Privacy-Optimal Strategies for Smart Metering Systems with a Rechargeable Battery

Simon Li
Department of Electrical and Computer Engineering
University of Toronto
Email: simonli@ece.utoronto.ca

Ashish Khisti
Department of Electrical and Computer Engineering
University of Toronto
Email: akhisti@ece.utoronto.ca

Aditya Mahajan
Department of Electrical and Computer Engineering
McGill University
Email: aditya.mahajan@mcgill.ca

***Abstract*—In smart grids, a smart meter communicates fine-grained information about a user's energy demand to the utility provider. A user's energy demand can be used to infer its private activities. Therefore, smart meters post a risk of violating privacy. This risk may be mitigated by using a rechargeable battery to obfuscate the user's demand. We investigate battery charging (and discharging) strategies that minimize the amount of information leaked to the grid, where information leakage is measured using mutual information.**

We model the energy demand as a Markov process and, after a series of simplifications, show that the problem of determining privacy-optimal charging strategies can be recast as a Markov decision process; the optimal strategy and the minimum leakage rate are given by the solution of a dynamic program. For the special case of i.i.d. demand, we explicitly characterize the optimal strategy and show that the associated minimum information leakage rate is given by a single-letter mutual information expression.

I. Introduction

Smart electricity meters deliver fine-grained information about a user's energy consumption to the utility provider. The utility provider may use this information to improve the efficiency of the grid [1]. However, this efficiency comes at the risk of privacy violation: the utility provider, or an eavesdropper, may employ data mining algorithms to infer user's private activities [2]. This risk may be mitigated by using a rechargeable battery to obfuscate the user's demand [3]. In particular, in the presence of a rechargeable battery, the energy consumed from the grid need not equal the user's demand. The consumed energy could be larger (where the excess energy is stored in the battery) or smaller (where the remaining energy is provided by the battery). Therefore, only part of the information about the user's demand can be inferred from the energy consumed from the grid. In this paper, we consider the problem of designing privacy-optimal charging strategies for such rechargeable batteries.

Several variations of privacy-aware charging strategies for smart metering systems have been considered in the literature. The information leakage rate of a best effort strategy—one that holds the energy consumed from the grid to its most recent value, if possible—was investigated in [3]. The privacy-optimal strategies for the binary model (i.e., binary (on/off) i.i.d. loads and a battery capacity equal to the load) was considered in [4]. This approach was generalized to include local energy generation in [5], [6]. It should be noted that in [4]–[6], the optimal charging strategy was identified by a brute force search over a subclass of all strategies; the performance of a particular charging strategy was evaluated by explicitly computing the joint distribution on the user's demand and consumption over a long time period. A generalization to the case when there is a constraint on the average and peak energy consumed from the grid is considered in [7]; explicit solutions are provided for the case of zero and infinite battery capacity. An information theoretic characterization of the privacy-utility trade-off was investigated in [8], [9].

In this paper we generalize our recent results [10], where the optimal strategy for the binary model was identified. We first characterize a structural property of optimal strategies that allows us to write information-leakage rate in an additive form. We then identify an equivalent sequential problem and show that the optimal strategy is given the solution of a dynamic program. This dynamic program is similar to that for partially observable Markov decision processes but, in our case, the per-step cost is not linear in the information state. For the special case of i.i.d. demand, we guess the identify the optimal strategy and show that it satisfies the dynamic program. The corresponding minimum information leakage rate is given by a single-letter mutual information expression. In a subsequent paper [11], we show that the solution for the i.i.d. demand can also be obtained by information theoretic arguments. After the present work was completed, we became aware of [12], where a similar dynamic programming framework is used. However, no explicit solutions are established in [12].

Notation: Uppercase letters X, Y, etc. denote random variables, the corresponding lowercase letters x, y, etc. denote their realization, and the corresponding script letters $\mathcal{X}$, $\mathcal{Y}$, etc. denote their alphabets. $\mathcal{P}_X$ denotes the space of probability distributions on the set $\mathcal{X}$ and $\mathcal{P}_{X|Y}$ denotes the space of coditional probability distributions on $\mathcal{X}$ given $\mathcal{Y}$. X_a^b is a short hand notation for $(X_a, \dots, X_b)$; X^b is a short hand for X_1^b. $\mathbb{P}(\cdot)$ denotes the probability of an event, $\mathbb{E}[\cdot]$ denotes expectation of a ranom variable, and $\mathbb{1}_{\mathcal{A}}\{\cdot\}$ denotes the indicator of the set $\mathcal{A}$.

Given random variables X and Y, $H(X)$ denotes entropy of X, $H(X|Y)$ denotes the conditional entropy of X given Y, and $I(X;Y)$ denotes the mutual information between X and

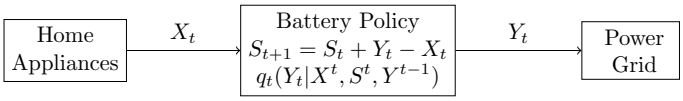

Fig. 1: An energy management system.

Y. Sometimes we need to explicitly indicate the dependence of these quantities on the underlying probability distributions. In particular, given a joint distribution $P_{X,Y}(x, y) = q(x|y)P_Y(y)$, we expression the conditional entropy $H(X|Y)$ as $H(q|P_Y) = -\sum_{x,y} P(x, y) \log q(y|z)$.

