Towards Provable Privacy Guarantees Using Rechargeable Energy-Storage Devices

Fabian Laforet, Erik Buchmann and Klemens Böhm
Karlsruhe Institute of Technology (KIT)
76131 Karlsruhe, Germany
{fabian.laforet, erik.buchmann, klemens.boehm}@kit.edu

ABSTRACT

The global energy transition requires the availability of energy-consumption data with high resolution. Smart meters record such data in real time. This however endangers privacy: Time series of energy-consumption data contain different kinds of private information, such as the employment status of the residents. We address this problem by proposing a consumption-perturbation approach that relies on energy-storage devices (aka. batteries). The energy (dis-)charged to them perturbs the actual data describing the consumption. So-called charging strategies specify the (dis-)charging behavior. A main objective of this article is to come up with privacy guarantees for such strategies. To this end, the strategies we propose rely on a generalization of the Irwin-Hall distribution, which facilitates closed-form analyses. For these strategies, we derive (ϵ, δ)-differential privacy guarantees. Next, we propose a new measure, which is statistical in nature, to quantify the risk of confusing the assignment of features to the time series they are computed on. We then develop a specific charging strategy that combines the properties required to provide the guarantees proven earlier with trend preservation to shield against filtering approaches. All in all, our strategies increase the failure probability of approaches inferring private information from the data.

1. INTRODUCTION

The power-supply system is changing significantly. This includes the ongoing installation of so called smart meters. In contrast to traditional analogous meters, they measure the power consumption in much shorter time intervals, e.g., every 15 minutes. They transmit these records to a central system, e.g., a grid operator or an energy provider. In this article, a record is the amount of energy consumed in a time interval. Such data facilitates applications such as monitoring and billing: Demand-response scenarios with flexible pricing strategies incentivize the consumers to shift their demand to off-peak hours. To provide an accurate billing, they rely on time-related consumption data [25].

On the other hand, metering energy-consumption data with high frequency endangers privacy [23]. Previous work has shown that smart meter data allows inferring private information: It is not only possible to observe daily routines [24], but also the relationship status, the employment and the social class of the residents [6, 7]. In addition, disaggregation approaches are able to break down the energy consumption to individual devices [5, 33]. It is even possible to identify the TV program currently followed [13].

To overcome this antagonism between necessity and hazard of data, several proposals to ensure privacy exist. Since smart meter data allows to generate finger prints which re-identify households [8], pseudonymization, i.e., removing the identifier, is not sufficient. Instead, current approaches perturb the transmitted data itself. We refer to them as *data-oriented*. However, data perturbation is subject to legal constraints. For instance, Article 13 of the directive 2006/32/EC of the European Parliament states that smart meters "accurately reflect the final customer's actual energy consumption" to enable correct billing.

We now sketch an alternative to data-oriented perturbation, namely *consumption perturbation*. To this end, observe that rechargeable energy-storage devices (aka. batteries) are proliferating at a rapid pace. They are common in households with photovoltaics installed [21] and will become even more popular with the rise of plug-in electric vehicles with batteries [22]. Their control is subject to *charging strategies*. A charging strategy specifies the amount of energy that is (dis-)charged to the energy-storage device in a certain situation. An example strategy that prevents peak loads during the day discharges the device if the consumption is larger than a given threshold, e.g., 1 kW, and charges during the night. The recorded data transmitted by the smart meter now consists of the actual consumption of the household and the (dis-)charged energy for each time interval, see Figure 1. On the one hand, the transmitted data reflects the amount of energy bought and is in line with current legislation. On the other hand, the actual consumption is perturbed by the energy (dis-)charged. Thus, rechargeable energy-storage devices are promising when it comes to protect privacy.

In our work, we analyze which privacy guarantees an energy-storage device can give, as a function of the charging strategy and the energy-storage device. A difficulty is that possible guarantees also depend on the underlying infrastructure, e.g., which data is transmitted to the central system.

To derive privacy guarantees that rely on closed-form anal-

e-Energy'16, June 21-24, 2016, Waterloo, ON, Canada

 ISBN 978-1-4503-4393-0/16/06…$15.00

DOI: http://dx.doi.org/10.1145/2934328.2934335

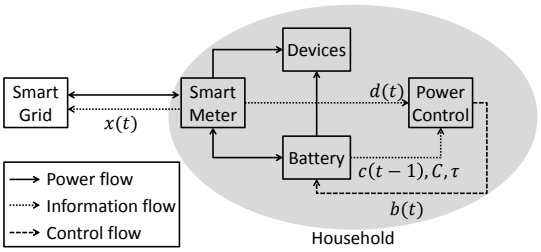


Figure 1: The system model

yses, we propose a generalization of the Irwin-Hall distribution (generalized Irwin-Hall distribution, *GIH*). The strategies we focus on are ones where the (dis-)charging rates follow this distribution. To provide more detail, we now turn to the infrastructures studied in related work.

Scenario (Part 1). *Each household has a rechargeable energy-storage device. Within each time interval, a certain amount of energy is (dis-)charged to this device. From a privacy perspective, the amount of energy (dis-)charged perturbs the actual consumption for this interval. We assume that all households use the same charging strategy. This will help to establish privacy guarantees, as we will see later.*

Literature [2, 5, 7, 8, 10] distinguishes between the following scenarios regarding the nature of the data transmitted:

Scenario (Part 2, Case 1). *Households are grouped into clusters, e.g., urban quarters. After perturbing an actual record, smart meters in a cluster communicate with each other, aggregate their consumption and send this information to the central system [2]. Here, the leaked data is the energy a group of households has consumed.*

Scenario (Part 2, Case 2). *After the (dis-)charging has perturbed an actual record, the smart meter sends the data to the central system, without communicating with other households. Here, the leaked data is the energy each household has consumed individually.*

In Case 1, the central system can analyze the aggregated data, without knowing the consumption of an individual household. However, that data can still leak privacy, e.g., if one knows some of the consumptions aggregated, if they are similar to each other, or if the aggregate enables negative disclosure. A well-established measure to quantify the uncertainty whether a record is part of the aggregation achieved by perturbing the data is differential privacy [10]. We show that there does not exist a charging strategy that achieves the hard requirements of ϵ-differential privacy. On the other hand, we prove that (ϵ, δ)-differential privacy [11] is feasible for strategies following our GIH distribution.

In Case 2, since the connection of the data to its household is explicit, it is obvious that private information such as daily routines or the employment status can be inferred from the data. The amount of inferable information depends on the consumption perturbation achieved by the charging strategy. Many approaches inferring private information do not rely on single records only, but compute features from sequences of records [6, 7, 8]. To our knowledge, a measure allowing to quantify the privacy a charging strategy provides in Case 2 does not exist. We propose a new measure that is statistical in nature, namely (σ, m)-confusability. It computes the risk of assigning privacy-relevant features to individuals correctly. To quantify (σ, m)-confusability, we propose a numerical solution scheme. It takes the particular distribution the (dis-)charging rates follow as well as the behavior of the energy-storage device into account.

Our next contribution is to propose a specific charging strategy, as follows. Certain charging strategies from the literature preserve characteristics of the data such as trends over several intervals. While they do this in a best-effort manner, i.e., they provide no provable guarantees and therefore are not related to the privacy measures discussed in this publication, this is still useful in practice. Our new charging strategy combines both the formal guarantees we have derived earlier, to deal with (ϵ, δ)-differential privacy and (σ, m)-confusability, as well as such best-effort characteristics. We evaluate our charging strategy on a large real energy-consumption dataset. Our evaluation shows that the strategy protects against well-known privacy-relevant problems such as prediction of private information and re-identification.

2. RELATED WORK

2.1 Differential Privacy for Smart Meters

Differential privacy [10] is an important and established notion for provable privacy guarantees. There exist various proposals on how to implement it for the smart grid, using secure aggregation approaches [12, 28, 29]. For our work, we assume that such an aggregation approach exists which allows smart meters to aggregate their records without the need for a trusted third party. Note that our approach only uses a rechargeable energy-storage device to perturb the consumption, and we leave aside altering the data itself during the aggregation process. [2] is a purely data-oriented approach that proposes to add noise following a Gamma distribution to each record. The resulting aggregation follows a Laplace distribution and therefore fulfills differential privacy. However, the deployment of energy-storage devices cannot result in exactly this perturbation. This is because they have a bounded capacity which restricts the range of the noise that one can add, while the Gamma distribution is unbounded.

2.2 Information Hiding on Time Series

Previous work on smart meter data has shown that inferring sensitive information is possible [5, 6, 7, 8, 13, 23, 24, 33]. To prohibit such inferences, many approaches to perturb time-series data have been proposed: Earlier ones add noise resistant to reconstruction attacks. This distorts the original records, but does not give any guarantees comparable to differential privacy [26]. More sophisticated approaches consider the mutual information between the original and the perturbed time series to quantify privacy [27]. However, all these approaches do not give any guarantee that sensitive information is hidden. Other work enables individuals to define privacy relevant patterns that must not appear in the perturbed data [19]. Other work allows to define secrets that can be hidden, following the definition of differential privacy [18]. This is achieved by computing Laplacian noise that prevents to decide whether a secret property occurs or not. Again, since Laplacian noise is unbounded, no energy-storage device with limited capacity can implement this. In the following, we identify a distribution that energy-storage devices can mimic, and that provides provable guarantees.

2.3 Privacy via Energy-Storage Devices

The use of energy-storage devices to protect privacy has become popular in recent years. The first approaches feature relatively simple charging strategies, with the goal to preserve a constant consumption level [16, 17]. Extensions of this idea reduce the standard deviation between the perturbed data and the overall average consumption [31]. However, these approaches do not give any privacy guarantee. Whenever the capacity of the energy-storage device is exhausted, or the difference between the consumption of two consecutive records is larger than the maximum (dis-)charging rate, sensitive information is leaked. In addition, if the consumption is about to be constant over several time intervals, e.g., when no residents are at home, the strategies do not alter the time series. To arrive at a better understanding on how energy-storage devices can hide sensitive information, current research also has studied how to prevent identifying the consumption of single devices. First approaches feature so called "power mixing algorithms" [15]. Such algorithms try to smoothen the consumption of each individual device by allocating a certain amount of the energy-storage capacity to each of them. Further approaches extend these ideas by considering measures such as mutual information to compute the perturbation achieved [9, 30]. In addition, there exist approaches that implement differential privacy to give provable guarantees regarding the uncertainty whether a single device is switched on or off [4, 32]. Instead of considering a closed system where the charging strategy can determine the optimal power level for an observable cumulative consumption of several devices, our approach considers scenarios where households perturb their actual consumption independently of each other, and a perturbation on the aggregation level is not possible.

