Smart Meter Privacy: A Utility-Privacy Framework

S. Raj Rajagopalan*, Lalitha Sankar†, Soheil Mohajer†, H. Vincent Poor†
*HP Labs, Princeton, NJ 08544. raj.raj@hp.com
†Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544. lalitha,smohajer,poor@princeton.edu

***Abstract*— End-user privacy in smart meter measurements is a well-known challenge in the smart grid. The solutions offered thus far have been tied to specific technologies such as batteries or assumptions on data usage. Existing solutions have also not quantified the loss of benefit (utility) that results from any such privacy-preserving approach. Using tools from information theory, a new framework is presented that abstracts both the privacy and the utility requirements of smart meter data. This leads to a novel privacy-utility tradeoff problem with minimal assumptions that is tractable. Specifically for a stationary Gaussian Markov model of the electricity load, it is shown that the optimal utility-and-privacy preserving solution requires filtering out frequency components that are low in power, and this approach appears to encompass most of the proposed privacy approaches.**

I. Introduction

Information collection and dissemination, some of it using smart meters, are critical to the smart grid. But information about electricity consumption that is collected and harnessed for a more efficient and multi-faceted grid may be used for purposes beyond electricity consumption, thereby making it potentially dangerous to individual privacy. The privacy consequences of smart grid development are hard to understand for two principal reasons: (i) the full range of technological capabilities and information extraction possibilities have not been laid out, and (ii) our concept of privacy in this space are yet poorly defined and shifting. Smart meters are an indispensable enabler in the context of smart grids, which deploy advanced information and communication technology to control the electrical grid.

The main motivations for high-resolution energy usage data collection are to forecast load demand and to provide optimized service to consumers in the form of pricing structure [1]. An electricity provider can use this information to facilitate more efficient network management, peak load reduction, load shaping, and a number of other such uses. However, it has been known for some time that the information of appliance use can be reconstructed from the overall real-time load using libraries of appliance load signatures that could be matched to signals found within the noise of a customer's aggregated electricity use and a large amount of detail concerning customer usage habits can be discerned [2]. [1] cites a list of privacy-sensitive characteristics that may be inferred from electricity load data ranging from house occupancy to personal habits and routines.

The research was supported in part by the National Science Foundation under Grants CCF-10-16671 and CNS-09-05398, in part by the Air Force Office of Scientific Research MURI Grant FA-9550-09-1-0643, and in part by DTRA under Grant HDTRA1-07-1-0037.

The NIST Smart Grid Interoperability Panel has also underlined risk to privacy of personal behavior because new types of energy use data are created and communicated by smart meters, such as unique electric signatures for consumer electronics and appliances, thereby opening up further opportunities for general invasion of privacy. [3] suggest that there will always "be the temptation to sell such information such as energy usage or appliance data, either in identifiable customer level, anonymized or aggregate form to third parties such as marketers seeking commercial gain.". Thus, a desired feature of privacy design in the smart grid would be "positive-sum, not zero-sum" in that it seeks to accommodate all legitimate interests and objectives in a fair manner without completely sacrificing privacy for utility or vice-versa.

A typical approach to privacy in smart meter data is aggregation along dimensions of space (using neighborhood gateways, e.g. [4]), time (using battery storage, e.g. [5]), or precision (by noise addition, e.g. [6]). These solutions seek to support utility and privacy in different ways; however, they do not have a robust theoretical basis for both privacy and utility. Such a basis is important for several reasons. First, a theoretical abstraction allows us to recast the problem in a technology-independent manner – we need a privacy framework that not only addresses the capabilities of current non-intrusive load monitoring (NALM) techniques but is also extensible to future ones. Second, a theoretical framework enables us to examine the costs of lost privacy against the benefits of data dissemination, namely, the tradeoff between privacy and utility. It would be desirable to give each customer the ability to decide that tradeoff and also to give the electricity provider the ability to incentivize the customer to participate in such a bargain by offering interesting points of tradeoff. Finally, a theoretical framework for privacy and utility may expose points of tradeoff that are unexpected.

We propose a general theoretical framework that brings most current treatments of the privacy-utility tradeoff into a single model – it enables us to look at a spectrum of abstract privacy-utility choices and enables us to find maximal points on such a tradeoff curve. It also suggests new possible ways of achieving this tradeoff that have not been considered thus far.

