Smart Meter Privacy in the Presence of an Alternative Energy Source

Deniz Gündüz¹ and Jesús Gómez-Vilardebó²

¹Department of Electrical Engineering, Imperial College London, London, UK
²Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain

Abstract— A smart-meter (SM) measures and reports the energy consumption of a user at frequent time intervals, revealing critical private information about user’s energy consumption behavior. In this paper, privacy in a SM system is studied in the presence of an alternative energy source (AES). The privacy-power function is introduced to study the trade-off between the achievable information theoretic privacy and the average power that can be provided by the AES. A single-letter information theoretic expression is provided for the privacy-power function, and its correspondence with the rate-distortion function is established. It is shown that the output alphabet can be restricted to be equal to the input alphabet without loss of optimality, which simplifies the numerical analysis significantly. Some numerical results are provided for various input alphabets and distributions.

I. INTRODUCTION

To increase the efficiency of energy distribution networks utility providers (UPs) need to match the energy generation to the energy consumption. If the UP knows the temporal distribution of the energy demand, it can adapt its energy generation to the demand, or apply dynamic pricing mechanisms to provide incentives for the users to shift their consumption. However, traditional meters provide only very low resolution information on the energy consumption patterns of the users. Smart meters (SMs) are intelligent devices that measure, in almost real-time, the energy consumption of users, and communicate these measurements to the UP. Due to their potential to increase the efficiency of energy networks, adoption of SMs is increasing rapidly [1]. However, this brings along a growing concern on the consumer side regarding privacy [2]. By employing non-intrusive load monitors, energy consumption patterns of users can be identified even when the SM can only read the aggregated household energy consumption [3]. Through SM readings, the UP can have direct access to user’s daily life habits, such as the times he is at home, the equipments he uses, or even the TV channel he watches [4].

There has been a growing literature on advanced mechanisms to protect the privacy of the users’ energy profile. Most of the work on SM privacy assume that the user has access to the SM readings and can manipulate them before forwarding to the UP. Bohli et al. [5] propose sending the aggregated energy consumption of a group of users, while [6] proposes compression of smart-meter data. The main limitation of this line of research is the assumption that the UP depends solely on the SM readings to measure the user’s energy consumption profile. However, the UP can easily put another non-intrusive energy monitor, and keep track of a single or multiple users’ energy consumption directly. Here, we assume that the SM readings cannot be tempered; hence, the UP can perfectly track the energy it provides to the user over time. To protect user’s privacy in such a scenario, [7] propose using energy storage devices to filter out the real energy consumption. Potential privacy benefits of storage devices has also been studied in [8] and [9] from an information theoretic perspective.

In this paper, we propose using an alternative energy source (AES), which can be another UP or an energy harvesting device available to the user, to reduce the information revealed about user’s energy profile. It is assumed that the UP does not have access to the energy provided by the AES. Obviously, if the AES is sufficient enough to provide all the required energy by the appliances, the privacy problem can be resolved in a straightforward manner. However, in general, the AES will be limited, in terms of average or maximum power it can provide, and as we show in this paper, how the user utilizes the energy provided by the AES is critical from the privacy perspective.

We use information theoretic techniques to identify the fundamental limits on privacy. We consider a discrete time system and model the energy demand of the user at each time instant as a random variable that needs to be kept as secret as possible from the UP. We measure privacy with the amount of leaked information about user’s energy consumption to the UP, which is quantified by the mutual information between the user’s real energy consumption and the energy provided by the UP. As opposed to our previous work in [9], here we do not consider any storage device, which will allow us to obtain analytical results for the fundamental limits of privacy achievable through an AES.

The privacy is achieved by deciding at each time instant how much of the energy demand should be requested from the UP and how much should be provided from the AES. We are constrained by the average and the peak values of the power that can be provided by the AES. We characterize the optimal trade-off between the available energy from the AES and the level of privacy that can be achieved. We establish a direct connection with the SM privacy problem and the rate-distortion problem in information theory. We show that the
achievable privacy for a given energy source can be written as a convex optimization problem in the discrete case. While closed-form solutions are available for only some simple models, we illustrate the numerical solution for various cases.