3. FUNDAMENTALS

In this work, we aim for charging strategies that provide provable privacy guarantees. We now describe how the system is modeled. We also propose a probability distribution the (dis-)charging rates of the charging strategies adhere to. This is the distribution that we will analyze subsequently.

3.1 System Model

We now describe the components of the system model in Figure 1. An energy-storage device has a bounded capacity C, i.e., the load level must be in $[0, C]$. $c(t)$ denotes the load level at the end of time interval t. Note that t stands for a time interval and not for a point of time. This is because the data transmitted by smart meters is recorded for a period of time. In an interval t, the energy-storage device is (dis-)charged by $b(t) = c(t) - c(t-1)$. If $b(t) < 0$ the energy-storage device is discharged, if $b(t) > 0$ it is charged. The maximum (dis-)charging rate τ restricts the amount of energy (dis-)chargeable in a time interval, i.e., $|b(t)| \leq \tau$.

A Power Control component, one per household, regulates the amount of energy (dis-)charged. It controls the energy-storage device by determining $b(t)$ following a given charging strategy. Our explanations focus on the case where it uses information on the current consumption of the devices $d(t)$ and the load level $c(t-1)$ as well as properties C and τ. In practice, the Power Control component obtains the consumption information during a time interval continuously and immediately controls the energy storage based on this information. This is because the energy-storage device cannot be (dis-)charged with $b(t)$ within the last margin of the time interval, right before the consumption is transmitted.

The resulting consumption of a household $x(t)$ is composed of the consumption of the devices $d(t)$ and the (dis-)charged energy $b(t)$. In the following, we call $d(t)$ the *actual consumption* and $x(t)$ the *perturbed consumption*. In Case 1, all households within a cluster I aggregate their perturbed consumptions and transmit $x_I(t) = \sum_{i \in I} x_i(t)$ to the central system. In Case 2, each household sends $x(t)$ directly to the central system.

3.2 Requirement

We now present the requirement that the consumption perturbation achieved by a randomized charging strategy follows a certain distribution. This is because such charging strategies provide provable privacy guarantees, as we will show. We propose the (dis-)charged energy to follow a specific distribution. However, note that the general idea behind our analyses is also applicable to different distributions. We generalize the Irwin-Hall distribution [14] that adds k i.i.d. random variables that are uniformly distributed on $[0, 1]$. Thus, the distribution can have the properties of a Uniform distribution for $k = 1$, a Triangular distribution for $k = 2$ or an approximately Gaussian distribution for large k. Our generalized Irwin-Hall (GIH) distribution in turn is obtained by adding random variables that are uniformly distributed on the interval $[-\frac{a}{k}, \frac{a}{k}]$. In consequence, it is defined on the interval $[-a, a]$ where a can be chosen arbitrarily out of $\mathbb{R}^+$.

Definition 1 (GIH distribution)**.** *The GIH probability density function (pdf) for $b(t) \in [-a, a]$ with $a \in \mathbb{R}^+$ and $k \in \mathbb{N}$ is*

$$f_{k,a}(b(t)) = \frac{1}{(k-1)!} \sum_{i=0}^{\lfloor \frac{b(t)+a}{2a} \cdot k \rfloor} (-1)^i \binom{k}{i} \left(\frac{b(t)+a}{2a} \cdot k - i \right)^{k-1}$$

and its cumulative distribution function (cdf) is

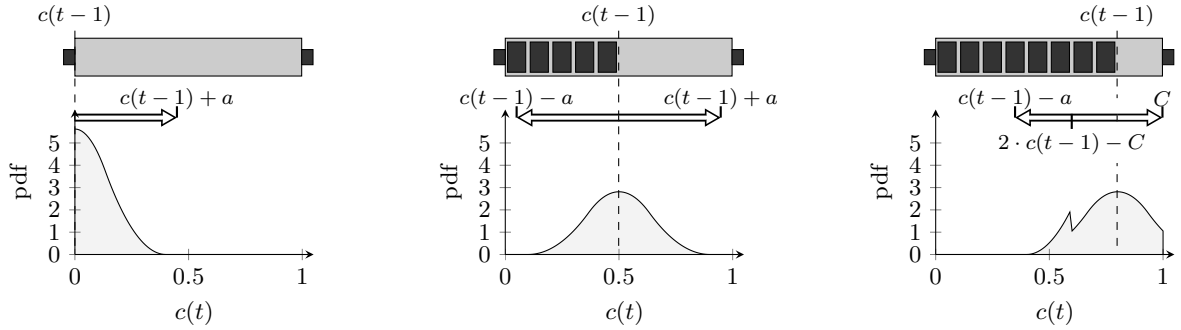
$$F_{k,a}(b(t)) = \frac{1}{k!} \sum_{i=0}^{\lfloor \frac{b(t)+a}{2a} \cdot k \rfloor} (-1)^i \binom{k}{i} \left(\frac{b(t)+a}{2a} \cdot k - i \right)^{k}$$

The maximum (dis-)charging rate τ restricts the range $[-a, a]$ of this distribution, i.e., $a \leq \tau$. Recall that energy-storage devices have limited capacity. In consequence, they cannot implement random values that are drawn independently of each other. In particular, the energy-storage device cannot be discharged under a load level of zero. The same holds for the reverse case if the load level is close to C. In what follows, we extend Definition 1 to a conditional pdf w.r.t. the load level of the previous time interval.

Lemma 1. *The function*

$$f_{k,a}(b(t)|c(t-1)) = \begin{cases} 0 & \text{if } c(t-1) + b(t) < 0 \\ & \text{or } c(t-1) + b(t) > C, \\ 2 \cdot f_{k,a}(b(t)) & \text{if } c(t-1) - a < 0 \\ & \text{and } c(t-1) + b(t) \geq 2 \cdot c(t-1), \\ 2 \cdot f_{k,a}(b(t)) & \text{if } c(t-1) + a > C \\ & \text{and } c(t-1) + b(t) \leq 2 \cdot c(t-1) - C, \\ f_{k,a}(b(t)) & \text{otherwise.} \end{cases}$$

is a conditional pdf given the load level $c(t-1)$ that fits the GIH distribution to the capacity bounds $[0, C]$ of the energy-storage device.



(a) $c(t-1) = 0$ (b) $c(t-1) = 0.5$ (c) $c(t-1) = 0.8$

Figure 2: Visualization of the load level change applying Lemma 1 ($\mathbf{C = 1}$ kWh, $\mathbf{a = 0.4}$ kWh, $\mathbf{k = 3}$)

The appendix of this work contains proofs of all lemmas.

To give an external observer the impression that the (dis-) charging rates follow a GIH distribution, the inequality $2a \leq C$ must be fulfilled. Figure 2 visualizes the load-level change between two time intervals applying Lemma 1. Note that the adaption of the GIH distribution is symmetric for the two cases "$c(t-1)$ is too low" and "$c(t-1)$ is too large", i.e., an observer who does not have any information on the current load level will not recognize the adaption.

Note that the standard deviation of the GIH distribution is $\sqrt{(a^2/3k)}$. Thus, the GIH distribution allows to determine the average amount of energy (dis-)charged to the energy-storage device. This gives way to the computation of other criteria [17] such as life time or energy loss of the storage device without performing additional experiments.

Having a charging strategy that provides a consumption perturbation relying on a GIH distribution, Section 4 shows that it guarantees (ϵ, δ)-differential privacy in Case 1. For Case 2, we present a way to quantify the confusion risk of different individual households in Section 5. In Section 6, we present a charging strategy that fulfills the requirements discussed here. In addition, it features best-effort characteristics such as trend preservation that prohibits removing the perturbation by filtering approaches in Case 2.

4. ACHIEVING DIFFERENTIAL PRIVACY

In this section, we analyze Case 1, i.e., all smart meters apply an aggregation approach and send the sum of all consumptions within the time interval to the central system. We first show that ϵ-differential privacy [10] cannot be fulfilled by any charging strategy. We then focus on (ϵ, δ)-differential privacy [11]. This allows to quantify achievable privacy guarantees for charging strategies whose (dis-)charging rate follows a GIH distribution.

4.1 Insights on Differential Privacy

Differential privacy gives provable privacy guarantees on the result of a statistical query, i.e., a query whose result is perturbed by a random variable. The perturbation ensures that the influence of each individual record on the query result is limited with stochastic guarantees. Thus, it is unable to distinguish whether an individual has published his private information or not. Differential privacy achieves this by adding noise to the query result. This guarantees that the probability of each result S differentiates by a maximum factor of e^{ϵ} between the cases where an individual object is part of the database and where it is not.

Definition 2 (ϵ-differential privacy [10])**.**
A query q is ϵ-differentially private if for all data sets DB_1 and DB_2 where DB_1 and DB_2 differ by at most one element, and for all subsets of possible answers $S \subseteq Range(q)$:

$$P(q(DB_1) \in S) \leq e^{\epsilon} \cdot P(q(DB_2) \in S)$$

In our scenario, each individual uses an energy-storage device that hides his actual consumption. The consumption perturbation addable is restricted by characteristics of the device. We now show that ϵ-differential privacy cannot be achieved in this scenario.

Theorem 1. *Let n be the number of individual households. They use a charging strategy that adds noise in the range $[-a, a]$ to the individual consumption. Then there does not exist any charging strategy that facilitates ϵ-differential privacy for the sum of the consumptions over all households.*

To give meaningful guarantees on the privacy achieved by charging strategies nevertheless, we apply a relaxed version of ϵ-differential privacy that features a further parameter δ. Here, $\delta > 0$ is an upper bound for the probability that the requirement of ϵ-differential privacy is not fulfilled. Thus, this parameter allows to give guarantees by excluding rare cases that occur on the tails of the distribution.

Definition 3 ((ϵ, δ)-differential privacy [11])**.**
A query q is (ϵ, δ)-differentially private if for all data sets DB_1 and DB_2 where DB_1 and DB_2 differ by at most one element, and for all subsets of possible answers $S \subseteq Range(q)$:

$$P(q(DB_1) \in S) \leq e^{\epsilon} \cdot P(q(DB_2) \in S) + \delta$$

The parameter δ facilitates that the fraction $P(q(DB_1) \in S_{\text{strict}})/P(q(DB_2) \in S_{\text{strict}})$ does not have to be smaller than e^{ϵ} for results S_{strict} that are unlikely (or even impossible) for $q(DB_2)$. Meaningful values of ϵ and δ should be chosen in combination: If δ is large, i.e., there exist many cases where the requirement of ϵ-differential privacy may be violated, ϵ can have small values, i.e., there exist strong privacy guarantees for the remaining cases. In addition, ϵ and δ depend on the global sensitivity Δq. The global sensitivity of a query is the maximum possible difference of two results based on DB_1 and DB_2 which differ in at most one individual record.