What we have found is that suppressing low power components would be consistent with intuitive notions of privacy in smart meter data. At the same time, our utility constraints guarantee that the bulk of the energy consumption information in the load measurements is retained in the revealed data. This suggests that it may indeed be possible to reveal significant energy consumption information without also revealing a lot of personal information and the resulting tradeoff can be tuned. This would be an interesting avenue for further exploration. The paper is organized as follows. In Section II, we outline current approaches to smart meter privacy. In Section III, we develop our model, metrics, and the privacy-utility tradeoff framework and illustrate our results in Section IV.

II. Related Work

The advantages and usefulness of smart meters in general is examined in a number of papers; see for example [7] and the references therein. [5] presents a pioneering view of privacy of smart meter information: the authors identify the need for privacy in a home's load signature as being an inference violation (resulting from load signatures of home appliances) rather than an identity violation (i.e. loss of anonymity). Accordingly, they propose home electrical power routing using rechargeable batteries and alternate power sources to moderate the effects of load signatures. They also propose three different privacy metrics: relative entropy, clustering classification, and a correlation/regression metric. However they do not propose any formal utility metrics to quantify the utility-privacy trade-off.

Recently, [8] proposes additional protection through the use of a trusted escrow service, along with randomized time intervals between the setup of attributable and anonymous data profiles at the smart meter. [9] shows, somewhat surprisingly, that even without *a priori* knowledge of household activities or prior training it is possible to extract complex usage patterns from smart meter data such as residential occupancy and social activities very accurately using off-the-shelf statistical methods. [4] and [9] propose privacy-enhancing designs using neighborhood-level aggregation and cryptographic protocols to communicate with the energy supplier without compromising the privacy of individual homes. However, escrow services and neighborhood gateways support only restricted query types and do not completely solve the problem of trustworthiness. [10] presents a formal state transition diagram-based analysis of the privacy afforded by the rechargeable battery model proposed in [5]. However, [10] does not offer a comparable model of utility to compare the risks of information leakage with the benefits of the information transmitted.

In, [6] the authors present a method of providing *differential privacy* over aggregate queries modeling smart meter measurements as time-series data from multiple sources containing temporal correlations. While their approach has some similarity to ours in terms of time-series data treatment, their method does not seem generalizable to arbitrary query types. On the other hand, [11] introduces the notion of partial information hiding by introducing uncertainty about individual values in a time series by perturbing them. Our method is a more general approach to time series data perturbation that guarantees that the perturbation cannot be eliminated by averaging.

III. Our Contributions

The primary challenge in characterizing the privacy-utility tradeoffs for smart meter data is creating the right abstraction – we need a principled approach that provides quantitative measures of both the amount of information leaked as well as the utility retained, does not rely on any assumptions of data mining algorithms, and provides a basis for a negotiated level of benefit for both consumer and supplier [3]. [10] provides the beginnings of such a model – they assume that in every sampling time instant, the net load is either 0 or 1 power unit represented by the smart meter readings X_k, $k = 1, 2, ...,$ are a discrete-time sequence of binary independent and identically distributed values. They model the battery-based filter of [5] as a stochastic transfer function that outputs a binary sequence $\hat{X}_k$ that tells the electricity provider whether the home is drawing power or not at any given moment. The amount of information leaked by the transfer function is defined to be the mutual information rate $I(X; \hat{X})$ between the random variables X and $\hat{X}$. By modeling the battery charging policy as a 2-state stochastic transition machine, they show that there exist battery policies that result in less information leakage than from the deterministic charging policy of [5]. Though [10] does not provide a general utility function to go with the chosen privacy function and the modeling assumptions are extremely simplistic, it nevertheless provides a good starting point for our framework.

