

An Overview of Privacy-Preserving Data Aggregation in Smart Metering Systems

Z. Erkin, *Member, IEEE*, J. R. Troncoso-Pastoriza, *Member, IEEE*, R. L. Lagendijk, *Fellow, IEEE*, and F. Pérez-Gonzalez, *Senior Member, IEEE*

Abstract

Growing energy needs forces governments to look for alternative resources and ways of better energy grid management and load balancing. As a major initiative, many countries including the UK, the USA and China have already started deploying smart grids. One of the biggest advantages of smart grids compared to traditional energy grids is the ability to remotely read fine-granular measurements from each smart meter, which enables the grid operators to balance load efficiently and offer adapted time-dependent tariffs. However, collecting fine-granular data also poses a serious privacy threat for the citizens as illustrated by the decision of the Dutch Parliament in 2009 that rejects the deployment of smart meters due to privacy considerations. Hence, it is a must to enforce privacy rights without disrupting the smart grid services, like billing and data aggregation. Secure Signal Processing aims at protecting the sensitive data by means of encryption and provides tools to process them under encryption, effectively addressing the smart metering privacy problem.

In this paper we present recent and ongoing research in the field of privacy protection for smart grids, where individual smart meter measurements are kept secret from outsiders, including the utility provider itself, while processing private measurements under encryption is still feasible. We focus particularly on data aggregation, which demonstrates the major research challenges in privacy protection for smart grids.

Index Terms

Privacy protection, smart metering, signal processing, cryptography.

I. INTRODUCTION

The Energy Independence and Security Act of 2007 defines smart grid as the modernization of the electricity delivery system that monitors, protects, and automatically optimizes the operation of its interconnected elements from generator to end users. Smart grids offer indisputable advantages over traditional power grids including remote readings and load balancing. Consequently, many countries including the UK, the USA and China have already

Z. Erkin and R. L. Lagendijk are with the Information Security and Privacy Lab, Department of Intelligent Systems, Delft University of Technology, 2628 CD, Delft, The Netherlands.

E-mail: {z.erkin, r.l.lagendijk}@tudelft.nl.

J. R. Troncoso-Pastoriza and F. Pérez-Gonzalez are with the Signal Theory and Communications Dept., University of Vigo, Spain.

E-mail: {troncoso,fperez}@gts.uvigo.es.

F. Pérez-Gonzalez is also with the Electrical and Computer Engineering Dept., University of New Mexico, USA.

started building smart electric grids. The European Commission has foreseen¹ the implementation of smart electric grids by the Member States, requiring that 80% of consumers be equipped with smart metering systems by 2020, after a viability assessment in 2012.

A smart grid consists of three segments: power generation, transmission-distribution network, and smart meters. In each segment, there are several challenges: power generation is highly related with wind turbines and solar panels, which are not as predictable as traditional power sources since energy production relies on environmental factors. Transmission-distribution network deals with efficiency problems, especially in the case of bi-directional energy transmission and distribution. And smart meters present a number of challenges in sensing, analyzing, and communication. Therefore, digital signal processing has found application in smart grid systems including specific hardware and software for sensing, processing digital signals, and low-cost communication.

Smart meters introduce new opportunities for the market as well. The traditional (analogue) metering systems rely on tamper-proof devices located at the households and they are *physically* read by the utility provider monthly. Smart meters, however, are anticipated to be read periodically in shorter intervals that range from minutes to milliseconds *remotely*, thus open up a wide range of new business opportunities for the utility providers. For instance, fine-granular remote readings can be used for performing statistical analyses that lead to effective consumption forecasting and profiling, which contribute to the prevention of power shortages and to apply load balancing. At the same time, the fine-grained readings will assist users in achieving a more efficient energy use and adapting to the network status and supply by choosing an appropriate and advantageous tariff.

Unfortunately, smart grid systems have a number of serious threats including security, safety, fraud, and privacy [5]. A virus or a DoS (Denial of Service) attack can severely damage the power infrastructure of a country. A remote switch-off button can be an appealing target for cyber-warfare. And manipulating smart meter readings can cause severe financial losses. Even though the research on security, safety and fraud prevention are attracting great attention from the governments, industry and academia, privacy aspects are not addressed sufficiently. A proof of how much privacy-sensitive data a smart meter reveals is shown by a Dutch student on bwired.nl. It is clear that the actions of the residents can be easily tracked by analyzing the smart meter data (gas, water, and electric consumption). It is even possible to determine the presence/absence of residents, the number of people living in a household; even their religion can be identified [5], [12]. Obviously, fine-granular smart meter measurements constitute a serious privacy and, in some cases, security threat for the citizens.

Many privacy related considerations in other on-line systems, such in as social networks, can be tackled by raising awareness among people on how to avoid revealing privacy-sensitive data. In the case of smart grids, raising awareness does not help the users sufficiently since reporting fine-granular consumption measurements is an essential part of an automated system. Therefore, security technologies, as well as law and regulations, are necessary to cope with privacy issues in smart grids. Not surprisingly, at the end of serious discussions, the Dutch Parliament refused the bill for smart grid deployment in 2009 on grounds of data protection concerns. Until a solution is found on the

¹Directives 2009/72/EC (electricity) and 2009/73/EC (gas) of the European Parliament Council of 13 July 2009.

basis of technology, it will be challenging to convince the governments and citizens in favor of deploying smart grids.

In the past years, solutions have been developed for privacy-preserving billing and data aggregation in smart grids based on security technology, in particular *Secure Signal Processing* (SSP). SSP is a powerful mechanism that, on one hand, protects the privacy-sensitive smart meter data and, on the other hand, enables the utility provider to still perform data analysis for the management of the grid. The main idea in SSP is to prevent the untrustworthy entities, including the utility provider, from accessing the private data, while providing tools to process the smart meter measurements, e.g. for billing and data analysis. To achieve this goal, cryptographic tools like homomorphic encryption and secure multi-party computation techniques are being used [19]. In particular, instead of reading measurements in plain text, the utility provider receives encrypted measurements from the smart meters. Without the decryption key, the utility provider cannot access the content of the encryptions; this guarantees the privacy of the residents. To perform the usual smart grid operations such as billing, the utility provider interacts with the smart meters according to a pre-defined protocol [24], [17], [15].

In this paper, we give an overview of recent and ongoing research in the field of privacy protection for smart grids. The paper serves as an introduction for the signal processing researchers, and thus explains the existing approaches, corresponding building blocks and current challenges. Our focus will be particularly on the computation of aggregated consumption, which has been addressed in a number of recent works. Architectural, hardware and technological limitations in privacy-preserving data aggregation are also valid research challenges for realizing other smart grid functions such as forecasting. To help the signal processing community in getting familiar with the privacy protection research in smart grids, we structured this paper as follows. We reserve Section II for a discussion on privacy considerations, architecture and the roles of stake-holders in smart grids. We present recent research on the computation of aggregated data in Section III. We discuss the existing challenges in SSP research in smart grids and present future work for the signal processing community in Section IV. We conclude the paper in Section V.