II. Model and Problem Formulation

We consider an energy management system shown in Fig. 1. Let $\{X_t\}_{t\geq 1}$, $X_t \in \mathcal{X}$, denote the (exogenous) demand process, $\{Y_t\}_{t\geq 1}$, $Y_t \in \mathcal{Y}$, denote the energy consumed from the grid, and $\{S_t\}_{t\geq 1}$, $S_t \in \mathcal{S}$ denote the energy stored in the battery. It is assumed that $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{S}$ are finite sets where $\mathcal{X} \subseteq \mathcal{Y}$ (so that the energy demand can always be met by the grid). We We assume that $\{X_t\}_{t\geq 1}$ is a first-order time-homogeneous Markov chain with transition probability Q and initial state X_1 distributed according to probability mass function P_{X_1}; the initial charge S_1 of the battery is distributed according to probability mass function P_{S_1}.

We assume an ideal battery that has no conversion losses or other inefficiencies. Therefore, the following conservation equation must be satisfied at all time instances:

$$S_{t+1} = S_t + Y_t - X_t. \tag{1}$$

Note that the battery is assumed to be ideal for the ease of presentation. Our results extend to a battery with conversion losses as well, as long as the battery update is deterministic.

The energy management system observes the demand and battery charge and consumes energy from the grid according to a randomized *charging strategy* $\mathbf{q} = (q_1, q_2, \dots)$. In particular, at time t, given (x^t, s^t, y^{t-1}), the history of demand, battery charge, and past consumption, then the current consumption Y_t is y with probability $q_t(y \mid x^t, s^t, y^{t-1})$. For a randomized charging strategy to be feasible, it must satisfy the conservation equation (1). Let $\mathcal{Y}_\circ(w) = \{y \in \mathcal{Y} : w + y \in \mathcal{S}\}$; then

$$\begin{aligned} &q_t(\mathcal{Y}_\circ(s_t - x_t) \mid x^t, s^t, y^{t-1}) \\ &\quad = \sum_{y \in \mathcal{Y}_\circ(s_t - x_t)} q_t(y \mid x^t, s^t, y^{t-1}) = 1. \end{aligned}$$

The set of all such feasible strategies is denoted by $\mathcal{Q}_A$.

The information leaked under a strategy $\mathbf{q}$ is given by the mutual information $I^{\mathbf{q}}(S_1, X^T; Y^T)$ that is evaluated according to the joint probability distribution on (X^T, S^T, Y^T) induced by $\mathbf{q}$:

$$\begin{aligned} \mathbb{P}^{\mathbf{q}}(S^T = s^T, X^T = x^T, Y^T = y^T) &= P_{S_1}(s_1)P_{X_1}(x_1) \\ \times q_1(y_1 \mid x_1, s_1) \prod_{t=1}^{T-1} \Big[&\mathbb{1}_{s_{t+1}}\{s_t - x_t + y_t\} Q(x_{t+1}|x_t) \\ &\times q_{t+1}(y_{t+1} \mid x^{t+1}, s^{t+1}, y^t) \Big]. \end{aligned}$$

Note that since the charging strategies are causal (i.e., Y_t does not depend on the future observations X_{t+1}^T), the mutual information $I^{\mathbf{q}}(S_1, X^T; Y^T)$ is equal to the directed information $I(S_1, X^T \to Y^T)$.

We use information leakage *rate* as a measure of the quality of a charging strategy. For a finite horizon, the information leakage rate of a strategy $\mathbf{q} = (q_1, \dots, q_T) \in \mathcal{Q}_A$ is given by

$$L_T(\mathbf{q}) = \frac{1}{T} I^{\mathbf{q}}(X^T, S_1; Y^T) \tag{2}$$

while for an infinite horizon, the worst-case information leakage rate of a strategy $\mathbf{q} = (q_1, q_2, \dots) \in \mathcal{Q}_A$ is given by

$$L_\infty(\mathbf{q}) = \limsup_{T\to\infty} L_T(\mathbf{q}). \tag{3}$$

We are interested in the following optimization problems:

Problem A. *Given the alphabet $\mathcal{X}$ of the demand, the initial distribution P_{X_1} and the transition matrix Q of the demand process, the alphabet $\mathcal{S}$ of the battery, the initial distribution P_{S_1} of the battery state, and the alphabet $\mathcal{Y}$ of the demand:*

1) *For a finite planning horizon T, find a battery charging strategy $\mathbf{q} = (q_1, \dots, q_T) \in \mathcal{Q}_A$ that minimizes the leakage rate $L_T(\mathbf{q})$ given by* (2)*.*
2) *For an infinite planning horizon, find a battery charging strategy $\mathbf{q} = (q_1, q_2, \dots) \in \mathcal{Q}_A$ that minimizes the leakage rate $L_\infty(\mathbf{q})$ given by* (3)*.*

III. The main-results

A. Simplification of optimal charging strategies

Let $\mathcal{Q}_B \subset \mathcal{Q}_A$ denote the set of charging strategies that choose consumption based only on the consumption history and the current demand and battery state. Thus, for $\mathbf{q} \in \mathcal{Q}_B$, at any time t, given history (x^t, s^t, y^{t-1}), the consumption Y_t is y with probability $q_t(y \mid x_t, s_t, y^{t-1})$. Then the joint distribution on (X^T, S^T, Y^T) induced by $\mathbf{q} \in \mathcal{Q}_B$ is given by

$$\begin{aligned} \mathbb{P}^{\mathbf{q}}(S^T = s^T, X^T = x^T, Y^T = y^T) &= P_{S_1}(s_1)P_{X_1}(x_1) \\ \times q_1(y_1 \mid x_1, s_1) \prod_{t=1}^{T-1} \Big[&\mathbb{1}_{s_{t+1}}\{s_t - x_t + y_t\} Q(x_{t+1}|x_t) \\ &\times q_{t+1}(y_{t+1} \mid x_{t+1}, s_{t+1}, y^t) \Big]. \end{aligned}$$

