

Security Is Not Enough! On Privacy Challenges in Smart Grids

Florian Skopik*

AIT Austrian Institute of Technology, 2444 Seibersdorf, Austria

Abstract

Smart Grid solutions are being rolled out at large scale these days. While security concerns are addressed on various levels and sophisticated technical measures applied, privacy issues are still not fully understood yet. The generation of high-frequency meter readings exposes customers to new threats. Various politic initiatives thus aim towards a strong regulatory framework to enable an appropriate protection of the customers' privacy. We argue that finally the only effective measure is to avoid the production of personalized fine-grained meter readings. However, it is important to study the impact of available technical privacy protection concepts on smart grid services, including demand management, load forecasting and energy theft protection, which typically rely on these data. Therefore, in this paper we outline the privacy challenge and the concept of privacy-by-design. In particular, we discuss privacy protection methods and compare them in terms of applicability and impact on smart grid services.

Keywords: Smart grid privacy, meter readings exploitation, data obfuscation, privacy-by-design

1. Introduction

Today, the electric power grid is facing a major paradigm shift. In order to cope with new arising challenges, the well-established static grid is transformed to a more flexible, intelligent and modern utility. This enables the establishment of a decentralized energy distribution system, which allows for more efficient energy usage and the massive integration of distributed renewable energy sources [1], [2]. Facing a world-wide economic crisis and the depletion of fossil fuels, these features are of paramount importance to ensure sustainable economic growth in the next decades. While there is no doubt in the value of this technological change, there remain numerous security and privacy concerns.

The centerpiece of the smart grid is a working smart metering infrastructure. This feature is indispensable to allow the fine-grained monitoring of the grid status which is required to implement efficient energy distribution algorithms, flexible load management, and dynamic pricing models – just to name a few. Since the electric power grid is by far the most important technical infrastructure, and literally the backbone of our economy and society, many studies have been performed to investigate potential security issues. Elaborating the design and architecture of a secure smart grid is in fact a highly non-trivial challenge that has been addressed by various key players of smart grid technologies in the recent years. Although, security and privacy have attracted major attention from both academia and industry, *privacy implications* are not fully understood yet.

We argue that implementing effective security mechanisms only, is not sufficient to protect the privacy of the customers. Besides secure operations in terms of availability and reliability, security also deals with potentially illegitimate users of data and keeps them out of the loop, while allowing legitimate smart grid stakeholders access to sensitive information in order to offer their services, including utility providers who need information about the grid status, the energy producers who need to estimate energy demands, and the billing companies which implement dynamic pricing. However, sensitive data is still widely produced and exchanged in the network. For that reason, politics in the EU as well as the US is currently

* Manuscript received July 13, 2012; revised August 13, 2012.

Corresponding author. Tel.: +43 664 8251495; fax: +43 50550 2813; *E-mail address:* florian.skopik@ait.ac.at .

investing huge efforts in solving data security and privacy issues with a strong regulatory framework. They commit smart grid stakeholders to ‘appropriately’ (e.g., in the US this means in a manner consistent with Federal Fair Information Practice (FIP) principles) deal with sensitive user data and implement strong privacy protection processes on organizational levels. Nevertheless, two issues are widely unsolved: (i) There is no consensus what data can potentially compromise a customer's privacy and to which degree. While for some types, such as individual fine-grained meter readings, it is more obvious, for others it is not. (ii) Even the most sophisticated technical and organizational measures to protect sensitive data cannot guarantee the privacy protection in case of successful malicious attacks [3] to the storage backend, e.g., carried out by a disgruntled employee. With regard to that, past events of large-scale credit card and customer data thefts have demonstrated the vulnerability of large companies to targeted attacks. Thus, we argue that besides the important regulatory framework, also technical means are required to effectively protect the privacy of customers. Here, the only effective way is to avoid the generation of sensitive user data. However, a careful balance needs to be arranged, i.e., on the one side to create as less sensitive data as possible, and on the other side to prevent major impact on the grid operations which might render its added value null and void.

The contributions in this paper are twofold:

- *Privacy Fundamentals in the Smart Grid.* We highlight the various dimensions of the privacy challenge in the smart grid and argue in detail why it is such an important objective to address.
- *Concepts for establishing Privacy-by-Design.* There are already some proposed concepts to establish privacy-by-design, i.e., making privacy protection mechanisms an integral part of smart grid architectures. Here, we compare these approaches and discuss their actual applicability.