II. PRIVACY MODEL AND SMART METERING ARCHITECTURES

We have argued that privacy is a crucial issue in smart metering; we can show it with a specific example of the privacy breach produced when collecting fine-grained readings from a household power consumption. Figure 1, taken from [14], represents a set of readings for a time span of forty minutes, where the consumption of each of the electric appliances of the household can be easily identified just by eye inspection. There are even more powerful techniques that take the aggregated measurements, and by using determined appliance signal models, they can disaggregate the measurements and provide an accurate estimation of the moment when each appliance is turned on and off [11]. These methods are usually called NIALM (Non-Intrusive Appliance Load Monitoring), and they are based either on transient or harmonic analysis, noise pattern recognition, or generic optimization algorithms for multiple-matching. With these methods and fine-grained readings, it is very easy to determine when the individual living in a house is at home, when he/she is having lunch, sleeping, watching TV, taking a shower, etc.

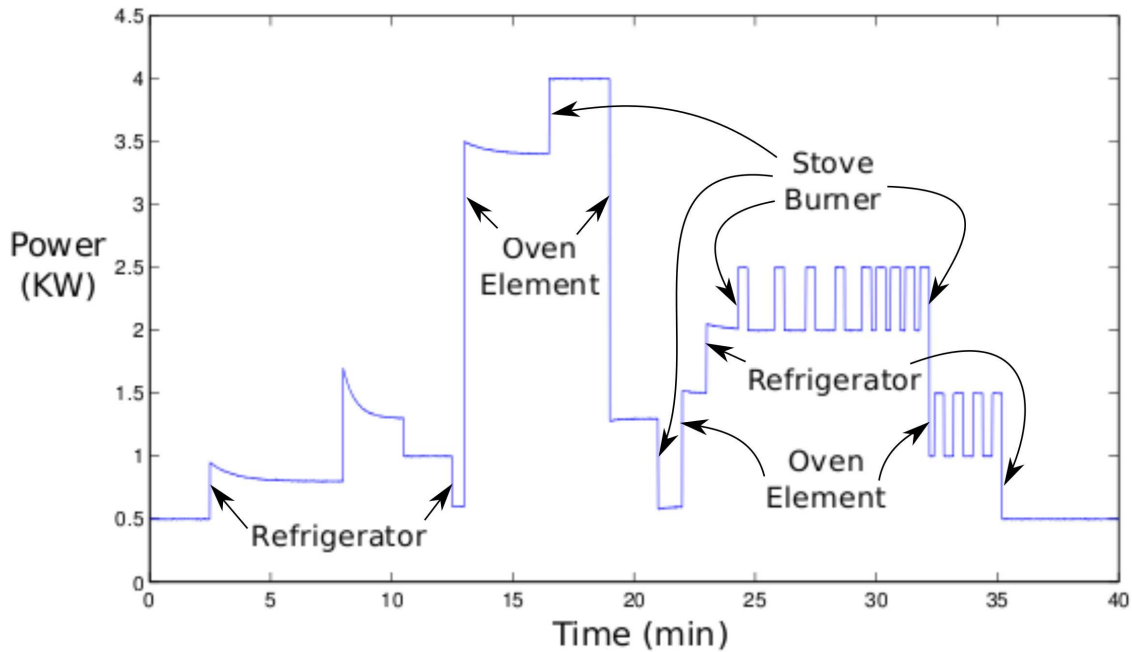


Fig. 1. Example of consumption readings for a short time period in a household (from [14]).

This reflects how important the privacy protection will be when smart metering is widely deployed. As mentioned before, this has already led some parliaments to paralyzing the adoption of smart metering infrastructures due to the violation of privacy regulations, despite the economic benefits and the energy savings it may produce. Hence, smart metering cannot be widely adopted until there are technological means to conceal the readings and therefore protect citizen's privacy. Before going into the details of these technological solutions, let us depict in this section the players in the smart metering scenario, the architectures in which a smart metering electricity network can be materialized, and their trust models.

A. Involved Parties in a Private Smart Metering Scenario

For the sake of completeness, we will now briefly describe the stakeholders in a Smart Metering scenario. We present a *functional* classification of roles, depicted in Figure 2. It is possible, though, that some of the "actual" stakeholders can simultaneously play several roles (i.e., producer/operator, aggregator/owner of the communication network, producer/aggregator). In fact, most of the works dealing with privacy in Smart Metering consider only two or three parties, each of them adopting several simultaneous roles, as a simplified representation of the problem.

- *Consumers/Customers*: The *end-users* that receive the power supply, either households or industrial users. The consumption patterns and specific individual fine-grained consumption information belonging to each user are sensitive data that must be protected for preserving the consumers' privacy. Commonly, customers have access to the metered data, either aggregated or not, in order to select an appropriate and advantageous tariff and be

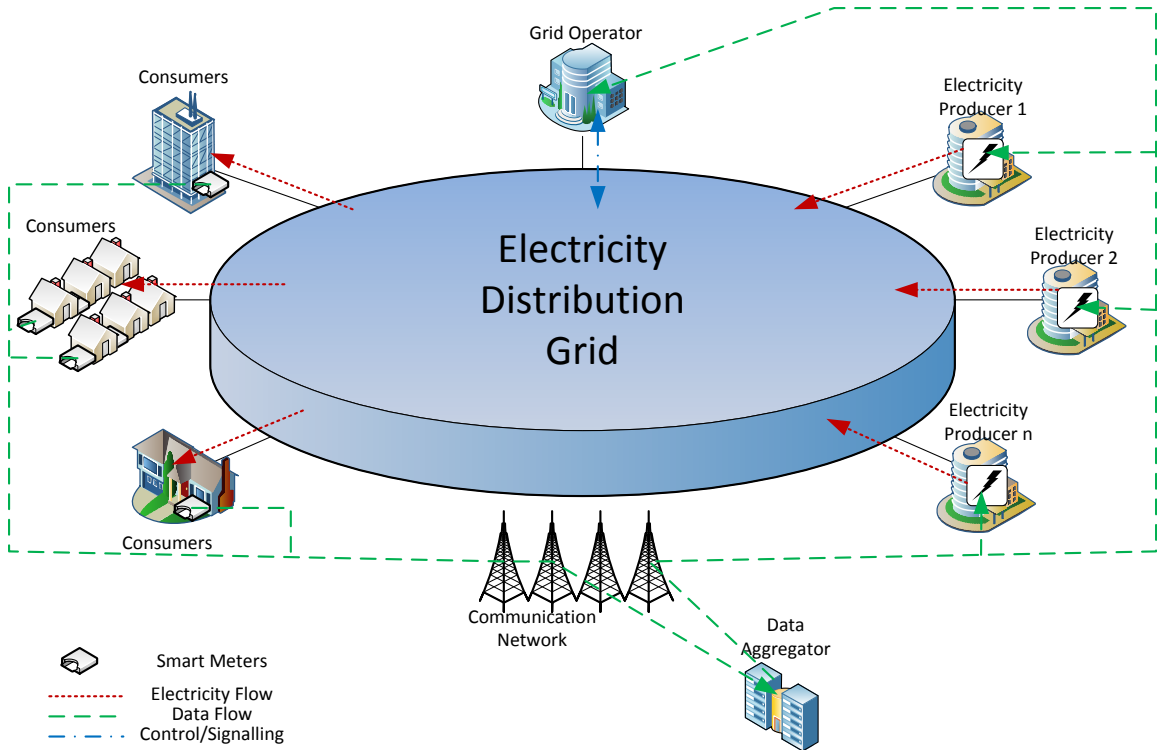


Fig. 2. Smart Metering scenario and its stakeholders.

able to suitably administer their consumption habits and electric appliances.