Proposition 1. *In Problem A, there is no loss of optimality in restricting attention to charging strategies $\mathbf{q} \in \mathcal{Q}_B$. Moreover, the objective function takes an additive form i.e., for $\mathbf{q} \in \mathcal{Q}_B$,*

$$L_T(\mathbf{q}) = \frac{1}{T} \sum_{t=1}^{T} I^{\mathbf{q}}(X_t, S_t; Y_t \mid Y^{t-1})$$

where

$$\begin{aligned} &I^{\mathbf{q}}(X_t, S_t; Y_t \mid Y^{t-1}) \\ &\quad = \sum_{\substack{x_t \in \mathcal{X}, s_t \in \mathcal{S} \\ y^t \in \mathcal{Y}^t}} \mathbb{P}^{\mathbf{q}}(X_t = x_t, S_t = s_t, Y^t = y^t) \\ &\qquad \times \log \frac{q_t(y_t \mid x_t, s_t, y^{t-1})}{\mathbb{P}^{\mathbf{q}}(Y_t = y_t \mid Y^{t-1} = y^{t-1})}. \end{aligned}$$

Proof. The proof relies on showing that for any strategy $\mathbf{q}_A \in \mathcal{Q}_A$, there exists a strategy $\mathbf{q}_B \in \mathcal{Q}_B$ such that

$$I^{\mathbf{q}_A}(S_1, X^T; Y^T) \geq \sum_{t=1}^{T} I^{\mathbf{q}_A}(S_t, X_t; Y_t \mid Y^{t-1})$$
$$= \sum_{t=1}^{T} I^{\mathbf{q}_B}(S_t, X_t; Y_t \mid Y^{t-1}) = I^{\mathbf{q}_B}(S_1, X^T; Y^T).$$

The details are omitted due to space limitations. □

B. An equivalent sequential problem

In this section, we develop a sequential optimization problem equivalent to Problem A. A similar approach was proposed in [13] for computing the capacity of Markov channels with feedback.

Consider a system with state process $\{X_t, S_t\}_{t\geq 1}$ where $\{X_t\}_{t\geq 1}$ is an exogenous Markov process as before and $\{S_t\}_{t\geq 1}$ is a controlled Markov process as specified below. At time t, a decision maker observes Y^{t-1} and chooses a distribution valued action $A_t \in \mathcal{A}$, where

$$\mathcal{A} = \left\{ a \in \mathcal{P}_{Y|X,S} : a(\mathcal{Y}_\circ(s-x) \mid x, s) = 1, \ \forall x \in \mathcal{X}, s \in \mathcal{S} \right\}.$$

Based on this action, an auxiliary variable $Y_t \in \mathcal{Y}$ is chosen according to the conditional probability $a_t(\cdot \mid x_t, s_t)$ and the state S_{t+1} evolves according to (1). Then for any history (x^t, s^t, y^{t-1}), $a^t \in \mathcal{A}$, and $s_{t+1} \in \mathcal{S}$,

$$\mathbb{P}(S_{t+1} = s_{t+1} \mid X^t = x^t, S^t = s^t, Y^{t-1} = y^{t-1}, A^t = a^t)$$
$$= \sum_{y_t \in \mathcal{Y}} a_t(y_t \mid x_t, s_t) \mathbb{1}_{s_{t+1}}\{s_t + y_t - x_t\}$$
$$= \mathbb{P}(S_{t+1} = s_{t+1} \mid X_t = x_t, S_t = s_t, A_t = a_t)$$

Thus, $\{X_t, S_t\}_{t\geq 1}$ is a controlled Markov process with control action $\{A_t\}_{t\geq 1}$ chosen as follows:

$$A_t = f_t(Y^{t-1}, A^{t-1})$$

where $\mathbf{f} = (f_1, f_2, \dots)$ is called the decision strategy. A decision strategy $\mathbf{f} = (f_1, \dots, f_T)$ induces a joint probability distribution on (X^T, S^T, Y^T) that is given by

$$\mathbb{P}^{\mathbf{f}}(S^T = s^T, X^T = x^T, Y^T = y^T) = P_{S_1}(s_1) P_{X_1}(x_1)$$
$$\times a_1(y_1 \mid x_1, s_1) \prod_{t=1}^{T-1} \Big[\mathbb{1}_{s_{t+1}}\{s_t - x_t + y_t\} Q(x_{t+1}|x_t)$$
$$\times a_{t+1}(y_{t+1} \mid x_{t+1}, s_{t+1}) \Big]$$

where $a_t = f_t(y^{t-1}, a^{t-1})$.

At each stage, the system incurs a per-step cost given by

$$c_t(x_t, s_t, a_t, y^t; \mathbf{f}) = \log \frac{a_t(y_t \mid x_t, s_t)}{\mathbb{P}^{\mathbf{f}}(Y_t = y_t \mid Y^{t-1} = y^{t-1})} \quad (4)$$

The objective is to choose a strategy $\mathbf{f} = (f_1, \dots, f_T)$ to minimize the total finite horizon cost given by

$$\tilde{L}_T(\mathbf{f}) = \frac{1}{T} \mathbb{E}^{\mathbf{f}} \left[\sum_{t=1}^{T} c_t(x_t, s_t, a_t, y^t; \mathbf{f}) \right]$$

where the expectation is evaluated with respect to $\mathbb{P}^{\mathbf{f}}$.

