Privacy-Relevant Smart Metering Use Cases

Günther Eibl, Dominik Engel and Christian Neureiter
Josef Ressel Center for User-Centric Smart Grid Privacy, Security and Control
Salzburg University of Applied Sciences
Urstein Sued 1, A–5412 Puch/Salzburg, Austria
Email: {guenther.eibl, dominik.engel, christian.neureiter}@en-trust.at

Abstract—Privacy in Smart Metering has been discussed extensively, as have privacy enhancing technologies (PETs). However, neither of these items has been put into the perspective of the actual use cases at hand. This perspective is crucial to (i) map the correct PETs to each use case and to (ii) identify gaps, i.e., use cases which are privacy-relevant, but not yet covered by a PET. Beside the construction of such a set of privacy-relevant smart metering use cases, some open research questions have been found. Most importantly, the Smart Metering systems must be described in more detail in order to facilitate a sound development of PETs.

I. INTRODUCTION

Smart Metering, as part of the Smart Grid, is a step towards modernizing our electrical grids. However, the discussion of how to achieve optimal roll-outs of smart meter technology has been accompanied by a – sometimes ferocious – debate on privacy concerns. Numerous contributions have pointed out that the load consumption data produced by a household is privacy-sensitive data, as it allows to deduce behavioral patterns of its inhabitants (e.g., [1], [2], [3]).

Privacy enhancing technologies (PETs) have been proposed to strike a balance between the functional requirements of Smart Metering and the requirement of preserving individual privacy. This paper focuses on methods that are applied near to the customer and that aim at providing the minimum amount of information needed to external parties. An excellent overview of PETs in Smart Metering [4] shows that PETs typically focus on either one of two use cases: billing and aggregation (Table I).

<table>
<thead>
<tr>
<th>PET</th>
<th>Billing</th>
<th>Aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymization</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Cryptographic Computation</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Perturbation</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Verifiable Computation</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Trusted Computation</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

TABLE I
MATCHING OF PETs AND USE CASES [4]

The goal of the billing use case is a rather infrequent calculation of the bill for a single customer using the consumption values and the tariff as input. Verifiable computing (VC) methods are cryptographic methods that enable the computation of a function by another, untrusted party. The output consists of the result and a zero knowledge proof (ZKP) that the calculation has been done properly. In the Smart Metering case the bill could be computed at the customer’s site as a function of the consumption of the customer and the tariff provided by the energy provider [5]. Since the consumption does not leave the customer’s site, privacy is preserved. On the other hand, the energy provider can be sure that the bill has been calculated correctly.

The aggregation use case aims at a frequent calculation of consumption values that are averaged over either space, typically a neighborhood, or time. Average consumption is believed to be sufficient for use cases concerning the network operator (NO) like load monitoring and prediction. Cryptographic computation methods typically employ homomorphic encryption methods, which have the special property that the product of the encrypted load values yields, after decryption, the sum of the load values. This property is exploited for the calculation of the average value [6], [7], [8]. Cryptographic methods can be combined with perturbation methods that add a well defined amount of noise to each single measurement in a distributed way such that the sum of these noise values is just sufficient for reaching differential privacy [9], [10], [11]. Differential privacy is a guarantee that the value of a single customer cannot influence the sum too much. As a consequence the sum cannot provide information about single customers.

In the PET literature, other use cases occur as evaluation of practical properties of the presented PET solution. To give an example, it has been studied that some PETs need extensive communication between smart meters and a communication structure that is organized as a tree. The high amount of connections needed in turn affects e.g. the resistance to failure of synchronization or failures of single smart meters [4]. On the other side, there are extensive collections of Smart Metering use cases [12], [13], [14], [15] that contain much more than two use cases. This observation motivates the question: are the use cases billing and aggregation enough or do other Smart Metering use cases need to be considered?

To best of the authors’ knowledge no collection of use cases exists that is accurate enough to be used as a basis for the development of PET’s. One contribution of this paper is a first step to the composition of a consistent set of Smart Metering use cases which is formulated in a way that is suitable for the development of privacy enhancing technologies for Smart Metering. During the composition of this set of Smart Metering use cases several topics remain open, which is another contribution.

The rest of this paper is organized as follows: Section II
shows how the use case collection was derived from existing use case collections. The resulting use case collection is described in section III. Section IV contains a first attempt to estimate the privacy relevance of use cases, and thus also their relevance for PETs, based on the data items that are transferred. Finally, Section V concludes the paper.

