarrow \mathsf{KeyGen}(1^\lambda)$$

[4] Set-like notation eases understanding of the decomposition of a message as mathematical notions like intersection and union become applicable.

Sign. *The algorithm* Sign *gets as input the secret key* sk *and the message* $\mathcal{M} = \{m_1, \ldots, m_\ell\}$, $m_i \in \{0,1\}^*$*:* $(\mathcal{M}, \sigma) \leftarrow \mathsf{Sign}(1^\lambda, \mathrm{sk}, \mathcal{M})$

Verify. *The algorithm* Verify *outputs a decision* $d \in \{\mathtt{true}, \mathtt{false}\}$*, indicating the validity of the signature* σ*, w.r.t.* pk*, protecting* $\mathcal{M} = \{m_1, \ldots, m_\ell\}$, $m_i \in \{0,1\}^*$*:* $d \leftarrow \mathsf{Verify}(1^\lambda, \mathrm{pk}, \mathcal{M}, \sigma)$

Redact. *The algorithm* Redact *takes as input the message* $\mathcal{M} = \{m_1, \ldots, m_\ell\}$, $m_i \in \{0,1\}^*$*, the public key* pk *of the signer, a valid signature* σ *and a set of elements* $\mathcal{R}$ *to be redacted. It returns a modified message* $\mathcal{M}' \leftarrow \mathcal{M} \setminus \mathcal{R}$ *(or* $\perp$*, indicating an error):* $(\mathcal{M}', \sigma') \leftarrow \mathsf{Redact}(1^\lambda, \mathrm{pk}, \mathcal{M}, \sigma, \mathcal{R})$

We require the correctness properties for $\mathcal{RSS}$s to hold: Hence, every genuinely signed or redacted message will verify. A formal definition is given in [5].

Security of $\mathcal{RSS}$ This section describes the required security properties and models on an informal level, the formal properties are described and proven in [5, 6, 14, 25]. A secure $\mathcal{RSS}$ must be unforgeable and private to be meaningful [5]. Unforgeability allows detecting Integrity violations, e.g., only the genuine signed message or a valid redaction thereof can bear a valid signature created by the owner of the secret signing key. A public verification key linked to a a known entity and an unforgeable signature allows authentication of origin.

Unforgeability. No one should be able to compute a valid signature on a message not previously issued without having access to any private keys [5].
This is analogous to the unforgeability requirement for standard signatures [13], except excluding all valid redactions from the set of forgeries.The attacker can generate genuinely signed messages using an oracle, but has no access to the secret key. He has breached unforgeability if and only if he is able to compute a signature on a 'fresh' message, which is valid under the corresponding public verification key fixed at the beginning. A message is considered 'fresh' if it either has not previously queried from the oracle and if it can not have been created by one or more redaction(s) from a message queried from the oracle.

Privacy (weakly and a strongly) A *private* $\mathcal{RSS}$ prevents everyone except the signer from recovering any information (esp. the original value) about elements redacted, given the redacted $\mathcal{M}'$ and a valid signature σ' over $\mathcal{M}'$.
Note that information leakage through the modified message itself is out of scope. A weakly private $\mathcal{RSS}$ allows a third party to derive that elements have been redacted without gathering more information about their contents. Assume that each redacted element's value being replaced with $\square$ remains a visible element of $\mathcal{M}'$ [16]. The definition of a strongly private $\mathcal{RSS}$ is very similar, but redacted elements are considered *not* being visible as elements of $\mathcal{M}'$.

6 Solution: Signing a range of values with an $\mathcal{RSS}$

Solution towards problem #1. We allow the SMGW to provide the Smart Grid stakeholders like DSO and DSM with signed and henceforth trustable SM