Proposition 2. *The sequential decision problem described above is equivalent to Problem A. In particular,*

1) *Given* $\mathbf{q} = (q_1, \dots, q_T) \in \mathcal{Q}_B$, *let* $\mathbf{f} = (f_1, \dots, f_T)$ *be*
$$f_t(y^{t-1}, a^{t-1}) = q_t(\cdot \mid \cdot, \cdot, y^{t-1}).$$
Then $\tilde{L}_T(\mathbf{f}) = L_T(\mathbf{q})$.
2) *Given* $\mathbf{f} = (f_1, \dots, f_T)$, *let* $\mathbf{q} = (q_1, \dots, q_T) \in \mathcal{Q}_B$ *be*
$$q_t(y_t \mid x_t, s_t, y^{t-1}) = a_t(y_t \mid x_t, s_t)$$
where $a_t = f_t(y^{t-1}, a^{t-1})$, *for* $t \in \{1, 2, \dots, T\}$.
Then $L_T(\mathbf{q}) = \tilde{L}_T(\mathbf{f})$.

Proof. The proof relies on the following observations. (i) under the transformations described above, $\mathbb{P}^{\mathbf{f}}$ and $\mathbb{P}^{\mathbf{q}}$ are identical probability distributions; and consequently, (ii) $\mathbb{E}^{\mathbf{f}}[c_t(X_t, S_t, A_t, Y^t; \mathbf{f})] = I^{\mathbf{q}}(S_t, X_t, Y_t \mid Y^{t-1})$. □

C. A dynamic program for the sequential problem

The model described in Section III-B above is similar to a partially observable Markov decision process: the system state (X_t, S_t) is partially observed by a decision maker who chooses A_t. However, in contrast to the standard cost model used in Markov decision processes, the per-step cost depends on the observation history and *past strategy*. Nonetheless, if we consider the belief state as the information state, the problem can be formulated as a standard Markov decision process.

For that matter, for any realization y^{t-1} of past observations and any choice a^{t-1} of past actions, define the belief state $\pi_t \in \mathcal{P}_{X,S}$ as follows: For any $s \in \mathcal{S}$ and $x \in \mathcal{X}$,

$$\pi_{t+1}(x, s) = \mathbb{P}^{\mathbf{f}}(X_t = x, S_t = s | Y^{t-1} = y^{t-1}, A^{t-1} = a^{t-1}).$$

Although, we use $\mathbb{P}^{\mathbf{f}}$ as the probability distribution in the above expression, when (y^{t-1}, a^{t-1}) are given, the probability does not depend on the strategy $\mathbf{f}$. If Y^{t-1} and A^{t-1} are random variables, then the belief state is a $\mathcal{P}_{X,S}$-valued random variable denoted by Π_t.

The belief state Π_t evolves in time in a state-like manner. In particular,

$$\Pi_{t+1} = \varphi(\Pi_t, Y_t, A_t) \quad (5)$$

where φ is a non-linear filter given by

$$\varphi(\pi, a, y)(x_+, s_+)$$
$$= \frac{\sum_{\hat{x}\in\mathcal{X}, \hat{s}\in\mathcal{S}} \mathbb{1}_{s_+}\{\hat{s} - \hat{x} + y\} Q(x_+|\hat{x}) a(y|\hat{x}, \hat{s}) \pi(\hat{x}, \hat{s})}{\sum_{\tilde{x}\in\mathcal{X}, \tilde{s}\in\mathcal{S}} a(y|\tilde{x}, \tilde{s}) \pi(\tilde{x}, \tilde{s})}.$$

The per-step cost defined in (4) can be expressed in terms of the belief state. In particular, for any strategy $\mathbf{f} = (f_1, f_2, \dots)$,

$$\mathbb{E}^f[c_t(X_t, S_t, A_t, Y^t; \mathbf{f}) \mid Y^{t-1} = y^{t-1}, A^t = a^t]$$
$$= \sum_{\substack{x\in\mathcal{X}, s\in\mathcal{S}, \\ y\in\mathcal{Y}}} \pi_t(x, s) a_t(y \mid x, s) \times \log \frac{a_t(y|x, s)}{\sum_{(\tilde{x},\tilde{s})\in\mathcal{X}\times\mathcal{S}} \pi_t(\tilde{x}, \tilde{s}) a_t(y \mid \tilde{x}, \tilde{s})} \quad (6)$$
$$= I(a_t; \pi_t)$$

Note that $I(a_t; \pi_t)$ does not depend on the strategy f.

For ease of notation, for any $a \in \mathcal{A}$ define the Bellman operator $\mathscr{B}_a \colon [\mathcal{P}_{X,S} \to \mathbb{R}] \to [\mathcal{P}_{X,S} \to \mathbb{R}]$ as follows:

$$[\mathscr{B}_a V](\pi) = I(a;\pi) + \sum_{\substack{x\in\mathcal{X}, s\in\mathcal{S},\\ y\in\mathcal{Y}}} \pi(x,s) a(y \mid x, s) V\left(\varphi(\pi, y, a)\right), \ \pi \in \mathcal{P}_{X,S}.$$

Eq. (5) implies that $\{\Pi_t\}_{t\geq 1}$ is a controller Markov process with control action A_t. In addition, equation (6) implies that the expected per-step cost can be expressed in terms of the *state* Π_t and the action A_t. Therefore, from standard results in Markov decision theory, we have the following:

Theorem 1. *The optimal solution of Problem A is as follows:*

1) *For the finite horizon T, iteratively define the following* value functions $V_t \colon \mathcal{P}_{X,S} \to \mathbb{R}$*: For any $\pi \in \mathcal{P}_{X,S}$, $V_{T+1}(\pi) = 0$, and for $t = T, T-1, \dots, 1$,*

$$V_t(\pi) = \inf_{a\in\mathcal{A}} [\mathscr{B}_a V_{t+1}](\pi). \tag{7}$$

Let $f_t^*(\pi)$ denote the arg min of the rhs of (7)*. Then, the strategy $\mathbf{f}^* = (f_1^*, \dots, f_T^*)$ is optimal for the finite horizon version of the sequential decision problem of Sec. III-B. By using the transformation presented in Proposition 2, an optimal strategy for the finite horizon version of Problem A is obtained. Moreover, let $\pi_1(x,s) = P_{X_1}(x) P_{S_1}(s)$. Then, the optimal (finite horizon) leakage rate is given by $V_1(\pi_1)/T$.*