In our model, we assume that the load measurements are sampled (at an appropriate frequency) from a smart meter, that they are real-valued, and can be correlated (models the temporal memory of both appliances and human usage patterns). Rather than assume any specific transfer function, we assume an abstract transfer function which maps the input load measurements X into an output sequence $\hat{X}$. As in [10], we assume a mutual information rate as a metric for privacy leakage; however, we allow for the fact that a large space of (unknown to us) inferences can be made from the meter data – we model the inferred data as a random variable Y correlated with the measurement variable X. Thus, the privacy leakage is the mutual information between Y and $\hat{X}$. We also provide an abstract utility function which measures the fidelity of the output sequence $\hat{X}$ by limiting the Euclidean distance (mean square error) between X and $\hat{X}$. Using these abstractions and tools from the theory of rate distortion we are able to meet all our requirements for a general but tractable privacy-utility framework: the privacy and utility requirements provide opposing constraints that expose a spectrum of choices for trading off privacy for utility and vice-versa.

A. Model

We write x_t, $t = 1, 2, \ldots, n$, to denote the sampled load measurements from a smart meter. In general, x_t are complex valued corresponding to the real and reactive measurements and are typically vectors for multi-phase systems [2]. For simplicity and ease of presentation, we model the meter measurements as a sequence of real-valued scalars (for example, such a model applies to two-phase 120 V appliances for which one of the two phase components is zero).

For appropriately small sampling intervals, the smart meter time-series data that result from sampling the underlying continuous-time continuous-amplitude processes can be viewed as being generated by a random source with memory. The memory models the continuity and the effect of both short-term and long-term correlations in the load measurements. The short term correlations typically model the effect of the set of appliances in use over the said duration while the long term correlations model the long term power usage pattern of the human user. We model the continuous valued smart meter data as a sequence $\ldots, X_{k-1}, X_k, X_{k+1}, \ldots$, of random variables $X_k \in \mathcal{X}$, $-\infty < k < \infty$, generated by a stationary continuous valued source with memory. Specifically, we model the continuous valued discrete-time smart meter data as a sequence $\ldots, X_{k-1}, X_k, X_{k+1}, \ldots$, of Gaussian random variables $X_k \in \mathcal{X}$, $k = 0, \pm 1, ...$, generated by a stationary Gaussian source with memory captured via the autocorrelation function

$$c_{XX}(m) = E\left[X_k X_{k+m}\right], m = 0, \pm 1, \pm 2, ... \quad (1)$$

The assumption of normal distribution for total load is a simplification from empirical observations [12] that the power consumption pattern of a typical appliance in the on state is approximately Gaussian.

B. Utility and Privacy Metrics

Since continuous amplitude sources cannot be transmitted losslessly over finite capacity links, a sampled sequence of n load measurements X^n is compressed before transmission. In general, however, even if the sampled measurements were quantized *a priori*, i.e., take values in a discrete alphabet, there may be a need to perturb (distort) the data in some way to guarantee a measure of privacy. However, such a perturbation also needs to maintain a desired level of fidelity.

Intuitively, utility of the perturbed data is high if any function computed on it yields results similar to those from the original data; thus, the utility is highest when there is no perturbation and goes to zero when the perturbed data is completely unrelated to the original. Accordingly, our utility metric is an appropriately chosen average 'distance' *distortion function* between the original and the perturbed data.

Privacy, on the other hand, is maximized when the perturbed data is completely independent of the original. Our privacy metric measures the difficulty of inferring any private information of the data collector's choice, defined as a sequence $\{Y_k\}$ of random variables $Y_k \in \mathcal{Y}$, $-\infty < k < \infty$, which is correlated with and can be inferred from the revealed data. The random sequence $\{Y_k\}$ for all k along with the joint distribution $p_{X^n Y^n}$ mathematically captures the space of all inferences that can be made from the measurements. We quantify the resulting privacy loss as a result of revealing perturbed data via the *mutual information* between the two data sequences.

As an aside, we note here that our model of privacy is between a single user (household) and the electricity provider. It does not consider the leakage possibilities of comparing the perturbed data from two or more different users. On the other hand it can address the possibility of side-information such as income level of the user that may cause further information leakage. If we know the statistics of the side-information that we can incorporate the possible leakage into the model and derive the consequent modified privacy-utility tradeoff. For simplicity we ignore the side-information aspect in this paper.