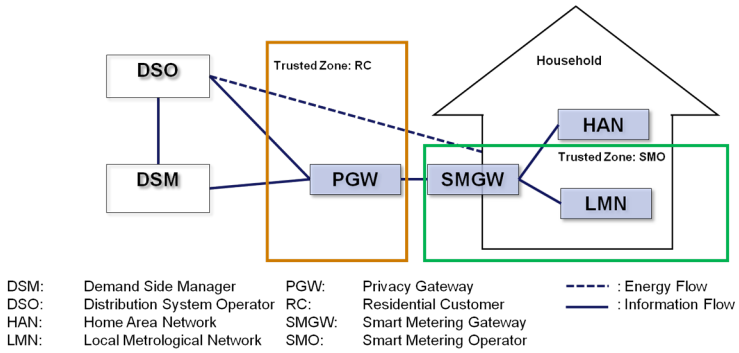

Fig. 4. System Structure with PGW

values, e.g., energy consumption values. At the same time, we allow the customer to achieve a desired level of privacy, by allowing the energy consumption value to be tampered with, e.g., adding noise. The party running PETs to achieve the consumer's privacy is termed Privacy Gateway (PGW). Our solution is depicted in Fig. 4. We assume that all information between the SMGW and the DSO and the DSM are running over the curator termed 'Privacy Gateway' (PGW).

Note that it is the SG stakeholder who knows and requests a desired level of data utility. This means in case of perturbation by noise to limit the maximum allowed noise. Of course, the SMGW could run privacy preserving algorithms directly and add noise to keep the customer's differential privacy. However this solution would require that the residential customer trusts the SM operator (SMO) to protect her privacy. The same problems occurs if the PGW is placed before the SMGW and would directly tamper with the readings from the SM. However, our solution allows the party doing the addition of noise to be trusted to preserve the customer's privacy, as the customer remains in full control. The task of the PGW is to tamper energy consumption values in order to protect the privacy of residential customers. The task of the SMGW is to sign the energy consumption values and the maximum tolerable perturbation in order to protect the integrity and trustworthiness of the SM readings. Both parties act on behalf of different stakeholders and hence are in different trust zone. Our solution uses redactable signatures to solves this conflict.

Solution towards problem #2. For brevity, we will now focus only on the transmission of a consumption value, other information that the SMGW sends alongside, like timestamps, are not considered.

The SMGW must make sure that values are not tampered in an unauthorized malicious way. Depending on the application DSO and DSM can tolerate a certain level of inaccuracy, e.g., allow that a certain amount of noise degrades their data utility. We denote the maximum amount of noise that can be added to an accurate reading by δ_{max}. Assuming SM measures the actual consumption

value v DSO/DSM will accept any reading in the range $[v - \delta_{max}, v + \delta_{max}]$ as valid. If the SMGW applies a classical signature scheme on v PGW can not tamper with data signed by SMGW without invalidating the signature. An invalid signature would indicate towards the DSO/DSM that the received value is not trustworthy, as it could have been maliciously tampered with in an arbitrary way. Henceforth, we assume that the SMGW will be instructed by the SMGW's operator about the tolerable noise, on behalf of the SG stakeholder. The tolerable noise depends on the required accuracy level for SG stakeholder's application. The actual values depend on the DSO or DSM application needs.

Note that fixing $\Delta = 2\delta_{max}$ in definition 1 allows calculating the maximum differential privacy that can be achieved. The PGW must be instructed by the consumer which level of privacy is tolerable for which optional applications. In this paper we assume that the consumer is free to not participate in an application for which his own personal privacy preference can not be achieved, i.e., PGW will not sent privacy-invasive data to a requesting SG stakeholder. However, we are fully aware that some communication must always be allowed for mandatory applications, e.g., net stability. For those mission critical mandatory SG applications we assume that the tolerable perturbation should be fixed by regulators.

6.1 Protocol Description

We propose the following phases: Setup, Signing, Adding Noise and Verification.

Setup:

1. Let $\mathcal{RSS} := (\mathsf{KeyGen}, \mathsf{Sign}, \mathsf{Verify}, \mathsf{Redact})$ be a secure (unforgeable and weakly private) redactable signature scheme.
2. After running KeyGen distribute the keys: SMGW gets a secret signing key sk and verification key vk, PGW and DSO/DSM get just the public SMGW's verification key vk.
3. SMGW is instructed by SMO which amount of noise it tolerates, and which accuracy is required.

Signing:

1. On receiving the actual consumption value v the SMGW calculates a range of discrete noisy values $\mathcal{M} = \{v - \delta_{max}, \ldots, v, \ldots, v + \delta_{max}\}$.
2. SGM signs $\mathcal{M}$ with an $\mathcal{RSS}$: $(\mathcal{M}, \sigma) \leftarrow \mathsf{Sign}(1^\lambda, sk, \mathcal{M})$.
3. SMGW sends $(\mathcal{M}, \sigma)$ to PGW.