2) *Suppose there exists a $J \in \mathbb{R}$ and $v : \mathcal{P}_{X,S} \to \mathbb{R}$ that satisfy the following fixed point equation:*

$$J + v(\pi) = \inf_{a\in\mathcal{A}} [\mathscr{B}_a v](\pi), \ \forall \pi \in \mathcal{P}_{X,S}. \tag{8}$$

Let $\mathbf{f}^*(\pi)$ denote the arg min of the rhs of (8)*. Then, the time-homogeneous strategy $\mathbf{f}^{*,\infty} = (f^*, f^*, \dots)$ is optimal for the infinite horizon version of the sequential decision problem of Sec. III-B. By using the transformation presented in Proposition 2, a time-homogeneous optimal strategy for the infinite horizon version of Problem A is obtained. Moreover, the optimal (infinite horizon) leakage rate is given by J.*

Remark 1. *The results of Theorem 1 can be generalized to k-th order Markov processes $\{X_t\}_{t\geq 1}$ by considering belief states, actions and per-stage costs of the form*

$$\pi_t(x_{t-k+1}^t, s_t) = \mathbb{P}^{\mathbf{q}}_{X_{t-k+1}^t, S_t \mid Y^t}(x_{t-k+1}^t, s_t \mid y^{t-1})$$
$$a_t(y_t \mid x_{t-k+1}^t, s_t) = q_t(y_t \mid x_{t-k+1}^t, s_t, y^{t-1})$$
$$I(a_t; \pi_t) = I(X_{t-k+1}^t, S_t; Y_t \mid Y^{t-1} = y^{t-1}).$$

The steps leading to the dynamic program follow in a similar manner as before.

IV. The special case of i.i.d. demand

In this section, we make the following assumption:

(A) The demand $\{X_t\}_{t\geq 1}$ is i.i.d. with distribution P_X.

Under this assumption, the belief state π_t can be decomposed into a product form

$$\pi_t(x,s) = P_X(x) P(S_t = s \mid Y^{t-1} = y^{t-1}, A^{t-1} = a^{t-1}).$$

A. Simplification of the dynamic program

Define an auxiliary state variable $W_t = S_t - X_t$ that takes values in $\mathcal{W} = \{s - x : s \in \mathcal{S}, x \in \mathcal{X}\}$. Let $\mathcal{P}_W$ denote the space of probability distributions on $\mathcal{W}$. For any realization (y^{t-1}, a^{t-1}) of past observations and actions, define $\xi_t \in \mathcal{P}_W$ as follows: for any $w \in \mathcal{W}$,

$$\xi_t(w) = \mathbb{P}^{\mathbf{f}}(W_t = w \mid Y^{t-1} = y^{t-1}, A^{t-1} = a^{t-1}).$$

As was the case for π_t, ξ_t does not depend on the choice of the strategy $\mathbf{f}$. If Y^{t-1} and A^{t-1} are random variables, then ξ_t is a $\mathcal{P}_W$-valued random variable that we denote by Ξ_t.

Let us define for any $w \in \mathcal{W}$,

$$\mathcal{D}(w) = \{(x,s) \in \mathcal{X} \times \mathcal{S} : s - x = w\}.$$

Lemma 1. *Under (A), Ξ_t is equivalent to Π_t. In particular,*

1) $\Xi_t(w) = \sum_{(x,s)\in\mathcal{D}(w)} \Pi_t(x,s)$.
2) $\Pi_t(x,s) = P_X(x)\theta(s)$ *where* $\theta(s) = (P_X * \Xi)(s)$.

Lemma 1 implies that $\{\Xi_t\}_{t\geq 1}$ is also a Markov process and that under (A), the dynamic programs of Theorem 1 may be written in terms of Ξ_t rather than Π_t. We now show that an additional simplification of the action space is possible.

Define $\mathcal{B}$ as follows:

$$\mathcal{B} = \left\{ b \in \mathcal{P}_{Y|W} : b(\mathcal{Y}_\circ(w) \mid w) = 1, \ \forall w \in \mathcal{W} \right\}.$$

Lemma 2. *Given $a \in \mathcal{A}$ and $\pi \in \mathcal{P}_{X,S}$ and its associated $\xi \in \mathcal{P}_W$, let us define $b \in \mathcal{B}$ as follows: for all $y \in \mathcal{Y}, w \in \mathcal{W}$*

$$b(y \mid w) = \frac{\sum_{(\hat{x},\hat{s})\in\mathcal{D}(w)} a(y \mid \hat{x}, \hat{s}) \pi(\hat{x}, \hat{s})}{\sum_{(\hat{x},\hat{s})\in\mathcal{D}(w)} \pi(\hat{x}, \hat{s})},$$

and $\tilde{a} \in \mathcal{A}$ as follows: for all $y \in \mathcal{Y}, x \in \mathcal{X}, s \in \mathcal{S}$

$$\tilde{a}(y|x,s) = b(y|s-x).$$

Then under (A), we have

1) *Invariant Transitions: $\varphi(\pi, y, a) = \varphi(\pi, y, \tilde{a}), \ \forall y \in \mathcal{Y}$*
2) *Lower Cost: $I(a;\pi) \geq I(\tilde{a};\pi) = I(b;\xi)$.*

Therefore, in the sequential problem of Sec. III-B, there is no loss of optimality in restricting attention to actions $b \in \mathcal{B}$.