Definition 4 (Global sensitivity Δq [10])**.**
The global sensitivity Δq of a query q for all data sets DB_1 and DB_2 which differ in at most one element is

$$\Delta q = \max_{DB_1, DB_2} |q(DB_1) - q(DB_2)|$$

In our scenario where the central system receives the consumption aggregate of several households within one time interval, Δq is the maximum possible consumption of a single household within one time interval.

4.2 Differential Privacy by GIH Distributions

We now show that charging strategies whose (dis-)charging rate follows a GIH distribution give provable (ϵ, δ)-differential privacy guarantees.

In a first step, we derive the distribution that occurs if n households add their i.i.d. perturbed consumption. In general, such a distribution is achieved by convoluting the underlying pdf n times. The convolution $f(x)$ of two pdfs $f_1(x)$ and $f_2(x)$ is given by

$$f(x) = \int_{-\infty}^{\infty} f_1(y) \cdot f_2(x-y)\,\mathrm{d}y$$

Since the analytic computation of the convolutions of n arbitrary distributions is difficult to impossible, we propose to apply a GIH distribution. This is because the GIH distribution is defined by the sum of n i.i.d. random variables. I.e., Definition 1 provides all the information necessary to describe the distribution that results from adding the perturbed consumption of n households.

Lemma 2. *The pdf of the distribution resulting from n households summing up their consumptions individually perturbed by a GIH distribution for $b(t) \in [-a \cdot n, a \cdot n]$ with parameters $a \in \mathbb{R}^+$ and $k \in \mathbb{N}$ is*

$$f_{k,a,n}(b(t)) = \frac{1}{(kn-1)!} \sum_{i=0}^{\lfloor \frac{b(t)+an}{2a} k \rfloor} (-1)^i \binom{kn}{i} \left(\frac{b(t)+an}{2a}k - i\right)^{kn-1}$$

To explain this formula, we adapt the pdf of Definition 1 by increasing the number of uniformly distributed random variables from k to $k \cdot n$ and by expanding the interval bounds from $[-a, a]$ to $[-a \cdot n, a \cdot n]$.

To quantify the extent of (ϵ, δ)-differential privacy, we compare the case where n households sum up their consumptions, and the n-th household has an impact of Δq on the query result to the case where the n-th household is left aside.

Theorem 2. *Let q be the query for the sum of consumptions in time interval t over several households that perturb their data individually by a GIH distribution with parameters k and a. Suppose that the requester does not know whether n or $n-1$ households are part of the data set. The result of q is (ϵ, δ)-differentially private with*

$$\epsilon = \max\left(\ln\left(\frac{f_{k,a,n-1}(\mathit{left})}{f_{k,a,n}(\mathit{left}-\Delta q)}\right), \ln\left(\frac{f_{k,a,n}(\mathit{right}-\Delta q)}{f_{k,a,n-1}(\mathit{right})}\right)\right)$$

$$\delta = \max\left(F_{k,a,n-1}(\mathit{left}), 1 - F_{k,a,n}(\mathit{right}-\Delta q)\right)$$

where

$$\mathit{left} = \Delta q - a \cdot n + x \cdot \frac{n}{2n-1} \cdot (a \cdot (2n-1) - \Delta q)$$

$$\mathit{right} = a \cdot (n-1) - x \cdot \frac{n-1}{2n-1} \cdot (a \cdot (2n-1) - \Delta q)$$

and x is any value in $(0, 1]$.

For a given instantiation of k, a, n and Δq, there exist different combinations of ϵ and δ. We can adapt their values with the parameter $x \in (0, 1]$ which defines the positions 'left' and 'right'. The smaller x becomes, the larger becomes ϵ, and the smaller becomes δ. We give some trade-off examples in Section 7.

5. ACHIEVING CONFUSABILITY

For Case 2 where all households send their individual consumption data directly to the central system, we see the following privacy risks: Approaches inferring private information use background information such as the employment status of some households to make predictions for other households for which such information is not available [6, 7]. One might think that it would be sufficient to remove the identifier of the households. However, re-identification approaches can restore the assignment in many cases [8]. Thus, one must rely on the consumption perturbation achieved by the energy-storage device to increase the failure probability of that kind of attacks. Since the respective state-of-the-art approaches work on features computed out of the time-series data (such as the consumption between 6 and 10 p.m., to give an example), we propose a new privacy measure that quantifies the probability that two households are confused with each other based on the features applied. We now present the privacy measure and then say how it is computed.

5.1 Fundamental Privacy Guarantee

In a first step, we quantify the confusability based on features of the consumption data. We consider a query q that computes a feature of a time series, such as the minimum consumption or the consumption over several given time intervals. We now assume that it is possible to deduce certain information from the query result. We illustrate this assumption with the following examples:

Example 1. *The query $q(X_i)$ returns the consumption of Household i at 10 p.m. Now think of a setting with two households. On weekdays, no residents of Household 1 are at home at 10 p.m. Thus,*

$$q(X_1) = c_1(10\ p.m.) = 0.5kWh$$

consist of stand-by consumptions only. At this time, residents of Household 2 are at home in turn and watch television. This results in

$$q(X_2) = c_2(10\ p.m.) = 1.0kWh.$$

By considering the consumption at 10 p.m., it is possible to deduce whether a time series belongs to Household 1 or 2. In addition, by considering the background information that employees have a low consumption of $0.5kWh$ on weekdays at 10 p.m. while unemployed persons have a high consumption of $1.0kWh$, it is possible to deduce the employment status of the residents.

Next, we consider time series whose actual consumption $d(t)$ is perturbed with $b(t)$ for each time interval t, where $b(t)$ follows a GIH distribution, as explained in Section 3.2. This perturbation is random in nature. In consequence, there is a chance to assign a time series to the wrong household or to predict the wrong employment status.

Example 2. *Consider the scenario described in Example 1. Suppose that both households apply a uniformly distributed consumption perturbation over the interval $[-1kWh, 1kWh]$ for each time interval t, i.e., $b(t)$ follows a GIH distribution with $k = 1$ and $a = 1kWh$. The result of query q for X_1 is uniformly drawn from $[-0.5kWh, 1.5kWh]$ and the one for X_2 from $[0.0kWh, 2.0kWh]$. Thus, if a consumption in $[0.0kWh, 1.5kWh]$ is generated, it is impossible to decide whether this consumption belongs to Household 1 or 2, or whether the residents are employed or unemployed.*

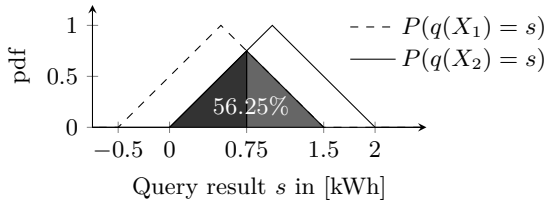


Figure 3: Exemplary visualization of the confusability of a single record with k = 2 and a = 1 kWh

Our goal here is to quantify the probability that a query result is confusable with the query result of another household. We define the confusability of a query w.r.t. two time series as follows:

Definition 5 (Confusability risk σ)**.**
The risk that two time series X_1 and X_2 are confusable w.r.t. a query q over all possible query results $s \in S$ is

$$\sigma = \int_{s \in S} \min\{P\left(q(X_1) = s\right), P\left(q(X_2) = s\right)\} \, \mathrm{d}s$$

Confusability quantifies the probability to be unable to distinguish whether a query result belongs to one of the time series X_1 or X_2. I.e., by considering the result to make the assignment, the accuracy is as good as with a probability of σ if there is no query result, i.e., the assignment is random. Transferring this concept to statistical hypothesis testing, a confusability risk of 100% means that the type 1 error, i.e., the probability to assign a query result to the wrong household ("false positive") is equal to the probability to assign the result correctly. Let us illustrate this insight with the following example:

Example 3. *We continue Example 2. For any result s in the interval $[0.0kWh, 1.5kWh]$, the pdf is 0.5 for both time series X_1 and X_2. Outside of this interval, each result has a pdf of zero for (at most) one time series. Thus, the probability that both time series are confusable is $1.5 \cdot 0.5 = 75\%$. If the households instead do a consumption perturbation following a GIH distribution with $k = 2$, they achieve a confusability of 56.25%, see Figure 3. In this example, if the query result s is in $[0\ kWh, 0.75\ kWh]$ the probability to assign the data correctly is higher for Household 1 than for Household 2. However, with a probability of 28.125%, this decision is as good as if one had tossed a coin. This is illustrated by the left darker part of the colored area in Figure 3. The same holds for $s \in [0.75\ kWh, 1.5\ kWh]$ for choosing Household 2, as illustrated with the brighter area.*

Since we want to give guarantees for data sets with more than two elements, we compute the confusability between pairs of data objects that are labeled with different private information. This is because our goal is to give guarantees that approaches which predict such information fail. In the following, we call a value $\Theta(X_i) \in \Theta$ that labels a data object X_i *prediction target*. For instance, the employment status splits the data set into objects that are labeled with 'employed' and ones labeled with 'unemployed', i.e., $\Theta := \{\text{'employed'}, \text{'unemployed'}\}$. In the re-identification scenario, the number of different prediction targets is the number n of data objects, i.e., $\Theta := \{1, \ldots, n\}$. Thus, we can compute probabilities that an object is confusable with at least a certain number of objects m having another prediction target.

Definition 6 ((σ, m)-Confusability)**.**
A query q is (σ, m)-confusable if each data object X_i has a probability of σ to be confused with at least m data objects belonging to a prediction target $\Theta(X_j)$ different from $\Theta(X_i)$:

$$\forall X_i \in DB : m \leq \left| \left\{ X_j \in DB \,\middle|\, \Theta(X_i) \neq \Theta(X_j) \wedge \int_{s \in S} \min\{P\left(q(X_i) = s\right), P\left(q(X_j) = s\right)\} \, \mathrm{d}s \geq \sigma \right\} \right|$$

In contrast to Example 1, most queries used for state-of-the-art approaches inferring private information do not refer to individual records only. In combination with constraints given by the energy-storage device, such as conditional load levels as presented in Lemma 1, it is difficult to impossible to derive an analytic solution of the confusability risk. Thus, we propose a numerical solution to compute (σ, m)-confusability in the following.