Adding Noise:

1. On receiving $(\mathcal{M}, \sigma)$ PGW uses its database of historic values and the actual consumption value, which must be at the center of the range in $\mathcal{M}$, PGW runs the differential privacy algorithms to identify the value n in $\mathcal{M}$ which should be sent to DSO/DSM in order to satisfy $\frac{Pr(k(D_1) \in S)}{Pr(k(D_2) \in S)} \leq e^\epsilon$ where ϵ is a user predefined minimum required privacy parameter. The application execution is denied, if ϵ can not be reached.
2. PGW calculates $\mathcal{R} = \mathcal{M} \setminus n$.

3. PGW obtains a signature on $\mathcal{M}' = n$: $(\mathcal{M}', \sigma') \leftarrow \mathsf{Redact}(1^\lambda, pk, \mathcal{M}, \sigma, \mathcal{R})$.
4. PGW sends $(\{n\}, \sigma')$ to the DSO/DSM.

Verification:

1. On receiving $(\{n\}, \sigma')$, DSO/DSM uses the SMGW's verification key vk to verify if the signature on n is valid.

The amount of elements in $\mathcal{M}$ depends on the maximum noise and the accuracy, as $\mathcal{M}$ must contain concrete values, e.g., $\mathcal{M} = \{0.99, 1.00, 1.01, 1.02, 1.03, \ldots, 1.48, 1.49, 1.50, \ldots, 1.96, 1.97, 1.98, 1.99\}$ for an accuracy of two decimals, $\delta_{max} = 0.50$ and $v = 1.49$. The $\mathcal{RSS}$ limits the PGW only to redactions based on provided values, e.g., for $\mathcal{M} = \{1.11\}$. The PGW could generate a valid signature facilitating the algorithm Redact. However, the PGW can not generate valid signatures on values outside the range, e.g., $\mathcal{M} = \{0.98\}$ or $\mathcal{M} = \{2.00\}$. To do so would be as hard as forging the signature scheme of the $\mathcal{RSS}$, e.g., breaking the signature scheme like RSA-PSS [24, 4]. To counter replaying or repressing messages, the SMGW can just add a timestamp as an additional element into $\mathcal{M}$ requiring this to be fresh and present during verification.

6.2 Security and Privacy Properties

We assume: SM is trusted to perform correct readings, can not be attacked, and transmits the reading securely to SMGW.

Theorem 1. *Our protocol is unforgeable, if the $\mathcal{RSS}$ is unforgeable.*

SG stakeholders can detect any subsequent malicious manipulation of information while it is travelling through the network. Additionally they can use the SMGW's verification key to identify the origin of noisy data.

Theorem 2. *Our protocol achieves the highest differential privacy possible for $\Delta = 2\delta_{max}$, if the $\mathcal{RSS}$ is at least weakly private.*

Proof Intuition for Th.1 If the $\mathcal{RSS}$ applied by the SMGW is unforgeable, than neither PGW nor attackers can forge a valid signature on a value $n^* \notin \mathcal{M}_i$, where $\mathcal{M}_i$ denotes all sets signed and sent by the SMGW. Any such forgery would be a forgery in the $\mathcal{RSS}$.

Proof Intuition for Th.2 Assume all communication from SMGW will always pass through PGW, see Fig. 4. The $\mathcal{RSS}$ allows PGW to be a separate entity acting as instructed by the residential customer. PGW is limited by the range defined within the SMGW's signature but can run the algorithm Redact to select any suitable value out of the range. So seeing a valid $(\mathcal{M}, \sigma)$, which verifies using Verify under the trusted public verification key of a SMGW, that no malicious modification has taken place. Privacy of the underlying $\mathcal{RSS}$ guarantees that attackers can not identify the actual value of removed elements. Hence attackers can not know the actual consumption. We distinguish two cases:
(1) If the $\mathcal{RSS}$ is strongly private, i.e., elements are completely removed during

redaction, then the attacker sees a set $\mathcal{M}$ with exactly one element, i.e., $|\mathcal{M}| = 1$. (2) If $\mathcal{RSS}$ is weakly private, i.e., original values are hidden behind a special symbol ($\square^r$), then the attacker sees a set $\mathcal{M}$ with exactly one element being an actual value and $2\delta_{max}$ symbols, i.e., $|\mathcal{M}| = 2\delta_{max} + 1$.
Hence, if $\mathcal{RSS}$ is weakly private attackers can infer δ_{max}. However, attackers do never learn the actual values of removed elements, nor their position because its a set. Using the differential privacy mechanism described in Sect. 4, PGW adds noise within the range guaranteeing a differential privacy of ϵ.