- *Smart Metering Devices*: They are installed at the customer side of the network; their function is to sense the consumed energy at every time slot (from milliseconds to minutes) and send the measurements to the consumer and/or the aggregator. One meter must be present at each consumer, so they are typically small and cheap devices with limited computational power and transmission capabilities.
- *Grid Operator/Supplier*: A company that controls the electricity distribution and transportation infrastructure. Operators may employ electricity usage data and distribution needs in order to optimally dimension and structure their resources; load balancing is a critical issue.
- *Communication Network*: It deals with the communication among all the parties in the smart meter scenario. If sensitive data are interchanged in plain text (i.e., individual consumption data coming from the meters), the communication channels must be secured.
- *Electricity Producer*: A company that sells the electricity to customers through the supplier's infrastructure. The price of the supplied electricity is agreed according to one or more tariffs. The producer must take into account a) the demanded power in order to adjust the produced electricity, and also b) know the total individual

consumptions for billing each consumer applying the contracted tariff.

- *Aggregator*: This party takes the metered consumption data and *aggregates* it, producing the relevant and needed figures, like individual and average total power consumption, estimation of power demand or average user profiling. This role is typically played by the same company that operates the grid.

B. Smart Metering Architectures

The interrelations among the stakeholders of the smart metering scenario differ depending on the implemented architecture, so the latter highly impacts the trust model, as we will see. There are two main choices of smart metering architecture, namely *centralized* and *distributed*.

A fully *centralized* management relegates the meters to just the sensing function, sending the measurements of short periods to a central data storage that acts as a hub (aggregator head end) and communicates with each smart meter. The aggregator database is then used for consumption calculation, load balancing and billing; each user may access the stored data in order to get information about his/her consumptions. This approach was the initial trend for smart metering implementation proposals, and all the computations are performed at the central aggregator, that has a high computational power compared to the meter devices.

For small grids, like self-sufficient grids in rural areas, a *distributed* (also known as *de-centralized* or *peer-to-peer*) energy management is usually adopted. In this case, the meters play the role of aggregators, and all the calculations over the metered data are distributed among the consumers, that jointly play the role of grid operators; the meters perform a partial data aggregation themselves (in-network aggregation [22], [6]), calculating the total energy consumption in each billable period, and they communicate the results to the appropriate parties (energy producer) typically once per billable period. Grid management and load balancing are performed collaboratively by the users, through dedicated interfaces under their control, and possibly assisted by a subcontracted company.

C. Trust Models

In any privacy-aware scenario, not specifically related to smart metering, there is an inherent interdependency between trust and privacy: those entities, parties and infrastructure elements of a smart metering system that are trusted will need no further privacy protection, and those elements in which privacy is enforced through a secure protocol will not need to be trusted. Hence, the definition of the trust model is of high relevance for properly and effectively preserving users' privacy. In this sense, untrusted parties can be considered mainly semi-honest (they follow the established protocols, but may try to infer information from the interchanged values) or malicious (they may deviate from the protocol, forging the interchanged messages to gain more information or to alter the output of the protocol).

Going back to the electricity metering case, the main trust relationships are established between the consumers and the suppliers/operator/aggregator. Trust from the consumers is directly related with privacy of the metered data: which stakeholders can access these data for a legitimate purpose. Conversely, the trust from the supplier/operator/aggregator focuses on the correctness of the data that the meters provide, so that "trusted

consumers” are assumed to provide the actual consumption values without trying to forge these measurements and the corresponding bills. The traditional sealed meters readable only at the customer’s home/facilities represented the mutual trust between the supplier/operator and the consumers, in such a way that consumers could not forge the measurements without tampering the meter and the operator could only access coarse measurements. The adoption of smart metering reshapes the trust model depending on the choice of architecture:

A centralized management and data-collection imposes a universal trust on the grid operator; this party would play the role of the aggregator, concentrating also the authentication and storage functionalities, and having access to all the fine-grained measurements, stored out of context at a central database; furthermore, the grid operator itself may have access in this scenario to the update and remote modification of the meters, hence the “universal trust”: users will be concerned not only with privacy, but also with the correctness of the meter usage and tariff calculation. This scenario is the prototypical example of privacy invasion that infringes the data protection directives; it is also a challenging scenario, for it poses many technical difficulties for the provision of an actual privacy-preserving solution.

A certain level of decentralization, together with the possibility of collaborative calculations among the meters, possibly grouped into cells, can facilitate the development of an effective mechanism that provide an actual privacy protection and correctness guarantees. Consequently, a partial decentralization is commonly assumed by works in the field, in such a way that the trust of the users is distributed amongst other users of the same cell, that are less likely to mutually collude, while the trust from the suppliers/operators still resides on the tamper-proofness of some of the meter elements like the sensors, timing devices, secure storage and secure cryptographic modules. Nevertheless, distributing data and calculations among several customers introduces also new challenges related with managing trust relationships and privacy protection not only between consumers and providers, but also among users.

D. Functions of Interest: Private Utility

Once we have established the trust model for each architecture, we can devote some space to the description of the figures and statistics that the grid operators or energy producers, untrustworthy for the consumers, may want to calculate from the private metered data.

Grid operators are not usually willing to openly disclose how they perform the grid management and which statistics they calculate. Any fine-granular data that could allow the grid operators to obtain useful statistics would be an asset for the business. Nevertheless, obtaining exact consumption data would be a breach of customers’ privacy. Furthermore, there are also legal bases that restrict this behavior: Data Protection Directives [1], [2] clearly state that the amount of collected sensitive data must follow the principles of proportionality and purpose. Hence, collecting the whole set of measurements without an adequate and rigorous justification would be in breach of these principles.

Consequently, as a first step for a correct management of private data, the needed statistics and figures for the proper operation of the Electricity Producer and the Grid Operator should be completely specified, determining also the processing that the metered data will undergo by the Aggregator.

The most obvious needed statistics are total consumption $C_{total}(t)$ and billing $B(t)$ for a given time period t , both needed by the Electricity Producer. These two figures can be represented as a general summation GS of the readings $m_{i,t}$, $GS(t) = \sum_{\mathcal{M}_s} f(m_{i,t})$, where $f(\cdot)$ is the identity function in the case of total consumption, or a given cost function in the case of billing (typically, a linear or piecewise linear function), and \mathcal{M}_s represents the set of involved measurements, either through time slots t , through space (the meter index i), or through both variables.

The sensitivity of these measurements creates a need for technical privacy preserving solutions that protect them from the grid operator, the electricity producer or the aggregator itself. This solutions should not to hinder the ability of the aggregator to calculate the needed $GS(t)$ and, at the same time, avoid the possibility of fraud (electricity theft).

Finally, it is worth mentioning that the general summation $GS(t)$ can represent many functions of interest for either the grid operator or the electricity producer (i.e., statistical measures or consumption forecasts). We will present the foundations of private solutions to some of them (mainly related to consumption calculation) in the next section; we must highlight that there are other private calculations on the metered data that may pose additional problems that fall out of the scope of this paper and will be briefly discussed in Section IV.

III. PRIVACY-PRESERVING COMPUTATION OF TOTAL CONSUMPTION

We now focus on the aggregation of measurements in order to show the recent privacy-preserving approaches presented to date. For a certain time instant t , the total consumption is defined as:

$$C_{total}(t) = \sum_{\mathcal{M}_s} f(m_{i,t}) = \sum_i m_{i,t} , \quad (1)$$

where $m_{i,t}$ is the measurement of the i th smart meter. As argued in the previous section, individual measurements are very privacy-sensitive, and thus should be protected.

