Privacy-aware Smart Metering: A Survey

Sören Finster and Ingmar Baumgart

***Abstract*—The increasing share of renewables creates new challenges for the existing electrical grid. To deal with these challenges, various efforts are being made to transform the existing power grid into a so-called smart grid. Part of this process is the deployment of an *advanced metering infrastructure*, which provides novel high-frequency two-way communication between consumers and producers. But as useful as the access to high-frequency measurements may be for energy suppliers, this also poses a major threat to the privacy of the customers. In this survey we present approaches to the problem of customer privacy-protection in the smart grid. We show that the privacy problem in smart grids can be further divided into the problems *metering for billing* and *metering for operations*. For each of these problems we identify generic approaches to solve them.**

***Keywords*—Smart Grid Communications, Trust, Privacy, Advanced Metering Infrastructure**

I. Introduction

TODAY'S electrical grid is changing rapidly to address the demands of distributed power generation. The production of energy from a large number of small scale sources, rather than a few hundred power plants, requires adjustments and at least partial redesign of the grid itself. One important feature of the new *smart grid* will be an *advanced metering infrastructure*. It enables distribution system operators and energy suppliers to optimize their existing services and even provide new services for their customers.

Among other functions, the most important one of the advanced metering infrastructure is the possibility of two-way communication offered by smart meters. Distribution system operators can use this functionality to monitor their energy grid at a much higher sampling rate and granularity than before. Energy suppliers can use the near real-time consumption data to control their production more efficiently and to offer their customers pricing schemes based on current offer and demand.

But this new functionality also has risks. The close monitoring of single smart meters provides deep insight into the energy consumption of customers. And with detailed knowledge of their energy consumption surprisingly accurate conclusions about their private life can be drawn. The privacy of, for example, working hours, vacations, habits and even religious beliefs is at risk.

Legal measures to prevent privacy intrusion are taken in most countries. But legal measures may not be enough, especially when theft of privacy relevant data is considered. Therefore, technical measures to protect customers privacy are equally important. In this survey, we take a look at the current state of research in the area of privacy protection in the advanced metering infrastructure of the new smart grid.

Manuscript received April 1, 2013; revised January 10, 2014; accepted April 9, 2014.

S. Finster and I. Baumgart are with the Institute of Telematics, Karlsruhe Institute of Technology (KIT), Germany, (e-mail: {finster, baumgart}@kit.edu)

The remainder of this paper is as follows: In Section II, we provide some background. We shortly introduce smart metering and show some work that proves how smart metering can be used to intrude customer privacy. We formulate the two important problems that have to be solved concerning privacy and smart metering: *metering for billing* and *metering for operations*. We close the background section with a list of terms we will use in the remainder of the paper.

The following sections are categorized into works that examine metering for billing (Section III), works that examine metering for operations (Section IV) and works that concentrate on both problems using a single approach (Section V).

Each section contains subcategories for each generic approach to solving the specific problem. Interesting features of single works are also discussed. Each subcategory closes with a short summary.

Finally, we compare the generic approaches in Section VI and close with concluding remarks in Section VII.

II. Background

A. What is Smart Metering

Smart metering is a crucial part for the realization of the vision "Smart Grid" [1]. In its most basic form, it describes a deployment of electric meters that enable two-way communication between meter and distribution system operator. This is often called the *Advanced Metering Infrastructure (AMI)*.

The two-way communication enables several services for the distribution system operator that were difficult or impossible to realize without smart metering. Detection of power outages, for example, was solved via phone calls from customers without power. Now, the distribution system operator can detect the power outage faster and without interaction with the customer. Reporting the quality of power delivery (voltage, frequency) is another service enabled by smart metering.

One of the most anticipated features of smart metering is monitoring of power flows within the distribution system. Before smart metering, this information was only available at the substation level. With smart metering, a detailed view on power flows is possible. This is especially important with the rising number of renewable energy sources connected to the grid.

Accurate monitoring of power flows enables energy suppliers to react rapidly on changes in consumption levels. If prices for energy are variable and react on current power flow information, we talk of *real-time pricing*. This is one possibility for *demand side management (DSM)*, a smart grid technology where the energy supplier influences the consumption of energy directly and immediately.

But not only distribution system operators and energy suppliers benefit from smart metering. Customers benefit from smart meter deployment by receiving timely information about their power consumption. This information is often processed by the energy supplier and presented to the customer as a web service or as an application for a mobile phone. It is also possible to process the data at the customer's site and present it as a local web service.

For the remainder of this paper, when we talk about smart metering, we mean the provision of meter measurements to the distribution system operator or the energy supplier.

B. Why Smart Metering should be privacy-aware

Nonintrusive load monitoring (NILM) is the interpretation of power load signatures with the intent to obtain information about the appliances causing the load. The first NILM devices were built in 1985 [2]. They could distinguish certain power events in the load signature and assign them to individual appliances. For example, they reported when the dryer ran or the toaster was switched on. This effectively revealed that the residents were at home and hungry. A feasibility study conducted in 1991 concluded that the generic isolation of single appliance loads from an aggregated load is possible [3]. Early results of the technology led NILM pioneer G. W. Hart to conclude that legal measures should be taken *"so that electric power usage is considered as private as any phone conversation"* [4].

Hart used power loads that were measured every five seconds. Other works concentrated on improving NILM using better algorithms and with measurements at a higher time resolution (e.g., [5]–[10]). But there are also efforts to provide accurate NILM with lower time resolutions up to one hour (e.g., [11]–[13]).

Millions of smart meters capable of such or even higher resolutions are already installed [14]. Often with questionable security features [15]. These are already collecting data and enable anyone who has access to this data to invade the privacy of residents. The dimension of privacy intrusion should not be underestimated: from working hours to religious beliefs, a plethora of immensely private data is readily available [12]. In its most harmless form, people might receive advertisements for appliances they do not already own or that might break in the near future. Thieves knowing exactly when a single resident or even a whole neighborhood is at work are a much more scary vision.

The potentially harmful effect of exposing this data makes it a valuable target for data thieves. Besides the question of what the electricity supplier or distribution system operator does with the metering data, it also rises the question how well this data is protected against data theft.

Given the potentially harmful effect of exposing this data, Hart's advice to seize legal measures to protect it might be not enough. In this paper, technical approaches to solve the problem of privacy in smart metering are discussed.

C. Two problems of Smart Metering

To guarantee perfect privacy for residents, no data about their energy load should have to be measured at all. This is clearly not feasible for two reasons. First, the energy bill has to be paid. The energy supplier needs measurement data to write that bill. Second, for the vision of the smart grid to come true, distribution system operators need high-frequency metering data to monitor where power in their grid is produced and where it is consumed. This enables a better match between production and consumption and is a prerequisite for real-time pricing.

The first reason poses a problem that we will call *"metering for billing"*. A solution to this problem would be the possibility for a household to pay for the used energy without revealing too much about itself. The second reason results in *"metering for operations"*. A solution for this problem must provide measures for monitoring the power grid while not monitoring individuals too closely.

Concerning data privacy, there are three important factors: *sampling frequency*, *attribution* and *exactness*. The lower the sampling frequency, the less information is deducible from the data. An example for an extremely low sampling frequency would be the yearly meter reading that was common before smart meters emerged. If data is not attributable to a specific household, its privacy is protected. Privacy can also be protected by not reporting exact measurements but the usefulness of measurements must be somehow retained.

Metering for billing needs perfect attribution and exactness but is flexible with sampling frequency—as long as the billing is correct, the frequency does not matter too much. Metering for operations needs a high sampling frequency and at least some exactness, but is flexible with attribution. For monitoring the grid, it is often sufficient to collect aggregated data about specific parts of the power grid. For example a block or all households connected to a substation.

The important difference between those problems is that metering for billing is concerned about data that has to be directly accountable to the end user. Metering for operations deals with data about a part of the power grid. As a consequence, both problems are solvable independently.

Metering for billing got much less attention in research than metering for operations. In most works concerning smart metering, the billing problem is not even covered. The reason for this is that there is an obvious approach to the problem that solves it acceptably: sampling frequency is reduced to the billing interval. As long as this interval is long enough, the data does not pose a privacy threat any more. For example, an extremely simple tariff that does not take into account the time at which energy gets consumed can be billed by using only the total consumption for the billing period. If the billing period is long enough (e.g., a year) the privacy of the end-user is not threatened since the meter already aggregates the readings over the year.

But with shorter billing periods or *time of use (TOU)* tariffs, the data needed to bill the end-user may pose a risk concerning privacy. Since time based tariffs and flexible pricing models are much wanted features of the smart grid, there is a need for better solutions to the problem. In Section III approaches to this problem are discussed.