7 Related Work

Techniques like group signatures [18] are based on the idea to hide the identity of household within a group. This prevents to address customers individually and thus limits potential SG applications to provide energy efficiency recommendations [2]. Another approach applies modifications inside the customers power circuit, e.g., consuming additional or less power from the grid by using a re-chargeable battery [3]. The downside of this approach are sever costs of the battery purchase as well as the maintenance effort. Those types are not optimal, due to the loss of addressing customers individually or the very high costs.

The concept of $\mathcal{RSS}$ was introduced by *Steinfeld* et al. [27] as "content extraction signatures" and almost at the same time by *Johnson* et al. as "homomorphic signatures" [19]. From their initial work many RSS constructions emerged in the last years [8, 21, 22]. Extensions working on more complex structures, e.g., trees [5], have been proposed, but a set is enough for the solution discussed in this paper. In [5] *Brzuska* et al. presented a formal security model. Note that according to this model many schemes are not secure, as they do not fulfil their notion of *Privacy* [5, 25]. Also note, that many schemes proposed are also only weakly private, i.e., one can see that a third party redacted something [16, 19, 22, 27, 29]. This generally gives more information to an outsider as already noted in [21]. In this paper we will not require transparency, thus we leak the range of noise, but the actual values of redacted elements stay private.

Several works try to identify which privacy relevant information can be inferred by analyzing energy consumption values [20, 23, 11]. it is shown that appliances, how the appliances are used and the behavior of the residential customers can be deduced by the energy consumption values. DR Application data holds additionally information about the incentive sensitivity. PET have been developed to minimize the amount of information which is sent by the SM [18, 26]. To the best of our knowledge only pseudonymization is considered to be applied. The minimization of information is either spatial or temporal [7, 18]. Temporal data minimization techniques provide only gross granular data, while spatial based data minimization do not allow to allocate energy consumption values to certain single households. While pseudonymisation allows to address single households, it is shown that this technique can be sidestepped by linkage attacks [17]. Data perturbation do not minimize data, but tamper it to protect privacy. The downside is the direct and severe impact on the data utility. This concept allows to

obtain the differential privacy guarantee for consumption values $[9, 1]$ as well as addressing customers individually.

8 Discussion and Open Questions

For any application of smart metering it is vital that the SG stakeholders receive *reliable* and trustworthy information. In this case *reliable* means that the SG stakeholder, e.g., a power grid provider, gets this information as (1) timely and as (2) accurate as needed for the SG application. The exact level of accuracy and timeliness will vary depending on the application itself, but also on the actual contractual, regulatory and installation setting, and is beyond the scope of this paper. In our construction the SM operator (SMO) limits the range in which data perturbation, in our case the addition of noise, is considered acceptable by applying a redactable signature ($\mathcal{RSS}$) at the SMGW over a range of the SMO's choosing. Knowing the allowed level of accuracy allows the customer's privacy gateway (PGW) to calculate the differential privacy guarantee that it could achieve using the data perturbation mechanisms it could deploy. With this information the PGW can independently judge if the allowed perturbation is enough to keep a sophisticated level of privacy for the customer. If not, it can withhold the information until the customer explicitly consents to this leaking of PII. If the PGW has enough freedom it will adjust the data accordingly and forward it after the modification. A $\mathcal{RSS}$ allows this alteration of signed data and the SG stakeholder can verify if the change was within his defined limits.

The presented idea differs slightly from the general idea of differential privacy. In differential privacy ϵ is chosen under the perspective to protect privacy. Our idea is to regard the application side and limit the noise to its needs. This allows calculating the ϵ depending on the maximum amount of noise that the data perturbation mechanism k is allowed to apply. The amount of noise is defined by the max. error which is acceptable for a SG application.