C. Perturbation: Encoding and Decoding

Encoding: We assume that a meter collects $n \gg 1$ measurements in an interval of time prior to communication and that n is large enough to capture the source's memory. The encoding function is then a mapping of the resulting *source sequence* $X^n = (X_1 \; X_2 \; X_3 \; \ldots \; X_n)$, where $X_k \in \mathbb{R}$, for all $k = 1, 2, \ldots, n$, to an index $W_n \in \mathcal{W}_n$ given by

$$F_E : \mathcal{X}^n \to \mathcal{W}_n \equiv \{1, 2, \ldots, M_n\} \quad (2)$$

where each index represents a quantized sequence.

Decoding: The decoder (at the data collector) computes an output sequence $\hat{X}^n = \left(\hat{X}_1 \; \hat{X}_2 \; \hat{X}_3 \; \ldots \; \hat{X}_n\right)$, $\hat{X}_k \in \mathbb{R}$, for all k, using the decoding function

$$F_D : \mathcal{W} \to \hat{\mathcal{X}}^n. \quad (3)$$

The encoder is chosen such that the input and output sequences achieve a desired utility given by an average distortion constraint

$$D_n = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[\left(X_k - \hat{X}_k\right)^2\right] \quad (4)$$

and a constraint on the information leakage about the desired sequence $\{Y_k\}$ from the revealed sequence $\left\{\hat{X}_k\right\}$ is quantified via the leakage function

$$L_n = \frac{1}{n} I\left(Y^n; \hat{X}^n\right) \quad (5)$$

where $\mathbb{E}[\cdot]$ denotes the expectation over the joint distribution of X^n and $\hat{X}^n$ given by $p_{X\hat{X}}(x^n, \hat{x}^n) = P^n_{X^n}(x^n)\, p_t(\hat{x}^n | x^n)$ where $p_t(\hat{x}^n|x^n)$ is a conditional pdf on $\hat{x}^n$ given x^n. The mean-square error (MSE) distortion function chosen in (4) is typical for Gaussian distributed real-valued data as a measure of the fidelity of the perturbation (encoding).

Note that D_n and L_n are functions of the number of measurements n and for stationary sources converge to limiting values [13]. Let D and L denote the corresponding limiting values for utility and privacy, respectively, i.e.,

$$D \equiv \lim_{n\to\infty} D_n \quad \text{and} \quad L \equiv \lim_{n\to\infty} L_n. \quad (6)$$

D. Utility-Privacy Tradeoff Region

Formally, the utility-privacy tradeoff region $\mathcal{T}$ is defined as follows.

Definition 1: The smart meter utility-privacy tradeoff region $\mathcal{T}$ is the set of all (D, L) pairs for which there exists a coding scheme given by (2) and (3) with parameters

$(n, M_n, D_n+\epsilon, L_n+\epsilon)$ satisfying (4) and (5) for n sufficiently large and $\epsilon > 0$.

Rate-Distortion-Leakage: The above utility-privacy tradeoff problem does not explicitly bound the number M_n of encoded (quantized) sequences. An explicit constraint on

$$M_n \leq 2^{n(R_n+\epsilon)} \tag{7}$$

results in a rate-distortion-leakage (RDL) tradeoff problem for which the feasible region is defined as follows. Let $R = \lim_{n\rightarrow\infty} (\log M_n)/n$.

Definition 2: The rate-distortion-leakage tradeoff region $\mathcal{R}_{RDL}$ is the set of all (R, D, L) tuples for which there exists a coding scheme given by (2), (3), and (7) with parameters $(n, M_n, D_n+\epsilon, L_n+\epsilon)$ satisfying (4) and (5) for n sufficiently large and $\epsilon > 0$. The function $\lambda(D)$ quantifies the minimal leakage achievable for a feasible distortion D such that the set of all $(R, D, \lambda(D))$ are the boundary points of $\mathcal{R}_{RDL}$.

Theorem 1: $\mathcal{T} = \{(D, L) : (R, D, L) \in \mathcal{R}_{RDL}, D \in [0, D_{\max}], L \geq \lambda(D)\}$.

Proof sketch: The crux of our argument is the fact that for any feasible utility vector D, choosing the minimum rate $R(D, \lambda(D))$, ensures that the least amount of *information* is revealed about the source via the reconstructed variable. This in turn ensures that the minimal leakage $\lambda(D)$ of the correlated sequence Y^n is achieved for that utility. For the same utility constraint, since such a rate requirement is not a part of the utility-privacy model, the resulting maximal privacy achieved is at most as large as that in $\mathcal{R}_{RDL}$.

