Smart Meter Privacy with Renewable Energy and a Finite Capacity Battery

Giulio Giaconi and Deniz Gündüz

Department of Electrical and Electronic Engineering, Imperial College London, London, UK

{g.giaconi, d.gunduz}@imperial.ac.uk

Abstract—We address the *smart meter* (SM) privacy problem by considering the availability of a renewable energy source (RES) and a battery which can be exploited by a consumer to partially hide the consumption pattern from the utility provider (UP). Privacy is measured by the mutual information rate between the consumer's energy consumption and the renewable energy generation process, and the energy received from the grid, where the latter is known by the UP through the SM readings, and the former two are to be kept private. By expressing the information leakage as an additive quantity, we cast the problem as a stochastic control problem, and formulate the corresponding Bellman equations.

I. INTRODUCTION

An essential component of a smart grid is the *smart meter* (SM), a device that records minutely the electricity consumption of a household. The adoption of SMs is a key advantage for both utility providers (UPs), who would be able to better monitor the consumption and trade energy with users, and the distribution system operators, who would better manage and run the network. The adoption of SMs is also favourable for consumers, since it allows a time-of-usage pricing with the consequent possibility to reduce electricity costs by choosing less expensive time slots for power-consuming appliances.

However, the high resolution of the data collected by the SMs also makes it possible to infer a consumer's energy consumption *load profile*, i.e., the time series of energy usage collected with regularity from a household. These profiles are extremely valuable, since it is possible to extrapolate sensitive information from them, such as users' habits and presence at home, illnesses or disabilities, the equipments being used, and even which TV channel is being watched [1].

Approaches to address SM privacy in the literature can be broadly classified into two groups: those that modify the SM readings before being sent to the UP, and those that modify the actual user energy demand. In the first group, obfuscation [2], anonymization [3] and aggregation techniques [4] are included. Regarding the second approach, user consumption can be filtered through a storage device, as described in [5], [6], [7], [8] and [9], or also by including an alternative energy source, e.g., a renewable energy source (RES), as in [10], [11] and [12]. In particular, in [8], privacy is evaluated through a Bayesian detection setting and the problem is formulated
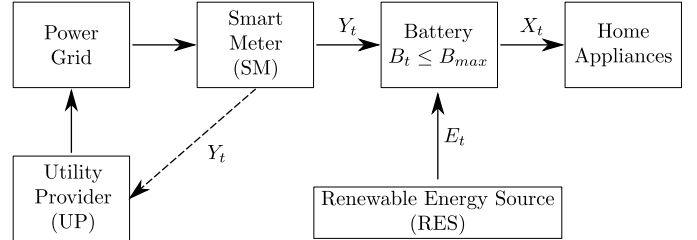
Figure 1. System model. X_t, Y_t, E_t and B_t are the user energy demand, the output load, the amount of renewable energy obtained from a renewable energy source, and the state of the battery at time t, respectively. The dashed line highlights the meter readings being reported to the UP.

as a Markov decision process (MDP). The smart grid state estimation problem is addressed in [13] and the trade-off between differential privacy and the mean distortion of the state estimate is studied. Despite all these efforts, SM privacy is an area of ongoing active research, and wide consensus is still to be reached even on fundamental questions, such as how to measure privacy.

In this paper, we study the SM privacy problem in the presence of a RES together with a rechargeable energy storage device, i.e., a battery. We adopt an information theoretic approach, by minimizing the mutual information between the input load and the renewable energy process, and the output load of the system. Our main contribution here is to cast this problem as an MDP, by finding an additive formulation for the information leakage. We note that a similar approach has been followed in [9], where a battery has been considered. Here, we also consider the presence of a RES to further hide user's consumption. Finally, the corresponding dynamic program (DP) is formulated, which can be solved numerically in order to identify the optimal energy management policy.

The remainder of this paper is organized as follows. In Section II the system model is introduced. In Section III it is assumed that the UP does not know the realizations of the renewable energy process, whereas in Section IV it does. For both cases, the information leakage minimization problem can be cast as an MDP. Conclusions are drawn in Section V.