II. Method

The approach here is to gather and combine as much information about Smart Metering use cases as possible from already existing documents [12], [13], [14], [15]. In section IV the privacy-sensitive use cases will then identified for further investigation.

A. Combination of Existing Use Cases

The resulting set of use cases should satisfy three main criteria. First, the set of use cases should be as complete as possible. This criterion is aimed to be fulfilled by combination of a number of different use case collections for Smart Metering. The use cases of the European Smart Metering Industry Group (ESMIG, [12]) is the most extensive set of use case collections and is therefore used as a starting point. Since it focuses on the business part ending at the Head End System (HES) as domain boundary, these use cases were enriched or specified in more detail by use cases from other use case collections that focus more on the customer side [13], [14].

B. Simplification of Involved Actors

As a second criterion, the use case description should be suitable for providing a quick overview over the use cases to people working on PETs. From a privacy point of view it is most important whether the user gives her data away or not. Together with the fact that this paper’s focus lies in PETs that are applied near the customer’s premise, this motivates the decision not to distinguish different parties outside the customer’s premise. Thus, only two parties, called “customer” and “service provider/utility”, respectively, occur. Note that the description of the billing and aggregation use cases for PETs each also only involve only a single service provider.

For privacy and security purposes a (trusted) third party (TTP) is typically needed. Such a third party could be an actor that distributes keys or acts as the trusted third party for trusted computation (TC) methods. Since the way a TTP comes into play and interacts with customers and the utility is resulting from the chosen privacy and security method employed and not directly from the use case, the TTP is omitted for the description of the use cases.

This rather crude simplification has two benefits. On one hand it considerably simplifies the description of the resulting use cases. On the other hand differences in architectures occurring for different use case collections vanish. This in turn enables a focus on a more concrete description of use cases. In fact, strictly speaking both the billing and the aggregation use cases are not use cases but even more specific functions that must be computed by the system. As a drawback, regarding all actors outside the household as equally trustworthy, PETs that describe how data are distributed outside the customer’s premise cannot be treated.

C. Addition of Data Items

As a third criterion, the use case description should be detailed enough to enable a privacy analysis. This criterion especially implies a description of the collected data items. Since the data that need to be transmitted are not specified in detail and instead modeled by placeholders called, e.g., MeterData, data items needed were gathered from additional sources [16] containing so-called baseline data required for the delivery of benefits for network operators.

D. Visualization and Clustering

In order to consider possible dependencies between use cases, a single use case is visualized as a rectangle. If a use case leads to the call of a second use case, an arrow is drawn from the first to the second use case leading to a graph that has to be visualized and brought into an adequate layout: The inner use cases occurring inside the household were placed in the middle and separated from the others by a big, dashed rectangle (Figure 1). Then, the outer use cases were manually regrouped around the inner use cases such that the intersection of the use case arrows is minimized. It turned out that this worked particularly well, if use cases were clustered using clusters and sub-clusters similar to the ones described in [14]. These clusters were added leftmost and rightmost, respectively (Figure 1).

E. General Changes of Use Cases

In the aforementioned use case collections, the installation process is not described in detail, thus the corresponding use cases are mainly listed here. Although the Asset-monitoring & error handling module could have been sorted to the Monitoring sub-cluster, too, it remained at the Maintenance cluster due to its different scope. Here, the focus lies in detection of failures in assets, i.e., either in failures of meters, the communication line or the other endpoint, which is called Meter Data Management (MDM) for simplicity. Customer move-in/out and Supplier Change are use cases that are likely to call other use cases.

Many of the use cases need to change the configuration of the meter, for this reason a corresponding module Meter configuration is introduced and placed within the household. Similarly, many use cases need to inform the user via a Customer information module. Since the use cases (Dis)connect energy supply and (Dis)connect devices act at the household, they are also placed inside the household area. Energy consumption behavior can only change due to the tariff, if an additional module exists inside the household that changes the consumption behavior. The new Local energy control module represents either the inhabitant of the household or an automatic Customer Energy Management System (CEMS).
F. Privacy-Related Adoptions

In the last step it was tried to make the use cases more concrete in order to incorporate privacy-relevant behavior. For this purpose, some additional rectangles or modules were introduced.