Existing solutions in the literature focusing on the protection of individual measurements while computing the total consumption obfuscate the individual measurements collected from the smart meters by means of encryption and obtain the total by processing the data under encryption. With this approach, also called Secure Signal Processing [19], it is feasible to protect the privacy of the citizens and perform the tasks required to run the smart grid.

There are three common assumptions in the literature for privacy-preserving aggregation in smart metering systems. The first assumption is that there is a communication network available. While a wired communication link to the utility provider is required, smart meters are also assumed to be able to communicate with each other, which can be possible using technologies like Bluetooth and ZigBee. A second assumption is the possession of a valid certificate *per* smart meter. This is required as a proof of identity so that the inputs from a smart meter with a valid certificate are accepted by the other parties. Therefore, a role for a Certification Authority exists. The third assumption is the capability of performing cryptographic operations, mostly in a hardware environment with limited computational power and memory. The type of such operations differs in every proposal but in general hash functions, pseudo-random number generators, symmetric (e.g. AES) and asymmetric encryption (e.g. RSA, Paillier, El Gamal), and Elliptic Curve Cryptography (ECC) are used.

In this paper, we explain four approaches from the literature to compute the total consumption. While the proposed protocols are designed for an arbitrary number of smart meters, we prefer to build a story around three customers, namely Alice, Bob and Charles. Assume that a Utility Company (UC), which plays the roles of energy producer, grid and network operator, wants to compute the total energy consumption of these three customers:

$$C_{total}(t) = m_{1,t} + m_{2,t} + m_{3,t} , \quad (2)$$

where $m_{1,t}$, $m_{2,t}$ and $m_{3,t}$ are Alice's, Bob's and Charles' measurements, respectively. Our goal is to enable the UC to compute the total consumption without revealing the individual measurements. The measurements are mostly kept secret by means of encryption. For the aggregation of the encrypted measurements, additively homomorphic encryption schemes such as Paillier [21] (see Box I) seems suitable. However, for the aggregation using homomorphic encryption, the same key has to be used. In the case of aggregation of measurements from different smart meters, using the same key for encryption alone does not provide privacy protection, and thus additional techniques have to be considered as explained in the following sections.

In the following, it is assumed that all involved parties act according to the semi-honest security model as described in Section II-C.

A. Using Homomorphic Encryption and Secret Sharing

Garcia and Jacobs propose a privacy-preserving protocol based on *secret sharing* (see Box II) in [12]. This protocol described below defines two roles: 1) the UC as the aggregator, and 2) customers with smart meters. The proposal completely hides the measurements from the UC since it receives encrypted measurements that it cannot decrypt, and random shares of the total consumption. At the same time, neither of the participants can retrieve meaningful information on the consumption of others as they only see the random shares.

The protocol starts with each user splitting their measurements into random shares, one share for each person:

$$\begin{aligned} \text{Alice: } m_{1,t} &= m_{1,t}(1) + m_{1,t}(2) + m_{1,t}(3) \bmod \eta , \\ \text{Bob: } m_{2,t} &= m_{2,t}(1) + m_{2,t}(2) + m_{2,t}(3) \bmod \eta , \\ \text{Charles: } m_{3,t} &= m_{3,t}(1) + m_{3,t}(2) + m_{3,t}(3) \bmod \eta , \end{aligned} \quad (3)$$

where η is a large integer. Keeping $m_{1,t}(1)$ for herself, Alice sends $m_{1,t}(2)$ and $m_{1,t}(3)$ to the UC after encrypting them with Bob's and Charles' public keys, respectively. Bob and Charles also repeat the same steps with their shares.

Assuming that the UC receives encrypted shares from Alice, Bob and Charles, it adds the shares, which are encrypted by the same key, using the homomorphic property of the encryption scheme as follows:

$$\mathcal{E}_{pk_i}(m'_{i,t}) = \prod_{j \neq i} \mathcal{E}_{pk_i}(m_{j,t}(i)) = \mathcal{E}_{pk_i}\left(\sum_{j \neq i} m_{j,t}(i)\right) , \quad (4)$$

where pk_1 , pk_2 and pk_3 are Alice's, Bob's and Charles' public keys, respectively. The UC then sends $\mathcal{E}_{pk_1}(m'_{1,t})$ to Alice, who can decrypt it using her secret key. She adds her share $m_{1,t}(1)$ to $m'_{1,t}$, obtaining $m_{1,t}(1) + m_{2,t}(1) + m_{3,t}(1)$ in clear text, and sends it to the UC. Bob and Charles also do the same. Upon receiving sums in plain text, the UC adds all inputs and obtains the total consumption.

Being very simple, the proposed scheme perfectly achieves the privacy goal as the UC cannot access the private individual measurements. Unfortunately, the cryptographic protocol relies on secret sharing, which increases the amount of data (note the modulo η , which is a large integer). Table I shows the complexity analysis with respect to the homomorphic operations. The protocol is not scalable since the total number of homomorphic encryptions and modular multiplications is in either case $\mathcal{O}(N^2)$, where N is the number of smart meters. As each cipher text is in the order of thousands of bits, this amount of encryption is also communication-wise expensive.

B. Using Masking and Brute Forcing

The second approach we consider in detail is proposed by Kursawe *et al.* in [18]. The authors propose two ways to efficiently compute the total consumption in a smart metering system with limited hardware resources. In the first one, called aggregation protocols, Alice, Bob and Charles mask their measurement in such a way that when inputs from all parties are summed, masking values cancel each other out and the aggregator obtains the total consumption. In the second approach, named comparison protocols, authors make an assumption that the aggregator (UC) roughly knows the total consumption. In this approach, Alice and the others compute $g_i^{m_{1,t}+r_1}$, $g_i^{m_{2,t}+r_2}$ and $g_i^{m_{3,t}+r_3}$, respectively, where g_i is computed as the hash of a unique identifier, e.g. a serial number or time and date of the measurement. The random numbers r_1 , r_2 , and r_3 are generated in such a way that they sum to zero and are used for masking the measurements. It is clear that the UC can easily aggregate the inputs from Alice and the others:

$$\prod_{j=1}^3 g_i^{m_{j,t}+r_j} = g_i^{\sum_{j=1}^3 m_{j,t}+r_j} \text{ mod } p, \quad (5)$$

where p is a large prime number.

Obviously, the UC cannot obtain the actual sum since this requires solving a discrete-log problem, which is infeasible. As the UC is given g_i and has an approximation of the total consumption $\tilde{C}_{total}(t)$, it can compute values and test for equality, thus brute-forcing values of $g_i^{\tilde{C}_{total}(t)}$, $g_i^{\tilde{C}_{total}(t)-1}$, $g_i^{\tilde{C}_{total}(t)+1}$, ... until a match is found.

The authors propose four protocols that provide different ways for a number of smart meters to derive r_j and g^{r_j} : one based on secret sharing and other three on Diffie-Hellman key exchange protocol and bi-linear map. In the following, we only summarize one of the protocols based on Diffie-Hellman key-exchange protocols to generate random numbers.