The remainder of the paper is organized as follows. In Section 2, we outline important related work in the privacy domain. Section 3 highlights the fundamental concepts of privacy and shows potential consequences of a lack of privacy protection. Then, Section 4 discusses already available approaches for integrating privacy protection mechanisms as integral parts into the design of smart grid architectures. Section 5 deals with an evaluation regarding the applicability of the introduced approaches and potential interference with existing or future services. Finally, Section 6 concludes the paper.

2. Related Work

Threats to and vulnerabilities of smart metering systems are widely discussed topics [4]-[6]. Data communication security controls (e.g., cryptographic functions such as encryption, message authentication codes, and digital signatures) provide standard security services in terms of confidentiality, integrity, and accountability of messages and their origin [7]. However, aspects of privacy and potential threats [8] to customers through smart meter data exploitation are not fully covered yet [5]. An important official first step towards a privacy-enabled smart grid has been made by NIST [9], where they defined problems related to privacy protection and legal constraints. The overall question is basically how to disconnect the identities from smart meter owners from meter readings. Various models have been proposed, such as the anonymization of metering data and inclusion of third-party escrow companies [10], the aggregation of data coming from numerous meters before readings reach the utility [11], [12], and even obfuscation systems that hinder the exact identification of the actual power consumption by using intelligent power routers and battery buffers in homes [13], [14]. SmartPrivacy (‘privacy-by-design’) [15] focuses on a more holistic view, instead of technical aspects only. This concept is an umbrella that offers the complete suite of protections to ensure data privacy. It consists of multiple measures ranging from regulatory protections to education and awareness.

3. Privacy Fundamentals in the Smart Grid

3.1. Definition of Privacy

Defining the notion of *privacy* in a consistent and holistic manner is challenging. A recent NIST report [9] outlines privacy to comprise the following aspects: 1) *Privacy of personal information* which is any

information relating to an individual, who can be identified, directly or indirectly, by that information and in particular by reference to an identification number or to one or more factors specific to his or her physical, physiological, mental, economic, cultural, locational or social identity. Privacy of personal information involves the right to control when, where, how, to whom, and to what extent an individual shares their own personal information, as well as the right to access personal information given to others, to correct it, and to ensure it is safeguarded and disposed of appropriately. 2) *Privacy of the person* is the right to control the integrity of one's own body. It covers such things as physical requirements, health problems, and required medical devices. 3) *Privacy of personal behavior*. This is the right of individuals to keep any knowledge of their activities, and their choices, from being shared with others. 4) *Privacy of personal communications*. This is the right to communicate without undue surveillance, monitoring, or censorship. In context of smart metering, the privacy of personal information (e.g., ownership of specific appliances) and the privacy of personal behavior (e.g., daily habits) are particularly at stake.

3.2. Smart meter data exploitation}

An illustrative example of smart meter data exploitation is given in a recent NIST report [9]. Here, the authors impressively demonstrate how electric appliances leave special marks on the wire that can be tracked to learn about used devices and thus infer people's habits. There are various parties who are particularly interested in meter readings; however, besides legitimate organizations using the data for billing or network services, many other use cases are not desirable, as summed up by Table 1. Several works dealt with potential questions that could be answered, just by studying meter readings. Especially [8] investigated real scenarios and proved that one might be able to find out, if a person (i) was at home during sick leave; (ii) got a good night's sleep; (iii) watched a particular sports game; (iv) left late for work; (v) left children alone at home; and (vi) prefers hot or cold breakfast.

Table 1. Exploitation opportunities for smart metering data (as presented in [9]).