ESMIG’s original use case Obtain Meter Reading was renamed into Transfer of Meter Reading. This name better describes the original intention of this use case which is the distribution of the data to the different actors outside, including the specification which actor gets which granularity of the data. It also includes issues such as storage, re-use, deletion and correction of data and linkage to other data sets which is of special importance. Note that this use case has no connections to other use cases which can be explained by the fact that it describes topics that cannot be treated due to the simplifications of actors (Section II-B).

Three additional processes, with thicker outer boundaries emphasizing their importance for privacy, were introduced in order to explicitly describe the flow of data from inside the household to an external actor. The first is called Meter sends consumption and describes the highly privacy-relevant process where the meter sends its consumption data outside the household. In order to emphasize the flow of consumption data, the arrows point outwards the household. The meter also sends power quality events, which are generally viewed as not privacy relevant. The third rectangle is describing the sending of all other events.

In order to emphasize the necessity to perform cryptographic computations, an additional rectangle Sign and encrypt messages was introduced which offers the possibility to sign and encrypt messages.

The use case Key exchange could be seen as part of the use case Communication setup, however due to the high relevance for privacy and security it was held outside.

The use cases of the NO are only coarsely described in [12]. The original use case “network operation optimization” is seen here as a third sub-cluster because it is likely that it will contain use cases like Load prediction or Determine energy balance which are based on consumption data. Although in [12] the use case Tampering detection is described to be based on metering events only, it could be possible that also for this use case some form of energy balance based on consumption data is calculated. The same holds for Outage Detection. Therefore arrows coming from Determine energy balance are pointing to these use cases.

Since both of the two original use cases Billing and Prepayment need a tariff, the handling of the tariff was introduced as a new use case since it will likely be treated as a single Tariff module. As a second benefit, such a module shows more clearly how demand response can be
achieved by changing the tariff. The original two use cases were named Bill Calculation and Credit Module. While most of the smaller billing and prepayment sub-use cases happen outside the household, the original sub-use case Request credit is triggered from within the household and was thus modeled by a Local credit module. Here, the information about the credit is modeled to be flowing via the use cases Meter configuration and Customer information to the Local credit module.

III. DISCUSSION OF THE USE CASE DESCRIPTION

The resulting description of use cases is shown in Figure 1. The use cases are aligned along the 3 columns in the middle. The big, dashed box denotes the outer boundary of the household, the middle column represents the respective use cases or modules inside the customer’s premise. The modules corresponding to use cases running primarily outside the household are the two columns to the left and right of the dashed household boundary.

The figure gives a good impression about the high number of use cases, a PET must be compatible with. Especially the Installation and Maintenance use cases can turn out to be critical. If timestamps are used for cryptographic protocols, Clock Synchronization turns out to be a critical issue. The use case Asset Monitoring & Error-handling includes detection of failures (of meters, communication, data receiver) and remedies such as firmware updates, which, as Key Exchange is highly important for privacy due to its security relevance.

The figure immediately shows which use cases (the ones with thicker outer boundaries) lead to a flow of data items outside the household. As it is planned by ESMIG’s use cases, here the calculation of the bill is performed outside the household via the module Bill Calculation. Thus, there is a need to send the consumption data outside the household. However, using verifiable computing methods, the Bill Calculation module could be placed inside the household which immediately removes the need to send consumption data outside. Instead of the consumption data, which would in this case even need to be attributable to a household, only the computed price and the corresponding zero knowledge proof would need to be sent outside.

Note that the description of the use cases shows data items that are transmitted outside but it does not indicate a specification in which form e.g. consumption data are sent. In order to study the impact of PETs such as aggregation using homomorphic encryption the data items need to be specified in more detail.

IV. PRIVACY RELEVANCE OF DATA ITEMS AND USE CASES

In this paper use cases are considered as privacy-relevant, if privacy-relevant data of customers would be transferred outside the household if no PETs are employed. In this section, it is attempted to estimate the privacy relevance of use cases. Since there is a gap between the information needed and the information available in the use case collections, the result can only be regarded as a first estimation.

A. Privacy Relevance of Data Items

Due to its definition the privacy-relevance of use cases is based on the estimation of the privacy-relevance of data which is the topic of this sub-section. The so-called baseline data of [16] are the basis for the list of data items. Meter configuration data were extended by items that are obviously necessary for use cases like e.g. cryptographic keys.