The protocol assumes that each customer has a unique ID j and a secret key R_j . To generate the r_j values, a generator of a Diffie-Hellman group g_i is computed using a hash function, with i being the time slot for computing the total consumption. Then, each smart meter computes the public key $g_i^{R_j}$ and distributes it to others with valid

certificates. After verification of the public keys, everyone computes,

$$g_i^{r_j} = \prod_{k \neq j} \left(g_i^{R_k} \right)^{(-1)^{k < j} R_j}, \quad (6)$$

where $k < j$ is 1 if the index of meter k is smaller than the index of meter j , and zero otherwise. Clearly the sum of all r_j is zero:

$$\sum_j r_j = \sum_j \sum_{k \neq j} (-1)^{k < j} R_k \cdot R_j = 0. \quad (7)$$

Once these random numbers are generated, Alice and the others can continue with the computation of the total consumption as explained before.

Table I shows the complexity of the above described protocol. The number of messages to be exchanged is $\mathcal{O}(N^2)$, as each smart meter has to access a new Diffie-Hellman key for the aggregation of the measurements in the most secure form of the protocol. The number of modular multiplications is $\mathcal{O}(N)$ and the number of exponentiations is $\mathcal{O}(1)$. Notice that the computations are on a Diffie-Hellman group, for which the key length is suggested to be 256 bits in the original work. Compared to previous work from Section III-A, that suggests to use the Paillier cryptosystem, which relies on very large key sizes, the small size of the key presents a significant advantage in performance.

C. Using Modified Homomorphic Encryption

The third approach we consider is a cryptographic protocol by Erkin and Tsudik that computes the total consumption in a smart metering system using a modified version of the Paillier cryptosystem [8]. Based on this modification, the authors propose three schemes for: 1) computing the aggregated consumption of a number of customers for a specific time slot (spatial case), 2) computing the total consumption of a single customer for a time interval (temporal case), and 3) computing total consumption in a neighborhood for a specific time slot and the total consumption of each customer for a time interval by using only one encryption per smart meter in each time slot (spatio-temporal case).

In the spatial computation case, Alice and the others compute the total consumption on their own, hence there are only customers, one of which acts as an aggregator. In the temporal computation of a single customer, which is interesting for billing purposes, the authors introduce a UC as an aggregator. The UC receives encrypted measurements from the smart meters but it is unable to decrypt without help from the smart meters. Only by receiving the last encryption, the UC is able to obtain the total consumption in plain text. In the spatio-temporal case, Alice, Bob and Charles disseminate their encrypted measurements in every time slot and each of them is able to easily compute the total consumption for that time slot. Similar to the temporal case, the UC relies on the last input from the smart meter to compute the total consumption of that single customer. Figure 3 summarizes these three scenarios.

The idea in [8] relies on the use of Paillier cryptosystem, where the modulo n is split into random shares (Box II). Assume that Alice, Bob and Charles have three random numbers such that $n_1 + n_2 + n_3 = n$. In such a case,

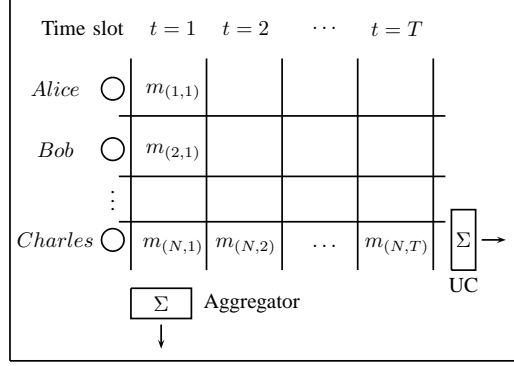


Fig. 3. Spatio-Temporal Consumption from [8].

Alice and the others can encrypt their measurements as follows:

$$\begin{aligned}
 \text{Alice: } \mathcal{F}_{pk}(m_{1,t}) &= g^{m_{1,t}} \cdot r^{n_1} \bmod n^2, \\
 \text{Bob: } \mathcal{F}_{pk}(m_{2,t}) &= g^{m_{2,t}} \cdot r^{n_2} \bmod n^2, \\
 \text{Charles: } \mathcal{F}_{pk}(m_{3,t}) &= g^{m_{3,t}} \cdot r^{n_3} \bmod n^2,
 \end{aligned} \tag{8}$$

where function \mathcal{F} denotes the modified encryption function. Here we use the modified Paillier cryptosystem for its homomorphic property and thus, the decryption key is also available to everyone. It is clear that even with the *public* decryption key, no one can decrypt the encryptions since the random numbers cannot be removed as they are not in the form of r^n . However, an aggregator, anyone in the group, can collect the encryptions and form a proper encryption of the total consumption:

$$\begin{aligned}
 \prod_i \mathcal{F}_{pk}(m_i) &= g^{\sum_i m_{i,t}} \cdot r^{\sum_i n_i} \bmod n^2 \\
 &= g^{\sum_i m_{i,t}} \cdot r^n \bmod n^2 := \mathcal{E}_{pk}\left(\sum_i m_{i,t}\right).
 \end{aligned} \tag{9}$$

Certainly, individual measurements are kept secret but the sum can be obtained by everyone easily. For this technique to work, Alice, Bob and Charles should use the same random number r . This is achieved by generating a hash value of the time-stamp, which also associates the encryption to a specific measurement in a certain time slot.

While Erkin and Tsudik propose very simple protocols for spatio-temporal total consumption, the schemes still require smart meters to be able to perform Paillier encryption, hash functions, random number generation and, most importantly, to communicate with each other. The complexity of the protocol is lower than the scheme in Section III-A as shown in Table I: only one encryption by each smart meter in each time slot and $\mathcal{O}(N)$ modular multiplication for the aggregation.

D. Using Masking and Differential Privacy

The last approach we consider is proposed by Ács and Castelluccia in [4]. The proposed solution also relies on an additively homomorphic encryption scheme but does not make use of expensive secret sharing or public cryptosystems such as Paillier. Private measurements from smart meters are encrypted with a simple and yet efficient cryptosystem, where the encryption is defined as $\mathcal{E}_{pk}(m, k, n) = m + k \bmod n$, where m is the measurement, k is the encryption key and n is a large number. Since addition over a modulo is significantly faster to perform compared to other encryption functions, for example the Paillier encryption function requires computing powers of large numbers, this simple cryptosystem is significantly efficient. It is additively homomorphic as well:

$$\begin{aligned} \mathcal{E}_{k_1}(m_1) + \mathcal{E}_{k_2}(m_2) &= m_1 + k_1 + m_2 + k_2 \bmod n \\ &= \mathcal{E}_{k_1+k_2}(m_1 + m_2) . \end{aligned} \quad (10)$$

The aggregation protocol presented in [4] defines an aggregator, which can be any customer or the UC. The protocol starts with smart meters choosing a set of other smart meters randomly. For the sake of simplicity, assume that Alice chooses Bob and Charles. The coupling between any two smart meter is bi-directional, so Bob and Charles choose Alice as well. Once Alice and Bob are coupled with each other, they generate a random number, $r_{1,2}$, by feeding a Pseudo Random Function (PRF) [13] with their shared keys. Alice adds $r_{1,2}$ to her measurement while Bob subtracts it from his own. Alice generates another random number with Charles, $r_{1,3}$, and adds it to her measurement too. Alice finally encrypts the resulting sum using a key, K_{AU} , which is shared with the UC:

$$\mathcal{E}_{K_{AU}}(\tilde{m}_{1,t}) = m_{1,t} + r_{1,2} + r_{1,3} + K_{AU} \bmod n . \quad (11)$$

Similarly, Bob and Charles encrypt their masked measurements,

$$\begin{aligned} \text{Bob: } \mathcal{E}_{K_{BU}}(\tilde{m}_{2,t}) &= m_{2,t} - r_{1,2} + r_{2,3} + K_{BU} \bmod n \\ \text{Charles: } \mathcal{E}_{K_{CU}}(\tilde{m}_{3,t}) &= m_{1,t} - r_{1,3} - r_{2,3} + K_{CU} \bmod n , \end{aligned} \quad (12)$$

where K_{BU} is the shared key between Bob and the UC, and K_{CU} between Charles and the UC. Alice, Bob and Charles then send the encryptions to the UC.