II. SYSTEM MODEL

We consider the discrete time system depicted in Figure 1. The input load $X_t \in \mathcal{X}$ is the energy requested by the user in time slot t, where $\mathcal{X} = \{0, 1, \ldots X_{\max}\}$. The

output load $Y_t \in \mathcal{Y}$ is the energy received from the UP, where $\mathcal{Y} = \{0, 1, \ldots, Y_{\max}\}$. We assume $\{X_t\}_{t=1}^{\infty}$ to be a first-order time-homogeneous Markov chain with transition probability q_X, whose initial state X_1 is distributed according to p_{X_1}. In time slot t, $E_t \in \mathcal{E} = \{0, 1, \ldots E_{\max}\}$ units of energy are generated by the RES, which becomes available at the beginning of time slot t. $\{E_t\}_{t=1}^{\infty}$ is also a first-order time-homogeneous Markov chain with transition probability q_E and initial state E_1, distributed according to p_{E_1}. We further assume the availability of a battery of capacity $B_{\max}$, whose state of charge at the beginning of time slot t is $B_t \in \{0, 1, \ldots B_{\max}\}$. The initial state of the battery B_1 has distribution p_{B_1}. We assume that all the involved random processes are defined over finite alphabets and that there is a minimum unit of energy such that all the aforementioned quantities are integer multiples of this unit. Moreover, we assume $Y_{\max} \geq X_{\max} + B_{\max}$.

In our model, user demands have to be satisfied at all times:

$$B_t + E_t + Y_t \geq X_t, \qquad \forall t. \tag{1}$$

We do not allow intentional energy waste, or selling energy back to the grid; that is, we impose

$$0 \leq Y_t \leq \left[B_{\max} + X_t - B_t - E_t\right]^+, \qquad \forall t, \tag{2}$$

where $[a]^+ = a$ if $a > 0$, and 0 otherwise. Similarly, we do not allow wasting of renewable energy when the battery is not full. These actions could provide additional privacy to the user, albeit at a significantly higher energy cost. However, energy can still be wasted inevitably, for example, when the battery is full and the input load is smaller than the generated energy. The battery state is updated as

$$B_{t+1} = \min\left\{B_t + E_t - X_t, B_{\max}\right\} + Y_t, \qquad \forall t. \tag{3}$$

Let $W_t \triangleq B_t + E_t - X_t$. The feasible set for Y_t, given $W_t = w_t$, is given by

$$\bar{\mathcal{Y}}(w_t) \triangleq \left\{y_t \in \mathcal{Y} : [-w_t]^+ \leq y_t \leq [B_{\max} - w_t]^+\right\}.$$

An energy management policy $f = (f_1, f_2, \ldots)$ decides on the amount of energy to request from the UP at each time t, given the previous values of input load X^t, renewable energy E^t, battery state B^t, and output load Y^{t-1}. We consider randomized policies, that is, each f_t is a conditional probability $f_t(y_t | x^t, e^t, b^t, y^{t-1})$, with $f \in \mathcal{F}$.

Our goal is to minimize the leakage to the UP of information about user's energy consumption as well as the amount of energy generated by the RES. Accordingly, the information leakage rate induced by policy f over n time slots is

$$\mathcal{I}(n, f) \triangleq \frac{1}{n} I^f(X^n, E^n, B_1; Y^n). \tag{4}$$

Thus, our optimization problem can be written as

$$\mathcal{I}(n) \triangleq \inf_{f \in \mathcal{F}} \frac{1}{n} I^f(X^n, E^n, B_1; Y^n), \tag{5}$$

where the minimization is over all the feasible policies

$f \in \mathcal{F}$. Note that, in the above form, this is an extremely complex optimization problem as we need to specify $p(Y_t | X^t, E^t, B_1, Y^{t-1})$ for every t, and every possible realization $(X^t, E^t, B_1, Y^{t-1}) = (x^t, e^t, b_1, y^{t-1})$.

In the following, we identify two different scenarios based on the information available at the UP regarding the renewable energy process. For both scenarios, we will rewrite (5) in an additive form, which consequently can be formulated as a DP. Given $q_X, q_E, p_{X_1}, p_{E_1}$ and p_{B_1}, our goal is to find an optimal battery charging policy $f = (f_1, f_2, \ldots) \in \mathcal{F}$ that solves (5), for both finite and infinite horizon settings.

III. RENEWABLE ENERGY NOT KNOWN BY THE UP

In this section E^n is treated as a random sequence whose realizations are known causally only to the user. This scenario may occur if, for example, E^n originates from light, vibration, thermal, or biological sources, which could be extremely difficult, if not impossible, for the UP to track.