Metering for operations does not have such an obvious, good enough solution and therefore got more attention in the research community. An extremely wide field of approaches to the problem emerged and is covered in Section IV.

D. Terminology

There are a lot of terms used throughout the works concerning smart metering. Sometimes, they are specific to a certain country, like the German "Messdienstleister" which is an entity resulting from the legally required liberalization of the energy market. But in general, most terms are translatable to a limited set of terms. We try to use only this limited set defined in the following section. If, for some reason, this set of terms does not suffice, we will specifically introduce and define the new term. See also Figure 1 for a graphical representation of the scenario.

We use *customer (C)* to describe the target of smart metering. This includes every entity that partakes in the electrical grid as consumer or producer. It is best thought of as a household with appliances and decentralized energy production like a *combined heat and power (CHP)*. But the

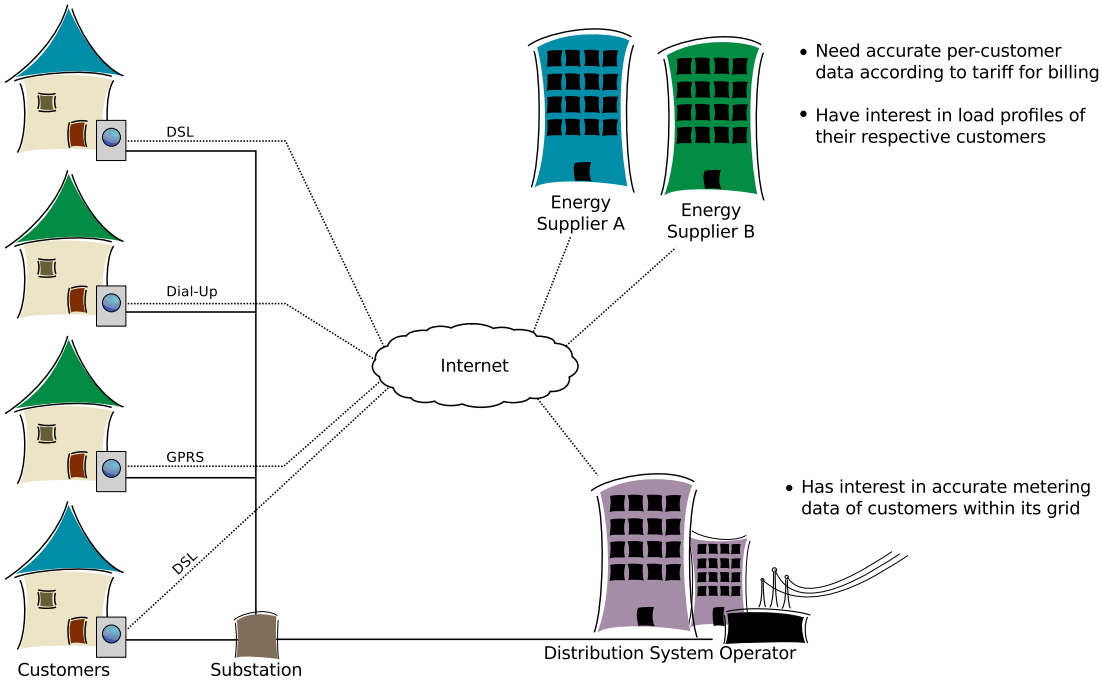

Fig. 1. Exemplary scenario with customers within a power grid which do not have the same energy supplier. As communication network, the Internet is used. Internet access for customers can be realized with different technologies resulting in different data rates. While the distribution system operator is interested in metering data from all customers within its power grid, energy suppliers are only interested in data from their direct customers. This can also include customers within other power grids operated by other distribution system operators (not depicted).

term is not limited to this. It also includes businesses or industry. If it consumes or produces energy and gets billed for it, we will call it a customer.

At a customer's site, a *smart meter (SM)* is installed. This is a metering device that is capable of metering at a fast sampling rate. It also provides means to communicate meter readings to another entity. It is usually tamper proof and has only a limited set of capabilities. Especially for works that involve advanced cryptography, these capabilities may be too limiting. If another device is needed at the customers site to fulfill all requirements, we will call that device *meter extension (ME)*.

The *distribution system operator (DSO)* manages the electrical grid. It is concerned with balancing consumption and production within the power grid. Therefore it has an interest in accurate and timely meter readings for planning and to react on changes in the power grid.

An *energy supplier (ES)* is an entity that sells or buys energy to customers. The energy is then transported by the power grid which is managed by the DSO. Its function is comparable to a trader and less technical than the DSO. Its main reason to access smart meter readings is to bill the customer. But for strategic planning and for complex tariffs, the ES can also be interested in high-frequency meter readings. In some works, the DSO is not included and the ES takes over both roles.

A *trusted third party (TTP)* is introduced in several works. It is an additional entity which often communicates with all other entities. As the name implies, it is trusted by at least one of the other entities. In most cases, all involved entities trust the TTP.

Demand side management (DSM) is the name of the scheme that tries to improve efficiency and reliability of the power grid by influencing the energy consumption. This can be done in several ways. From educating customers to remotely controlling devices in the household, e.g., air conditioning. In most smart grid scenarios it is present in the form of real-time pricing. Raising prices in peak hours provides incentives for customers to lower their energy consumption. The monitoring capabilities of a wide roll-out of smart meters helps to improve DSM. A less flexible form of DSM are *time of use (TOU)* tariffs that have different, but stable prices for energy throughout the day according to a timetable.

The definitions of privacy terms is in accordance with the terminology for talking about privacy by Pfitzmann and Hansen [16]. In example, anonymity means that the subject is not identifiable within a set of subjects and pseudonymity means that a pseudonym is used as identifier instead of the real identifier.

E. Standardization

The establishment of industrial standards for smart metering techniques is an important part for the feasibility of the future smart grid. Currently, there are extensive activities in standardizing components and communication between components of the advanced metering infrastructure. Tasks like remote meter reading are well covered through several standard smart meter communication protocols like DLMS/COSEM or SML (see [17] for a comparison) and standard protocols for house automation like ZigBee, M-Bus or KNX (see, e.g., [18]).

And yet, with standardization still focused on making the smart grid feasible, security features have higher priority and privacy enabling features are scarcely covered. Most works in this survey, especially approaches using direct communication paths between smart meters, use non-standardized functionality to enable privacy. For further reading on the current state of standardization [19]–[22] are suggested.

III. Approaches to Metering for Billing

The problem of metering for billing is often considered solved because of the very coarse-grained measurements that are sufficient. For a very simple tariff, where only the total consumption per billing period is necessary to calculate the bill, this is indeed true[1]. Even simple time-based tariffs can be billed by using meters equipped with several registers and register-switching based on the time of day. If, for example, energy consumption during normal workdays is added to one register and energy consumption during the night and the weekend to another, the customer can be billed monthly based on those two registers.

But this solution only holds, if tariffs remain fairly simple and aggregate over very coarse-grained consumption values. As soon as tariffs are more complex privacy is at risk.

A true solution to metering for billing enables the energy supplier to use complex tariffs and to bill the customer correctly without violating his privacy.

There are three generic approaches to this problem:

- Let a trusted third party calculate the bill
- Let the customer calculate the bill and ensure correctness via trusted computing
- Let the customer calculate the bill and ensure correctness via cryptography

A. Billing via a Trusted Third Party

Calculating the bill using a TTP is an easy and elegant solution. In this approach, SMs send all measurements to the TTP. The TTP aggregates them per smart meter and over a specified period of time before it sends the aggregated consumption to the ES. But it has a huge drawback: the trust which both parties need to put into the TTP is enormous. The customer trusts the TTP with sensitive private data and the correct calculation of his bill. The ES trusts the TTP with the foundation of his billing, a central aspect of his business. Why both parties should put that much trust in the same TTP simultaneously is an open question.

Bohli et al. [23] provide an extension to the basic variant. They suggest that ES and SM sidestep the TTP to agree upon an arbitrary identifier for the SM. The TTP then operates only with those arbitrary identifiers and has no identifying information about the SM. Essentially, this is a case of pseudonymization. This concept is illustrated in Figure 2.

One problem of this approach is the transfer of metering data from the SM to the TTP. Depending on the used communication network, the TTP gains knowledge of the network address (e.g., IP address) of the SM. This alone does not pose a threat, but if the TTP combines this knowledge with other information sources (e.g., visits to certain websites) the identity of the SM is at risk.

Further, the advantages of pseudonymization are doubted by Jaruwek et al. [24]. Using anomaly detection and behavior pattern matching, they show that pseudonymization leaves multiple attack vectors open. Even though their approach needs accurate secondary data sources (e.g., working hours, vacations), it poses a threat to the effectiveness of pseudonymization.