Upon receiving the encryptions, the UC computes the total consumption by aggregating them. When added, the random numbers that are mutually generated cancel each other out and the UC obtains the aggregated sum in plain text by subtracting K_{AU} , K_{BU} and K_{CU} as he knows these values.

Clearly, the UC cannot observe the actual measurements of anyone since each measurement is masked by a set of random numbers, which cancel each other out only when they are all added. Therefore, individual measurements are kept completely hidden from the UC. However, for this scheme to work, each smart meter has to share keys with the UC and exchange pseudo-random numbers with many other smart meters.

One difference in [4] compared to previously discussed approaches is that Ács and Castelluccia obscure the aggregated data prior to encryption using *Differential Privacy* (Box III) [7]. In particular, the UC is assumed to

access only $\sum_i m_{i,t} + \mathcal{L}(\alpha)$, where $\mathcal{L}(\alpha)$ is the Laplacian noise generated according to the ϵ parameter. Assuming that aggregator is a different entity from the UC, the Laplacian noise could be added by the aggregator prior to sending the encrypted total to the UC. However, for this scheme to work, the aggregator should be trustworthy. Instead of relying on a aggregator, the authors prefer to jointly generate the Laplacian noise by the smart meters. This is made possible by using a lemma that states that the Laplacian noise is divisible and can be constructed as the sum of i.i.d. gamma distributions: $\mathcal{L}(\alpha) = \sum_i (\mathcal{G}_1(i, \alpha) - \mathcal{G}_2(i, \alpha))$, where $\mathcal{G}_1(i, \alpha)$ and $\mathcal{G}_2(i, \alpha)$ are i.i.d. random variables having gamma distribution with a probability distribution function (PDF) as specified in [4].

Given that the gamma distributions with this specification are generated locally, each smart meter adds gamma noise to its measurement to obtain $m_{i,t} + \mathcal{G}_1(i, \alpha) - \mathcal{G}_2(i, \alpha)$ before encrypting it. When aggregated, the sum yields to $\sum_i m_{i,t} + \mathcal{L}(\alpha)$. Note that with this construction, the UC and aggregator can be the same entity.

The overall complexity of the protocol given in Table I is significantly lower than the previously mentioned methods due to the very simple encryption function. Communication, however, is dominated by the exchange of random numbers. Note that the computational complexity is linear in the number of smart meters as the previous approaches. However, due to the ϵ parameter for differential privacy, the number of smart meters directly influences the level of privacy. The original work suggests to have a large cluster of smart meters as the noise is calibrated according to the maximum consumption. Interested readers can find a thorough discussion on the number of smart meters and the average privacy achieved in [4].

TABLE I
COMPLEXITY ANALYSIS OF DESCRIBED APPROACHES FOR A SMART METER (SM) AND AN AGGREGATOR (A).

Operations	Garcia&Jacobs [12]		Kursawe <i>et. al</i> [18]		Erkin&Tsudik [8]		Ács&Castelluccia [4]	
	SM	A	SM	A	SM	A	SM	A
	Paillier (2048 bits)		DH Group (256 bits)		Paillier (2048 bits)		HE (32 bits)	
Encryption	$\mathcal{O}(N)$	-	-	-	$\mathcal{O}(1)$	-	-	-
Decryption	$\mathcal{O}(1)$	-	-	-	-	$\mathcal{O}(1)$	-	-
Multiplication	-	$\mathcal{O}(N^2)$	-	$\mathcal{O}(N)$	-	$\mathcal{O}(N)$	-	-
Exponentiation	-	-	$\mathcal{O}(1)$	-	-	-	-	-
Addition	-	-	-	-	-	-	$\mathcal{O}(1)$	$\mathcal{O}(N)$
Subtraction	-	-	-	-	-	-	-	$\mathcal{O}(1)$
Communication	$\mathcal{O}(N)$	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$	$\mathcal{O}(N^2)$	$\mathcal{O}(1)$	$\mathcal{O}(N)$	$\mathcal{O}(1)$	$\mathcal{O}(N)$

IV. DISCUSSION

We have presented several examples of technological privacy-preserving solutions to the computation of total consumption. These solutions try to tackle the privacy issues of current smart metering infrastructures, in order to avoid the need for a universal trust on the grid operator and to comply with data protection regulations. As we have seen, the peculiarities of each model (i.e., different management, authentication and storage, and their corresponding trusted elements) highly influence the optimal way to guarantee privacy with the minimal interference into the service provision and utility. At the same time, it is also desirable to minimize the extra hardware requirements that, if excessive, can negate the benefits of the smart metering paradigm, discouraging its deployment and implementation.

From this starting point, we can identify and foresee several crucial challenges for addressing the privacy-aware smart metering scenario, taking into account the strengths and weaknesses of the compared approaches.

A. Challenges derived from hardware limitations

1) *Scalability and flexibility*: A cluster in a grid can provide supply to a number of users that ranges from a few hundreds (in distributed networks) to tens of thousands. For a proper operation of the grid, it is essential that all the implemented systems scale well. For the protocols reviewed in the previous section, Table I gives a glimpse of their scalability properties with respect to the number of consumers in the same cluster. While both [12] and [18] present a communication complexity that is quadratic for the aggregator and linear for each meter, both [8] and [4] present a linear complexity for the aggregator and constant for the meters. Hence, the latter two protocols will scale much better than the former. A similar conclusion can be drawn for the computational complexity, for which [4] presents a really lightweight protocol. It must be noted, though, that [4] needs that the number of customers per cluster be large enough, or the differentially private protocol will introduce an excessive noise power into the results.

2) *Communication bottleneck*: Meter devices have to communicate with the aggregator or with the utility company to send the (authenticated) measurements, either the whole sequence or a partially aggregated summary. This is the minimum communication that the meters must perform. In terms of energy consumption for embedded devices, using a wireless link for sending a bit is equivalent to carrying out around one hundred microcontroller instructions [23], so data transmission should be kept to a minimum, and only triggered when strictly necessary.

Nevertheless, the proposed simplified homomorphic encryption schemes and key exchanges between meters, as well as the distributed noise generation processes [4] involve many interaction rounds among the meters. Thus, the concealment of the private values through reduced-computation mechanisms is achieved through a collaborative process among meters, at the cost of increasing the communication complexity in the network. This overhead can be the true bottleneck of the smart metering system. Hence, actually optimizing communication, computation and power drain at the meters is a hard task that has not been fully addressed in privacy-preserving approaches.