The classification of data items with respect to their privacy relevance is not a trivial task. For example, personal behavior and current circumstances determine activities in the household which in turn can lead to the use of appliances whose summed consumption is measured by smart meters. Thus, inferring personal behavior from consumption is not simple. However, several studies suggest that active energy data are highly privacy relevant [17], [18], [19]. The information gained from consumption data highly depends on the granularity of the measurements in time. This fact is reflected by the presence of several entries for active energy in Table II. The estimation of the privacy relevance for different measurement intervals is based on results of [20].

For reactive energy the situation is not so clear because

<table>
<thead>
<tr>
<th>Data Item</th>
<th>Privacy relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active energy per household [s]</td>
<td>high</td>
</tr>
<tr>
<td>Active energy per household [day]</td>
<td>high</td>
</tr>
<tr>
<td>Active energy per household [year]</td>
<td>high</td>
</tr>
<tr>
<td>Reactive energy [s]</td>
<td>low</td>
</tr>
<tr>
<td>Generated energy</td>
<td>low</td>
</tr>
<tr>
<td>Consumption data per household [s]</td>
<td>low</td>
</tr>
<tr>
<td>Consumption per household</td>
<td>low</td>
</tr>
<tr>
<td>Tariff</td>
<td>low</td>
</tr>
<tr>
<td>Credit</td>
<td>high (?</td>
</tr>
<tr>
<td>Billing data</td>
<td>high</td>
</tr>
<tr>
<td>Mean voltage</td>
<td>no</td>
</tr>
<tr>
<td>Voltage sags and swells</td>
<td>no</td>
</tr>
<tr>
<td>Voltage alarms</td>
<td>no</td>
</tr>
<tr>
<td>Power quality data</td>
<td>no</td>
</tr>
<tr>
<td>Voltage events</td>
<td>no</td>
</tr>
<tr>
<td>Maximum demand in 30min &gt; Threshold</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Average energy &gt; Threshold</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Reactive average power &gt; Threshold</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Energy consumption returned below limit</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Meter events</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Supply disabled/restored</td>
<td>low/high (?)</td>
</tr>
<tr>
<td>Device enabled/disabled (external)</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Device enabled/disabled (home automation)</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Operating conditions</td>
<td>low/medium (?)</td>
</tr>
<tr>
<td>Contracted power/flow</td>
<td>low</td>
</tr>
<tr>
<td>Meter status</td>
<td>no</td>
</tr>
<tr>
<td>Meter access log</td>
<td>no (?)</td>
</tr>
<tr>
<td>Device information</td>
<td>low</td>
</tr>
<tr>
<td>DER parameters</td>
<td>no</td>
</tr>
<tr>
<td>Keys</td>
<td>no</td>
</tr>
<tr>
<td>Configuration parameters</td>
<td>low (?)</td>
</tr>
</tbody>
</table>

TABLE II

PRIVACY RELEVANCE OF DATA ITEMS WITHOUT PETs APPLIED. (?): FURTHER INVESTIGATIONS NEEDED
in typical non-intrusive appliance load monitoring studies reactive energy is used together with active energy but not alone. Since motors have a reactive component [21], reactive energy could be used to identify a subset of appliances such as the garage door opener or the water pump. Since the garage door opener could deliver valuable information about leaving and arrival times, privacy relevance could be considered as medium. Generated energy has low privacy impact because it mainly depends on external factors like the weather.

Data items arising from billing scenarios are clearly personal and have therefore high potential of getting sensitive. Since the current tariff is more influenced by the service provider than by the customer, it is likely to have low privacy impact. However, the current credit could be considered as highly sensitive: the fact that the credit is zero could indicate that the customer is bankrupt. Dependent on the kind of tariff and on the frequency of its calculation, the bill itself could provide indirect information about the consumption.

Power Quality Data consist of mean voltage values, voltage sags and swells or voltage alarms. Since voltage does not depend on the household, all these data items are not privacy-relevant.

The privacy-relevance of meter events is likely to be low. However, this is not really clear for all data items. Data items that are not influenced by the customer such as incoming supply failure detected are not privacy-relevant. Many events are created when a physical quantity is compared with a threshold value. A comparison of a quantity like power with many different thresholds can be viewed as a quantization. The number and values of the threshold levels determine the information contained. As an example, a single comparison of the active power with the value of 1 kW can provide information about the usage of ohmic high-power appliances which are typically used for cooking. Through the timestamps of the events, the on/off pattern can be measured which could be used to determine the appliance more precisely.