This approach can be criticised for its weak privacy protection. Very small noise will allow appliance detection and behaviour deduction. It remains unclear to which extent this small noise prohibits use mode detection. Due to the need of very fine grained data to get an acceptable quality level for use mode detection we assume the reduced accuracy by noise will limit invasiveness. It can also be argued that as the SG stakeholder controls the amount of noise, it can limit the privacy protection by setting a too low limit. Further investigation and discussion for concise applications with known data quality needs is required.

However, our approach creates a metric for the privacy loss, which can be used to compare privacy invasiveness of different applications from different SG stakeholders. The metric is to compare the maximum differential privacy (ϵ) that can be achieved if the allowable noise, and by this the data utility, has been fixed. As in general, several applications will be provided in the smart grid, each application and each application provider can in theory require a different degree of data utility, e.g., data precision. With the given metric, the consumer is able to compare the privacy invasiveness of any given application. Henceforth,

we envision the customer to exercise an informed choice and either accepts or rejects to participate in the application. To illustrate the idea consider a SM which solely gathers consumption values for net stability. This is an essential and required basic application in SG. The data quality needed for those mission critical mandatory SG applications must follow data protection's principal of data minimization, probably under a close watch by regulators. Here, the user needs to accept this privacy loss, there is no real choice other than to participate. Given the maximum amount of noise for this application leads to a worst case privacy loss of $\epsilon_{netstab}$. Now, a new demand-response (DR) application is advertised to the customer. We assume the DR application tolerates only a smaller amount of noise, the worst case privacy loss is denoted as ϵ_{DR}. The consumer is now able to use the calculated worst case privacy losses for comparison. For example, a comparison value $\frac{\epsilon_{DR}}{\epsilon_{netstab}}$ greater than 1 will indicate that the optional DR-application will result in a decrease of privacy compared to net stability. This comparison can also be done to choose from different DR applications. Further research must show and define the needed accuracy for certain SG applications. While the value proposed for a privacy metric in this paper itself is still abstract, further research could use it to compare the privacy guarantees for concrete applications. Furthermore, user studies could help to show which loss of privacy is accepted by users and craft privacy endangerment statements depending on several ϵ, e.g., a traffic light system. Finally, we remark that current research barely considers the privacy impact of the input channel to the household.

References

1. G. Ács and C. Castelluccia. I Have a DREAM!: Differentially Private Smart Metering. In *Proc. of IH'11*, pages 118–132. Springer-Verlag, 2011.
2. H. Allcott. Social norms and energy conservation. *Journal of Public Economics*, 95(9–10):1082 – 1095, 2011. Special Issue: The Role of Firms in Tax Systems.
3. M. Backes and S. Meiser. Differentially private smart metering with battery recharging. *IACR Cryptology ePrint Archive*, 2012:183, 2012.
4. M. Bellare and D. Micciancio. A new paradigm for collision-free hashing: incrementality at reduced cost. In *Eurocrypt'97*, pages 163–192. Springer-Verlag, 1997.
5. C. Brzuska, H. Busch, O. Dagdelen, M. Fischlin, M. Franz, S. Katzenbeisser, M. Manulis, C. Onete, A. Peter, B. Poettering, and D. Schröder. Redactable Signatures for Tree-Structured Data: Definitions and Constructions. In *Proc. of the 8th ACNS'10*, pages 87–104. Springer, 2010.
6. C. Brzuska, M. Fischlin, T. Freudenreich, A. Lehmann, M. Page, J. Schelbert, D. Schröder, and F. Volk. Security of sanitizable signatures revisited. In *Proc. of PKC'09*, pages 317–336. Springer-Verlag, 2009.
7. T.-H. H. Chan, E. Shi, and D. Song. Privacy-preserving stream aggregation with fault tolerance. In *Financial Cryptography*, volume 7397 of *LNCS*, pages 200–214. Springer, 2012.
8. Ee-Chien Chang, Chee Liang Lim, and Jia Xu. Short Redactable Signatures Using Random Trees. In *Proc. of CT-RSA*, CT-RSA '09, pages 133–147. Springer, 2009.
9. Cynthia Dwork. Differential privacy. In *ICALP (2)*, pages 1–12, 2006.
10. Robert Earle, Edward P. Kahn, and Edo Macan. Measuring the capacity impacts of demand response. *The Electricity Journal*, 22(6):47 – 58, 2009.