3) *Limited resources of smart meters and efficient homomorphic encryption*: For the smart metering scenario to be economically viable from the point of view of Grid Operators, smart meters must be cheap and easily replaceable and/or reconfigurable devices. This need responds to a scalability principle, for which the cost of deploying all the meter devices in all the households and facilities of the served users must be manageable and covered by the energy savings and consumption reduction that the “smart” use of the grid and the optimal load balancing will provide. Hence, smart meters cannot be fully-fledged PCs, but small embedded devices with very limited computation resources and, obviously, small power consumption.

Due to these fundamental constraints, some of the proposals for privacy-preserving smart metering consumption have targeted the use of simple homomorphic encryptions, like the modular addition of Ács and Castelluccia [4], or the use of light secret sharing schemes.

It is worth noting that existing current meters do not comprise trusted elements capable of performing complex homomorphic encryption; if any, they use symmetric cryptography [16], [9], usually supporting “light” cryptographic functions like hashes and secret-key encryption/decryption and HMAC signatures. Most of the proposed privacy-preserving solutions require [12] the inclusion at the meters of tamper-proof cryptographic modules (similar to smart cards); these modules must handle integrity, distributed authentication and heavy public key data encryption and signatures. Furthermore, if homomorphic processing is used, the meters must also cope with homomorphic operations that involve large modular additions, multiplications and exponentiations. It is worth noting also that the encryption techniques used by privacy-preserving protocols (like Paillier) are not a widely used standard like RSA, and they are not present by default in typical cryptographic modules, so they have to be recompiled and optimized; this may be a problem in the short term, while there is no consensus in the encryption methods needed for an integral privacy-preserving solution. Nevertheless, in the long term this will not be a major problem, as the massive adoption of smart metering will lower the production costs of the chosen solution.

B. Challenges related to secure cryptographic protocols

1) Malicious parties and tampering: All the presented solutions to private smart metering are devised for a semi-honest adversarial model, in which none of the parties will deviate from the established protocol or forge any results. This is a very optimistic model, unlikely for a real scenario with malicious parties. These malicious parties will find more opportunities to compromise the correct operation of the system as the communication needs of the used protocol grow, so these needs have to be minimized. Additionally, meter tamper-proofness is essential to prevent forgeries and deviations. In a non-private system, the tamper-proof section of the device will comprise only the sensors and timing, but when privacy-preserving protocols come into play, the cryptographic module in charge of producing and receiving the needed transcripts has to be also tamper-proof so that the user cannot modify the correct behavior. As a man-in-the-middle attack is practically unavoidable, tamper-proofness is not enough, and other additional specific cryptographic mechanisms have to be considered also for the protocol to be valid against malicious adversaries.

It is remarkable, though, the strength shift in the smart metering roles when tackling privacy constraints: in a non-private system the utility company has all the control and concentrates the need of trust from the consumers, that must blindly assume that they will be billed correctly for their consumptions. Conversely, if consumer privacy is guaranteed, then it is the utility company who must trust that the measurement aggregation and billing calculation are correctly performed, as it will not have access to specific individual measurements. This is the main reason for the grid operators to be reluctant to adopt a privacy preserving solution if it does not come together with a fraud detection mechanism and technical guarantees that cheating customers will not succeed.

2) Key management: For private protocols based on homomorphic processing, it is a common requirement that all the encrypted values be produced with the same key in order to be homomorphically “combinable” [12], [8], in such a way that the secret key is shared among several customers and even the utility company. In an ordinary setting, this key disclosure would imply losing the possibility of correct authentication, and pose other problems

related to the possibility of forgeries by dishonest users with decryption capabilities. The solution to these problem passes through unusual key distribution mechanisms, like the sub-key generation process by Erkin and Tsudik [8], or the peer-to-peer key establishment by Ács and Castelluccia [4], in which each two coupled users share a uniquely generated key for each iteration of the private consumption calculation protocol. Hence, it is not yet possible to have fixed unique individual secret keys without having to resort to too costly strategies like proxy-reencryption or encryption delegation.

3) *Securing Billing Calculations*: Billing $B(t)$ is one of the private utility functions needed by Electricity Producers. But the peculiarities of the billing process pose additional issues [17]. A simple privacy-preserving protocol could calculate a certified private bill combining the encrypted measurements and the appropriate tariffs; but an integral privacy protection mechanism would also include secure deposits and anonymous payments. Furthermore, when the used protocol is differentially private, like those in [17] (See Box III), the output billing is *fuzzy*, and the noise added to the calculations involves producing inaccurate invoices. There will be a noisy *fake* consumption (positive or negative for each client) that the protocol itself has to compensate by providing secure mechanisms of in advance payments and *a posteriori* rebates, together with an assertion protocol for avoiding fraud or abuse by the noise addition procedures [17]. While there are already proposed solutions for this scenario [17], it is likely that some customers will not be comfortable with paying in advance for a fake consumption, so there is still room for improvement and further research in this area.

C. Challenges related to Signal Processing

1) *Complex utility functions*: Throughout this article, we have only presented and discussed the problem of private total consumption calculation, for which the general summation function $GS(t)$ takes a very simple form. More complex functions may be desirable from the utility company point of view, ranging from billing with non-linear tariffs, to more complex statistical calculations related to profiling, load forecasting, state estimation, adaptive frequency estimation, or network modeling². Devising privacy-preserving protocols that deal with these complex functions while keeping a low overhead that does not exceed the capabilities of smart meters is a challenging task; this cannot be handled by homomorphic encryption alone, and using further secure interactive protocols increases communication and computation complexity.

Furthermore, restricting the possible utility functions that the providers may privately obtain from the measurements has the effect of limiting the functionality of the grid operators and bounding the information that they might otherwise get from an indiscriminate access to fine-grained consumption data. This is not desirable for providers, but it is beneficial from a privacy point of view, as it forces the providers to explicitly specify which computations and which results they want to obtain from the private data at every moment, in fulfillment of the Data Protection Directives.

But this problem is not exclusively a cryptographic issue; signal processing can also play a fundamental role in solving this challenge: the signal processing algorithms for performing complex calculations like forecasting

²We refer the interested reader to [3] for a current view of complex signal processing related tasks and challenges in future smart grids.

(e.g., predictive filtering) or profiling (e.g., maximum likelihood estimation) have been originally developed without privacy in mind, and without the restrictions that current cryptographic privacy-preserving techniques pose. There is a very interesting challenge in finding approximate signal processing analogous protocols that conform to the limitations of the cryptographic techniques while providing, with a bounded error, similar results to those complex forecasting and profiling algorithms.

2) *Accuracy loss*: For *fuzzy* mechanisms that add noise in order to guarantee differential privacy [7], there is an accuracy loss for the information that providers might get as the outcomes of these mechanisms. There is a direct relationship [7] between the induced noise power (measurement accuracy) and the ϵ level of differential privacy that the mechanism achieves. This tradeoff has to be carefully considered and evaluated for each utility function, as it might be the case in some scenarios that the obtained results get lost in noise and become useless if the needed privacy level is too high (i.e., too noisy billing data). This is closely related to the use of approximate algorithms to achieve strict efficiency goals.