5.2 Computing (σ, m)-Confusability

To compute (σ, m)-confusability in our scenario, we have to consider the limitations of the energy-storage device.

Example 4. *Think of a query which computes the overall consumption over a sequence of time intervals, e.g., between 6 and 10 p.m. If one assumed that for each time interval the (dis-)charging rate was determined independently of the previous (dis-)charging rates, he would add independent random values to specify the resulting propabilities $P(q(X) = s)$. According to the Bienaymé formula [20], the variances of these random variables sum up. In consequence, there would exist a number of random variables where the resulting standard deviation of $P(q(X) = s)$ would be larger than the capacity of the energy-storage device. However, the capacity bounds possible deviations: The deviation is maximal if the energy-storage device is empty (or full) at 6 p.m. and full (or empty) at 10 p.m. Thus, we cannot assume that the random variables are independent of each other.*

Summing up Example 4, the sum of the perturbations of subsequent records cannot be larger than the capacity of the energy-storage device. More precisely, the perturbation achievable depends on the load level of the first time interval t: If the energy storage is empty at t, the maximal perturbation is the capacity C, if it is half-filled, the maximal perturbation is $0.5 \cdot C$. Thus, in a first step, we identify the density function of the probability that the energy-storage device has a certain load level. The result is independent of the current time interval t. I.e., we assume that if one has no background information on the load levels of previous time intervals, he will always observe the same load-level distribution. This effect occurs if the charging strategy is in use long enough. If the probabilities between two time intervals do not alter, we call the load-level distribution *stable*.

Definition 7 (Stable load level pdf)**.**
Let an energy-storage device with a load level $c(t) \in [0, C]$ at any time interval t and its (dis-)charging rates $b(t)$ that follow a conditional pdf $f(b(t)|c(t-1))$ be given. $g(c(t))$ is the load-level pdf that describes the density to observe a load level $c(t)$ in time interval t. A load level is stable if the following holds for any $c(t) \in [0, C]$:

$$\begin{aligned} g(c(t)) &= \int_0^C f(c(t) - c(t-1)|c(t-1)) \cdot g(c(t-1)) \, \mathrm{d}c(t-1) \\ &= \int_0^C f(c(t+1) - c(t)|c(t)) \cdot g(c(t)) \, \mathrm{d}c(t) = g(c(t+1)) \end{aligned}$$

Roughly speaking, the pdf is stable if the probability that a load level $c(t)$ is observed at t is identical to the probability at $t+1$ for all load levels. Note that $c(t) - c(t-1) = b(t)$. In consequence, the conditional pdf $f(b(t)|c(t-1))$ that describes the likelihood to (dis-)charge the energy-storage device by $b(t)$ when observing a load level $c(t-1)$ is $f(c(t) - c(t-1)|c(t-1))$. In our work, we consider (dis-)charging rates that follow a GIH distribution and apply the conditional pdf described in Lemma 1.

However, we are not aware of any approach for a closed-form computation of g. Thus, we propose to apply a numeric approach that divides the load-level range into small intervals. The density of those intervals is updated by a repeated calculation, as described in Definition 7, until the changes are marginal. We describe a respective algorithm in Appendix E.

We visualize results for an energy-storage device with capacity 1 kWh and a GIH distribution with $a = 0.25$kWh and $k \in [1, 25]$ in Figure 4. We observe that for instantiations of k with a low value, i.e., (dis-)charging rates with a high standard deviation, load levels around $0.5 \cdot C$ are most probable. For increasing values of k, the pdf flattens around $0.5 \cdot C$ and drops down at the capacity bounds 0 and C. Our results cover the insights presented in [3].

Having the pdf of the stable load level at hand, we can now compute the confusability for queries that take subsequent records into account: For each time interval t the query refers to, we can compute the probability to observe a (dis-)charging rate w.r.t. the load-level pdf of the previous time interval $t-1$. The first time interval t_0 referred to by the query applies the stable load-level pdf g to determine the initial load-level probabilities. The (dis-)charging rate of later time intervals t then depends on the (dis-)charging rates of previous time intervals t' with $t > t' \geq t_0$. With this procedure, the capacity bounds $[0, C]$ of the energy-storage device cannot be violated. This is because the load level at time interval t is composed of the initial load level and the (dis-)charging rates of previous time intervals. In consequence, the probability to observe a (dis-)charging rate resulting in a load level smaller than 0 or larger than C is zero. Thus, the scenario in Example 4 where the standard deviation would be larger than C cannot occur.

To show how to compute the pdf of a query which refers to several time intervals, we take the consumption over a sequence of intervals as example.

Lemma 3. *The pdf $P(q(X) = s)$ of the result s of the query q "What is the consumption during the period $[1, T]$?" is*

$$\int_{-C}^{C} \cdots \int_{-C}^{C} \int_{0}^{C} f_{k,a}\left(s - \sum_{t=1}^{T} d(t) - \sum_{t=1}^{T-1} b_t \,\middle|\, c + \sum_{i=1}^{T-1} b_i\right) \cdot \prod_{t=1}^{T-1} f_{k,a}\left(b_{T-t} - \sum_{i=1}^{T-t-1} b_i \,\middle|\, c + \sum_{i=1}^{t-1} b_i\right) \cdot g(c)\,\mathrm{d}c\,\mathrm{d}b_1 \ldots \mathrm{d}b_{T-1}$$

Since we do not have a closed-form computed solution of g, an analytic computation of this query-result pdf is impossible. Thus, we rely on a numerical solution similar: We divide the conditional load-level pdf f into many small intervals. For each interval, we iteratively compute the pdfs over all time intervals $t \in [1, T]$.

Section 7 presents experimental results for (σ, m)-confusability for different features over several time intervals. The computations of the pdfs of those features follow the explanations in this section. I.e., they all consider the stable load-level pdf to describe the initial load-level probabilities and then iteratively compute the conditional (dis-)charging-rate pdfs over the intervals referred to by the queries.

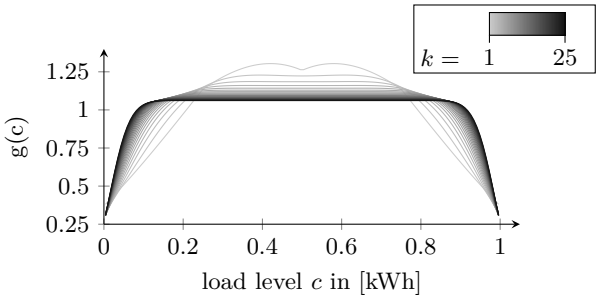


Figure 4: Visualization of the stable load level pdf ($a = 0.25$ kWh and $C = 1$ kWh)

6. CHARGING STRATEGY

We now present a charging strategy that fulfills the requirement that the (dis-)charging rates follow a GIH distribution described in Section 3.2. In addition, we identify the following privacy-relevant requirements a charging strategy should fulfill:

1. [26] has shown that, in Case 2, adding consecutive random values independently of each other to a time series results in a perturbation that can easily be filtered. To avoid this problem, a charging strategy should choose the (dis-)charging rates in a way that the information on the actual consumption is hidden reliably.
2. State-of-the-art charging strategies aim at hiding changes in the time series [9, 15, 16, 17, 30, 31]. Their experiments show that smoothening the time series can help to increase the failure probability of approaches inferring private information. Thus, a charging strategy should smoothen the perturbed consumptions.

To fulfill these requirements, we add a trend preservation in a best-effort manner to our charging strategy. Here, trend preservation means that our charging strategy aims at preserving the current linear trend that results from the perturbed records of the last two time intervals. Such a trend preservation helps to perturb the actual consumption reliably: As long as the differences between consecutive records remain the same, a filtering approach is unable to identify wavelet coefficients that result from perturbing the time series [26]. In consequence, the filtering approach cannot reconstruct the actual consumption values. In addition, such trends smoothen the time series, in line with the second requirement. Note that the trend preservation is carried out in a best-effort manner, i.e., it might fail in the same way as it does for other charging strategies: The trend must be adapted if the capacity bounds are reached, or if the resulting (dis-)charging distribution does not correspond to the GIH distribution. This idea of testing whether the distribution is preserved can also be combined with objectives featured by related charging strategies, both privacy-centered ones such as minimizing the distance to the average consumption and economic ones such as minimizing operating costs [31]. In case the Power Control component works continuously, we propose to apply a priority mechanism such as the one proposed in [25].

Algorithm 1 GIH Charging

Private: Array[#bins] devCount
Trend $\alpha \cdot t + \beta$
Input: $c(t-1), d(t), C, k, a, \gamma$
Output: $b(t)$

1: **if** $t = 0$ **then**
2: $b(t) = r$, where r is generated from $f_{k,a}(b(t)|c(t-1))$
3: $\beta = b(t) + d(t)$
4: **else if** $t = 1$ **then**
5: $b(t) = r$, where r is generated from $f_{k,a}(b(t)|c(t-1))$
6: $\alpha = b(t) + d(t) - \beta$
7: **else**
8: $b(t) = \alpha \cdot t + \beta - d(t)$
9: **if not**$(0 \leq c(t-1) + b(t) \leq C$ **and** $-a \leq b(t) \leq a$ **and** devCount$[\lfloor F_{k,a}(b(t)) \cdot \#\text{bins} \rfloor] \leq \frac{t}{\#\text{bins}} \cdot (1+\gamma))$ **then**
10: Determine for each bin that is inside the capacity bounds $[\max(-c(t-1), -a), \min(C - c(t-1), a)]$ and whose count devCount[bin] $\leq \frac{t}{\#\text{bins}} \cdot (1+\gamma)$ a weight $\frac{t}{\#\text{bins}} \cdot (1+\gamma) -$ devCount[bin]
11: Draw random bin B according to the weights
12: Draw uniform distributed value $u \in [\frac{B}{\#\text{bins}}, \frac{B+1}{\#\text{bins}}]$
13: $b(t) = r$, with $F_{k,a}(r) = u$
14: $\alpha = b(t) + d(t) - (\alpha \cdot (t-1) + \beta)$
15: $\beta = b(t) + d(t) - \alpha \cdot t$
16: devCount$[\lfloor F_{k,a}(b(t)) \cdot \#\text{bins} \rfloor]$++
17: **return** $b(t)$

Algorithm. Algorithm 1 is the charging strategy. As an input, it receives the load level $c(t-1)$ of the end of the last interval, the individual consumption $d(t)$ of the current interval, the capacity C of the energy-storage device, the parameters k and a of the GIH distribution and an accuracy factor γ. γ is used to test whether a perturbation that follows the current trend would violate the requirement that the perturbation follows a GIH distribution. If γ is small, the GIH distribution is observable for each time interval, but the trend preservation must be adapted frequently. The output is the amount of energy $b(t)$ that is (dis-)charged in the current time interval.