The rest of the paper is organized as follows. In Section II we introduce the system model and the privacy-power function. In Section III we characterize the privacy-power function when the appliances have independent and identically distributed (i.i.d.) energy demand over time. In Section IV we prove that the privacy-power function can be written as the solution of a convex optimization problem. Section V presents numerical results. Finally, Section VI concludes the paper.

II. System Model

We consider a discrete time system model and denote by \( x_t \) the energy demand of the appliances at time instant \( t \). This energy demand has to be satisfied at each time instant, either from the UP or from an AES. We denote by \( Y_t \in \mathcal{Y} \) the energy that the user receives from the UP, i.e., from the energy grid, at time \( t \). In short, we call \( x_t \) and \( Y_t \) the input and output load, respectively. We consider a discrete input load \( \mathcal{X} = \{ x_1, \ldots, x_K \} \), where \( 0 \leq x_1 < \cdots < x_K \) and a continuous output load \( \mathcal{Y} = [0, x_K] \). Observe, that we do not allow the user to ask more energy from the UP than required by the appliances, i.e., we have \( Y_t \leq x_t \), \( \forall t \).

The remainder of the power required by the appliances at each time instant is provided from an AES. We have a peak power constraint of \( \bar{P} \) on the power that can be provided by the AES, i.e., \( x_t - Y_t \leq \bar{P} \), \( \forall t \). The SM measures the output load, \( Y_t \), at each time instant, and reports it to the UP.

We measure the privacy by the average information leaked to the UP about the input load. Assuming that the statistical behavior of the energy demand of the appliances is known by the UP, its initial uncertainty of the real energy consumption is given by \( H(X^n) \). This uncertainty is reduced to \( H(X^n|Y^n) \) after the UP observes the output load. Hence, the information leaked to the UP through the energy management policy of the user can be measured by the reduction in the uncertainty, or equivalently, by the mutual information between the input and output loads, \( I(X^n; Y^n) = H(X^n) - H(X^n|Y^n) \).

We consider energy management policies that decide on the amount of power that will be received from the AES at each time instant \( t \) based on the input load up to time \( t \), \( X^t \), and the output load up to the previous time instant, \( Y^{t-1} \). We allow stochastic policies that satisfy the peak power constraint.

**Definition 1:** A length-\( n \) energy management policy is composed of, possibly random, power allocation functions

\[
f_t : \mathcal{X}^t \times Y^{t-1} \to \mathcal{Y},
\]

for \( t = 1, \ldots, n \), such that \( 0 \leq x_t - Y_t \leq \bar{P} \). The privacy of this policy is given by the information leakage rate defined as

\[
I_n = \frac{1}{n} I(X^n; Y^n),
\]

while the required average power from the AES is given by

\[
P_n = \mathbb{E} \left[ \frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t) \right],
\]

where the expectation is taken over the joint distribution of the input and output load.

**Definition 2:** An information leakage rate - average power pair \((I, P)\) is said to be achievable if there exists a sequence of energy management policies of duration \( n \) with \( \lim_{n \to \infty} I_n \leq I \) and \( \lim_{n \to \infty} P_n \leq P \).

**Definition 3:** The information leakage rate-average power region is the closure of the set of all achievable information leakage rate - average power pairs \((I, P)\).

The privacy-power function characterizes privacy that can be achieved for an AES. We consider here only the average and peak power constraints on the AES rather than the instantaneous management of the available power as in [9]. This is based on the assumption that the AES, which can be an energy harvester, has its own storage unit, and we are only concerned about the long-term stability of this storage.

III. Privacy-Power Function

Our goal is to give a mathematically tractable expression for the privacy-power function, and identify the optimal energy management policy that achieves the highest level of privacy for a given AES. In the rest of the paper, we consider, for simplicity, an i.i.d. input load. In the next theorem, we characterize the privacy-power function in a single-letter format for an i.i.d. input load distribution.