[1]There is still a leakage of private information if the billing interval is short enough. With weekly billing, for example, a vacation is easily detectable. Solving this problem would need fiscal escrow services, which is out of scope of this paper.

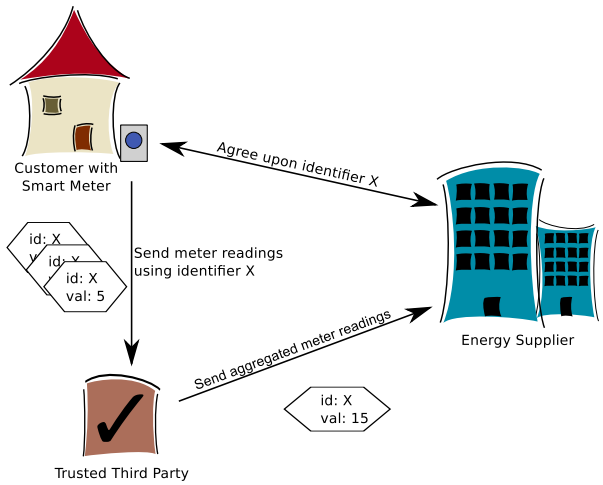

Fig. 2. Billing via a trusted third party: The customers Smart Meter and its energy supplier agree upon an arbitrary identifier. The Smart Meter then transmits meter readings to the trusted third party using this identifier. The trusted third party aggregates those meter readings and submits them to the energy supplier using the provided identifier.

B. Billing using a trusted platform

A different approach to metering for billing is to use a *Trusted Platform Module (TPM)* [25] within the smart meter. This is suggested by Petrlic [26] and LeMay et al. [27].

By isolating the calculation of the bill within the TPM, its integrity and confidentiality is ensured as long as the TPM is secure. In essence, trusted computing seals off the software and data involved into bill calculation from the rest of the system. It provides access only to data and software that the TPM deems trustworthy.

This approach is often used in a more generic approach to security in the smart grid. By rendering the smart meter into a trusted platform, a diversity of applications are enabled to run locally at the customers site. Even old fashioned meters are, in essence, trusted platforms since they are designed to be tamper proof.

The main problem of this approach lies in the fact that the TPM provides security for the ES but does not necessarily provide privacy for the customer. What exactly is executed within the TPM is the responsibility of the ES. If the ES issues code that transfers sensible information, the whole TPM does not help. In conclusion, either the ES needs to be trusted with the private data or some auditing process is necessary to check if the software to be run is deemed trustworthy. Since customers typically can not do this themselves, a third party has to provide trust in the software. This boils down to a TTP solution.

C. Billing secured via cryptography

To provide a solution where no preliminary trust is needed, several works suggest the usage of cryptographic commitments [28]. Using those commitments, the SM calculates the bill and additionally provides proof that it calculated it correctly.

A commitment is a cryptographic tool that enables one party to provide a commitment[2] $c = Commit(x, r)$. Given only c, it is hard to compute x. A commitment can be opened using c, x and r which returns true or false. Given c, x and

[2]Given both parties agreed upon the generators of the specific commitment scheme.

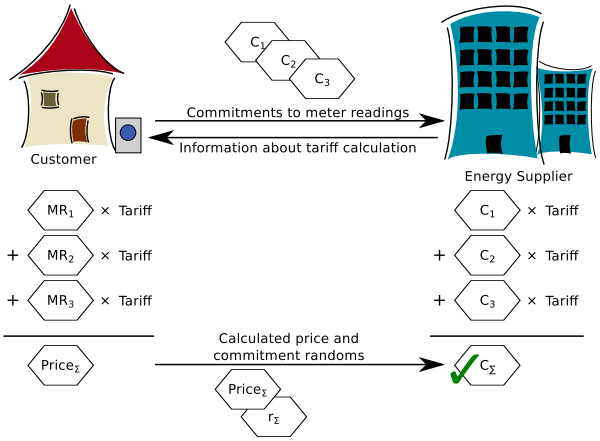

Fig. 3. Billing through Commitments. For each measurement, the SM generates a commitment which is sent to the ES. With the commitments alone, the ES does not gain any insight into the energy consumption of the customer. The SM uses tariff information and meter readings to calculate a final price and performs the same calculation in parallel with the random numbers used for the initial commitments. The final price and randoms are transferred to the ES. The ES can open the result of the calculation on the received commitments with the provided final price and randoms and therefore can be sure that the price was calculated correctly.

r it is hard to compute an $x' \neq x$ and matching r' so that $Open(c, x', r')$ returns true. So, once a party transmitted a commitment c, it is bound to the used values x and r.

The most often used cryptographic commitment scheme in smart metering is by Pedersen [29]. Additionally to the above mentioned characteristics, it offers homomorphic features. By multiplying two commitments, their parameters are added. By exponentiation, their parameters can be scaled. See equation (1).

$$\begin{aligned} Commit(x,r) \cdot Commit(y,s) &= Commit(x+y, r+s) \\ Commit(x,r)^k &= Commit(x \cdot k,\ r \cdot k) \end{aligned} \tag{1}$$

This property is exploited to perform billing (see Figure 3). The SM or a trusted meter extension (ME) creates commitments $c_i = Commit(x_i, r_i)$ for every measurement x_i using a random number r_i. These commitments are then signed with a key the SM got via a *Public Key Infrastructure (PKI)*. Then, they are transferred to the ES. Using the PKI, the ES can now verify that the commitments were signed by the trusted component and indeed are commitments to actual measurements of the SM.

The ES provides the ME with tariff information for a time period. The tariff information consists of a list of prices for a specific time frame. The ME uses the tariff to calculate the billing price by multiplying each measurement x_i with its corresponding price from the tariff information and summing them up. Additionally, it does the same with the random numbers r_i used in the earlier commitments. It then transfers the final price and the calculation of the random numbers to the ES.

The ES checks if the price is correct by performing the calculation on the received commitments c_i. It ends up with a new commitment that must open to true if the correct price and the correct random number is used. Both are provided by the ME.

This scheme only works for basic tariffs where every measurement simply has to be multiplied by a factor. More complex tariffs, for example with additional costs for going over a certain consumption threshold, are not possible.

This is the basic principle used by Jawurek et al. [30] and Molina-Markham et al. [9].

Rial and Danezis [31] provide a solution that is also based on commitments. But in addition to the basic case, they also provide a protocol that is able to perform complex tariffs. They use an extensive toolbox of cryptographic functions and zero knowledge proofs that are out of scope for this paper.

D. Summary

Since metering for billing effectively is about handling money, it is mainly concerned with trust. All three generic approaches to this problem differ mainly in who they trust.

The solutions of Section III-A solve the problem by inserting a trusted third party into the chain. Since someone or something needs to be trusted, the introduction of such a trustee is an easy solution. But the enormous trust put into this party makes it difficult to decide who should be this party.

The approach of using trusted computing (III-B) essentially renders the metering infrastructure into a trusted party. But the security of the walled TPM garden is up to discussion and often doubted to withstand profound attacks [32]. Essentially, the providers of the TPM platform have to be trusted to provide a secure platform.

Using cryptography (III-C) to ensure correct billing offers the possibility to put the trust in science. The ES has to trust, that the SM emits and transmits correct commitments. Since this is not a complex operation, it can be done within the SM and be trusted as much as the emitted meter data itself is trusted. Since the customer only transfers commitments and a final price, his privacy remains protected. This approach looks especially promising if simple tariffs are used. More complex tariffs are difficult and require a much bigger set of cryptographic primitives. To support them, the complexity of software is immense and thus the probability of errors grows. This limits acceptance of such solutions by the industry.

IV. Approaches to Metering for Operations

The problem of metering for operations leaves more room for different approaches than the billing problem. With its main application to provide monitoring capabilities for the distribution system operator, it is up to the definition of that task to decide when exactly the problem is considered as solved.

A solution to metering for operations enables the DSO to monitor the grid without violating the privacy of customers. What capabilities the task of “monitoring the grid” exactly contains differs from work to work. A general consensus seems to be found in the fact that the DSO only needs to monitor sections of the power grid as a whole which enables aggregation as a form of privacy enhancement. The granularity of aggregation, i.e. how many customers are aggregated together, is a trade-off between privacy and utility that is not further examined in the covered approaches.

Since there are plenty of different approaches, this section is structured into several subsections. Each subsection discusses a different approach to the problem. First the approach is described in a generic way, then the works that use the examined approach are discussed.

A. Provide anonymization or pseudonymization without aggregation

A simple approach to solve the problem is the removal of identifying features from all meter readings. This results

either in completely anonymous meter readings or in meter readings with pseudonyms.