11. M. Enev, S. Gupta, T. Kohno, and S. N. Patel. Televisions, video privacy, and powerline electromagnetic interference. In *ACM CCS*, pages 537–550. ACM, 2011.
12. Bundesamt für Sicherheit in der Informationstechnik. BSI TR-03109 @ONLINE, 2011.
13. S. Goldwasser, S. Micali, and R. L. Rivest. A Digital Signature Scheme Secure Against Adaptive Chosen-Message Attacks. *SIAM Journal on Computing*, 17:281–308, 1988.
14. J. Gong, H. Qian, and Y. Zhou. Fully-secure and practical sanitizable signatures. In *Information Security and Cryptology*, volume 6584 of *LNCS*, pages 300–317. Springer-Verlag, 2011.
15. U. Greveler, B. Justus, and D. Löhr. Identifikation von Videoinhalten über granulare Stromverbrauchsdaten. In *Sicherheit*, volume 195 of *LNI*, pages 35–45. GI, 2012.
16. S. Haber, Y. Hatano, Y. Honda, W. G. Horne, K. Miyazaki, T. Sander, S. Tezoku, and D. Yao. Efficient signature schemes supporting redaction, pseudonymization, and data deidentification. In *ASIACCS*, pages 353–362, 2008.
17. Marek Jawurek, Martin Johns, and Konrad Rieck. Smart metering depseudonymization. In *ACSAC*, pages 227–236, 2011.
18. T. Jeske. Privacy-preserving smart metering without a trusted-third-party. In *SECRYPT*, pages 114–123. SciTePress, 2011.
19. R. Johnson, D. Molnar, D. Song, and D.Wagner. Homomorphic signature schemes. In *Proceedings of the RSA Security Conference - Cryptographers Track*, pages 244–262. Springer, Feb. 2002.
20. M. A. Lisovich, D. K. Mulligan, and S. B. Wicker. Inferring personal information from demand-response systems. *IEEE Security and Privacy*, 8(1):11–20, 2010.
21. K. Miyazaki, G. Hanaoka, and H. Imai. Digitally signed document sanitizing scheme based on bilinear maps. In *Proc. of ASIACCS '06*, pages 343–354, New York, NY, USA, 2006. ACM.
22. K. Miyazaki, M. Iwamura, T. Matsumoto, R. Sasaki, H. Yoshiura, S. Tezuka, and H. Imai. Digitally Signed Document Sanitizing Scheme with Disclosure Condition Control. *IEICE Transactions*, 88-A(1):239–246, 2005.
23. A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin. Private memoirs of a smart meter. In *Proc. of 2nd ACM BuildSys '10*, pages 61–66. ACM, 2010.
24. R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 26(1):96–99, 1983.
25. K. Samelin, H. C. Pöhls, A. Bilzhause, J. Posegga, and H. de Meer. Redactable signatures for independent removal of structure and content. In *ISPEC*, volume 7232 of *LNCS*, pages 17–33. Springer, 2012.
26. E. Shi, T.-H. H. Chan, E. G. Rieffel, R. Chow, and D. Song. Privacy-preserving aggregation of time-series data. In *NDSS*. The Internet Society, 2011.
27. R. Steinfeld, L. Bull, and Y. Zheng. Content extraction signatures. In *Proc. of 4th ICISC '01*, volume 2288, pages 163–205. Springer, 2002.
28. David von Oheimb. IT security architecture approaches for Smart Metering and Smart Grid. In *SmartGridSec*, pages 1–25, 2012.
29. Z.-Y. Wu, C.-W. Hsueh, C.-Y. Tsai, F. Lai, H.-C. Lee, and Y. Chung. Redactable Signatures for Signed CDA Documents. *J. of Med. Systems*, pages 1795–1808, 2012.
30. H. Ziekow, C. Goebel, J. Strüker, and H.-A. Jacobsen. The potential of smart home sensors in forecasting household electricity demand. In *SmartGridComm*, 2013.