The update of Ξ_t in terms of $b \in \mathcal{B}$ can be written as

$$\Xi_{t+1} = \tilde{\varphi}(\Xi_t, Y_t, B_t)$$

where $\tilde{\varphi}$ is a non-linear filter given by

$$\tilde{\varphi}(\xi, y, b)(w_+) = \frac{\sum_{\hat{w}\in\mathcal{W}} P_X(y + w - w_+) b(y \mid \hat{w}) \xi(\hat{w})}{\sum_{\tilde{w}\in\mathcal{W}} b(y \mid \tilde{w}) \xi(\tilde{w})}.$$

For any $b \in \mathcal{B}$ and $\xi \in \mathcal{P}_W$, let us define the Bellman operator $\tilde{\mathscr{B}}_b : [\mathcal{P}_W \to \mathbb{R}] \to [\mathcal{P}_W \to \mathbb{R}]$ as follows:

$$[\tilde{\mathscr{B}}_b V](\xi) = I(b;\xi) + \sum_{y\in\mathcal{Y}, w\in\mathcal{W}} \xi(w) b(y \mid w) V\big(\tilde{\varphi}(\xi, y, b)\big).$$

Theorem 2. *Under (A), the optimal solution of Problem A is given as follows:*

1) *For the finite horizon case, iteratively define the following* value functions $\tilde{V}_t \colon \mathcal{P}_W \to \mathbb{R}$*: For any* $\xi \in \mathcal{P}_W$, $\tilde{V}_{T+1}(\xi) = 0$*, and for* $t = T, T-1, \dots, 1$,

$$\tilde{V}_t(\xi) = \inf_{b \in \mathcal{B}} [\tilde{\mathscr{B}}_b \tilde{V}_{t+1}](\xi). \tag{9}$$

Let $f_t^\circ(\xi)$ *denote the arg min of the rhs of* (9)*. Then, the strategy* $\mathbf{f}^\circ = (f_1^\circ, \dots, f_T^\circ)$ *is optimal for the finite horizon version of the sequential decision problem of Sec. III-B. By using the transformation presented in Proposition 2 and Lemma 1, an optimal strategy for the finite horizon version of Problem A is obtained. Then, the optimal (finite horizon) leakage rate is given by* $\tilde{V}_1(\xi_1)/T$*, where* $\xi_1(w) = \sum_{(x,s) \in \mathcal{D}(w)} P_X(x) P_{S_1}(s)$.

2) *Suppose there exists a* $\tilde{J} \in \mathbb{R}$ *and* $\tilde{v} : \mathcal{P}_S \to \mathbb{R}$ *that satisfy the following fixed point equation:*

$$\tilde{J} + \tilde{v}(\xi) = \inf_{b \in \mathcal{B}} [\tilde{\mathscr{B}}_b \tilde{v}](\xi), \ \forall \xi \in \mathcal{P}_S. \tag{10}$$

Let $\mathbf{f}^\circ(\xi)$ *denote the arg min of the rhs of* (10)*. Then, the time-homogeneous strategy* $\mathbf{f}^{\circ,\infty} = (f^\circ, f^\circ, \dots)$ *is optimal for the infinite horizon version of the sequential decision problem of Sec. III-B. By using the transformation presented in Proposition 2 and Lemma 1, a time-homogeneous optimal strategy for the infinite horizon version of Problem A is obtained. Moreover, the optimal (infinite horizon) leakage rate is given by* $\tilde{J}$.

B. The solution of the dynamic program

In this section, we identify an optimal strategy and leakage rate under (A). We begin by stating our main theorem then we provide a series of intermediate results that lead to its proof.

Theorem 3. *Define*

$$J^* = \min_{\theta \in \mathcal{P}_S} I(S - X; X) \tag{11}$$

where $X \sim P_X$ *and* $S \sim \theta$*. Let* θ^* *denote the arg min in* (11)*,* $\pi^*(x,s) = P_X(x)\theta^*(s)$ *and* $\xi^*(w) = \sum_{(x,s) \in \mathcal{D}(w)} \pi^*(x,s)$*. Then*

1) J^* *is the optimal (infinite horizon) leakage rate.*
2) *The charging strategy,* $\mathbf{q}^* = (q_1^*, q_2^*, \dots)$ *defined by*

$$q_t^*(y \mid x_t, s_t, y^{t-1}) = b^*(y \mid s_t - x_t), \ \forall t. \tag{12}$$

where

$$b^*(y|w) = \begin{cases} \frac{\pi^*(y, y+w)}{\xi^*(w)} & \text{if } y \in \mathcal{X} \cap \mathcal{Y}_\circ(w) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

achieves the optimal leakage rate in Problem A. □

Remark 2. *The optimal* θ^* *(and therefore the optimal* b^**) in Theorem 3 may be computed using the Blahut-Arimoto algorithm [14].*

We separate the proof into 2 steps: We first obtain a lower bound to the infinite horizon average cost problem by solving the associated average cost optimality inequality. Then we verify that b^* given in the statement of the Theorem achieves this lower bound.

Step 1: Converse - A lower bound on optimal leakage rate

The main idea is to identify a $J \in \mathbb{R}$ and a bounded function $v : \mathcal{P}_W \to \mathbb{R}$ that satisfy the average cost optimality inequality, i.e.,

$$J + v(\xi) \le \min_{b \in \mathcal{B}} [\tilde{\mathscr{B}}_b v](\xi), \ \forall \xi \in \mathcal{P}_W. \tag{14}$$

Then, such a J is a lower bound on the optimal leakage rate.