V. CONCLUSION

Deployment of smart grids is progressing fast in many countries despite several challenges in the legal, business and technology point of view. One related research challenge for the signal processing community is the protection of private smart meter measurements from the untrustworthy stakeholders while the core smart grid functions stay intact. Secure Signal Processing, which aims for computing a signal processing function with private signal inputs, presents itself as a powerful technological solution which can make the deployment of smart grids more acceptable for the end-users. The distributed setting of the smart meters, different involved parties, and the functions to be realized in a privacy-preserving manner with hardware constraints constitute an appealing problem domain for the signal processing research community, which can take advantage of experiences in distributed computing, optimization and efficient communication. Certainly, for privacy protection, the researchers should also invest in cryptography, getting familiar with its utility and limitations. With this paper, we identify the privacy problems in smart grids, summarize the recent research on data aggregation, and present an overview of existing research challenges for SSP.

BOX I: HOMOMORPHIC ENCRYPTION

A plain text message m is encrypted in Paillier [21] with the following function:

$$\mathcal{E}_{pk}(m) = g^m \cdot r^n \bmod n^2, \quad (13)$$

where n is a product of two large primes, p and q , g is a random number in $\mathbb{Z}_{n^2}^*$ with an order n , meaning that $g^n \bmod n^2 = 1$. The tuple g, n is the public key. The random number r is chosen such that $\gcd(r, n) = 1$. Using a different random value for every encryption guarantees that the cipher text for the same plain text will be different in each case, hence the encryption scheme is called probabilistic. Note that the recipient of the cipher text does not need to know r to decrypt the message since $(r^n)^\lambda \bmod n^2 = 1$, where λ is the secret key and it is given by

$\text{lcm}(p-1, q-1)$. Therefore, any r that is co-prime to n can be removed easily if it is raised first to the power of n and then λ .

The Paillier encryption scheme is additively homomorphic, meaning that multiplication of cipher texts of two messages results in an encryption of the sum of these two messages:

$$\begin{aligned} \mathcal{E}_{pk}(m_1) \cdot \mathcal{E}_{pk}(m_2) &= g^{m_1} \cdot r_1^n \cdot g^{m_2} \cdot r_2^n \bmod n^2 \\ \mathcal{E}_{pk}(m_1 + m_2) &:= g^{m_1+m_2} \cdot (r_1 r_2)^n \bmod n^2 . \end{aligned} \quad (14)$$

We refer readers to [21] for more details on the decryption function. Interested readers can find more information on Homomorphic Encryption in [10] and its usage in encrypted signal processing in [19]. For other cryptographic notions, we refer the reader to [20].

BOX II: SECRET SHARING

The main idea in secret sharing is dividing a secret s into m pieces called shares. Each of these shares are then sent to a user in a secure way. A coalition of some of users is later able to reconstruct the secret. Shamir explains a threshold secret sharing scheme in [25], where any combination of k shares out of m can be used to reconstruct the secret. The proposed method is based on generating random points on a polynomial of degree k whose constant term is the secret. Clearly, when any k points are combined, the polynomial can be reconstructed, and the secret can be revealed.

BOX III: DIFFERENTIAL PRIVACY

A function \mathcal{F} is ϵ -differentially private, if for all data sets D_1 and D_2 , where D_1 and D_2 differ on at most one element, and for all subsets of possible answers $S \subseteq \text{Range}(\mathcal{F})$,

$$P(\mathcal{F}(D_1) \in S) \leq e^\epsilon \cdot P(\mathcal{F}(D_2) \in S) . \quad (15)$$

The above definition says that a differentially private function produces indistinguishable outputs for inputs that differ by a single element, meaning that a modification in one data set changes the probability of any output by a factor of e^ϵ at most. Here, ϵ controls the level of privacy: the lower values of ϵ , the stronger privacy.

REFERENCES

- [1] Directive 95/46/EC of the European Parliament and of the Council. Official Journal L 281, 23/11/1995 P. 0031 - 0050, October 1995.
- [2] Directive 2002/58/EC of the European Parliament and of the Council. Official Journal L 201, 31/07/2002 P. 0037 - 0047, July 2002.
- [3] Special Issue on: The Future Smart Grid, Signal Processing Challenges. *Signal Processing Magazine, IEEE*, 29(5), sept. 2012.
- [4] G. Ács and C. Castelluccia. I have a DREAM! (Differentially PrivatE smart Metering). In *Information Hiding Conference*, LNCS, pages –. Springer, May 18-20 2011.
- [5] R. Anderson and S. Fuloria. On the security economics of electricity metering. In *The 9th Workshop on the Economics of Information Security*, 2010.

- [6] C. Castelluccia, E. Mykletun, and G. Tsudik. Efficient aggregation of encrypted data in wireless sensor networks. In *Proceedings of the The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 109–117, Washington, DC, USA, 2005. IEEE Computer Society.
- [7] C. Dwork. Differential privacy. In *Automata, Languages and Programming*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin / Heidelberg, 2006.
- [8] Z. Erkin and G. Tsudik. Private computation of spatial and temporal power consumption with smart meters. In *International Conference on Applied Cryptography and Network Security*, pages 561–577. Springer-Verlag, June 26-29 2012.
- [9] ETSI. Open smart grid protocol (osgp), 2012.
- [10] C. Fontaine and F. Galand. A survey of homomorphic encryption for nonspecialists. *EURASIP Journal on Information Security*, 2007, 2007.
- [11] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel. Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1):28–39, Jan.-March 2011.
- [12] F. D. Garcia and B. Jacobs. Privacy-friendly energy-metering via homomorphic encryption. In J. C. et al., editor, *6th Workshop on Security and Trust Management (STM 2010)*, volume 6710 of *Lecture Notes in Computer Science*, pages 226–238. Springer Verlag, 2010.
- [13] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *J. ACM*, 33(4):792–807, Aug. 1986.
- [14] G. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, Dec 1992.
- [15] M. Jawurek, M. Johns, and F. Kerschbaum. Plug-in privacy for smart metering billing. In *PETS*, pages 192–210, 2011.
- [16] S. Keemink and B. Roos. Security analysis of dutch smart metering systems. Technical report, Universitait van Amsterdam, July 2008.
- [17] M. Kohlweiss and G. Danezis. Differentially private billing with rebates. In *Information Hiding Conference, LNCS*, pages –. Springer, May 18-20 2011.
- [18] K. Kursawe, G. Danezis, and M. Kohlweiss. Privacy-friendly aggregation for the smart-grid. In *PETS*, pages 175–191, 2011.
- [19] R. Lagendijk, Z. Erkin, and M. Barni. Encrypted signal processing for privacy protection. *IEEE Signal Processing Magazine*, 2013. to appear.
- [20] A. J. Menezes, S. A. Vanstone, and P. C. V. Oorschot. *Handbook of Applied Cryptography*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 1996.
- [21] P. Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In J. Stern, editor, *Advances in Cryptology — EUROCRYPT '99*, volume 1592 of *LNCS*, pages 223–238. Springer, May 2-6, 1999.
- [22] S. Peter, K. Piotrowski, and P. Langendoerfer. On concealed data aggregation for wireless sensor networks. In *4th IEEE Consumer Communications and Networking Conferences*, 2007.
- [23] K. Piotrowski, P. Langendoerfer, and S. Peter. How public key cryptography influences wireless sensor node lifetime. In *Proceedings of the fourth ACM workshop on Security of ad hoc and sensor networks*, pages 169–176, New York, NY, USA, 2006. ACM.
- [24] A. Rial and G. Danezis. Privacy-preserving smart metering. In *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, WPES '11, pages 49–60, New York, NY, USA, 2011. ACM.
- [25] A. Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, Nov 1979.