Who wants smart meter data?	How could the data be used?
Utilities	To monitor electricity usage and load; to determine bills
Electricity usage advisory companies	To promote energy conservation and awareness
Insurance companies	To determine health care premiums based on unusual behaviors indicating illness
Marketers	To profile customers for targeted advertisements
Law enforcers	To identify suspicious or illegal activity
Civil litigators	To identify property boundaries and activities on premises
Landlords	To verify lease compliance
Private investigators	To monitor specific events
The press	To get information about famous people
Creditors	To determine behavior that might indicate creditworthiness
Criminals	To identify the best times for a burglary or to identify high-priced appliances

3.3. Privacy threats for smart grid customers

A list of potential consequences when privacy in Smart Grid systems is compromised includes: identity theft, determining personal behavior patterns, determining specific appliances used, performing real-time surveillance, revealing activities through residual data, targeted home invasions (latch key children, elderly, etc.), providing accidental invasions, activity censorship, decisions and actions based upon inaccurate data, profiling, unwanted publicity and embarrassment, tracking behavior of renters/leasers, behavior tracking (possible combination with personal behavior patterns), or public aggregated searches revealing individual behavior. Some of these threats have been studied extensively, such as behavior profiling [16]. In order to mitigate these effects, the regulatory frameworks (e.g., NIST [9]) basically consist of two essential recommendations (which are supposed to be converted to national law later on): (i) Limit the collection of data to only that necessary for Smart Grid operations, including planning and management, improving energy use and efficiency, account management, and billing. (ii) Obtain the data by lawful and fair means and, where appropriate and possible, with the knowledge or consent of the data subject. The second recommendation also includes the customer's right to demand information on saved and processed data concerning him/her.

3.4. Non-technical challenges}

Various organizational and legal aspects of smart metering are still unsolved. Smart meter data will not only be used by the utility to offer more efficient services, but should also be made accessible to the customer. This is essentially in line with the definition of privacy as mentioned before, where a customer must have the right to know what data a company owns about him/her. Additionally, providing detailed meter readings to the concerned customer will eventually strengthen his/her awareness regarding energy consumption. The industry argues that this way a customer is able to identify inefficient electric appliances, such as old refrigerators or expensive heaters. While these are noble aims in the simple case, where the customer is the actual originator of smart meter data, the situation completely changes for leased apartments or offices. Imagine a case where a tenant or other consumer occupant of a building is not the utility's customer. What access, if any, should the tenant or consumer occupant have to information from the utility about their energy usage? On the other side, is the energy footprint of public buildings, financed through taxes, not of interest to the public? A further heavily discussed issue is that most customers might not be aware of these existing threats to their privacy. Fine-grained meter readings would be enabled by default, and the legal framework could offer them the opportunity to opt-out in case they feel uncomfortable with this situation. However, in case where personal privacy is at stake, many argue that it would be far better to offer an opt-in feature for services that require high-frequency readings (e.g., demand response management), and by default just enable low-frequency data for billing purposes. Finally, setting up an effective privacy protection scheme is not a technical *or* regulatory challenge only, but requires a strong interplay between both of these aspects. For instance, a regulatory framework can provide detailed guidelines for utilities on how to deal with sensitive data, and a legal framework can ensure that these guidelines must be followed. However, once under attack, an organization might not be able to properly protect this sensitive information. Past incidents of large-scale credit card information thefts clearly demonstrate that when sensitive data is widely produced and processed, there will eventually be cases of theft and misuse. Thus we argue, additional technical means are required to address the privacy challenges which – at least partly – avoid already the production of such data.

4. Establishing Privacy by Design

The main threat of smart metering concerning the customers is that high-frequency data which is required for efficient network operations, including advanced load management, demand side response and load shedding and shifting, may expose private information. Assuming that metering data needed for utility purposes, e.g. billing and account management, are communicated at lower frequencies, the main question becomes [10]. How can high-frequency data be anonymized, i.e. not be attributable to a specific smart meter or person respectively, without negatively affecting network operations or the availability of high-frequency metering data?

While current efforts mainly focus on setting up strong regulatory frameworks and adapting existing privacy laws to be applicable for the smart grid, the adoption of technical means are still underdeveloped. Although this fact, several sound solutions already exist which partly address the privacy challenge. However, the most obvious one, that is the limitation of the amount of data, i.e., increase of time intervals of transferred meter readings, is not an option since it would render novel grid services unusable. Fundamental efforts can be categorized into the following classes:

- *Anonymization of Metering Data* applies the idea of separating technical data (meter readings) from customer ids. For that purpose a third-party id escrow company is involved.
- *Metering Data Obfuscation* is about masking the own energy consumption profile with local battery buffers, e.g., from an electric car, so that one cannot infer detailed energy consumption profiles, while the overall consumption profile remains intact.
- *Privacy-preserving Metering Data Aggregation* deals with the online aggregation of data from geographically co-located consumers, i.e., before readings reach the data center, so that the utility provider can still get a clear picture about the grid's state in a street or district, but not for a single customer.