Lemma 3. *Let* (W_+, W, Y) *be random variables with joint distribution induced by* $b \in \mathcal{B}$ *and* $\xi \in \mathcal{P}_W$ *as such*

$$\mathbb{P}(W_+ = w_+, Y = y, W = w) = P_X(y + w - w_+) b(y|w) \xi(w),$$

and let $v^\circ : \mathcal{P}_W \to \mathbb{R}$ *be given by* $v^\circ(\xi) = H(\xi)$*. Then*

$$[\tilde{\mathscr{B}}_b v^\circ](\xi) = I(W;Y) + H(W_+ \mid Y).$$

Lemma 4. J^* *defined in Theorem 3 and* v° *defined in Lemma 3 satisfy the average cost optimality inequality* (14)*. Therefore,* J^* *is a lower bound on the optimal leakage rate.*

Proof. For any action $b \in \mathcal{B}$ and $\xi \in \mathcal{P}_W$, let (W_+, X_+, S_+, Y, W) be random variables with joint distribution given by

$$\begin{aligned} &\mathbb{P}(W_+ = w_+, X_+ = x_+, S_+ = s_+, W = w, Y = y) \\ &\quad = \mathbb{1}_{w_+}\{s_+ - x_+\} P_X(x_+) \mathbb{1}_{s_+}\{y + w\} b(y|w) \xi(w). \end{aligned}$$

Consider the following of inequalities. For any $\xi \in \mathcal{P}_W$, $b \in \mathcal{B}$,

$$\begin{aligned} [\tilde{\mathscr{B}}_b v^\circ](\xi) - v^\circ(\xi) &\overset{(a)}{=} -H(W|Y) + H(W_+|Y) \\ &\overset{(b)}{=} -H(S_+|Y) + H(S_+ - X_+|Y) \\ &\overset{(c)}{\ge} \min_{\hat{\theta}_+ \in \mathcal{P}_S} -H(\hat{S}_+) + H(\hat{S}_+ - X_+) = J^* \end{aligned}$$

where (a) follows from Lemma 3 and the fact that $I(W;Y) = H(W) - H(W \mid Y)$; (b) follows because $S_+ = Y + W$ and, therefore, $H(S_+|Y) = H(W|Y)$; (c) follows from $H(A|B) \ge \min_{P_{\hat{A}} \in \mathcal{P}_A} H(\hat{A})$ for any joint distribution on (A, B). □

Step 2: Achievability - A strategy that achieves the lower bound

In this section we show that b^* achieves the lower bound J^*. We first identify some properties of b^*.

Lemma 5. *For* b^* *and* ξ^* *defined in Theorem 3:*

1) $b^* \in \mathcal{B}$ *and* $b^*(y|w) > 0, \ \forall y \in \mathcal{X} \cap \mathcal{Y}_\circ(w), w \in \mathcal{W}$.
2) *For any* $y \in \mathcal{X}$, $\tilde{\varphi}(\xi^*, y, b^*) = \xi^*$.
3) $I(b^*; \xi^*) = J^*$.

Note that an immediate implication of Lemma 5 is that if we start at $\Xi_1 = \xi^*$ and use the constant action b^*, then, for all t, $\Xi_t = \xi^*$ and $I(b^*, \Xi_t) = J^*$.

Lemma 6. *Consider the time-homogeneous charging strategy* $\mathbf{f}^{\circ,\infty} = (f^\circ, f^\circ, \dots)$ *where* $f^\circ(\xi) = b^*$ *for all* $\xi \in \mathcal{P}_W$*. Then, under* $\mathbf{f}^{\circ,\infty}$*, for any initial state* ξ_1*, the process* $\{\Xi_t\}_{t \ge 1}$ *converges weakly to* ξ^**. Therefore, for any continuous function* $c : P_W \to \mathbb{R}$ *and initial belief state* ξ_1*, we have that*

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[c(\Xi_t) | \Xi_1 = \xi_1] = c(\xi^*).$$

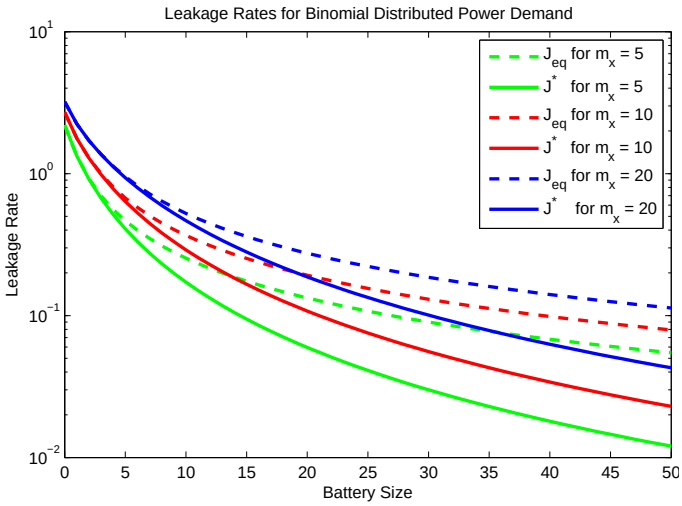

Fig. 2: Information leakage rate under the optimal strategy $\mathbf{q}^*$ and a benchmark strategy $\mathbf{q}_{eq}$ (given by (15)). The corresponding leakage rates are denoted by J^* and J_{eq}. It is assumed that $\mathcal{X} = \mathcal{Y} = \{0, \dots, m_x\}$ and $\mathcal{S} = \{0, \dots, m_s\}$ and that X_t is distributed according to binom$(m_x, 0.5)$. We plot J^* and J_{eq} versus m_s for values of $m_x \in \{5, 10, 20\}$.

Proof. This result relies on a result on convergence of partially observed Markov processes due to Kaijser [15]. □

Now we return to the proof of Theorem 3.