A. Additive Formulation for the Information Leakage Rate

To formulate the problem as an MDP we need to write the cost function, i.e., the leakage rate, in an additive form. Specifically, we aim at a formulation in which, conditioned on Y^{t-1}, the output load at time t, Y_t, leaks information only on X_t, E_t and B_t, but not on their past values. This will be achieved by restricting the set of possible policies to a smaller set $\mathcal{F}' \subseteq \mathcal{F}$, where each $f' \in \mathcal{F}'$ is a conditional probability of the form $f'_t(y_t | x_t, e_t, b_t, y^{t-1})$, i.e.,

$$f'_t : \mathcal{X} \times \mathcal{E} \times \mathcal{B} \times \mathcal{Y}^{t-1} \to \mathcal{Y}, \qquad \forall t.$$

The joint measure induced by f' is

$$p^{f'}(X^n = x^n, E^n = e^n, B^n = b^n, Y^n = y^n)$$
$$= p_{X_1}(x_1) p_{E_1}(e_1) p_{B_1}(b_1) f'_1(y_1 | x_1, e_1, b_1)$$
$$\times \prod_{t=2}^{n} \left[\mathbb{1}_{b_t}\left\{ \min\{b_{t-1} + e_{t-1} - x_{t-1}, B_{\max}\} + y_{t-1}\right\}\right.$$
$$\left. \times q_X(x_t | x_{t-1}) q_E(e_t | e_{t-1}) f'_t(y_t | x_t, e_t, b_t, y^{t-1})\right],$$

where $\mathbb{1}$ is the indicator function, i.e., $\mathbb{1}_{b_t}\{a\} = 1$ if $b_t = a$, and 0 otherwise. $b_t \in \{0, \ldots, B_{\max}\}$ holds since $Y_t \in \bar{\mathcal{Y}}(w_t)$.

The next theorem, whose proof is omitted due to space limitations, states that this restriction is without loss of optimality.

Theorem 1. *There is no loss of optimality in focusing only on charging strategies $f' \in \mathcal{F}'$, where $\mathcal{F}' \subseteq \mathcal{F}$. Moreover, the minimum information leakage rate can be written in the following additive form*

$$\mathcal{I}(n) = \inf_{f' \in \mathcal{F}'} \frac{1}{n} \sum_{t=1}^{n} I^{f'}(X_t, E_t, B_t; Y_t | Y^{t-1}). \tag{6}$$

It is possible to generalize Theorem 1 to the scenario in which X and E are Markov chains with order higher than 1, i.e., $p_{X_t | X^{t-1}} = p_{X_t | X_{t-m}^{t-1}}$ and $p_{E_t | E^{t-1}} = p_{E_t | E_{t-l}^{t-1}}$. If new policies $f'' \in \mathcal{F}''$ are defined such that

$$f''_t : \mathcal{X}^m \times \mathcal{E}^l \times \mathcal{B} \times \mathcal{Y}^{t-1} \to \mathcal{Y}, \qquad \forall t, \tag{7}$$

where $\mathcal{F}'' \subseteq \mathcal{F}$, then the following corollary holds.

Corollary 1. *Let X and E be Markov chains of order m and l, respectively. There is no loss of optimality in focusing only on charging strategies $f'' \in \mathcal{F}''$, and for those strategies the minimum information leakage rate can be written in the following additive form*

$$\mathcal{I}(n) = \inf_{f'' \in \mathcal{F}''} \frac{1}{n} \sum_{t=1}^{n} I^{f''}(X_{t-m+1}^{t}, E_{t-l+1}^{t}, B_t; Y_t | Y^{t-1}).$$

B. MDP Formulation

Our next goal is to cast the problem as a stochastic control problem, which can be formulated as a DP. For this, we need to specify the state space, the control actions and the instantaneous cost corresponding to state-action pairs. The per-step cost in (6) depends on past observations y^{t-1}, which could be considered as the state of the DP at time t. However, this would mean a state space growing with time. To avoid this, we follow the approach of [14] and [9], and introduce a belief state, which can be shown to replace the y^{t-1} sequence.

The state of the DP at time t is considered to be the causal posterior probability distribution over the triplet $S_t \triangleq (X_t, E_t, B_t)$, given the knowledge of past outputs Y^{t-1}:

$$\beta_t(s_t) \triangleq \begin{cases} p(s_1), & \text{if } t = 1, \\ p(s_t | y^{t-1}), & \text{otherwise.} \end{cases} \tag{8}$$

$\beta_t(s_t)$ can be considered as the *belief* that the UP has about s_t at time t, given its past observations Y^{t-1}.