To perform some form of authentication, most works introduce a *trusted third party (TTP)*. The TTP acts as an intermediary and has to be fully trusted by both, the DSO and the customer. While collecting meter readings from SMs, the TTP can authenticate the smart meter and remove identifying information or substitute it with pseudonyms.

In papers published by Petrlic [26] and Molina-Markham et al. [9], the TTP acts as an information relay and just passes measurements from smart meters to the DSO. The measurements are signed and encrypted by using certificates and a PKI. The TTP checks the validity of signatures from the smart meters, then removes them and any information about the smart meter. The remaining information is signed with the TTP's certificate, encrypted for and then sent to the DSO.

This simple scheme achieves privacy for the customer, but has side-channels of information depending on the capabilities of an adversary. If an adversary can listen to communication between SM and TTP, it can provide the DSO with a list of timestamps when which SM sent something to the TTP. Matching this list against the actually received measurements, the DSO achieves at least some insight into which measurement came from which SM. The adversary does not need to be able to decrypt any messages. Just the event of sending data to the TTP is linkable to the event of the TTP sending data to the DSO. For a discussion of the *traffic analysis problem*, see the works of Chaum [33].

Efthymiou and Kalogridis [34] provide a solution that avoids this problem by performing pseudonymization. In their approach, the SM sends its measurements directly to the DSO. A TTP is introduced but does not act as an intermediary. Instead it functions as a trust anchor. SMs use two sets of credentials while communicating with the DSO. One set is for low-frequency measurements and contains certificates signed by the DSO which identify the smart meter. The other set is for high-frequency measurements and contains anonymous certificates that do not identify the smart meter. Essentially, these are pseudonyms which are signed by the TTP. The setup process for those identities is fairly complex and thoroughly described in the paper. It includes several randomly chosen waiting times that help to provide unlinkability of events. The time intervals are in the region of days. For example, if a new meter is installed, the DSO could easily link high-frequency and low-frequency identities if they would be set up at the same time.

An approach to pseudonyms without a TTP is presented by Finster and Baumgart [35]. It uses blind signatures to achieve the distribution of pseudonyms to smart meters by the DSO without revealing the mapping between pseudonym and smart meter. The used pseudonyms are public keys which are generated by smart meters and for which only the generating smart meter knows the private key. Meter readings are encrypted for the DSO and signed by smart meters using their private key and the public key of the DSO. Afterwards, the smart meter's public key together with the signature on the public key by the DSO is attached to the meter reading. The DSO can then check if the public key and the meter reading are properly signed. To submit a meter reading from a smart meter to the DSO, a smart meter takes part in a peer-to-peer anonymity network and therefore sends meter readings through multiple other smart meters before they reach the DSO. This prevents the DSO from learning a mapping between network address and pseudonym which could later be used to break the pseudonym.

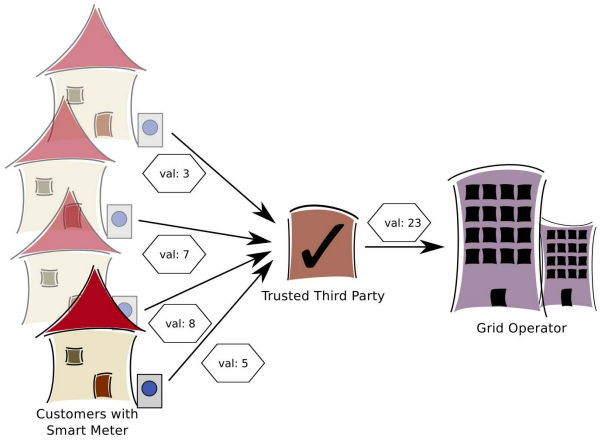

Fig. 4. Metering for operations via a trusted third party with aggregation: Smart Meters send meter readings to the trusted third party. The trusted third party aggregates all received meter readings and sends a single value to the energy supplier. As described in [23].

B. Aggregation with a trusted third party

Bohli et al. [23] avoid the side-channel problem through aggregation. The TTP, additionally to anonymization, aggregates meter readings. Instead of sending an incoming meter reading immediately, they wait until all SMs sent their reading, aggregate all those readings and then send only that aggregation to the DSO. This effectively unlinks events on both communication channels and provides privacy. But the DSO receives only aggregated data. This is illustrated in Figure 4.

Kim et al. [36] provide a related scheme which uses an obfuscation function to protect privacy. In this scheme, the state of the electrical grid is estimated using a function that operates on a vector of SM measurements. The measurements of SMs are then obfuscated in such a way that they are far different from their original values and that it is difficult to deduce an original value from the obfuscated one. This obfuscation is performed in a way that guarantees that the state estimation function results in the same state as with the original values. The obfuscated measurements are sent to a third party computing service that combines them and performs the state estimation function. The result is then returned to the DSO.

Vetter et al. [37] propose a hybrid approach which uses homomorphic encryption [38] in combination with a trusted third party that manages certificates and cryptographic keys for smart meters.

Homomorphic encryption enables the encryption of data in such a way that the encrypted data can still be used for calculations. A fully homomorphic encryption scheme allows to perform multiplication and addition of the ciphertexts. Such schemes are fairly new and complex. A partially homomorphic encryption scheme allows only one of those operations. Such schemes are more common. For aggregating meter readings, only addition is necessary.

A bihomomorphic encryption scheme is homomorph not only in data, but also in keys. Thus, when one value v is encrypted with key a, resulting in v_a and another value w is encrypted with key b (w_b), the sum $v_a + w_b = x_{a+b}$ can be decrypted with the key $a + b$.

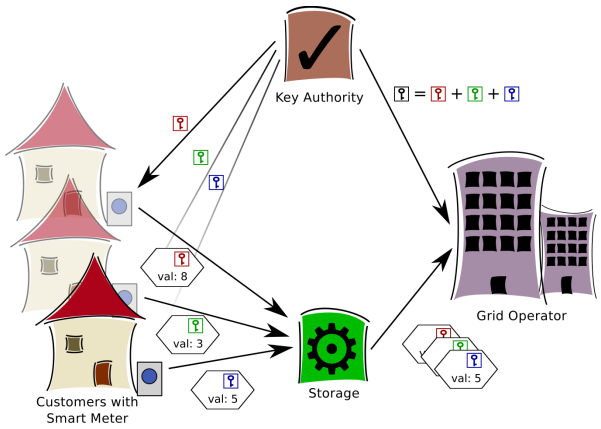

Fig. 5. Metering for operations using homomorphic encryption as described in [37]. SMs receive private keys for the homomorphic encryption scheme from the key authority. Using this key, they send their encrypted measurements to a central storage. The DSO requests a combined key for certain groups of smart meters. If this is in compliance with the key authorities privacy policy, it calculates the respective key and transmits it to the DSO. With the data received from the storage and the calculated key, the DSO can decrypt the aggregated meter readings.

In [37], depicted in Figure 5, every smart meter encrypts his measurements using a bihomomorphic encryption scheme and a private key issued by a trusted third party, the key authority (KA). The encrypted measurements are then sent to a central storage. This central storage can be queried from other parties, for example by the DSO. The central storage has no access to any keys and deals only in encrypted data. Therefore, it does not need to be trusted to keep data secret. It only has to be functionally trusted.

To work with the data which is retrieved from the central storage, the correct decryption key must be known. To protect the privacy of a single smart meter, the KA only hands out keys that can be used to decrypt the aggregation of measurements from a group of smart meters. This is achieved by exploiting the additive homomorphic property of the keys in the used encryption scheme. For example, the DSO can request the sum of measurements for a group of smart meters from the central storage. It retrieves the decryption key for that group from the KA and then decrypts the aggregated value.

This approach enables a separation of functions: a storage or database provider that can be untrusted and a trusted third party that only manages keys and certificates.

C. Aggregation without a trusted third party

The following section describes approaches, where the privacy of the customer is protected by aggregating the measured data with the data of other customers before the DSO gets access to them. And aggregation is done without the support of a trusted third party.

The main challenge in this scenario is how a single customer is able to add his data to the aggregate without revealing his data to other customers or the distribution system operator.

This problem can be described as “calculating with unknown data”. The tool used in most works is homomorphic encryption.

The following approaches all use homomorphic encryption and differ mainly in two points (see also: Table I).

- Who performs aggregation?
- How are keys managed?

An approach suggested by Li et al. [39], [40] uses a WLAN mesh network to connect smart meters. The DSO organizes the SMs in a tree with itself at the root. The approach uses a homomorphic scheme with asymmetric keys (Paillier and Pointcheval [45]) and distributes the DSO’s key through the network. Thus, every SM is capable of encrypting his measurements in such a way that only the DSO can decrypt them, but every SM can aggregate them. The leaf nodes of the tree start by passing their encrypted meter readings to their parents. All other nodes wait until they receive data from their children, aggregates their encrypted measurement with the received data and then send it further up the aggregation tree.