As mentioned, the challenge however is not to disturb the regular grid operation which might render many smart services useless. For instance, when a customer decides to feed back energy into the grid a direct real-time communication path to the billing company needs to be established (e.g., to avoid fraud). While basic functionality can be provided using readings in quite long intervals, for more advanced services, the user should have an explicit 'opt-in'-opportunity. Thus, in order to receive customized services, s/he explicitly should confirm his/her will to share more data (but not to a higher degree than absolutely required to offer the services). With respect to that, important work has investigated the trade-off between gained privacy and loss of utility through privacy preserving mechanisms [17]. Especially regarding billing processes that typically rely on accurate meter data, there is substantial proof that privacy-preserving protocols allow for both, protection of privacy and usage of added value services [18].

4.1. Anonymization of metering data

The key idea of this concept is the separation of smart meter readings from the customer id who possesses the meter [10]. The utility only collects smart meter readings linked to unique ids, however does not have the connection to real customers. For that purpose a further entity is integrated in the whole smart grid framework: an escrow organization, which manages the links between real customer ids and smart meter ids. Similar to key escrow companies, well-known from the IT sector, these companies would greatly contribute to the customers' establishment of trust in novel smart grid technologies.

The authors of [10] further propose to distinguish between low-frequency readings for billing purposes that do not threaten privacy (one reading per week or month); and high-frequency readings (below a minute) that are required for running the technical infrastructure only and do *not* necessarily need to be linked to a certain individual. While low-frequency data is sent directly to the utility and billing company respectively, high-frequency data can be processed in the next substation (where data is actually needed for load management algorithms) and are not being stored in the utility's backend. With this hybrid approach, basic billing services can be provided directly, while anonymized fine-grained meter readings contribute to technical services. This essentially realizes a clear distinction between customer-specific data and technical data.

4.2. Metering data obfuscation

The application of this approach basically does not require any changes of the overall smart grid architecture. The basic idea is to install intelligent power routers together with rechargeable batteries at home [13] in order to obfuscate the usage of specific electric appliances. With these technologies characteristic load peaks are smoothed and hidden from the outside [14]. Intelligent power management algorithms [19] are used to obfuscate the actual power consumption of a household. There, energy is either taken from a battery or from the grid. Especially [13] provide preliminary proof that integrating a battery that is loaded and discharged in non-periodic intervals greatly reduces the leakage of information about events going on in one's home. Since an important part of the smart grid initiative is also the wide application of e-mobility, soon households will own this required battery anyway, making this model feasible for fast adoption. Finally, it should be noted that obfuscation is not meant to fully smoothen energy consumption spikes or fully cover load signatures, but is utilized to significantly reduce information leakage, making it rather impossible to infer an individual's behavior (cf. Table 1).

4.3. Privacy-preserving metering data aggregation

The idea of smart meter data aggregation [11], [12] is essentially motivated by a substantial reduction of the amount of information, e.g., to reduce the computational effort in large-scale metering infrastructures or to minimize storage capacity requirements. As an additional advantage, aggregation also limits the threat to data exploitation. Basically we can distinguish between *aggregation by location*, i.e., over numerous smart meters, so that data still arrives at real-time at the utility, but instead for a single household data characterizes a larger grid segment (e.g., a whole apartment block or street); and *aggregation over time*, i.e., the aggregation of single readings from a particular meter over a longer

interval. In the latter case, information can still be correlated with a single customer but readings are smoothed over time before being transmitted to the utility's backend. This is particularly of interest when this aggregation is done in the nearest substation to the originating meter, so that high-frequency meter readings can be used for local grid management, but is not exploitable afterwards.