E. Rate-Distortion-Leakage Tradeoff

We now use Theorem 1 to precisely quantify the utility-privacy tradeoff via the RDL tradeoff region. The proof is a direct generalization of the RDL region for memoryless sources (see, for example, [14], [15]), and hence, is omitted for lack of space. Intuitively, the proof follows from upper and lower bounding the minimal communication rate R as a function of D and L and the minimal leakage rate λ as a function of D.

Theorem 2: The rate-distortion-leakage region for a source with memory subject to distortion and leakage constraints in (4) and (5) is given by the rate-distortion and minimal leakage functions

$$R(D, L) = \lim_{n\rightarrow\infty} \inf_{p(x^n,y^n)p(\hat{x}^n|x^n)} \frac{1}{n} I\left(X^n; \hat{X}^n\right) \tag{8}$$

$$\lambda(D) = \lim_{n\rightarrow\infty} \inf_{p(x^n,y^n)p(\hat{x}^n|x^n)} \frac{1}{n} I\left(Y^n; \hat{X}^n\right). \tag{9}$$

The utility-privacy tradeoff is captured by $\lambda(D)$ which is the minimal privacy leakage for a desired distortion (utility) D.

Remark 1: The Markov relationship $Y^n - X^n - \hat{X}^n$ is captured via the set of all distributions in (8) and (9) which minimize $R(D, L)$ and $\lambda(D)$.

Corollary 1: For $Y_k = X_k$, for all k, i.e., for the case in which the actual measurements need to be undisclosed, $\lambda(D) = R(D, L) = R(D)$ where $R(D)$ is the rate-distortion function for the source.

In general, the optimal distribution minimizing the rate subject to both the distortion and leakage constraints depends on the joint distribution of the measurement and inference sequences. Modeling this relationship is, in general, not straightforward or known *a priori*. Given this limitation, we consider a simple linear inference model given by

$$Y_k = \alpha_k X_k + Z_k, \text{ for all } k, \tag{10}$$

where $Z_k \sim \mathcal{N}(0, 1)$ is independent of X_k, and α_k are constants. In this paper, we limit our results to these models to simplify our analysis and develop the intuition that can eventually lead us to develop complete solutions for a more general inference model. The following theorem captures our result.

Theorem 3: The utility-privacy tradeoff for smart meter measurements modeled as a Gaussian source with memory with $Y_k = \alpha_k X_k + Z_k$, for all k, is given by the leakage function $\lambda(D)$ which results from choosing the distribution $p(\hat{x}^n|x^n)$ as the rate-distortion (without privacy) optimal distribution.

Proof: The proof follows directly from noting that, for a given jointly Gaussian distribution of the source and correlated hidden sequence, $p_{X^nY^n}$, the infimum in (8) and (9) is strictly over the space of conditional distributions of the revealed sequence given the original source sequence as a result of the Markov chain relationship $Y^n - X^n - \hat{X}^n$. Expanding the leakage as $I(Y^n; \hat{X}^n) = h(Y^n) - h(Y^n|\hat{X}^n)$, and using the fact for correlated Gaussian processes, $Y_k = \alpha_k X_k + Z_k$, for all k, where $\{Z_k\}$ is a sequence independent of $\{X_k\}$ and α_k is a constant for each k, one can show that the jointly Gaussian distribution of X^n and $\hat{X}^n$ which minimizes (8) also minimizes (9). ■

Remark 2: Theorem 3 simplifies the development of the RDL region for Gaussian sources with memory for which the rate-distortion function is known. For Gaussian sources with memory the rate-distortion function is known and lends itself to a straightforward practical implementation that we discuss in the following section.

F. Rate-Distortion for Gaussian Sources with Memory

In general, the rate distortion functions for sources with memory are not straightforward to compute. However, for Gaussian sources, the rate-distortion function $R(D)$ (without the additional privacy constraint) is known and can be obtained via a transformation of the correlated source sequence X^n to its eigen-space in which the transformed sequence $\tilde{X}^n$ is a collection of independent random variables with, in general, different variances.