Mármol et al. [41], [42] exploit a bihomomorphic encryption scheme (Armknecht et al. [46]) to perform aggregation. Instead of measurements, this approach aggregates keys. Each SM chooses a random key of a bihomomorphic encryption scheme. Measurements are encrypted with that key and then sent to the DSO using a secure channel that hides the identity of the SM. The keys are aggregated within a group of SMs by electing a random SM as aggregator. This SM sums up the keys and sends the result to the DSO. With this key, the DSO can decrypt the sum of encrypted measurements.

If the aggregating SM cooperated with the DSO, it could transfer individual keys to the DSO. Since a mapping between individual key and encrypted measurement is not available, the DSO would have to try all possible combinations. Although such an attack would be computationally expensive, an additional step is done to eliminate it.

To prevent this attack, the SMs of a group arrange themselves in a ring structure. Using this ring structure, the SMs cooperatively manipulate their keys in such a way, that the individual keys change, but the aggregated key stays the same. Therefore, the individual keys that were once transferred to the aggregator are no longer in use. Even if the aggregator shares those keys with the DSO, they can not be used to decrypt individual measurements. However, if a SM joins or leaves the group, the aggregated key must be recalculated.

In [43] by Garcia and Jacobs, aggregation is done by using some primitives from the Slice-Mix-Aggregate algorithm (SMART) published in [47]. The protocol described in [43] uses a concentrator at the neighborhood level. Every SM in the neighborhood has its own public key for the homomorphic encryption scheme. Using the concentrator as a hub, these private keys are exchanged within the neighborhood, so that every SM knows the public key of all other SMs within his neighborhood. Every smart meter then takes his measurement and splits it into one share per meter in the neighborhood (including himself). Every share but the one for himself is encrypted with the public key for the respective smart meter and then sent to the concentrator. The concentrator uses the properties of the homomorphic encryption scheme to sum up all encrypted shares destined for a smart meter and then sends that sum to the respective smart meter. All smart meters wait for that sum to arrive, decrypt it and add the share they kept back. This sum is then sent to the concentrator unencrypted, so that the concentrator can aggregate and result in a final sum that equals the sum of all measurements.

An important property of this approach is the usage of additional infrastructure, the central hub per neighborhood. Through the application of homomorphic encryption, this hub does not need to be trusted since it can not derive information from the data presented to it. There is also no special trust in

TABLE I. COMPARISON OF AGGREGATION WITHOUT TTP

Paper	Required cryptosystem	Aggregate	Aggregation performed by	Key management
[39], [40]	Homomorphic	Meter readings	Several SMs (in-network within aggregation tree)	DSO distributes its public key throughout the network
[41], [42]	Bihomomorphic	Cryptographic keys	Randomly chosen SM within group	Each SM uses a randomly generated key and submits it to the designated SM
[43]	Homomorphic	Meter readings	Every SM but only using data that is not yet complete (which provides privacy)	Each SM uses a randomly generated key, public keys are distributed by the neighborhood concentrator
[44]	None (SMART)	Meter readings	Every SM but only using masked readings	None

the smart meters of the neighborhood necessary except that at least two smart meters must be non-malicious to keep privacy. If all smart meters but one are malicious, the privacy of the non-malicious smart meter is at risk.

An approach suggested by Finster and Baumgart [44] relies solely on the Slice-Mix-Aggregate algorithm [47] for privacy compliant aggregation and does without homomorphic encryption. Smart meters are arranged in an aggregation tree that is built using a peer-to-peer overlay (e.g., Chord [48]). In this tree, the leafs consist of small groups of smart meters. Within a leaf, the Slice-Mix-Aggregate algorithm is used to mask the meter readings of all members in such a way that each single meter reading is randomly distorted but the sum of all meter readings stays the same. The members of a leaf then send their masked readings to the parent. All non-leaf smart meters receive the meter readings from their children, aggregate all incoming meter readings with its own meter reading and finally pass the sum up to the next smart meter in the aggregation tree. In a last step the root of the tree passes the sum to the DSO.

D. Submit imprecise data

The possibility to submit imprecise data is a largely ignored approach to the problem. The sampling interval of smart metering already limits its accuracy. Therefore, results of smart metering are only an approximation of the real state of the power grid. One could argue, that this impreciseness provides some leeway for privacy enhancement. As long as it does not hurt the overall precision too much, a single smart meter could submit data that is imprecise enough to conceal some privacy relevant features. The problem with this approach is that it needs some form of cooperation between smart meters. If all smart meters submitted imprecise data without coordination, overall precision would suffer tremendously.

Bohli et al. [23] propose, that smart meters add random noise to their measurements and submit that fuzzed data to the DSO. But this noise is not totally random. It is generated by a distribution with known variance and expectation. For a great number of smart meters, the sum of the noise is then known. If every SM adds that noise to its readings, the DSO simply has to subtract the expected value of the noise from the aggregated meter readings. It will receive the sum of the original meter readings up to some accuracy. The weak point of this approach is the large number of smart meters that have to participate. To achieve acceptable privacy, millions of smart meters are necessary, which makes the implementation of this specific approach impractical.

E. Summary

Metering for operations mainly deals with the problem of keeping the customer's information private while providing time-accurate monitoring of the grid. Unlike metering for billing, measurements can not be aggregated over time and therefore measurements are often aggregated over different smart meters.

TABLE II. COMPARISON OF TTP METERING

Paper	TTP Role	Provided Data
[9], [26]	Data gateway	Metering data from individual meters using pseudonymization or anonymization to provide privacy
[34]	Trust anchor	Directly transferred pseudonymized individual meter readings
[23]	Data gateway and aggregator	Metering data aggregated over a group of smart meters
[37]	Key authority	Metering data aggregated over a group of smart meters

For approaches in Section IV-A, pseudonymization or anonymization services are provided to smart meters. The approaches in this section do not aggregate meter readings. They provide the DSO with individual information about each meter, but try to keep the identity of the meter secret. This is either done by removing all information about the meter or substituting all identifying information with pseudonyms. This operation resembles closely the concept of network mixes for anonymous electronic mail where privacy relevant information is removed by an intermediate system. Special attention has to be paid to possible information side-channels. For example, the appearance or disappearance of a pseudonym could be traced to a new or leaving customer and therefore reveal the mapping between customer and pseudonym.

The approaches outlined in Section IV-B aggregate meter readings of several smart meters using a TTP. This removes some of the problems that a simple anonymization or pseudonymization solution has. An overview how TTPs are used is given in Table II.

The question why a TTP should be trusted and who should be the TTP are avoided by approaches in Section IV-C. Instead of relying on a TTP to provide privacy, these solutions perform aggregation among the smart meters. To prevent smart meters from threatening each others privacy, these approaches often use homomorphic encryption. If a third party is used, it does not need to be trusted.

Finally, Section IV-D describes an approach where an individual smart meter submits a value that is so inaccurate that it no longer threatens privacy. Metering is established by knowing in advance how inaccurate meter readings will be for a large number of meters.

V. AVOID GENERATING DATA THAT POSES A PRIVACY RISK

In this section, approaches are outlined that solve both problems stated in Section II-C at once. This is done by

avoiding the existence of privacy-relevant data in the first place. This can either be done by not generating privacy-relevant data or pose limitations on the ability to measure privacy-relevant data. Therefore, there are two basic concepts:

- Shape the resulting and reported load of a customer to be not privacy-relevant. We will call this concept *battery concept* and it is covered in Section V-A.
- Reduce the sampling rate of measurements of the SM to a rate that does no longer pose a threat to privacy. We will call this concept *sampling concept* and it is covered in Section V-B.

Both of these concepts only need modifications at the customers site and do not rely on cooperation with other customers or the energy supplier. Since they are orthogonal to other approaches covered in this paper, they can be combined with them to further improve privacy.

A. Battery concept

The battery concept is an approach where the energy consumption of a household is shaped in such a way that the time discrete observation of said load by the SM does reveal as little private information as possible. In this approach, the meter readings of the SM are not privacy-threatening, regardless of the sampling rate. It is proposed by Kalogridis et al. [49], [50] and McLaughlin et al. [51].

They use a fine-granular control over energy consumption and decentralized energy production to shape the resulting load. A combined heat and power for example, may be able to conceal the consumption of a stove if it is switched on and off at the right points in time. For a more detailed control, an in-residence energy storage (e.g., a rechargable battery) is introduced, that can be charged and discharged accordingly. Exemplary loads with the resulting overall load are shown in Figure 6.