**Theorem 1:** The privacy-power function \( \mathcal{I}(P) \) for an i.i.d. input load \( X \) with distribution \( p_X(x) \) is given by

\[
\mathcal{I}(P) = \min_{p_{Y|X}:\mathbb{E}[X-Y] \leq P} I(X; Y).
\]

Some basic properties of the privacy-power function \( \mathcal{I}(P) \) is characterized in the following lemma. The proof follows from standards techniques based on time-sharing arguments.

**Lemma 1:** The privacy-power function \( \mathcal{I}(P) \), given above, is a non-increasing convex function of \( P \).

Next we provide a proof of Theorem 1.

**Proof:** The achievability is trivial. Given a conditional probability density function \( p_{Y|X} \) that satisfies (3), we generate each \( Y_t \) independently using \( p_{Y|X}(y_t|x_t) \). The mutual information leakage rate is then given by \( I(X; Y) \) whereas the average and peak power constraints are satisfied.

For the converse, assume that there is a series of power allocation functions that satisfy the average and peak power constraints. The information leakage rate of the resulting output load series will satisfy the following chain of inequalities:

\[
\frac{1}{n} I(X^n; Y^n) = \frac{1}{n} \left[ H(X^n) - H(X^n|Y^n) \right],
\]

\[
= \frac{1}{n} \sum_{t=1}^{n} \left[ H(X_t) - H(X_t|X^{t-1}Y^n) \right],
\]
\[ \geq \frac{1}{n} \sum_{t=1}^{n} [H(X_t) - H(X_t|Y_t)] , \quad (6) \]
\[ = \frac{1}{n} \sum_{t=1}^{n} I(X_t;Y_t) , \quad (7) \]
\[ \geq \frac{1}{n} \sum_{t=1}^{n} \mathbb{I}[E[X_t - Y_t]] , \quad (8) \]
\[ \geq \mathbb{I} \left( \frac{1}{n} \sum_{t=1}^{n} E[X_t - Y_t] \right) , \quad (9) \]
\[ \geq \mathbb{I}(P) , \quad (10) \]

where (6) follows as conditioning reduces entropy; (8) follows from the definition of the privacy-power function \( \mathbb{I}(\cdot) \); and (9) follows from the convexity of function \( \mathbb{I}(\cdot) \) stated in Lemma 1 and Jensen’s inequality.

Remark 1: The achievability part of the proof reveals that the optimal privacy can be achieved by simply considering the instantaneous input load, and the energy management unit can ignore all the past input loads and actions. This is achieved by using a stochastic energy management policy rather than a deterministic one. We note here that the same performance could also be achieved by a deterministic block-based energy management policy if the user knew all the future energy demand over a block of \( n \) time instants.

We note here the similarity between the privacy-power function in (3) and the rate-distortion function [10]. The rate-distortion function characterizes the minimum required number of bits per sample, such that the receiver can reconstruct the source sequence within a specified average distortion based on the transmitted bits. Similarly, the goal here is to reconstruct \( Y^n \) such that the mutual information between \( X^n \) and \( Y^n \) is minimized while satisfying the constraints on the AES. There are two major differences between the two problems: i) We do not have a digital interface in the SM problem. Here \( Y^n \) is the direct output of the “encoder”, rather than the reconstruction of the decoder based on the transmitted index. ii) The energy management unit does not operate over blocks of input load realizations. Instead, the output load is decided instantaneously based on the previous input and output loads.

More rigorously, the privacy-power function in (3) is equivalent to the rate-distortion function with the following difference distortion measure:

\[ d(x,y) = \begin{cases} \infty & \text{if } 0 \leq x - y \leq \bar{P}, \\ x - y & \text{otherwise}. \end{cases} \]  

(11)

This correspondence allows us to use various tools from rate-distortion theory to study privacy in a SM system.

IV. OUTPUT LOAD ALPHABET

In the previous section we have given a single-letter expression for the privacy-power function of a SM system for given input load distribution and average and peak power constraints on the AES. Similar to the rate-distortion function, it is not always possible to give a closed-form analytical expression for the privacy-power function. An alternative approach is to evaluate it numerically.