The (dis-)charging rates of the first two time intervals are drawn independently from each other and create the initial linear trend (Lines 1-6). For each subsequent time interval, the algorithm tries to preserve the trend (Line 8). This fails in one of the following cases (Line 9):

- The (dis-)charging rate in combination with the previous load level would violate the capacity bounds $[0, C]$.
- The (dis-)charging rate would violate the range of the GIH distribution $[-a, a]$.
- The (dis-)charging rate would result in an overall distribution deviating from the desired GIH distribution by at most γ.

To identify whether the GIH distribution is violated, we propose to divide the interval $[-a, a]$ into bins and count the number of values that appear for each bin in the array devCount (Line 16). We apply an equal-frequency partitioning. Thus, we expect the same number of data-object occurrences ($t/\#$bins) for each bin after each time interval t. If a (dis-)charging rate violates this expectation by a factor of at most γ, we reject this rate and search for another one. We do so by weighing the remaining possible bins by a weight that prefers those that have occurred rarely in the past (Line 10). This is necessary to ensure that the (dis-)charging rates always follow the GIH distribution. We draw a bin B where bins with a higher weight have a higher chance to get chosen (Line 11). We now determine a $b(t)$ by an inverse transform sampling: We determine a uniformly distributed random value u from the interval $[B/\text{bins},(B+1)/\text{bins}]$ (Line 12). The (dis-)charging rate $b(t)$ where the GIH cdf $F_{k,a}(b(t)) = u$ is the result of the inverse transform sampling. Finally, we update the trend w.r.t. $b(t)$ (Lines 14-15).

Note that the focus of this charging strategy is on preserving the GIH distribution. This is because we do not follow the current trend if the observable distribution would not follow the GIH distribution any longer. In Section 7, we show that this charging strategy outperforms existing approaches by orders of magnitude w.r.t. its ability to increase the failure probability of approaches inferring private information from energy-consumption data.

7. EXPERIMENTS

We now quantify the provable privacy guarantees presented in this work on real-world data and use it to evaluate our GIH charging strategy. More specifically, we use all 2526 time series of the CER dataset [1] that are recorded from 14th September 2009 to 31st December 2010 and that are labeled with private information, as explained in [6]. Each time series has been metered with a sampling rate of 30 minutes. This amounts to about 65 million records.

We now quantify guarantees of (dis-)charging rates that follow a GIH distribution for (ϵ, δ)-differential privacy and (σ, m)-confusability. The results are valid for GIH charging and show that energy-storage devices can provide privacy guarantees. In what follows, we evaluate our charging strategy on the CER dataset by quantifying the failure probability of privacy-relevant approaches in Case 2. We expect that assigning query results to wrong households, as quantified by (σ, m)-confusability, lets GIH charging yield an increased failure probability, compared to related best-effort approaches. The use cases examined here where privacy is at risk are as follows: We consider the two orthogonal scenarios where the data is used to identify its owner and where the data is used to predict private information about its owner. We use the approaches presented in [8] and [7] as representatives of these use cases. In Appendix F, we show that the trend preservation protects against filtering approaches which try to reconstruct the actual records. In addition to such attacks, other reconstruction methods such as identifying the (dis-)charging patterns by applying a disaggregation approach [5, 33] are possible. Thus, the standard deviation of the GIH distribution must fulfill the guarantees presented in [32] to prevent inferences on the different states of the original energy consumption.

In addition to such attacks, other reconstruction methods such as identifying the (dis-)charging patterns by applying a disaggregation approach [5, 33] are possible. Thus, the standard deviation of the GIH distribution must fulfill the guarantees presented in [32] to prevent inferences on the different states of the original energy consumption.

7.1 (ϵ, δ)-differential privacy

Guarantees obtainable with (ϵ, δ)-differential privacy depend on the number of households n that aggregate their data and the sensitivity Δq, i.e., the maximum possible consumption of a household in one time interval. The tradeoff between ϵ and δ is adjustable by the parameter x, as explained in Theorem 2. Table 1 contains results for house-

n	x	$\Delta q = 1$ kWh ϵ	$\Delta q = 1$ kWh δ	$\Delta q = 2$ kWh ϵ	$\Delta q = 2$ kWh δ	$\Delta q = 4$ kWh ϵ	$\Delta q = 4$ kWh δ	$\Delta q = 8$ kWh ϵ	$\Delta q = 8$ kWh δ
100	0.7	1.01	$1.38\times E^{-5}$	1.90	$2.01\times E^{-5}$	3.63	$4.60\times E^{-5}$	6.99	$2.04\times E^{-4}$
100	0.9	0.31	$8.86\times E^{-2}$	0.60	$1.02\times E^{-1}$	1.18	$1.33\times E^{-1}$	2.31	$2.14\times E^{-1}$
500	0.9	0.31	$5.28\times E^{-13}$	0.61	$9.37\times E^{-13}$	1.20	$2.90\times E^{-12}$	2.39	$2.57\times E^{-11}$
500	0.95	0.15	$2.11\times E^{-4}$	0.30	$2.80\times E^{-4}$	0.60	$4.87\times E^{-4}$	1.19	$1.37\times E^{-3}$
1000	0.95	0.15	$4.88\times E^{-13}$	0.30	$8.44\times E^{-13}$	0.60	$2.48\times E^{-12}$	1.20	$2.01\times E^{-11}$
1000	0.97	0.09	$1.06\times E^{-5}$	0.18	$1.47\times E^{-5}$	0.35	$2.80\times E^{-5}$	0.72	$9.54\times E^{-5}$

Table 1: Exemplary parameters for (ϵ, δ)-differential privacy ($\mathbf{k = 1}$ and $\mathbf{a = 1}$ kWh)

Query	m	Θ = Household ID $C = 1$ $\tau = 0.25$	$C = 2$ $\tau = 0.5$	$C = 4$ $\tau = 1$	$C = 8$ $\tau = 2$	Θ = Employment $C = 1$ $\tau = 0.25$	$C = 2$ $\tau = 0.5$	$C = 4$ $\tau = 1$	$C = 8$ $\tau = 2$
Consumption in (6 p.m. - 10 p.m.)	1	57.6%	70.0%	80.9%	88.6%	31.9%	46.5%	62.9%	76.4%
	3	23.8%	39.5%	57.6%	75.0%	12.0%	21.3%	37.4%	57.3%
Weekend Consumption	1	20.7%	28.5%	36.1%	42.7%	4.3%	8.7%	15.9%	24.4%
	3	2.0%	7.7%	18.1%	28.3%	0.0%	0.1%	2.0%	6.5%
Maximum Consuption	1	54.3%	58.8%	65.9%	71.8%	46.8%	50.6%	57.9%	66.2%
	3	39.8%	44.2%	53.1%	62.9%	15.6%	29.4%	37.6%	49.5%
First time consumption > 1 kWh	1	71.1%	71.0%	77.1%	91.8%	60.6%	61.3%	64.9%	85.1%
	3	62.9%	63.8%	73.6%	88.1%	42.6%	43.1%	49.5%	79.4%

Table 2: Results for $(\sigma, \mathbf{m})$-confusability ($\mathbf{k = 1}, \mathbf{a = \tau}$)

holds that apply a charging strategy where the (dis-)charging rates follow a GIH distribution with $a = 1kWh$ and $k = 1$. Obviously, a larger number of households facilitates stronger privacy guarantees, i.e., smaller values for ϵ and δ. An increasing sensitivity in turn lowers the guarantees. An increasing value of x results in lower ϵ and increased δ. As illustrated in Table 1, there is a high degree of privacy in many cases where the probabilities may differ from each other by a factor of 2 ($\epsilon = 0.69$), and this requirement is violated (significantly) less frequently than $\delta = 1\%$.

7.2 (σ, m)-confusability

We evaluate (σ, m)-confusability for four queries, see first column in Table 2. The query results represent typical features that are applied for privacy-relevant approaches such as [6, 8]. We present σ values for $m = 1$ and $m = 3$, i.e., each query result must be σ-confusable with the ones of at least 1 and 3 households with different prediction targets Θ. A first prediction target is the label of its household, given by its ID. The second target is $\Theta :=$ {'employed', 'unemployed'}. In addition, we quantify the effect of the energy-storage device on the guarantees. We consider four energy-storage devices with increasing capacity C and increasing maximum (dis-) charging rates τ. The (dis-)charging rates follow a GIH distribution with $k = 1$ and $a = \tau$.

Obviously, for an increasing m, i.e., number of households each household must be confusable with, the values for σ decrease. If the prediction target partitions the households by their IDs, we obtain higher confusability compared to the case with the employment status. This is because in the first case, each household is compared to all other households of the dataset. In the second case, each household is compared to fewer households of a different employment status, i.e., if a time series is labeled with 'employed', confusability is calculated only for time series that are labeled with 'unemployed'. In consequence, the confusabilities for Θ = Employment can be at most as good as the one for Θ = Household ID.

For the first query that returns the consumption between 6 and 10 p.m., we obtain high values for σ that increase with the size of the energy-storage device. In contrast, the confusabilities for the second query that calculates the consumption during the weekend are lower. This is because the number of time intervals that are important to the query is much larger for the second query. Therefore, the differences in the consumption between the households are larger. Since the consumption perturbation achievable with an energy-storage device is capped by its capacity, the confusability achievable for queries over a long time obviously decreases. However, if one aggregates over many time intervals, he loses information on individual behavior patterns. Therefore, such queries endanger privacy less than queries over few time intervals, as illustrated in [23].

The third query regarding the maximum consumption has high confusabilities again. Although the query is computed on a large number of time intervals, only those records that have high values affect the pdf of the result. Thus, the result can be perturbed reliably, and we obtain high values for σ.