Proof of Theorem 3. By Lemma 6, b^* is an admissible action and by Lemma 7, $f^{*,\infty}$ achieves J^*, which was shown in Lemma 4 to be a lower bound on optimal leakage rate. Therefore, J^* is equal to the optimal leakage rate and $\mathbf{q}^* = (q^*, q^*, \dots)$ where $q^*(y|x,s) = b^*(y|s-x)$. □

C. An example: binomially distributed energy demand

Suppose there are m_x devices that have a probability p of being on or off. Then, $\mathcal{X} = \{0, \dots, m_x\}$ and the demand X_t at any time is binomial(p, m_x). We assume that $\mathcal{Y} = \mathcal{X}$. For a specific value of $\mathcal{S}$, the optimal charging strategy $\mathbf{q}^*$ is given by Theorem 3. In Fig. 2, we plot the information leakage rate under the optimal optimal strategy as a function of battery size for different values of m_x.

We also compare the performance of $\mathbf{q}^*$ with a benchmark strategy $\mathbf{q}_{eq} \in \mathcal{Q}_B$, which is defined as follows: at each t, for any $y \in \mathcal{Y}$, $w \in \mathcal{W}$,

$$q_{eq,t}(y|w) = \frac{\mathbb{1}_{\mathcal{Y}^\circ(w)}\{y\}}{|\mathcal{Y}^\circ(w)|} \tag{15}$$

The information leakage rate under strategy $\mathbf{q}_{eq}$ as a function of battery size is also shown in Fig. 2. Note that for small battery sizes, $\mathbf{q}_{eq}$ is close to optimal but it performs poorly compared to the optimal strategy for large battery sizes.

V. Conclusion

In this paper, we study a privacy-aware smart metering system that uses a rechargeable battery to minimize the information about the user's demand that is leaked to an eavesdropper that observes the energy consumed from the grid. We first identify a structural property of the optimal strategy that allows us to write the mutual information cost in an additive manner. We then identify a Markov decision process that is equivalent to the original system. The optimal charging strategy and the minimum leakage rate are given by the solution of an appropriate dynamic program. For the case of i.i.d. demand, we identify a charging strategy that satisfies the dynamic program.

In this paper, we did not consider the impact of obfuscating information on the utility provider. An important future direction is to consider a model with a cost function at the utility provider and consider the problem of jointly minimizing the cost and information leakage. Note that the results of Theorems 1 and 2 naturally extend to this setup. Extending the result of Theorem 3 will depend on the specific form of the cost function.

References

[1] A. Ipakchi and F. Albuyeh, "Grid of the future", *IEEE Power Energy Mag., vol. 7, no. 2, pp. 52-62*, Mar.-Apr. 2009.

[2] A. Predunzi, "A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel", *Proc. IEEE Power Eng. Society Winter Meeting, New York*, Jan. 2002.

[3] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: towards undetectable appliance load signatures," *Proc. IEEE Smart Grid Comm. Conf.*, Maryland, 2010.

[4] D. Varodayan and A. Khisti, "Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage", *Proc. IEEE Int. Conf. Acoust. Speech Sig. Proc.* Prague, Czech Republic, May 2011.

[5] O. Tan, D. Gunduz and H. V. Poor, "Smart meter privacy in the presence of energy harvesting and storage devices", *Proc. IEEE Int'l Conf. Smart Grid Comm., Tainan City, Taiwan*, Nov 2012.

[6] D. Gunduz and J.Gomez-Vilardebo, "Smart meter privacy in the presence of an alternative energy source", *IEEE Int'l Conf. on Comm. '15 (ICC)*, June 2013.

[7] G. Giaconi, D. Gunduz, H.V. Poor, "Smart meter privacy with an energy harvesting device and instantaneous power constraints," *Proc. '15 IEEE Int'l Conf. on Comm. (ICC)*, vol., no., pp.7216-7221, 8-12 June 2015

[8] L. Yang, X. Chen, J. Zhang and H. V. Poor, "Cost-Effective and Privacy-Preserving Energy Management for Smart Meters," *IEEE Trans. on Smart Grid*, vol. 6, no. 1, pp. 486-495, Jan. 2015.

[9] L. Sankar, S.R. Rajagopalan, and H.V. Poor, "An Information-Theoretic Approach To Privacy," *Proc. 48th Allerton Conf. on Comm., Cont'l, and Comp.*, Monticello, IL, Sep. 2010, pp. 1220-1227.

[10] S. Li, A. Khisti, and A. Mahajan, "Structure of optimal privacy-preserving strategies in smart-metered systems with a rechargeable battery", *Proc. IEEE Int'l Workshop on Sig. Proc. Adv. in Wireless Communications (SPAWC)*, Stockholm, Sweden, June 2015, pp. 1-5.

[11] S. Li, A. Khisti, and A. Mahajan, "Privacy preserving rechargeable battery strategies for smart metering systems", *Proc. Int'l Zurich Seminar on Comm.*, Zurich, Switzerland, March 2-4, 2016.

[12] J. Yao and P. Venkitasubramaniam, "On the Privacy of an In-Home Storage Mechanism", *52nd Allerton Conf. on Comm. Comp. and Cont'l*, Monticello, IL, October 2013.

[13] S. Tatikonda and S. Mitter, "The capacity of channels with feedback", *IEEE Trans. on Inform. Theory*, vol. 55, no. 1, pp. 323-349, January 2009.

[14] R. W. Yeung, "Information Theory and Network Coding," New York, NY: Springer, 2008.

[15] T. Kaijser, "A limit theorem for partially observed Markov chains", Ann. Probab., Vol.3, no. 4, pp. 667-696, 1975.