In the system model outlined in Section II, the user can ask any amount of power from the UP at each time instant as long as the conditions \( 0 \leq X_t - Y_t \leq \bar{P} \) are satisfied. This corresponds to a continuous output alphabet even though the input alphabet is discrete. Therefore, we have infinitely many variables when we try to optimize the conditional probability density function \( p_{Y|X}(y|x) \). In this section, we prove that we can bound the output load alphabet to be the same as the input load alphabet without loss of optimality. This will reduce the complexity of the optimization problem significantly.

Theorem 2: The output load alphabet can be constrained to \( \mathcal{Y} = \mathcal{X} \) without loss of optimality.

Proof: Assume that the optimal privacy-power function is achieved by the conditional probability density function \( p_{Y|X}(y|x) \) from \( \mathcal{X} \) to \( \mathcal{Y} \triangleq [0, x_K] \).

Consider any \( x_i \in \mathcal{X} \) for \( i = 1, \ldots, K - 1 \), or \( x_0 \triangleq 0 \). For any \( i = 0, \ldots, K - 1 \), using the optimal conditional density function \( p_{Y|X}(y|x) \) we define a new conditional probability function \( p_{\hat{Y}|X}(y|x) \) as follows:

\[ p_{\hat{Y}|X}(y|x) = \begin{cases} 0 & \text{if } x_i < \hat{y} < x_{i+1}, \\ p_{Y|X}(y|x) & \text{if } \hat{y} \leq x_i \text{ or } \hat{y} > x_{i+1}, \\ \int_{\Omega} p_{Y|X}(y|x)dy & \text{if } \hat{y} = x_{i+1}, \end{cases} \]

where we have defined \( \Omega \triangleq (x_i, x_{i+1}) \). The new conditional probability density function does not allow any output value within \( (x_i, x_{i+1}) \), by assigning an output of \( x_{i+1} \) whenever the original distribution \( p_{Y|X}(y|x) \) assigned an output in \( \Omega \).

We first note that the new conditional probability function will satisfy all the constraints regarding the power requirements. In particular, since the output load is not reduced for any input load value, the average power received from the AES can only decrease. Moreover, the output load at any time instant is still less than what is requested by the appliances. We will show that the new conditional distribution, \( p_{\hat{Y}|X}(y|x) \), leaks at most the same amount of information to the UP as well.

Since \( I(X;Y) = H(X) - H(X|Y) \), and the input load is not changed, we will focus on the conditional entropy term, and show that \( H(X|\hat{Y}) \geq H(X|Y) \). We have

\[ H(X|Y) = \int_{y \in \mathcal{Y}} H(X|Y = y)p_Y(y)dy, \]
\[ = H_o + \int_{\Omega} H(X|Y = y)p_Y(y)dy, \]

(12)

where we have defined

\[ H_o \triangleq \int_{y \in \mathcal{Y}, y \notin \Omega} H(X|Y = y)p_Y(y)dy. \]

(13)

For any \( j \in \{i + 1, \ldots, K\} \), we have

\[ p_{X|Y}(x_j, x_{i+1}) = \int_{y \in \Omega} p_{XY}(x_j, y)dy, \]
\[ = \int_{y \in \Omega} p_{X|Y}(x_j|y)p_Y(y)dy. \]

(14)
Then we can write

\[ p_{X|Y}(x_j|x_{i+1}) = \int_{\Omega} p_{X|Y}(x_j|y) \frac{p_Y(y)}{p_Y(x_{i+1})} dy. \]  

(14)

At this point, we will use the integral form of the Jensen’s inequality, which we rephrase in the following lemma.

\textbf{Lemma 2:} Let \( S \) be a measurable subset of the real line, and \( f(x) \) be a probability density function such that \( \int_S f(x) dx = 1 \). Then, for any real-valued function \( g(\cdot) \) and function \( h(\cdot) \) which is concave over the range of \( g(\cdot) \), we have

\[ h \left( \int_S g(x)f(x)dx \right) \geq \int_S h(g(x))f(x)dx. \]