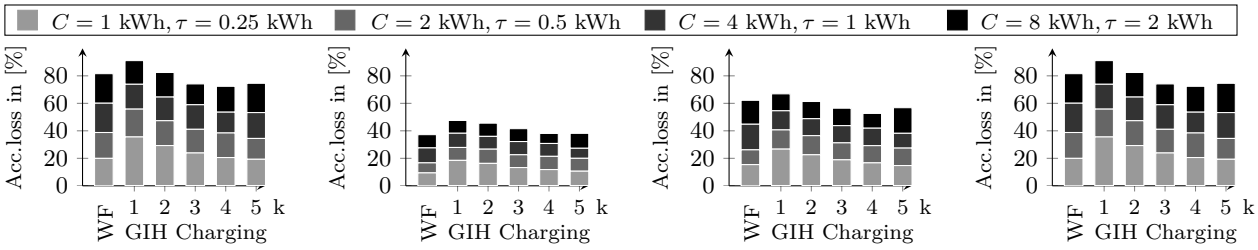
The fourth query that returns the first time interval where the consumption is larger than 1 kWh has high confusabilities. Somewhat surprisingly, we observe that for an increasing energy-storage-device size from $C = 1kWh$ to $2kWh$ for $m = 1$ and Θ = Household ID the confusabilities do not increase, i.e., drop from 71.1% to 71.0%. Let us illustrate the reason for this effect with an example:

Example 5. *Think of two time series X_1 and X_2 with two records each. X_1 consists of the values $[0.5, 1.5]$, X_2 of $[0.75, 1.5]$. For a GIH distribution with $a = 0.25$ and $k = 1$, a value larger than 1.0 occurs the first time with a likelihood of 100% for the second record for both time series. A bigger energy-storage device that allows the implementation of a GIH distribution with $a = 0.5$ now lets the first record of X_2 have values in $[0.25, 1.25]$. In consequence, there is a chance of 25% to observe a value larger than 1.0 for the first record and 75% for the second record for X_2, while the probabilities for X_1 remain the same. Thus, the confusability drops from 100% to 75% after enlarging the energy-storage device.*

We conclude that even for small energy-storage devices, (dis-)charging rates following a GIH distribution facilitate high (σ, m)-confusability with $\sigma > 50\%$ for queries that endanger privacy.

7.3 Re-Identification

A privacy-relevant problem is that smart-meter data allows the assignment to its owner. This is critical in cases where the actual consumption is recorded, but without an



(a) employment (b) devices (c) social class (d) retirement

Figure 6: Experimental results on inferring private information

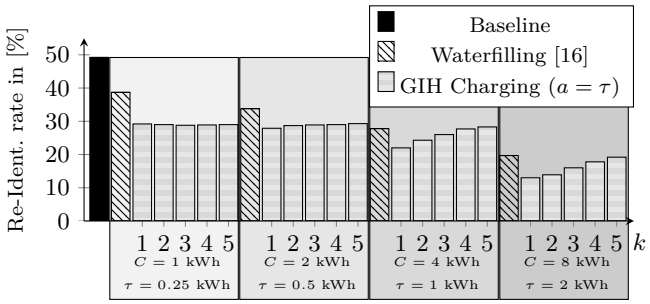


Figure 5: Experimental results on Re-Identification

identifier. Approaches trying to assign perturbed consumption data to its owner are called re-identification approaches. We now study the likelihood of success of the re-identification approach described in [8] for our GIH charging strategy and the competitor Waterfilling [16]. It aims at preserving a constant consumption level but does not give any provable privacy guarantee. We expect that perturbing the features the re-identification approach relies on increases the failure probability of the assignment. Figure 5 compares the accuracy of the assignment for energy-charging devices of different sizes, as visualized by the colored boxes. We randomly divide the data set into groups of about 100 households and average the results over all groups. The first bar of Figure 5 shows the re-identification rate if no energy-storage device is present. The hatched bars show the results of Waterfilling [16], and the remaining bars show the results of the GIH charging strategy with different values of k. As expected, the re-identification rate decreases with an increasing size of the energy-storage device. For bigger energy-storage devices in particular, the re-identification rate increases with the value of k for GIH charging. This is because the standard deviation of the consumption perturbation decreases with an increasing k. However, our GIH charging strategy outperforms Waterfilling for each size.

7.4 Classification

Approaches such as [7] use features of consumption data that is labeled with private information to train a classifier. The classifier then makes predictions for consumption data where the private information is not available. We expect that increasing the confusability of households belonging to different prediction targets decreases the accuracy of such predictions. We use the following measure that quantifies the fraction of accuracy lost:

$$\text{accuracyloss} = \frac{\text{acc}_{\text{actual}} - \text{acc}_{\text{charging}}}{\text{acc}_{\text{actual}} - \text{acc}_{\text{baseline}}}$$

where $\text{acc}_{\text{actual}}$ is the prediction accuracy of the actual consumption, $\text{acc}_{\text{charging}}$ is the one of the consumption perturbed by a charging strategy, and $\text{acc}_{\text{baseline}}$ is the fraction of the most frequent label. Namely, a classifier that is unable to learn any prediction rule will use the most frequent label for each household as prediction. From a privacy perspective, a charging strategy that results in the same accuracy as the actual consumption has the worst accuracy loss of 0%. A strategy in turn which does not let the classifier learn any rule has the best accuracy loss of 100%. Figure 6 shows our experimental results for the approach described in [7] for private information on the employment status, the number of devices, the social class and the retirement status. We compare Waterfilling [16] with GIH charging with $k = 1, \ldots, 5$ for four energy-storage devices with increasing size. We observe that the accuracy loss increases with larger sizes for all charging strategies. However, the accuracy loss differs between the prediction targets. We conclude that the loss depends on the features that are important for the prediction. For instance, the employment status might be predictable by considering whether the consumption in the morning is high, i.e., whether the residents are at home. For this feature, the confusability is quite high, cf. Section 7.2. On the other hand, the consumption over a long period of time might be useful to predict the number of devices. It is difficult to make such long-term features confusable, as shown for the weekend feature in Section 7.2. Comparing the charging strategies, GIH charging outperforms Waterfilling [16] for all predictions and all sizes for small values of k. We conclude that the ability of GIH charging to guarantee privacy reduces prediction accuracy, as expected.

8. CONCLUSIONS

In this work we have addressed the important problem of giving privacy guarantees for smart meter data with energy-storage devices. Related work studies two main cases, one where the data is aggregated over several households, the other one where the data of each household can be analyzed individually. We cover both cases, the first one with (ϵ, δ)-differential privacy, the other one by introducing (σ, m)-confusability. To quantify the guarantees achievable, we have focused on (dis-)charging rates that follow a generalization of the Irwin-Hall distribution. An energy-storage device with limited capacity can implement such charging strategies. Furthermore, we have proposed a specific charging strategy that combines this distribution with trend preservation in a best effort manner. Our experiments show that our strategy can give good privacy guarantees and outperforms conventional charging strategies in protecting against approaches that infer private information from the data.

Acknowledgement

We thank Rudolf Biczok for implementing the classifier to predict private information [7].

9. REFERENCES

[1] Customer behaviour trials findings report (CER11/080a). Technical report, Commission for Energy Regulation (CER), 2011.

[2] G. Ács and C. Castelluccia. I Have a DREAM! (DiffeRentially privatE smArt Metering). In *Information Hiding - 13th International Conference (IH)*, 2011.

[3] T. Antal and S. Redner. Escape of a Uniform Random Walk from an Interval. *Journal of Statistical Physics*, 123(6), 2006.

[4] M. Backes and S. Meiser. Differentially Private Smart Metering with Battery Recharging. In *Data Privacy Management and Autonomous Spontaneous Security*, volume 8247 of *Lecture Notes in Computer Science*. 2014.

[5] N. Batra, H. Dutta, and A. Singh. INDiC: Improved Non-intrusive Load Monitoring Using Load Division and Calibration. In *12th International Conference on Machine Learning and Applications (ICMLA)*, 2013.

[6] C. Beckel, L. Sadamori, and S. Santini. Towards Automatic Classification of Private Households Using Electricity Consumption Data. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '12, 2012.

[7] C. Beckel, L. Sadamori, and S. Santini. Automatic socio-economic classification of households using electricity consumption data. In *The Fourth International Conference on Future Energy Systems (e-Energy)*, 2013.

[8] E. Buchmann, K. Böhm, T. Burghardt, and S. Kessler. Re-identification of Smart Meter data. *Personal and Ubiquitous Computing*, 17(4), 2013.

[9] Z. Chen and L. Wu. Residential Appliance DR Energy Management With Electric Privacy Protection by Online Stochastic Optimization. *IEEE Transactions on Smart Grid*, 4(4), 2013.

[10] C. Dwork. Differential Privacy. In *Automata, Languages and Programming, 33rd International Colloquium (ICALP)*, 2006.

[11] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Advances in Cryptology (EUROCRYPT)*, volume 4004 of *Lecture Notes in Computer Science*. 2006.

[12] Z. Erkin, J. Troncoso-Pastoriza, R. Lagendijk, and F. Perez-Gonzalez. Privacy-preserving data aggregation in smart metering systems: an overview. *Signal Processing Magazine, IEEE*, 30(2), 2013.

[13] U. Greveler, P. Glösekötterz, B. Justusy, and D. Loehr. Multimedia content identification through smart meter power usage profiles. In *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*, 2012.

[14] P. Hall. The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika*, 1927.

[15] G. Kalogridis, R. Cepeda, S. Denic, T. Lewis, and C. Efthymiou. ElecPrivacy: Evaluating the Privacy Protection of Electricity Management Algorithms. *IEEE Transactions on Smart Grid*, 2(4), 2011.

[16] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda. Privacy for smart meters: Towards undetectable appliance load signatures. In *2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010.

[17] G. Kalogridis, Z. Fan, and S. Basutkar. Affordable Privacy for Home Smart Meters. In *Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops (ISPAW)*, 2011.

[18] S. Kessler, E. Buchmann, and K. Böhm. Deploying and Evaluating Pufferfish Privacy for Smart Meter Data. In *International Conference on Ubiquitous Intelligence and Computing (UIC 2015)*, 2015.

[19] F. Laforet, E. Buchmann, and K. Böhm. Individual privacy constraints on time-series data. *Information Systems*, 54, 2015.

[20] M. Loève. Probability Theory. *Graduate Texts in Mathematics*, 45(4), 1977.

[21] E. Martinot, C. Dienst, L. Weiliang, and C. Qimin. Renewable Energy Futures: Targets, Scenarios, and Pathways. *Annual Review of Environment and Resources*, 32(1), 2007.

[22] A. Masoum, S. Deilami, P. Moses, and A. Abu-Siada. Impacts of battery charging rates of Plug-in Electric Vehicle on smart grid distribution systems. In *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, 2010.