A standard approach to analyze correlated data is to project the data to an orthogonal basis in which the leakage and distortion constraints remain invariant. Since the data is random, we project on to the principal axes of the $n \times n$ correlation matrix G_{XX} whose entry in the i^{th} row and j^{th} column is $c(i-j)$ defined in (1) for which the mean-square error (Euclidean distance) function and the mutual information leakage are invariant. Thus, while the constraints for the original and

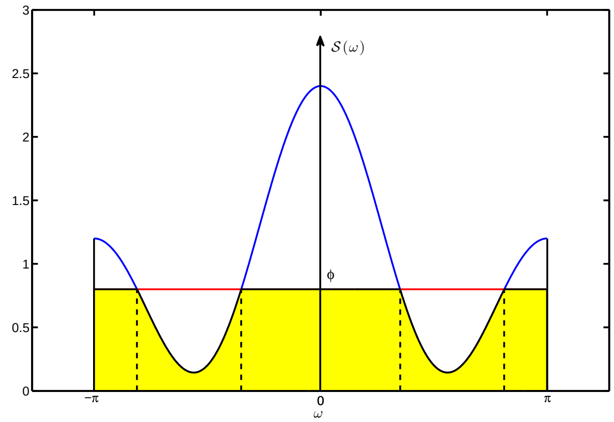

Fig. 1. The PSD of $\{X_k\}$. The area below the curve and the horizontal line is equal to D.

transformed measurements are the same, the advantage of the transformation is that the resulting measurements in any block of length n are statistical independent.

We write $S_X(f)$ denote the unitary transformation of the correlation matrix G_{XX}, i.e., $S_X(f)$ is the power spectral density (PSD) of the time series process $\{X(n)\}$, at discrete frequencies, $f = 0, 1, 2, ..., n-1$. We henceforth refer to the transform domain as the spectral domain in keeping with the literature. Similarly, let $\mathcal{S}_Y(\omega)$ and $\mathcal{S}_{XY}(\omega)$ denote the PSDs of the $\{Y_k\}$ and the $\{X_k Y_k\}$ processes where $\mathcal{S}_{XY}(\omega)$ is the transform of the cross-correlation function $c_{XY}(m)$ of the two sequences. Let ϕ denote the Lagrangian parameter for the distortion constraint (4) in the rate minimization problem. Explicitly denoting the dependence on the water-level ϕ, the rate-distortion function $R_\phi(D)$ and the average distortion function $D(\phi)$ are given by [16]

$$R_\phi(D) = \int_{-\pi}^{\pi} \max\left(0, \frac{1}{2}\log\frac{\mathcal{S}_X(\omega)}{\phi}\right)\frac{d\omega}{2\pi} \qquad (11)$$

$$D(\phi) = \int_{-\pi}^{\pi} \min\left(\mathcal{S}_X(\omega), \phi\right)\frac{d\omega}{2\pi}. \qquad (12)$$

Note that the water-level ϕ is determined by the desired average distortion $D(\phi) = D$. Thus, $R(D)$ for a Gaussian source with memory can be expressed as an infinite sum of the rate-distortion functions for independent Gaussian variables, one for each angular frequency $\omega \in [-\pi, \pi]$. The "water-level" ϕ captures the average time-domain distortion constraint across the spectrum such that the distortion for any ω is the minimum of the water-level and the PSD. The privacy leakage $\lambda(D(\phi))$ is then the infinite sum of the information leakage about $\{Y_k\}$ for each ω, and is given by

$$\lambda(D(\phi)) = \int_{-\pi}^{\pi} \frac{1}{2}\log\left(\frac{\mathcal{S}_Y(\omega)}{\mathcal{S}_{XY}(\omega)g(\omega) + \mathcal{S}_Y(\omega)}\right)\frac{d\omega}{2\pi} \quad (13)$$

where $g(\omega) \equiv (\min(\mathcal{S}_X(\omega), \phi) - 1)$.

Remark 3: The transform domain "waterfilling" solution suggests that in practice the time-series data can be filtered for a desired level of fidelity (distortion) and privacy (leakage) using Fourier transforms. The privacy-preserving rate-distortion optimal scheme thus reveals only those frequency components with power above the water-level ϕ. Furthermore, at every frequency only the portion of the signal energy which is above the water level ϕ is preserved by the minimum-rate sequence from which the source can be generated with an average distortion D.