Since controlling appliances and energy production is out of scope for a SM, these concepts introduce a ME. It is connected to controllable appliances like refrigerators or washing machines. Devices for decentralized energy production and rechargeable batteries (e.g., an electric vehicle) are also under its control.

There are three important questions to be addressed with this concept.

- How can be decided what kind of load is threatening privacy? Or more specifically, when the ME monitors the customer's load, what triggers it to shape the load and to what kind of load will it shape.
- Given the limited resources of a battery, what strategy will yield the most privacy?
- How can achieved privacy be quantified and measured?

All three papers provide the same answer to the first question. The answer is motivated by the techniques used by NILM (see Section II-B) to analyze loads. They try to reduce the number of features in the resulting load. A feature is defined as a change in the load in positive or negative direction. For example, if a light bulb is switched on, it generates a feature of $\{+100W\}$.

Following this argumentation, the answer to the second question is simple. The strategy that removes more features will yield more privacy. Yet, there are two different approaches:

In [49], [50] a water filling algorithm is used to decide if the battery is used or not. If it recognizes a positive feature, it tests if the battery has enough energy to compensate said feature through discharging. If that is the case, it does so. A negative feature is managed alike but with charging the battery.

The algorithm employed by [51] tries to maintain a target load by charging and discharging the battery. The initial target load is calculated from the capacity of the battery. If the battery has not enough energy or capacity to maintain the target load, the algorithm switches into recovery. While in recovery, a recovery target load is calculated from the current load in such a way that the battery can maintain it. In example, if the battery is empty and can not compensate a higher load than the target load, the recovery target load is even higher than the current load to allow the battery to charge. When recovery is finished, a new target load is calculated using a exponential, weighted, moving average and normal operation is resumed.

Finally, since the resulting load is still measured by the SM and communicated to the DSO, it is important to *measure* how much information is still contained in the load. The proposed metrics for evaluating the gained privacy are *relative and empirical entropy*, *cluster classification*, *regression analysis* and *residual features*.

Relative entropy (or *Kullback Leibler Distance*) is an information theoretical approach. It is used to compare two sources of information. It is used in [49], [50] by interpreting features in the load as a stochastic process. The load before and after load leveling forms a stochastic process. Those two processes are then compared with relative entropy.

Cluster classification and regression analysis are used only in [49]. The cluster classification metric simply classifies its input data into clusters. The data set is compromised of the power consumption values and the difference between power values is used as a metric. Cluster classification then groups the points in the data set into clusters while minimizing the distance between them. It is performed on the unshaped and the shaped load and the number of resulting clusters is then compared. If the shaped load yields less clusters than the unshaped load, this is interpreted as a privacy gain. Regression analysis is a combination of cross-correlation and regression. It shifts the unshaped load (in time) until it aligns with the shaped load at the point of their maximum cross-correlation. The aligned loads are then used in a simple linear regression. The quality of the resulting predictor is then used as a measure of the privacy protection.

In [51], two metrics are used. First, the number of features left in the shaped load are counted and compared to the number of features in the unshaped load (residual features). As a second metric entropy is used. But unlike relative entropy, the empirical entropy per shaped and unshaped load is calculated independently and then compared with each other.

An overview over the different approaches is presented in table III.

Through the usage of different metrics, it is not possible to reliably compare the two algorithms for gained privacy. All works conclude that the concept is effective but its efficiency depends on the availability of advanced battery technology.

It is also remarkable that both algorithms target a steady load and measure its success by observing features and entropy. What kind of privacy gains a randomized load would cause is not covered and metrics to measure that kind of privacy are also still missing. A more recent work on the trade-off between privacy protection and usefulness of the resulting data set by Rajagopalan et al. [52] could be a good

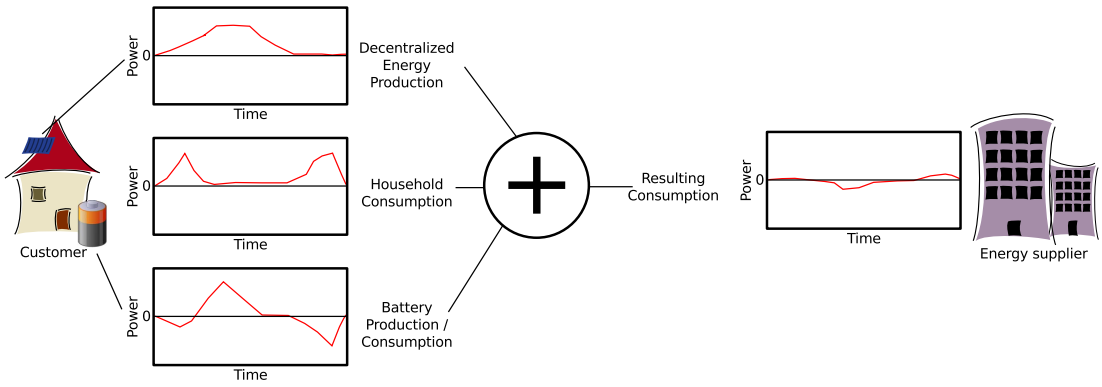

Fig. 6. Load levelling: By using decentralised energy production and in-resident energy storage like a battery, loads can be coordinated in such a way that the resulting load that is seen by the utility reveals only minimum private information.

TABLE III. COMPARISON OF WORKS USING A RECHARGABLE BATTERY FOR LOAD SHAPING

Paper	Battery control algorithm	Privacy evaluation metrics	Origin of metering data for evaluation	Battery sizes
[49]	Water filling algorithm	Relative entropy, cluster classification, regression analysis	24h measured appartment, simulation	0.5,1,2,4kWh
[50]	Water filling algorithm	Relative entropy	24h measured appartment	0.5,1,2,4kWh
[51]	Adaptive target load	Residual features, empirical entropy	1-2 months, 1 appartment 1 town house	0.6kWh

starting point for future approaches using the battery concept.

B. Sampling concept

Cardenás et al. [53] provide a new approach to privacy problems of smart metering by arguing that the sampling rate of smart meters is a design parameter of demand side management.

They concentrate on demand side management schemes as the limiting factor for smart metering sampling frequency. The higher the sampling frequency, the better a DSM scheme works. By modeling DSM schemes as discrete time control systems, the closed-loop system properties of this systems can be observed for varying sampling frequencies. Through this formalization it is possible to formulate an optimization problem that finds the largest sampling frequency so that the closed-loop property remains within the set of desirable performance goals.

In this brief paper, only the concept and examples are provided. But in future work, detailed studies of DSM schemes and their closed-loop properties are promised.

C. Summary

There are two basic concepts that try to inhibit the existence of data that threatens privacy.

One of them is shaping the load that is measured and billed by DSO and ES. It tries to fix the source of the data. But this concept heavily relies on the capability to shape the load. In current works, this is done via batteries. The feasibility of this approach is tightly connected to the availability of efficient and cheap batteries.

The other concept target the observer of the data. By limiting the rate at which smart meters sample, privacy is at least enhanced. But the sweet spot between privacy and functionality has yet to be found. And if that sweet spot will provide enough privacy is a highly anticipated result of future research.

VI. COMPARISON OF THE GENERIC APPROACHES

In this section, we provide a comparison of the generic approaches to metering for billing and metering for operations. This comparison is not on the level of specific papers, but on the level of the different categories used in this survey. Note, that this comparison does not include the approaches described in Section V. Their focus on the avoidance of data at the customer's site makes it impossible to compare them to the other approaches. The changes proposed by them change merely behavior at the customer's site. This, however, makes them great candidates to improve other approaches or to mitigate their shortcomings.

We will compare the different approaches on the following properties:

- *SM complexity:* computing complexity or otherwise necessary changes at the customer's site
- *Infrastructure complexity:* computing complexity or otherwise necessary changes outside the customer's site
- *Attack complexity:* the capabilities needed for an adversary to break privacy or the minimum complexity of a successful attack

We start by covering metering for billing and follow with metering for operations.

A. Comparison of approaches to metering for billing

We classified approaches in three groups:

- Trusted Third Party
- Trusted computing
- Cryptographic proofs

A comparison of these groups regarding SM complexity shows that the TTP solution has the lowest requirements concerning the smart meter. Only basic and industry standard cryptography is necessary to open trusted and secured connections to the TTP. The trusted computing approach needs additional hardware in smart meters and therefore has a higher SM complexity. The solution with cryptographic proofs has the highest SM complexity since it needs complex software and computations for the involved cryptography.