[23] E. McKenna, I. Richardson, and M. Thomson. Smart meter data: Balancing consumer privacy concerns with legitimate applications. *Energy Policy*, 41, 2012.

[24] A. Molina-Markham, P. J. Shenoy, K. Fu, E. Cecchet, and D. E. Irwin. Private memoirs of a smart meter. In *BuildSys'10, Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, 2010.

[25] P. Palensky and D. Dietrich. Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads. *IEEE Trans. Industrial Informatics*, 7(3), 2011.

[26] S. Papadimitriou, F. Li, G. Kollios, and P. S. Yu. Time Series Compressibility and Privacy. In *Proceedings of the 33rd International Conference on Very Large Data Bases (VLDB)*, 2007.

[27] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor. Smart meter privacy: A utility-privacy framework. In *IEEE Second International Conference on Smart Grid Communications (SmartGridComm)*, 2011.

[28] V. Rastogi and S. Nath. Differentially Private Aggregation of Distributed Time-series with Transformation and Encryption. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data (SIGMOD)*, 2010.

[29] C. Rottondi, G. Verticale, and C. Krauss. Distributed Privacy-Preserving Aggregation of Metering Data in Smart Grids. *IEEE Journal on Selected Areas in Communications*, 31(7), 2013.

[30] O. Tan, D. Gunduz, and H. Poor. Increasing Smart Meter Privacy Through Energy Harvesting and Storage Devices. *IEEE Journal on Selected Areas in Communications*, 31(7), 2013.

[31] L. Yang, X. Chen, J. Zhang, and H. Poor. Cost-effective and privacy-preserving energy management for smart meters. *IEEE Transactions on Smart Grid*, 6(1), 2015.

[32] J. Zhao, T. Jung, Y. Wang, and X. Li. Achieving differential privacy of data disclosure in the smart grid. In *INFOCOM, 2014 Proceedings IEEE*, 2014.

[33] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12), 2012.

APPENDIX

A. PROOF OF LEMMA 1

Lemma 1. The function

$$f_{k,a}(b(t)|c(t-1)) = \begin{cases} 0 & \text{if } c(t-1)+b(t)<0 \\ & \text{or } c(t-1)+b(t)>C, \\ 2\cdot f_{k,a}(b(t)) & \text{if } c(t-1)-a<0 \\ & \text{and } c(t-1)+b(t)\geq 2\cdot c(t-1), \\ 2\cdot f_{k,a}(b(t)) & \text{if } c(t-1)+a>C \\ & \text{and } c(t-1)+b(t)\leq 2\cdot c(t-1)-C, \\ f_{k,a}(b(t)) & \text{otherwise.} \end{cases}$$

is a conditional pdf given the load level $c(t-1)$ that fits the GIH distribution to the capacity bounds $[0, C]$ of the energy-storage device.

Proof. The conditional pdf removes the part of the GIH pdf that exceeds the capacity bounds $[0, C]$, mirrors it on the mean and adds it to $f_{k,a}(b(t))$. This means that the resulting function is a feasible pdf. This is because its integral over $[0, C]$ keeps having value 1, and it is greater than or equal to zero at each point. □

B. PROOF OF THEOREM 1

Theorem 1. Let n be the number of individual households. They use a charging strategy that adds noise in the range $[-a, a]$ to the individual consumption. Then there does not exist any charging strategy that facilitates ϵ-differential privacy for the sum of the consumptions over all households.

Proof. Differential privacy compares the two cases where n and $n-1$ households publish their aggregated perturbed consumption data D_n and D_{n-1}. Thus, the result of a query in these cases has the following ranges:

$$q(DB_n) \in [D_n - n\cdot a, D_n + n\cdot a]$$
$$q(DB_{n-1}) \in [D_{n-1} - (n-1)\cdot a, D_{n-1} + (n-1)\cdot a]$$

Independently of the values of D_n and D_{n-1}, the complement $\text{Range}(q(DB_n)) \setminus \text{Range}(q(DB_{n-1}))$ is never empty. This is because the interval length $|\text{Range}(q(DB_n))| = 2\cdot n\cdot a > 2\cdot(n-1)\cdot a = |\text{Range}(q(DB_{n-1}))|$. In consequence, there exist query results S_{distinct} that have a probability greater than zero for $q(DB_n)$ and that are equal to zero for $q(DB_{n-1})$. In these cases, the fraction

$$\frac{P(q(DB_n)\in S_{\text{distinct}})}{P(q(DB_{n-1})\in S_{\text{distinct}})}$$

is undefined, and no ϵ can be found that guarantees differential privacy. □

C. PROOF OF THEOREM 2

Theorem 2. Let q be the query for the sum of consumptions in time interval t over several households that perturb their data individually by a GIH distribution with parameters k and a. Suppose that the requester does not know whether n or $n-1$ households are part of the data set. The result of q is (ϵ, δ)-differentially private with

$$\epsilon = \max\left(\ln\left(\frac{f_{k,a,n-1}(\text{left})}{f_{k,a,n}(\text{left}-\Delta q)}\right), \ln\left(\frac{f_{k,a,n}(\text{right}-\Delta q)}{f_{k,a,n-1}(\text{right})}\right)\right)$$
$$\delta = \max\left(F_{k,a,n-1}(\text{left}), 1 - F_{k,a,n}(\text{right}-\Delta q)\right)$$

where

$$\text{left} = \Delta q - a\cdot n + x\cdot\frac{n}{2n-1}\cdot(a\cdot(2n-1)-\Delta q)$$
$$\text{right} = a\cdot(n-1) - x\cdot\frac{n-1}{2n-1}\cdot(a\cdot(2n-1)-\Delta q)$$

and x is any value in $(0, 1]$.

Proof. For our proof, we assume $a < \Delta q$, i.e., the energy-storage device cannot (dis-)charge more energy than the household with the most energy-intensive devices consumes within one time interval. This results in two pdfs $f_{k,a,n-1}(b(t))$ for n and $f_{k,a,n}(b(t)-\Delta q)$ for $(n-1)$ households that intersect each other (at most) one time, as illustrated in Figue 7. The proof for $a > \Delta q$ where they intersect two times then is straightforward and is omitted here.

In a first step, we divide the set of possible answers $\text{Range}(q)$ into three subsets LEFT, MIDDLE and RIGHT:

$$\text{LEFT} = [-a\cdot(n-1), left]$$
$$\text{MIDDLE} = [left, right]$$
$$\text{RIGHT} = [right, \Delta q + a\cdot n]$$

The lower bound $-a\cdot(n-1)$ of LEFT is the lowest value where $f_{k,a,n-1}(b(t))$ can achieve a pdf greater zero, the upper bound $\Delta q + a\cdot n$ of RIGHT is the largest value where $f_{k,a,n}(b(t)-\Delta q)$ can achieve a pdf greater zero. We initialize the inner bounds 'left' between LEFT and MIDDLE and 'right' between MIDDLE and RIGHT as follows, applying a variable $x \in (0, 1]$:

$$\text{left} = \Delta q - a\cdot n + x\cdot\frac{n}{2n-1}\cdot(a\cdot(2n-1)-\Delta q)$$
$$\text{right} = a\cdot(n-1) - x\cdot\frac{n-1}{2n-1}\cdot(a\cdot(2n-1)-\Delta q)$$

'left' and 'right' are mirror-inverted w.r.t. $f_{k,a,n-1}(b(t))$ and $f_{k,a,n}(b(t)-\Delta q)$, i.e., 'left' divides $f_{k,a,n-1}(b(t))$ into two parts that have the same relative range sizes as the two parts that arise from dividing $f_{k,a,n}(b(t)-\Delta q)$ by 'right'. We now assume that

- $\frac{f_{k,a,n-1}(b(t))}{f_{k,a,n}(b(t)-\Delta q)}$ receives the smallest value within LEFT for $b(t) = \text{left}$, i.e., the term is increasing for $b(t) < \text{left}$ and decreasing for $b(t) > \text{left}$ and that
- $\frac{f_{k,a,n}(b(t)-\Delta q)}{f_{k,a,n-1}(b(t))}$ receives the smallest value within RIGHT for $b(t) = \text{right}$, i.e., the term is increasing for $b(t) > \text{right}$ and decreasing for $b(t) < \text{right}$.

We now analyze the two cases where the left side of the inequation of Definition 3 describes the dataset containing $(n-1)$ households and where it contains n households.

First case: We start with the information need for $q(DB_1)$ stated in Definition 3. We do so by calculating the sum of $(n-1)$ households, i.e.,

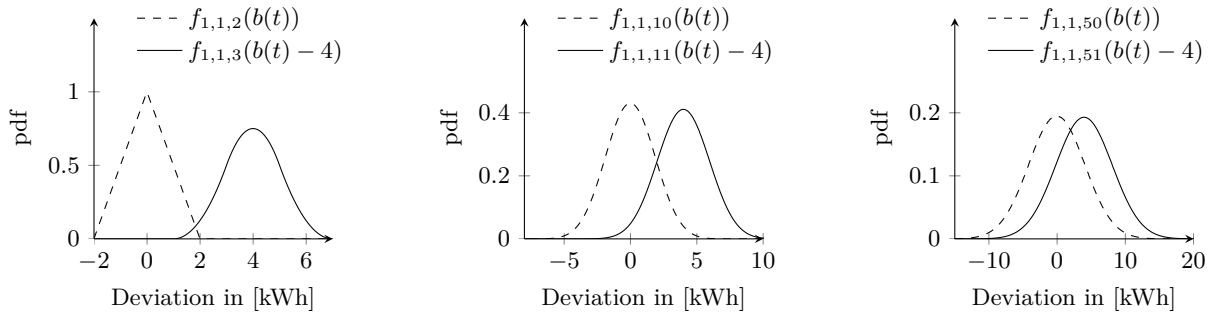
$$P(q(DB_{n-1})\in S) \leq e^{\epsilon}\cdot P(q(DB_n)\in S) + \delta$$

We now divide S into the two sets LEFT and $\neg$LEFT := MIDDLE $\cup$ RIGHT:

$$P(q(DB_{n-1})\in \text{LEFT}) + P(q(DB_{n-1})\in \neg\text{LEFT}) \leq$$
$$e^{\epsilon}\cdot(P(q(DB_n)\in \text{LEFT}) + P(q(DB_n)\in \neg\text{LEFT})) + \delta$$

We now decrease the right side by $e^{\epsilon}\cdot P(q(DB_n)\in \text{LEFT})$:

$$P(q(DB_{n-1})\in \text{LEFT}) + P(q(DB_{n-1})\in \neg\text{LEFT}) \leq$$
$$e^{\epsilon}\cdot P(q(DB_n)\in \neg\text{LEFT}) + \delta$$