IV. Illustration

The following example illustrates our results. We assume that the private information to be hidden is the measurement sequence itself, i.e., $Y_k = X_k$, for all k. For the meter measurements modeled as a stationary Gaussian time series $\{X_k\}$, we choose $X_k \sim \mathcal{N}(0,1)$ for all $k \in \mathcal{I}$, and an autocorrelation function

$$c_m = \mathbb{E}[X_k X_{k+m}] = \begin{cases} 1 & m = 0, \\ 0.3 & m = \pm 1, \\ 0.4 & m = \pm 2, \\ 0 & \text{otherwise.} \end{cases}$$

The power spectral density PSD (frequency domain representation of the autocorrelation function) of this process is given by

$$\mathcal{S}(\omega) = \sum_{m=-\infty}^{\infty} c_m \exp(im\omega) = 1 + 0.6\cos(\omega) + 0.8\cos(2\omega),$$
$$-\pi \le \omega \le \pi. \quad (14)$$

In order to obtain the rate-distortion function $R_\phi(D)$ for this source, for a given D we have to find the water-level ϕ satisfying (12).

Figure 1 shows the PSD function c_m. Determining ϕ is equivalent to determining the height of the horizontal line, such that the area below the curve and the line equals D as given by (12). Having determined ϕ, $R_\phi(D)$ is then given by (11). $\mathcal{S}(\omega)$ takes its minimum value at $\omega_0 = \arccos(\frac{-3}{16}) \simeq 1.7594$. Thus, for $D \le \mathcal{S}(\omega_0) \simeq 0.1437$, $\phi = D$ such that $R(D) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\log(\mathcal{S}(\omega)/D)\,d\omega$,which is the same as the rate-distortion function for a Gaussian source with variance $\tilde{\sigma}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\mathcal{S}(\omega)d\omega$, i.e., when the distortion falls below a certain threshold, the rate required to reproduce the source at the receiver with the desired fidelity is the same as that of a memoryless Gaussian source. Finally, since we have chosen to hide the original meter measurements, for this problem, the privacy leakage is the same as the rate distortion. The resulting tradeoff between is shown in Fig. 2.

V. Discussion and Concluding Remarks

The theoretical framework that we have developed here allows us to precisely quantify the utility-privacy tradeoff problem in smart meter data. Given a series of smart meter measurements X, we reveal a perturbation $\hat{X}$ that allows us to guarantee a measure of both privacy in X and utility in $\hat{X}$. The privacy guarantee comes from the bound on information leakage while the utility guarantee comes from the upper bound on the MSE distance between X and $\hat{X}$.

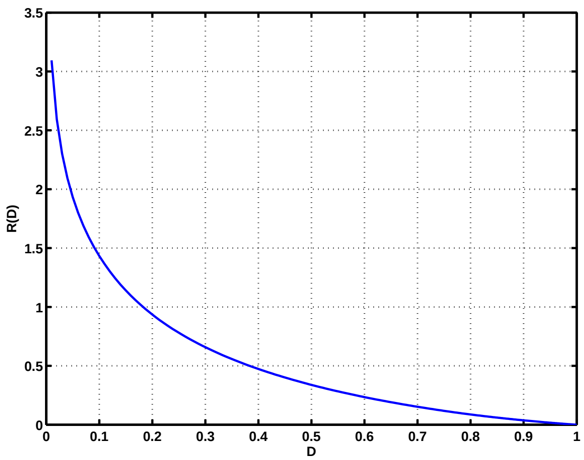

Fig. 2. Plot of $R_\phi(D) = \lambda_\phi(D)$ vs. average distortion D.

Our model of privacy, namely information leakage, does not depend on any assumptions about the inference mechanism (i.e. the data mining algorithms); instead it presents the least possible (on average) guarantee of information leakage about X, while the utility is preserved in an application-agnostic manner. Our framework is also agnostic about how the perturbation is achieved; for example it can be achieved using a filter such as a battery or by adding noise.