TABLE IV. COMPARISON OF APPROACHES TO METERING FOR BILLING

Approach	SM complexity	Infrastructure complexity	Attack complexity
Trusted Third Party	low	high	high
Trusted Computing	medium	medium	medium
Cryptographic Proofs	high	low-medium	high

This order is almost reversed when infrastructure complexity is observed. The TTP approach obviously needs a trusted third party to be set up and maintained. This is not only a technical overhead but also a strategical overhead. Since the trust in the TTP needs to be well-grounded, the question of who is responsible for it is hard to answer. Therefore, the infrastructure complexity is very high. With the trusted computing approach, the cryptography for handling signed and encrypted messages is necessary. But also, the trusted computing hardware has to be set up and managed. Infrastructure complexity is therefore considered medium. When using cryptographic proofs, the infrastructure complexity depends on the used cryptographic schemes. In most cases it is only necessary to check the provided data for correctness. This can be very simple (e.g., [30]) or more complex (e.g., [31]).

The capabilities needed to gain sensible information from the TTP approach are either to listen on the secured connection or to compromise the TTP. Assuming up to date functionality is used to secure connections and infrastructure, both needed capabilities are considered hard. For a trusted computing solution, an adversary can not get information by listening on the connection since the transmitted information is already considered privacy-aware. This transformation is performed by the trusted computing module within the smart meter and sensible data therefore does not leave the smart meter. For the adversary to get information, the software to be run within the trusted platform module has to be modified. The adversary therefore has to compromise whoever blesses this software with trustworthyness or the TPM itself. Considering that the TPM is a static target (compared to the moving target of up to date secure connections), the needed capabilities are less than for the generic TTP approach. The approach with cryptographic proofs is the hardest target for an adversary. Neither the software on the SM is variable, nor sensible data is transmitted. An adversary would have to compromise the smart meter itself or find information leaks in the cryptographic proofs.

A tabular overview is given in Table IV.

B. Comparison of approaches to metering for operations

We classified approaches in four groups:

- Anonymization or pseudonymization without aggregation
- Aggregation using a TTP
- Aggregation without TTP
- Submit imprecise data

Concerning smart meter complexity, the non-aggregating approach has the lowest requirements. An exception to this is [35]. Since this approach does not use a TTP, the SM complexity is higher. Aggregation using a TTP has in general a very low SM complexity. As shown by [37], a separation of data storage and data management can increase complexity enormously. But this is not a problem of the generic approach. Most of the outlined approaches of aggregation without TTP rely heavily on homomorphic encryption and, in part, complex coordination of smart meters with other smart meters. The complexity of this approaches is therefore high. An outsider to the generic approaches is the submission of imprecise data. Although SM complexity is very low in this approach, it shows problems concerning feasibility.

TABLE V. COMPARISON OF APPROACHES TO METERING FOR OPERATIONS

Approach	SM complexity	Infrastructure Complexity	Attack complexity
Anon- / pseudonymization without aggregation	low	high	high
TTP with aggregation	low	high	high
Aggregation without TTP	high	low	medium
Submit imprecise data	low	low	high

Infrastructure complexity is high for approaches using a TTP. Since, again, a TTP has to be set up and maintained. The additional effort for aggregation is only marginal. Therefor both categories, anonymization without aggregation and aggregation using a TTP are considered to have high infrastructure complexity. An exception is again [35] which does not need a TTP and has only low infrastructure complexity. For aggregation without TTP, infrastructure complexity is medium. Depending on the specific approach, the DSO only needs to decrypt received values using a homomorphic encryption scheme. Additionally a coordination of SMs is necessary in some approaches but this does not justify a higher rating in the infrastructure category since this can be done by the smart meters alone. The submission of imprecise data also provides the lowest complexity for infrastructure since the DSO only needs to calculate and remove the imprecise part to retrieve usable data.

For approaches using anonymization or pseudonymization without aggregation and for approaches using aggregation with a TTP, an adversary needs the same capabilities as in metering for billing via TTP: Listening on secured connections or compromising the TTP. Therefore, the ratings are the same. Aggregation without TTP provides more attack vectors for an adversary. In some approaches, the placement of a smart meter in an aggregation tree or its role within the system is crucial. If this can be influenced by the adversary, there is a higher risk of a privacy threat. Especially a sybil attack [54], where an attacker introduces a number of forged identities, has to be taken into account. Using this attack, it would be easier to place a malicious smart meter in a sensitive position. Additionally, most covered approaches rely on homomorphic encryption schemes. If an attack on the specific implementation is published, the smart metering protocol is immediately broken. Therefore, an adversary needs only medium capabilities. The submission of imprecise data has no viable attack vector except compromising the smart meter itself. Considering the software on the smart meter is fixed, this is deemed hard.

A summary can be found in Table V.

VII. CONCLUDING REMARKS AND FUTURE RESEARCH

In this survey, we presented various approaches to protect the privacy of customers while still using smart metering. We identified two problems that have to be solved to realize privacy aware smart metering: metering for billing and the metering for operations. We identified the approaches to these problems in research concerning privacy in smart metering. Using a categorization of the different generic approaches,

we provided an overview over the specific solutions and a comparison of the generic approaches.

Which of these approaches is best suited for the task depends on multiple factors. An important one will be the willingness of distribution system operators and energy suppliers to adopt new technology. Approaches that rely on already proven industry standards are in a clear advantage.

Finally, we close this survey with our thoughts on future research in the covered field:

Research in this field is currently at an early state. Most works merely state an idea to a solution of the specific problem. A thorough analysis of the behavior of the proposed solution under real world conditions is often left out. Especially the communication between smart meters themselves and smart meters with the distribution system operator or energy supplier are seldom analyzed. Considering that these connections are also subject to failures or attacks, a solution has to take in account what happens when those appear. Especially approaches that use interaction between smart meters must consider what happens if a smart meter suddenly becomes unavailable. A good solution must recover from such a state, signal problems to the data sink and still provide data. An important task will be to actually implement and test these ideas under real world conditions.

One important point of the early visions for the smart grid involved the simplified process of meter deployment. Focus on this feature got lost in most works while concentrating on privacy. But even a very good solution (considering privacy protection) will have troubles being accepted by industry if it involves complicated setup procedures and maintenance. Self-organization is an important topic that could ease the deployment process and therefore improve industry acceptance.

REFERENCES

[1] D. Hart, "Using AMI to realize the Smart Grid," *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pp. 1–2, July 2008.

[2] G. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[3] F. Sultanem, "Using appliance signatures for monitoring residential loads at meter panel level," *IEEE Trans. Power Delivery*, vol. 6, no. 4, pp. 1380–1385, Oct. 1991.

[4] G. Hart, "Residential energy monitoring and computerized surveillance via utility power flows," *IEEE Technol. Society Mag.*, vol. 8, no. 2, pp. 12–16, June 1989.

[5] C. Laughman, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *IEEE Power Energy Mag.*, vol. 1, no. 2, pp. 56–63, Mar. 2003.

[6] M. Baranski and J. Voss, "Genetic algorithm for pattern detection in NIALM systems," in *IEEE International Conf. Syst., Man Cybernetics*, Hague, Oct. 2004, pp. 3462–3468 vol. 4.

[7] H. Lam, G. Fung, and W. Lee, "A novel method to construct taxonomy of electrical appliances based on load signatures," *IEEE Trans. Consumer Electronics*, vol. 53, no. 2, pp. 653–660, 2007.

[8] M. A. Lisovich, D. K. Mulligan, and S. B. Wicker, "Inferring personal information from demand-response systems," *IEEE Security Privacy Mag.*, vol. 8, no. 1, pp. 11–20, Jan. 2010.

[9] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proc. 2nd ACM Workshop Embedded Sensing Syst. Energy-Efficiency Building - BuildSys '10*, p. 61, 2010.

[10] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging smart meter data to recognize home appliances," in *IEEE International Conf. Pervasive Comput. Commun.*, Mar. 2012, pp. 190–197.

[11] A. Prudenzi, "A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel," in *IEEE Power Engineering Society Winter Meeting*, vol. 2, NY, 2002, pp. 941–946.

[12] M. A. Lisovich and S. B. Wicker, "Privacy concerns in upcoming residential and commercial demand-response systems," in *Clemson Power Syst. Conf. 2008*, vol. 1, no. 1, Mar. 2008, pp. 1–10.

[13] J. Z. Kolter, S. Batra, and A. Y. Ng, "Energy disaggregation via discriminative sparse coding," in *Advances Neural Inf. Process. Syst.*, 2010, pp. 1153–1161.

[14] A. Cavoukian, J. Polonetsky, and C. Wolf, "SmartPrivacy for the smart grid: Embedding privacy into the design of electricity conservation," *Identity Inf. Society*, vol. 3, no. 2, pp. 275–294, Apr. 2010.