Data aggregation effectively protects privacy, however raises a multitude of new concerns. A first issue to deal with is how to keep smart grid services that rely on high-frequency metering data operational. While aggregation does not interfere with services that require data aggregation anyway (e.g., accounting and billing), the situation is different for dynamic load management. Here the nature and degree of aggregation is important, since aggregation by location is not hindering load management of a whole grid segment. A second issue concerns the integrity of aggregated data. Since there are no detailed statistics down to single meters, wrong readings or energy theft would be impossible to detect. For that purpose advanced integrity checks have been introduced, such as [20]. A third issue deals with data handling in general. It is widely agreed that data that leaves an individual's home need to be encrypted to avoid eavesdropping. Thus, decryption would be required at those entities in the architecture (e.g., substations) that perform aggregation operations. Since this would eventually break a security architecture, protocols have been invented that allow the aggregation of data without the need for prior decryption [11].

5. Discussion

With the smart grid a number of novel services become reality which heavily relies on fine-grained smart meter readings. We identified the most important ones through literature reviews:

Individual Demand Management: refers to letting the utility provider configure and control electric appliances of an individual's home. For instance, the time when a hot water tank is heated up can vary a bit. In the future, when electric cars will increasingly be used, batteries should be loaded when the overall energy consumption in the whole network is at minimum. This can be controlled by the utility based on individual meter readings.

Multi-cast Demand Management: is similar to individual demand management, however here, a multi-cast (or broad-cast) message is sent to the network and appliances can be configured to react on such messages. No individual relations, e.g., manifested by SLAs, are established, which however, avoids personalization and decreases predictability of energy consumption from the utility's perspective. For multi-cast demand management aggregated readings should be sufficient.

Energy Theft Detection: Roughly (and when neglecting losses at transmission), the sum of produced and distributed energy must match the sum of consumed energy, captured by meter readings. Additionally, households have a rather static consumption behavior over years. With individual meter readings, utilities can cross-check and detect large-scale energy thefts.

Load Forecasting: is useful to detect local overloads and enable load shifting and shedding. While these forecasts can be performed at local stations for whole network segments, individual meter readings might be useful to detect energy distribution issues on a very low level. Thus, individual meter readings are required.

Energy Feedback: is a feature of the smart grid, allowing participants not to act as consumers only, but also as producers. This service is essential in order to establish a real decentralized energy supply and integrate distributed alternative energy sources, such as privately owned wind mills and solar panels.

Table 2 lists these smart grid services together with the results of a compatibility evaluation regarding their realization with the introduced privacy protection mechanisms in place. The essential question here is, if – basically – a privacy preserving technique avoids the realization of one or more of these features. Notice the basic question here is if these services can be realized without huge extra effort (besides the implementation of privacy specific entities) or changing the overall smart grid architecture.

To sum up, *anonymization* keeps the individual fine-grained meter reading profiles intact, however, effectively decouples readings (managed by the utility) and customer ids (managed by an escrow organization). Together with a comprehensive legal framework that backs up this solution, anonymization does not seriously interfere with major grid services. However, privacy can be compromised in case the

utility and escrow organization exchange and correlate their information. Furthermore, larger structural changes of smart metering architectures are necessary. Next, *obfuscation* requires less changes in the grid management architecture, however, interferes with services that explicitly require individual meter readings. Finally, *aggregation* is the means that protect privacy best, since fine-grained meter readings are not stored at the back end and thus simply not available. Here, further mechanisms, such as advanced integrity checks when aggregating meter readings, are required to provide some of the services, including energy theft detection – at least for grid segments.

Table 2. Comparison of privacy preserving techniques.

Grid feature	Anonymization	Obfuscation	Aggregation
Individual demand management	supported	not supported	not supported
Multi-cast demand management	supported	supported	supported
Energy theft detection	Supported ^A	supported	not supported ^B
Detailed load forecasting	supported	not supported ^C	not supported ^D
Energy feedback	not supported ^E	supported ^F	not supported
Rated degree of privacy protection	medium	medium – high	very high

^A Requires de-anonymization

^B Only with advanced integrity checks

^C Depends on the degree of obfuscation

^D Only for whole network segments

^E With additional involvement of an escrow company

^F Indirectly through intermediate buffering in a battery

6. Conclusion and Future Work

In this work, we introduced and defined the overall privacy challenge in smart grids in detail. Although security concerns are receiving more and more attention, we argue that also privacy must not be neglected in this context. We surveyed existing privacy preserving approaches that aim at minimizing or obfuscating smart metering data. Future work includes studies about the adoption of widely proposed privacy preserving techniques in industry standards, their implementation, and applicability from a smart grid provider's as well as customer's perspective. We conclude that particularly the separation of technical metering data used for smoothly running the grid, and customer-specific data used to provide end-user services is a promising approach. Although the decoupling of metering data and customer ids makes the overall architecture more complicated, we expect the fast adoption of these mechanisms in the future.