Modeling a smart meter as a Gaussian source with memory and extending known results from rate distortion theory, we show that a utility-privacy tradeoff framework can be constructed that gives tight bounds on the amount of privacy that can be achieved for a given level of utility and vice-versa. The critical parameter of choice in the utility-privacy tradeoff is the water level ϕ, which in turn depends on the bound on the distortion that is acceptable. The choice of ϕ dictates the extent to which the original signal (meter measurements) can be distorted and the rate $R_\phi(D)$ is the maximum data precision allowed for which information leakage is at most $\lambda(D(\phi))$. In a practical context, the choice of ϕ is dictated by the choice of the privacy-utility tradeoff operating point, which in turn has to be negotiated between the energy provider and consumer.

Our distortion model can be viewed as a filter on the load signal X – it filters out all frequencies that have power below a certain threshold (determined directly by ϕ). This filter is novel and comes directly as a result of our model. From a practical point of view, it makes sense in the following way. From the appliance signature chart in [1], frequency components that have low power typically correspond to fluctuations in energy consumption that are short-lived, which in turn are caused by appliances such as kettles and television sets and transmit the bulk of information about underlying human behavior. Frequency components that have high power tend to caused by continuously running appliances such as air conditioning units and refrigerators that reveal much less about human behavior. Suppressing low power components would thus reduce or eliminate the components of the signal that are likely to be most revealing about human behavior and thus match our intuition on privacy protection in smart meter data. At the same time, our utility constraints guarantee that most of the useful energy consumption information is retained in the revealed load data. This holds out hope that we can reveal significant energy consumption information while at the same time protecting significant personal information in a tunable tradeoff. This would be an interesting avenue for further exploration. Another interesting avenue to explore would be to apply and demonstrate the power of these concepts in a practical context.

References

[1] *Smart Metering and Privacy: Existing Law and Competing Policies*, Colorado Public Utilities Commission, 2009, http://www.dora.state.co.us/puc/.

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

[3] A. Cavoukian, J. Polonetsky, and C. Wolf, "Smartprivacy for the smart grid: embedding privacy into the design of electricity conservation," *Identity in the Information Society*, vol. 3, pp. 275–294, 2010, 10.1007/s12394-010-0046-y. [Online]. Available: http://dx.doi.org/10.1007/s12394-010-0046-y

[4] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *1st IEEE Intl. Conf. Smart Grid Commun.*, Gaithersberg, MD, Oct. 2010, pp. 327–332.

[5] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *Proc. IEEE 1st Intl. Conf. Smart Grid Comm.*, Gaithersburg, MD, Oct. 2010, pp. 232–237.

[6] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proc. 2010 Intl. Conf. Data Management*, Indianapolis, Indiana, USA, 2010, pp. 735–746.

[7] G. Deconinck and B. Decroix, "Smart metering tariff schemes combined with distributed energy resources," in *Proc. 4th Intl. Conf. Critical Infrastructures*, Linköping, Sweeden, 2009, pp. 1–8.

[8] C. Efthymiou and G. Kalogridis, "Smart grid privacy via anonymization of smart metering data," in *Proc. IEEE 1st Intl. Conf. Smart Grid Comm.*, Gaithersburg, MD, USA, Oct. 2010, pp. 238–243.

[9] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proc. 2nd ACM Workshop Embedded Sensing Systems for Energy-Efficiency in Building*, ser. BuildSys '10. New York, NY, USA: ACM, 2010, pp. 61–66. [Online]. Available: http://doi.acm.org/10.1145/1878431.1878446

[10] D. Varodayan and A. Khisti, "Smart meter privacy using a rechargeable battery: minimizing the rate of information leakage," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Prague, Czech Republic, 2011.

[11] S. Papadimitriou, F. Li, G. Kollios, and P. S. Yu, "Time series compressibility and privacy," in *Proc. 33rd Intl. Conf. Very Large Databases*, Vienna, Austria, 2007, pp. 459–470.

[12] M. Marwah, M. Arlitt, G. Lyon, M. Lyons, and C. Hickman, "Unsupervised disaggregation of low frequency power measurements," HP Labs, Tech. Rep., 2010.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[14] H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 918–923, Nov. 1983.

[15] L. Sankar, S. R. Rajagopalan, and H. V. Poor, "A theory of utility and privacy in databases," Feb 2011, submitted to the *IEEE Trans. Inform. Theory*.

[16] T. Berger and J. Gibson, "Lossy source coding," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1980.