[15] I. Rouf, H. Mustafa, R. Miller, and M. Gruteser, "Neighborhood watch: Security and privacy analysis of automatic meter reading systems categories and subject descriptors," in *Proc. 2012 ACM Conf. Comput. Commun. Security*, Raleigh, Oct. 2012, pp. 462–473.

[16] A. Pfitzmann and M. Hansen, "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management," 2010.

[17] S. Feuerhahn, M. Zillgith, C. Wittwer, and C. Wietfeld, "Comparison of the communication protocols DLMS/COSEM, SML and IEC 61850 for smart metering applications," in *IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Brussels, Oct. 2011, pp. 410–415.

[18] K. D. Craemer and G. Deconinck, "Analysis of state-of-the-art smart metering communication standards," in *Proc. 5th Young Researchers Symp.*, Leuven, Mar. 2010, pp. 1–6.

[19] S. Rohjans, M. Uslar, R. Bleiker, J. Gonzalez, M. Specht, T. Suding, and T. Weidelt, "Survey of smart grid standardization studies and recommendations," in *First IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Gaithersburg, Oct. 2010, pp. 583–588.

[20] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: Communication technologies and standards," *IEEE Trans. Industrial Informatics*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[21] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart grid," *Comput. Netw.*, vol. 55, no. 15, pp. 3604–3629, Oct. 2011.

[22] Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, Z. Zhu, S. Lambotharan, and W. H. Chin, "Smart grid communications: Overview of research challenges, solutions, and standardization activities," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 21–38, Jan. 2013.

[23] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *IEEE International Conf. Commun. Workshops (ICC)*, Capetown, May 2010, pp. 1–5.

[24] M. Jawurek, M. Johns, and K. Rieck, "Smart metering depseudonymization," in *Proc. 27th Annual Comput. Security Applications Conf. - ACSAC '11*, Orlando, Dec. 2011, pp. 227–236.

[25] D. Challener, K. Yoder, R. Catherman, D. Safford, and L. Van Doorn, *A Practical Guide to Trusted Computing*, 1st ed. IBM Press, 2007.

[26] R. Petrlic, "A privacy-preserving concept for smart grids," *Sicherheit Vernetzten Systemen*, vol. 18, pp. B1–B14, 2010.

[27] M. LeMay, G. Gross, C. Gunter, and S. Garg, "Unified architecture for large-scale attested metering," in *40th Annual Hawaii International Conf. Syst. Sciences (HICSS'07)*, Waikoloa, Jan. 2007, pp. 115–115.

[28] G. Brassard, D. L. Chaum, and C. Crépeau, "Minimum disclosure proofs of knowledge," *J. Comput. Syst. Sciences*, vol. 37, pp. 156–189, 1988.

[29] T. P. Pedersen, "Non-interactive and information-theoretic secure verifiable secret sharing," in *Advances in Cryptology — CRYPTO '91, LNCS 576*. Springer Berlin Heidelberg, 1992, pp. 129–140.

[30] M. Jawurek, M. Johns, and F. Kerschbaum, "Plug-in privacy for smart metering billing," *Lecture Notes Comput. Science*, vol. 6794, pp. 192–210, 2011.

[31] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proc. 10th Annual ACM Workshop Privacy Electronic Society - WPES '11*, Chicago, Oct. 2011, p. 49.

[32] S. Vaughan-Nichols, "How trustworthy is trusted computing?" *Comput.*, vol. 36, no. 3, pp. 18–20, Mar. 2003.

[33] D. L. Chaum, "Untraceable electronic mail, return addresses, and

digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–90, Feb. 1981.

[34] C. Efthymiou and G. Kalogridis, "Smart grid privacy via anonymization of smart metering data," in *First IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Gaithersburg, Oct. 2010, pp. 238–243.

[35] S. Finster and I. Baumgart, "Pseudonymous smart metering without a trusted third party," in *Proc. 3rd IEEE International Symp. Anonymity Commun. Syst. Conjunction IEEE TrustCom*, Melbourne, July 2013, pp. 1723–1728.

[36] Y. Kim, E. C.-H. Ngai, and M. B. Srivastava, "Cooperative state estimation for preserving privacy of user behaviors in smart grid," in *IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Gaithersburg, Oct. 2011, pp. 178–183.

[37] B. Vetter, O. Ugus, D. Westhoff, and C. Sorge, "Homomorphic primitives for a privacy-friendly smart metering architecture," in *Proc. International Conf. Security Cryptography (SECRYPT 2012)*, Rome, July 2012, pp. 102–112.

[38] C. Gentry, "Computing arbitrary functions of encrypted data," *Commun. ACM*, vol. 53, no. 3, pp. 97–105, 2010.

[39] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *First IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Gaithersburg, Oct. 2010, pp. 327–332.

[40] ——, "Secure and privacy-preserving information aggregation for smart grids," *International J. Security Netw.*, vol. 6, no. 1, p. 28, 2011.

[41] F. Mármol, C. Sorge, O. Ugus, and G. Pérez, "Do not snoop my habits: Preserving privacy in the smart grid," *IEEE Commun. Mag.*, vol. 50, no. 5, pp. 166–172, May 2012.

[42] F. Mármol, C. Sorge, R. Petrlic, O. Ugus, D. Westhoff, and G. Martínez Pérez, "Privacy-enhanced architecture for smart metering," *International J. Inf. Security*, vol. 12, no. 2, pp. 67–82, 2013.

[43] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *Proc. 6th International Conf. Security Trust Management*, Surat, May 2011, pp. 226–238.

[44] S. Finster and I. Baumgart, "Elderberry: A peer-to-peer, privacy-aware smart metering protocol," in *IEEE INFOCOM Workshop Commun. Control Smart Energy Syst.*, Turin, Apr. 2013, pp. 3411–3416.

[45] P. Paillier and D. Pointcheval, "Efficient Public-Key Cryptosystems Provably Secure against Active Adversaries," in *Advances in Cryptology - ASIACRYPT'99*. Springer Berlin Heidelberg, 1999, pp. 165–179.

[46] F. Armknecht, D. Westhoff, J. Girao, and A. Hessler, "A lifetime-optimized end-to-end encryption scheme for sensor networks allowing in-network processing," *Comput. Commun.*, vol. 31, no. 4, pp. 734–749, Mar. 2008.

[47] W. He, X. Liu, H. Nguyen, K. Nahrstedt, and T. Abdelzaher, "PDA: Privacy-preserving data aggregation in wireless sensor networks," in *IEEE INFOCOM 2007 - 26th IEEE International Conf. Comput. Commun.*, Anchorage, May 2007, pp. 2045–2053.

[48] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord," in *Proceedings 2001 Conf. Applications, Technol., Architectures, Protocols Comput. Commun. - SIGCOMM '01*, San Diego, Aug. 2001, pp. 149–160.

[49] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *First IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Gaithersburg, Oct. 2010, pp. 232–237.

[50] G. Kalogridis, Z. Fan, and S. Basutkar, "Affordable privacy for home smart meters," in *IEEE Ninth International Symp. Parallel Distributed Process. Applications Workshops*, Busan, May 2011, pp. 77–84.

[51] S. McLaughlin, P. McDaniel, and W. Aiello, "Protecting consumer privacy from electric load monitoring," in *Proc. 18th ACM Conf. Comput. Commun. Security - CCS '11*, Chicago, Oct. 2011, pp. 87–98.

[52] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor, "Smart meter privacy: A utility-privacy framework," in *IEEE International Conf. Smart Grid Commun. (SmartGridComm)*, Brussels, Oct. 2011, pp. 190–195.

[53] A. Cárdenas, S. Amin, and G. Schwartz, "Privacy-aware sampling for residential demand response programs," in *Proc. 1st International ACM Conf. High Confidence Netw. Syst. (HiCoNS)*, Bejing, Apr. 2012.

[54] J. Douceur, "The sybil attack," in *Peer-To-Peer Systems*. Springer Berlin Heidelberg, 2002, pp. 251–260.

Sören Finster is a Ph.D. student at the Institute of Telematics at the Karlsruhe Institute of Technology (KIT). He received his diploma in computer science from University of Karlsruhe (now KIT) in 2008. His research interests include privacy in communication, peer-to-peer networks and smart grid communication.

Ingmar Baumgart received a diploma degree in computer science in 2005 from University of Karlsruhe, Germany and a Ph.D. in computer science in 2010 from Karlsruhe Institute of Technology (KIT), Germany. Currently he is leading a Young Investigator Group at Karlsruhe Institute of Technology. His research interests include security and privacy of peer-to-peer networks, distributed socials networks and mobile communications. Dr. Baumgart is also project manager of the open-source overlay framework *OverSim*.