References

- [1] Massoud AS and Wollenberg BF. Toward a smart grid: power delivery for the 21st century. *IEEE Power and Energy Magazine*, 2005; 3(5)34–41.
- [2] European Regulators Group for Electricity and Gas (ERGEG). (2010) ERGEG-public consultation: Position paper on smart grids no. E09-EQS-30-04. [Online]. Available: http://www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_CONSULT/CLOSED%20PUBLIC%20CONSULTATIONS/ELECTRICITY/Smart%20Grids/RR/smart%20grids_bne.pdf
- [3] Anderson RJ. *Security Engineering - A Guide to Building Dependable Distributed Systems*. 2nd ed. Wiley, 2008.
- [4] Metke AR and Ekl RL. Security technology for smart grid networks. *IEEE Transactions on Smart Grid*, 2010; 1(1):99–107.
- [5] Khurana H, Hadley M, Lu N, and Frincke DA. Smartgrid security issues. *IEEE Security & Privacy*, 2010; 8(1):81–85.
- [6] Wei D, Lu Y, Jafari M, Skare P, and Rohde K. An integrated security system of protecting smart grid against cyber attacks. In: *Proc. of Innovative Smart Grid Tech.*, Jan. 2010:1–7.
- [7] DeBlasio R and Tom C. Standards for the smart grid. In: *Proc. of IEEE Energy 2030 Conference*, 2008:1–7.
- [8] Molina-Markham A, Shenoy P, Fu K, Cecchet E, and Irwin D. Private memoirs of a smart meter. In: *Proc. of ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, 2010:61–66.
- [9] NIST. Nistir 7628: Guidelines for smart grid cyber security: vol. 2, privacy and the smart grid. Tech. Rep., 2010.
- [10] Efthymiou C and Kalogridis G. Smart grid privacy via anonymization of smart metering data. In: *Proc. of 2010 First IEEE International Conference on Smart Grid Communications*, 2010:238–243.

- [11] Kursawe K, Danezis G, and Kohlweiss M. Privacy-friendly aggregation for the smart-grid. In: *Proc. of International Conference on Privacy Enhancing Technologies*, 2011:175–191.
- [12] Li F, Luo B, and Liu P. Secure and privacy-preserving information aggregation for smart grids. *Int. J. Secur. Netw.*, 2011; 6(1):28–39.
- [13] Varodayan DP and Khisti A. Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage. In: *Proc. of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011:1932–1935.
- [14] Kalogridis G, Efthymiou C, Denic SZ, *et al.* Privacy for smart meters: Towards undetectable appliance load signatures. In: *Proc. of International Conference on Smart Grid Communications*, 2010:232–237.
- [15] Cavoukian A, Polonetsky J, and Wolf C. SmartPrivacy for the smart grid: embedding privacy into the design of electricity conservation. *Identity in the Information Society*, 2010; 3(2):275–294.
- [16] Kalogridis G and Denic SZ. Data mining and privacy of personal behaviour types in smart grid. In: *Proc. of ICDM Workshops*, 2011:636–642.
- [17] Rajagopalan SR, Sankar L, Mohajer S, and Poor HV. Smart meter privacy: A utility-privacy framework. CoRR, 2011.
- [18] Rial A and Danezis G. Privacy-preserving smart metering. In: *Proc. of ACM workshop on Privacy in the Electronic Society*, 2011:49–60.
- [19] Kalogridis G, Cepeda R, *et al.* Elecprivacy: Evaluating the privacy protection of electricity management algorithms. *IEEE Trans. Smart Grid*, 2011; 2(4):750–758.
- [20] Taban G and Gligor V. Privacy-preserving integrity-assured data aggregation in sensor networks. In: *Proc. of International Conference on Computational Science and Engineering*, 2009:168-175.