Another class of data items are operating conditions. The status “disabled” or “restored” could indicate a consumer move-in, a consumer move-out, network control operations or a lack of credit in the prepayment scenario (Figure 1). Together with a credit zero information, the information about supply disablement could get very sensitive. Enablement or disablement of devices during device control is a direct information about appliances and highly privacy-relevant. If a device is automatically controlled from outside, as it is envisioned in Figure 1, privacy relevance could be considered as low since the usage of the appliances is not triggered by concrete actions of persons living in the household. However, this is not the case, if devices are controlled from within the household (home automation). Since home automation via smart metering is unlikely to happen it is not considered here.

Finally, meter configuration data are generally likely to be of low privacy relevance since they typically do not depend on the persons living in the household.

The results of the discussion above are also summarized in Table II where the privacy relevance of the data class is set as the maximum privacy relevance over its data items. Summarizing, consumption data and billing data are highly privacy-relevant. Power quality data and configuration parameters are rather privacy-safe. Meter events are likely to be of low or medium privacy relevance. By themselves, operating conditions have rather low privacy-relevance. However, in combination with credit information they could turn out to be a useful side-information. These results should be seen as a first, preliminary assessment of privacy-relevance needing further investigations.

B. Privacy Relevance of Use Cases

In principle, the combination of the results of sections III and IV-A leads to the privacy-relevance of use cases in a straightforward manner (left column of Table III). Note that in Figure 1 no PETs like Verifiable Computing are applied.

However, remaining uncertainties input lead to uncertainties in the classification. Here, a conservative approach is taken assuming supply disablement as highly privacy relevant and assuming that the data used for the calculation of the energy balance is also used for tampering and outage detection.

Therefore, as long as the precise intended treatment is unclear, the use cases for energy net monitoring and optimization remain privacy-relevant with a question mark. The uncertainty about the use case also leads to a question mark for the PET method: if the use cases can be handled using aggregated consumption data, then homomorphic aggregation would be a PET that could be applied. For Tampering detection, in the case of fraud alarm, it could still be necessary to check single consumption profiles, too.

Since privacy and some PETs such as homomorphic encryption rely on security the corresponding use cases must also be considered in the sense of constraints on the PET method (Table IV). Furthermore, for protocols that employ timestamps [8], proper clock synchronization is a constraint.

There are three use cases that are out of focus of this paper but clearly privacy-relevant. The use case Transfer
of Meter Reading Data handles all topics arising after data have been collected. If locally controlled Home Automation data get outside, privacy is likely to be decreased. It is likely that through charging and the credit module the use cases concerning Electric Vehicles will be connected with the Smart Metering use cases considered here introducing information about the location of a person.

C. Tentative Application of PETs

Privacy impact of active energy not only depends on granularity in time but also depends on the spatial granularity, i.e. whether the data are available for each household or whether they are aggregated over households. There, the privacy relevance depends on the size of the aggregation set (Table V). In order to compute such aggregates, PETs need to be applied. If verifiable computing is used for the calculation of the bill, only the bill needs to be transferred instead of the consumption per household (Table V).

Privacy enhancing technologies are available for most of the use cases. The use case Home Automation can most easily be handled by locality which means that is likely that all tasks of this use case can be performed locally, based on parameters set from outside. Some of the use cases above have no privacy-preserving method assigned which does not mean that no methods exist. It rather seems plausible that methods from other fields such as the banking or the social network domains can be adapted.

V. CONCLUSION AND OUTLOOK

In the literature about privacy enhancing technologies for Smart Metering the two use cases billing and aggregation are considered. This paper constitutes a first step in answering the question if other use cases need to be considered, too. By combining and reorganizing use cases of existing use case collections a set of privacy relevant use cases was found. This set must be supplemented by another set of primarily security-relevant use cases.

While the results of the paper answered some questions, even more, new topics arised. There is a gap between the accuracy of the description of the use cases and the accuracy needed for the development of PETs. This holds especially for the data items that are transferred during the execution of use cases. It seems most important that this gap is bridged by the provision of more concrete systems and functions instead of use cases. An interesting topic for future research are possible privacy implications occurring for either new data items like e.g. the knowledge that the consumption is above or below a threshold or combinations of data items like e.g. the current credit together with the status of supply. A third topic for future research could be a generalization of the way the use cases were combined to general domains.

ACKNOWLEDGEMENTS

The financial support by the Austrian Federal Ministry of Economy, Family and Youth and the Austrian National Foundation for Research, Technology and Development is gratefully acknowledged.

REFERENCES