(a) $n = 3$ (b) $n = 11$ (c) $n = 51$

Figure 7: Visualization of cumulative pdf plots with k = 1, a = 1 kWh and Δq = 4 kWh

By requiring that

$$\delta \geq P(q(DB_{n-1}) \in \text{LEFT}) = F_{k,a,n-1}(\text{left})$$

we can reduce the inequation to

$$P(q(DB_{n-1}) \in \neg\text{LEFT}) \leq e^{\epsilon} \cdot P(q(DB_n) \in \neg\text{LEFT})$$
$$\frac{P(q(DB_{n-1}) \in \neg\text{LEFT})}{P(q(DB_n) \in \neg\text{LEFT})} \leq e^{\epsilon}$$
$$\frac{f_{k,a,n-1}(b(t))}{f_{k,a,n}(b(t)-\Delta q)} \leq e^{\epsilon}, \forall b(t) \in \neg\text{LEFT}$$

According to our previous assumption that $\frac{f_{k,a,n-1}(b(t))}{f_{k,a,n}(b(t)-\Delta q)}$ is decreasing with $b(t)$, we must take $b(t) =$ left. This is because 'left' is the lower bound of ¬LEFT. Summarizing, in our first case, we obtain

$$\epsilon = \ln\left(\frac{f_{k,a,n-1}(\text{left})}{f_{k,a,n}(\text{left}-\Delta q)}\right), \delta = F_{k,a,n-1}(\text{left}).$$

Second case: The procedure is symmetric to the first case, i.e., we start with

$$P(q(DB_n) \in S) \leq e^{\epsilon} \cdot P(q(DB_{n-1}) \in S) + \delta$$

We divide S into RIGHT and ¬RIGHT := LEFT ∪ MIDDLE:

$$P(q(DB_n) \in \text{RIGHT}) + P(q(DB_n) \in \neg\text{RIGHT}) \leq$$
$$e^{\epsilon} \cdot (P(q(DB_{n-1}) \in \text{RIGHT}) + P(q(DB_{n-1}) \in \neg\text{RIGHT})) + \delta$$

We decrease the right side by $e^{\epsilon} \cdot P(q(DB_{n-1}) \in \text{RIGHT})$ and require that

$$\delta \geq P(q(DB_n) \in \text{RIGHT}) = 1 - F_{k,a,n}(\text{right})$$

Thus, we can reduce our inequation to

$$\frac{P(q(DB_n) \in \neg\text{RIGHT})}{P(q(DB_{n-1}) \in \neg\text{RIGHT})} \leq e^{\epsilon}$$
$$\frac{f_{k,a,n}(b(t)-\Delta q)}{f_{k,a,n-1}(b(t))} \leq e^{\epsilon}, \forall b(t) \in \neg\text{RIGHT}$$

We now take $b(t) =$ right where the fraction receives the largest value in ¬RIGHT and obtain

$$\epsilon = \ln\left(\frac{f_{k,a,n}(\text{right}-\Delta q)}{f_{k,a,n-1}(\text{right})}\right), \delta = 1 - F_{k,a,n}(\text{right}).$$

Result Merging: We now achieve two (possible) slightly different values for ϵ and δ in the first and the second case. To give provable privacy guarantees, we have to take their maximum values to fulfill the inequation presented in Definition 3. In consequence, we can guarantee (ϵ, δ)-differential privacy with

$$\epsilon = \max\left(\ln\left(\frac{f_{k,a,n-1}(\text{left})}{f_{k,a,n}(\text{left}-\Delta q)}\right), \ln\left(\frac{f_{k,a,n}(\text{right}-\Delta q)}{f_{k,a,n-1}(\text{right})}\right)\right)$$
$$\delta = \max\left(F_{k,a,n-1}(\text{left}), 1 - F_{k,a,n}(\text{right}-\Delta q)\right)$$

□

D. PROOF OF LEMMA 3

Lemma 3. The pdf $P(q(X) = s)$ of the result s of the query q "What is the consumption during the period $[1, T]$?" is

$$\int_{-C}^{C} \cdots \int_{-C}^{C} \int_{0}^{C} f_{k,a}\left(s - \sum_{t=1}^{T} d(t) - \sum_{t=1}^{T-1} b_t \,\middle|\, c + \sum_{i=1}^{T-1} b_i\right) \cdot \prod_{t=1}^{T-1} f_{k,a}\left(b_{T-t} - \sum_{i=1}^{T-t-1} b_i \,\middle|\, c + \sum_{i=1}^{t-1} b_i\right) \cdot g(c)\,\mathrm{d}c\,\mathrm{d}b_1 \ldots \mathrm{d}b_{T-1}$$

Proof. It is well known that the probability density of the sum of two independent random variables described by the pdf f is given by

$$\text{sum}(x) = \int f(y) \cdot f(x-y)\,\mathrm{d}y.$$

In our scenario, we have to generalize this expression for an arbitrary number of random variables and conditional pdfs:

1. **Adding several random variables:**
   For each time interval, we observe a random deviation b_t that can take values from $[-C, C]$: If the energy-storage device is empty, the maximum observable deviation is C. If it is full, that deviation is $-C$. Thus, that interval defines the integral boundaries. Now remember that the query result s is the sum of all deviations b_t and the sum of all consumption values $d(t)$ for all $t \in [1, T]$. For the last time interval T, the (dis-) charging rate equals $s - \sum_{t=1}^{T} d(t) - \sum_{t=1}^{T-1} b_i$. For all previous $t \in [1, T-1]$, it is $b_{T-t} - \sum_{i=1}^{T-t-1} b_i$. By convoluting the respective probabilities, we obtain the pdf of the sum of the random variables.
2. **Applying conditional pdfs:**
   We take the probabilities of observing a certain load level into account by integrating over $g(c)$, where c is

Algorithm 2 Compute g

Input: conditional (dis-)charging rate pdf $f(b(t)|c(t-1)$
capacity C
Output: stable load level pdf g

1: Start at time interval $t = 0$
2: g_t is an array where each object has value $\frac{1}{|g|}$
3: **do**
4: $t = t + 1$
5: g_t is an array where each value is 0.
6: **for** loadLevel $c(t) \in [0, C]$ **do**
7: **for** loadLevel $c(t-1) \in [0, C]$ **do**
8: $g_t(c(t)) = g_t(c(t)) + f(c(t) - c(t-1)|c(t-1)) \cdot g_{t-1}(c(t-1))$
9: error = dist$(g_t - g_{t-1})$
10: **while** error is too large
11: **return** g_t

defined on $[0, C]$. The pdfs $f_{k,a}$ now consider this load level in the conditional term, as explained in Lemma 1. The load level relevant for each time interval t is the load level of the beginning plus all (dis-)charging rates b_i of all previous time intervals, i.e., $i \in [1, t-1]$.

□

E. ALGORITHM TO COMPUTE THE STABLE LOAD LEVEL PDF

We now present an iterative, numeric approach to compute the stable load-level pdf $g(c(t))$, as sketched in Section 5.2. Algorithm 2 is the pseudo-code of this approach. The algorithm requires $f(b(t)|c(t-1))$ and C as input. We start at time interval 0 (Line 1) with an arbitrary load-level distribution g, e.g., a uniform distribution on $[0, C]$ (Line 2). Then, we apply a numerical integration by dividing the range $[0, C]$ into many intervals (Lines 6-7) and repeatedly calculate the pdf of g as described in Definition 7 (Line 8). We iteratively adapt the pdf of g until the changes are marginal (Lines 11-12).

F. EXPERIMENTAL RESULTS ON FILTERING ATTACKS

We now evaluate how GIH charging performs against approaches that perform a filtering technique to reconstruct the actual consumption. Such approaches assume that the perturbation results from white noise, i.e., the random variables added to each record are uncorrelated. We expect the trend-preservation property of GIH charging to prevent reconstructing the actual consumption. We use the following measure that quantifies the fraction of removed perturbation:

$$\text{removedPerturbation} = \frac{\text{std}_{\text{perturbed}} - \text{std}_{\text{reconstructed}}}{\text{std}_{\text{perturbed}}}$$

$\text{std}_{\text{perturbed}}$ is the standard deviation of the differences between the perturbed and the actual records and $\text{std}_{\text{reconstructed}}$ is the standard deviation of the differences between the reconstructed and the actual records. From a privacy perspective, a charging strategy where a filtering approach reconstructs each actual record has the worst perturbation removal of 100%. A charging strategy in turn which does not let the filtering approach reconstruct any actual record has a perturbation removal of 0%. When the standard deviation of the differences between the reconstructed and the actual records is larger than the one between the perturbed and the actual records, the perturbation removal has negative values.

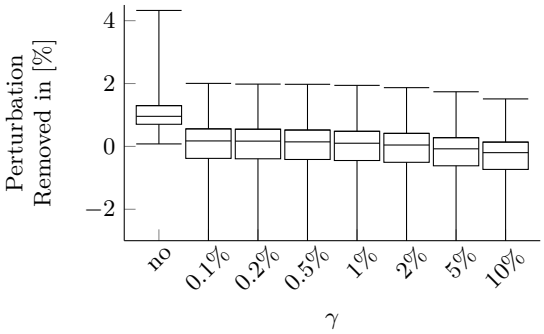


Figure 8: Experimental results on Filtering

For our evaluation, we apply the filtering approach based on Wavelet decomposition proposed in [26]. Figure 8 presents our results on the same data of the CER dataset [1] used for experiments in Section 7. It visualizes the percentage of perturbation removed by boxplots with whiskers from minimum to maximum. All minimum whiskers of those boxplots that lie outside the range of the y-axis have values of about −50%. The first boxplot visualizes a charging strategy without a trend-preservation feature, i.e., a random consumption perturbation according to Lemma 1 is used for each time interval. For this charging strategy, the filtering approach can reconstruct some information for each time series. The remaining boxplots show the results for GIH charging with different values of γ. As expected, the perturbation removed decreases slightly with increasing values of γ. If γ is small, the trend preservation must be adapted frequently, and therefore the chance of the filtering approach to succeed increases. However, for all instantiations of γ, the median is about zero. This means that the number of time series where the filtering approach succeeds is about the same as the number of time series where it fails. Thus, an attacker who does not know the actual records cannot decide whether the filtering was successful in removing the perturbation. In consequence, our experimental results show that the trend-preservation feature of the charging-strategy presented in Section 6 protects against filtering approaches.