The control action $U_t \in \mathcal{U}$ is the conditional probability $u_t(y_t | s_t)$. A *randomized history-dependent policy* $\pi = \{\pi_1, \pi_2, \ldots\}$ chooses control action at time t via $u_t = \pi_t(h)$, where h represents the history available to the controller. Thus, the time ordering of the events is $S_1, U_1, Y_1, S_2, U_2, Y_2, \ldots$.

Without loss of optimality, we can focus on randomized *Markov policies* that depend only on the current state β_t, i.e., $\pi_t(\beta_t) = \pi_t(h)$. This holds for both finite and infinite horizon problems under mild assumptions [15]. Policy π induces the following joint measure:

$$p^{\pi}(X^n = x^n, E^n = e^n, B^n = b^n, Y^n = y^n)$$
$$= p_{X_1}(x_1) p_{E_1}(e_1) p_{B_1}(b_1) u_1(y_1 | x_1, e_1, b_1)$$
$$\times \prod_{t=2}^{n} \left[\mathbb{1}_{b_t} \left\{ \min\{b_{t-1} + e_{t-1} - x_{t-1}, B_{\max}\} + y_{t-1} \right\} \right.$$
$$\left. \times q_X(x_t | x_{t-1}) q_E(e_t | e_{t-1}) u_t(y_t | x_t, e_t, b_t) \right].$$

The state can be updated recursively, i.e., $\beta_{t+1} = \phi(\beta_t, u_t)$

$$\beta_{t+1}(s_{t+1}) = p(s_{t+1} | y^t),$$
$$= \sum_{s_t} p(s_t, s_{t+1} | y^t),$$
$$= \sum_{s_t} \frac{p(s_t, s_{t+1}, y_t | y^{t-1})}{p(y_t | y^{t-1})},$$

$$\stackrel{(a)}{=} \frac{\sum_{s_t} p(s_t | y^{t-1}) p(y_t | s_t, y^{t-1}) p(s_{t+1} | s_t, y_t)}{\sum_{s_t, s_{t+1}} p(s_t | y^{t-1}) p(y_t | s_t, y^{t-1}) p(s_{t+1} | s_t, y_t)},$$

$$\stackrel{(b)}{=} \frac{\sum_{s_t} \beta_t(s_t) u_t(y_t | s_t) q_X(x_{t+1} | x_t) q_E(e_{t+1} | e_t)}{\sum_{s_t, s_{t+1}} \beta_t(s_t) u_t(y_t | s_t) q_X(x_{t+1} | x_t) q_E(e_{t+1} | e_t)}$$

$$\times \frac{\mathbb{1}_{b_{t+1}} \left\{ \min\{b_t + e_t - x_t, B_{\max}\} + y_t \right\}}{\mathbb{1}_{b_{t+1}} \left\{ \min\{b_t + e_t - x_t, B_{\max}\} + y_t \right\}}, \tag{9}$$

where (a) follows from Bayes rule and the Markov chain $Y^{t-1} \to (S_t, Y_t) \to S_{t+1}$; and (b) is due to the definitions of β_t and u_t.

Given $Y^{t-1} = y^{t-1}$, the per-step cost of taking action u_t when $s_t = (x_t, e_t, b_t)$ is

$$g_t(x_t, e_t, b_t, u_t, y^t) \triangleq \log \frac{u_t(y_t | x_t, e_t, b_t)}{p(y_t | y^{t-1})}. \tag{10}$$

It is possible to show that this new formulation is equivalent to the original problem, by considering the average n-horizon cost and the knowledge of y^{t-1} as in $\mathcal{I}(n, f')$

$$\mathcal{I}(n, \pi) = \frac{1}{n} \mathbb{E}^{\pi} \left[\sum_{t=1}^{n} g_t(x_t, e_t, b_t, u_t, y^t) \right],$$
$$= \frac{1}{n} \sum_{t=1}^{n} \sum_{s_t \in \mathcal{S}, y^t \in \mathcal{Y}^t} p(s_t, y_t | y^{t-1}) \log \frac{u_t(y_t | s_t)}{p(y_t | y^{t-1})},$$
$$= \mathcal{I}(n, f'), \tag{11}$$

where we remind that u_t is also a function of y^{t-1}, since $u_t = \pi_t(\beta_t)$. Given a policy π, β_t and u_t are determined by y^{t-1}. The average information leakage at time t is