Note that \( f(y) \triangleq \frac{p_Y(y)}{p_Y(x_{i+1})} \) is a probability density function over the set \( \Omega \) as it satisfies \( \int_\Omega f(y)dy = 1 \). We let \( g(y) \triangleq p_{X|Y}(x|y) \). Then, from the concavity of the entropy function, Lemma 2 and using (14), we get

\[ H(X|\hat{Y} = x_{i+1}) \geq \int_{y \in \Omega} H(X|Y = y) \frac{p_Y(y)}{p_Y(x_{i+1})} dy. \]  

(15)

Finally, we get

\[ H(X|\hat{Y}) = \int_{y \in \mathcal{Y}} H(X|\hat{Y} = \hat{y})p_{Y}(\hat{y})d\hat{y}, \]

\[ = H_o + H(X|\hat{Y} = x_{i+1})p_{Y}(x_{i+1}), \]

\[ \geq H(X|Y), \]

where the last inequality follows from (12) and (15).

We have shown that assigning \( x_{i+1} \) as the output load for all output loads between \( (x_i, x_{i+1}) \) reduces both the leaked information and the power load demanded from the AES. We can apply this operation for all \( i = 0, \ldots, K - 1 \). Similarly, if \( 0 \notin \mathcal{X} \) we can also show that we can remove 0 from \( Y \) without loss of optimality as well. The final conditional probability density function \( p_{Y|X} \) has \( p_{Y|X}(y|x) = 0 \) for any \( y \notin \mathcal{Y} \) and \( x \in \mathcal{X} \); and hence, we can consider \( \mathcal{Y} = \mathcal{X} \), and \( p_{Y|X} \) becomes a probability mass function. This concludes the proof of the theorem.

\textbf{Remark 2:} With this reduction in the size of the output alphabet, the characterization of the privacy-power function \( I(P) \) in (3) becomes a convex optimization problem since the mutual information is a convex function of the conditional probability values, \( p_{Y|X}(y_m|x_k) \), for \( y_m \in \mathcal{Y}, x_k \in \mathcal{X} \), and the constraints are linear.

Since the optimization problem in (3) is convex, the privacy-power function for a discrete memoryless input load can be evaluated efficiently in polynomial time. Moreover, we can use the well-known tool for numerical computation of rate-distortion functions, namely the Blahut-Arimoto algorithm [10], to compute the privacy-power function.

\textbf{V. NUMERICAL RESULTS}

It is desirable in general to find a closed-form expression for the privacy-power function for a given input load distribution. Unfortunately, similar to the rate-distortion function, a closed-form expression can be found only for some special input distributions. Below we provide an analytical expression for a binary input load. We then consider larger alphabet sizes, and evaluate the privacy-power function numerically using the Blahut-Arimoto algorithm.

\textbf{A. Binary Input Load}

Consider a binary input alphabet with a Bernoulli distribution, i.e., \( \mathcal{X} = \mathcal{Y} = \{0, 1\} \), and \( p_X(0) = p \); that is, there is either a constant power demand by the appliances or no demand, independently at each time instant. When there is power demand, the energy management policy fulfills this demand either from the UP or from the EH unit in a random manner. The maximum power constraint on the AES is \( P = 1 \). Then, the privacy-power function, \( I(P) \), is given by:

\[ P \log_2(P) - (1 - p) \log_2(1 - p) - (p + P) \log_2(p + P), \]

if \( P \leq 1 - p \), and \( I(P) = 0 \), otherwise. As expected, we have \( I(0) = h(p) \), where \( h(\cdot) \) is the binary entropy function.

We plot the \( I(P) \) function for the binary input load in Fig. 1 for different \( p \) values. As expected, the required average power from the AES is maximum when the user wants perfect privacy, and it is zero when no privacy is required. We also observe clearly that the privacy-power function is convex. Another interesting observation from the figure is the fact that the \( I(P) \) curves for two different input load distributions, i.e., different \( p \) values, might intersect. This means that, to achieve the same level of privacy a lighter input load might require lower or higher average power than a heavier input load.