$$\mathbb{E}^{\pi} \left[g_t(X_t, E_t, B_t, U_t, Y^t) \right]$$
$$= I(X_t, E_t, B_t; Y_t | Y^{t-1} = y^{t-1}),$$
$$= \sum_{\substack{x_t \in \mathcal{X}, e_t \in \mathcal{E} \\ b_t \in \mathcal{B}, y^t \in \mathcal{Y}^t}} p(x_t, e_t, b_t | y^{t-1}) p(y_t | x_t, e_t, b_t, y^{t-1})$$

$$\times \log \frac{u_t(y_t | x_t, e_t, b_t)}{p(y_t | y^{t-1})},$$

$$= \sum_{s_t \in \mathcal{S}, y_t \in \mathcal{Y}} \beta_t(s_t) u_t(y_t | s_t) \log \frac{u_t(y_t | s_t)}{\sum_{\tilde{s}_t \in \mathcal{S}} \beta_t(\tilde{s}_t) u_t(y_t | \tilde{s}_t)},$$

$$= I(X_t, E_t, B_t; Y_t | \beta_t, u_t), \tag{12}$$

where the last step confirms that $Y^{t-1} \to (\beta_t, u_t) \to (X_t, E_t, B_t, Y_t)$ is a Markov chain.

The following lemma summarizes the results of this section.

Lemma 1. *Without loss of optimality, the SM privacy problem (5) for the scenario in which the UP does not know the realizations of the renewable energy process can be modeled as an MDP, such that*

1) *the state at time t is given by (8),*
2) *the action at time t is specified by $u_t(y_t | x_t, b_t, e_t)$,*
3) *and the instantaneous cost is given by (12).*

In order to formulate the Bellman equations, it is convenient

to first define an operator T for the DP as follows

$$(TJ)(\beta) = g(s, \pi(\beta), \beta) + \sum_{s \in \mathcal{S}, y \in \mathcal{Y}} \beta(s)u(y|s)J(\phi(\beta, u)),$$

for $\beta \in \bar{\mathcal{B}}$, where $J : \bar{\mathcal{B}} \to \mathbb{R}$ is the value function.

For the finite horizon setting, let J_t denote the value function at time $t \leq n$, with $J_{n+1} = 0$. For $t \leq n$, we have

$$J_t(\beta) = \inf_{u \in \mathcal{U}}[TJ_{t+1}](\beta). \tag{13}$$

The minimization problem is solved by going backwards in time from $t = n$ to $t = 1$ in order to find the optimal policy $\pi^* = (\pi_1^*, \pi_2^*, \ldots, \pi_n^*)$ that minimizes (13) for every t.

In the infinite horizon scenario, since the total information leakage over an infinite number of stages is generally infinite, we minimize the average information leakage per stage, i.e.,

$$J(\beta_0) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}^\pi \left[\sum_{t=0}^{n-1} g_t(s_t, \pi_t(\beta_t), \beta_t) \right]. \tag{14}$$

The solution for the infinite horizon problem can be determined as the solution to the following Bellman equation

$$\lambda + J(\beta) = \inf_{u \in \mathcal{U}}[TJ](\beta), \tag{15}$$

where $\lambda \in \mathbb{R}$ is the optimal average information leakage, and the vector $J(\beta)$ is the relative or differential privacy leakage, i.e., the difference of the expected leakage to reach a conventional state and the cost that would be incurred if the cost per stage was equal to λ for all states. Via efficient dynamic programming algorithms, e.g., value iteration and policy iteration [15], (15) can be solved and an optimal stationary policy $\pi^* = (\pi^*, \pi^*, \ldots, \pi^*)$ can be found.

Proposition 1. *The value functions $\{J_t\}_{t=1}^n$ are concave.*

Concave value functions allow the use of convex optimization algorithms. Finally, the following corollary generalizes our result to an input load and a RES with larger memory.

Corollary 2. *Let X and E be Markov chains of order m and l, respectively. Let $S_t \triangleq (X_{t-m+1}^t, E_{t-l+1}^t, B_t)$. The previous steps follow also for this scenario, where*
  1) *the state is $\beta_t(s_t) \triangleq p(s_t|y^{t-1})$,*
  2) *the action is $u_t(y_t|s_t)$,*
  3) *and the cost is $g_t(s_t, u_t, y^t) \triangleq \log \frac{u_t(y_t|s_t)}{p(y_t|y^{t-1})}$.*

IV. RENEWABLE ENERGY KNOWN BY THE UP

Here we assume that the UP knows the realizations of the renewable energy process E^n, as shown in Figure 2. This scenario can occur if we consider solar energy as the RES, and the UP can accurately estimate the renewable energy produced from its own observations in nearby locations, weather forecast of the area, and the specifications of the solar panel.