\textbf{B. Discrete Uniform Distribution}

We next consider a model with a larger input alphabet size and different maximum power levels \( P \). We assume that the input load has a uniform distribution \( U(\mathcal{X}) \) with \( \mathcal{X} = \{0, c, 2c, \ldots, (N - 1)c\} \), where \( c \triangleq \frac{1}{N} \) is a constant used to impose a mean value of \( E[X] = 1 \). Based on Theorem 2, the output load alphabet can be limited to \( \mathcal{X} \) as well without loss of optimality. For a peak power constraint of \( P \), we define

\begin{align*}
\end{align*}

![Fig. 1. Privacy-power function for a binary input-output system with different \( p \) values.](image-url)
The load \( N_P = \frac{P}{c} + 1 \). Then, observe that the AES load satisfies \( X - Y \in \mathcal{X}_P \) with \( \mathcal{X}_P \triangleq \{0, c, 2c, \ldots, (N_P - 1)c\} \).

We set \( N = 21 \), and in Fig. 2 we plot the privacy-power function, \( \mathcal{I}(P) \), for different maximum power levels \( \bar{P} \in \{c, 6c, 11c, 16c, 20c\} \). As expected, we observe that the \( \mathcal{I}(P) \) is a non-increasing convex function of \( P \) for all \( P \) values, while the minimum achievable information leakage rate decreases as \( \bar{P} \) increases. When there is no AES available, i.e., \( P = 0 \), the UP can track the input load perfectly, and the information leakage rate is equal to the entropy of the input load, \( H(X) = 4.39 \). On the other hand, we also observe that, with a bound on the maximum instantaneous power that is less than the maximum input load, i.e., \( N_P < N \), the minimum information leakage rate is greater than zero, and can not be further decreased by increasing \( P \) beyond a certain value \( P_{\text{min}} \).

In certain scenarios, we may want to increase \( P \) beyond \( P_{\text{min}} \); for example, for pricing reasons it is possible that the energy from the AES is preferable to that from the UP. Let \( P_{\max} \) be the maximum possible average power that can be received from the AES while satisfying the power constraints, \( \bar{P} \geq X_0 - Y_0 \geq 0 \). We will find the information leakage rate \( \mathcal{I}_{P_{\max}} \) corresponding to this point of operation. The energy management policy assigns

\[
y = \begin{cases} 
x - \bar{P} & \text{if } x \geq \bar{P}, \\
0 & \text{if } x < \bar{P}.
\end{cases}
\]

Assuming a uniform input load, the maximum power received from the AES, \( P_{\max} \), is found to be

\[
P_{\max} = E[X - Y],
\]

\[
= \frac{1}{N} \sum_{i=0}^{N_P-2} ci + \frac{1}{N} \sum_{i=N_P-1}^{N-1} \bar{P},
\]

\[
= \bar{P} \left(1 - \frac{N_P}{2N}\right).
\]

However this strategy is suboptimal in terms of information leakage rate. Observe that for \( y > 0 \) we have \( X = y + \bar{P} \), that is, \( X \) is deterministic; thus, \( H(X|y) = 0 \). If \( y = 0 \), \( X \) has a uniform distribution, \( U(\mathcal{X}_P) \). Then,

\[
\mathcal{I}(P_{\max}) = H(X) - p_Y(0)H(U(\mathcal{X}_P)),
\]

\[
= \log(N) - \frac{N_P}{N} \log(N_P).
\]

In Fig. 2 for each each \( N_P \) value, we show the corresponding \( (P_{\max}, \mathcal{I}(P_{\max})) \) point with an x-mark. It can be seen from the figure that the user needs to sacrifice its privacy to better exploit the AES. In our future work we will explore the optimal privacy that can be achieved when each energy source has a cost and the user has a limited budget.

VI. CONCLUSIONS

We have proposed a mathematical model for studying privacy in a SM system in the presence of an AES. We have shown that the user can hide its energy consumption profile from the UP by utilizing this AES in a stochastic manner.

Using an information theoretic measure, we have characterized the optimal privacy that can be achieved for given average and peak power constraints on the AES. We have shown that, for i.i.d. input loads, the privacy-power function has a single-letter expression, and more importantly, can be evaluated numerically as the solution of a convex optimization problem. We have provided numerical results for certain generic input load distributions, and explored the effect of the average and peak values of the AES on the privacy-power function. We will extend our study to continuous input load distributions as well as multi-user systems in our future work.

REFERENCES