The goal is to find a battery charging policy $f \in \mathcal{F}$ that minimizes the following information leakage rate

$$\mathcal{I}(n, f) = \frac{1}{n}I(X^n, E^n, B_1; Y^n|E^n),$$
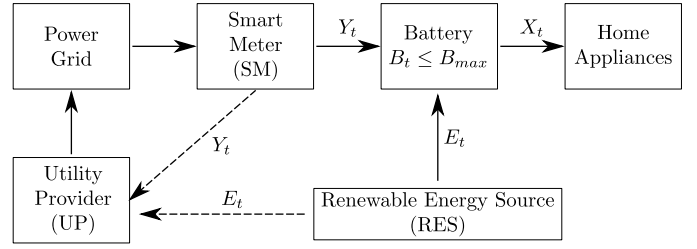$$= \frac{1}{n}I(X^n, B_1; Y^n|E^n), \tag{16}$$

Figure 2. System model. In this scenario, information about the realizations of the renewable energy process is available to the UP.

where the charging policy at time t is

$$f_t : \mathcal{X}^t \times \mathcal{E}^t \times \mathcal{B}^t \times \mathcal{Y}^{t-1} \to \mathcal{Y}, \qquad \forall t.$$

A. Additive Formulation for the Information Leakage Rate

Similarly to Section III, we want to express the problem in an additive form. We define policies $f_t'(y_t|x_t, e^t, b_t, y^{t-1})$ as

$$f_t' : \mathcal{X} \times \mathcal{E}^t \times \mathcal{B} \times \mathcal{Y}^{t-1} \to \mathcal{Y}, \qquad \forall t, \tag{17}$$

and state the following theorem.

Theorem 2. *There is no loss of optimality in focusing only on charging strategies $f' \in \mathcal{F}'$, where $\mathcal{F}' \subseteq \mathcal{F}$. Moreover, the minimum information leakage rate can be written in the following additive form*

$$\mathcal{I}(n) = \inf_{f' \in \mathcal{F}'} \frac{1}{n} \sum_{t=1}^n I^{f'}(X_t, B_t; E_t, Y_t|E^{t-1}, Y^{t-1}). \tag{18}$$

The proof follows similar steps to that of Theorem 1.

For the setting in which X and E are Markov processes of order m and l, respectively, we define policies f'' such that

$$f_t'' : \mathcal{X}^m \times \mathcal{E}^t \times \mathcal{B} \times \mathcal{Y}^{t-1} \to \mathcal{Y}, \qquad \forall t. \tag{19}$$

Then the following corollary holds.

Corollary 3. *Let X and E be Markov chains of order m and l, respectively. There is no loss of optimality in focusing only on charging strategies $f'' \in \mathcal{F}''$, and for those strategies the minimum information leakage rate can be written in the following additive form*

$$\mathcal{I}(n) = \inf_{f'' \in \mathcal{F}''} \frac{1}{n} \sum_{t=1}^n I^{f''}(X_{t-m+1}^t, B_t; E_t, Y_t|Y^{t-1}, E^{t-1}).$$

B. MDP Formulation

As in Section III-B we specify the state space, control actions and the instantaneous cost. The state of the DP at time t is considered to be the causal posterior probability distribution over $S_t \triangleq (X_t, B_t)$, given the knowledge of Y^{t-1} and E^{t-1}:

$$\beta_t(s_t) \triangleq \begin{cases} p(s_1), & \text{if } t = 1, \\ p(s_t|y^{t-1}, e^{t-1}), & \text{otherwise.} \end{cases} \tag{20}$$

$\beta_t(s_t)$ has again the interpretation of *belief* that the UP has about s_t at time t, given (Y^{t-1}, E^{t-1}).

The action $U_t \in \mathcal{U}$ is the conditional probability $u_t(y_t|s_t, e_t)$ given by $u_t = \pi_t(h)$. As before, we consider without loss of optimality Markov policies $\pi = \{\pi_1, \pi_2, \ldots\}$ that depend only on the current state β_t, i.e., $\pi_t(\beta_t) = \pi_t(h)$. As in (9), β is updated recursively, i.e., $\beta_{t+1} = \phi'(\beta_t, u_t)$. e_t is not in the belief as the UP has perfect knowledge about it.

We follow steps similar to those of Section III-B, and define the cost of taking action u_t when $s_t = (x_t, b_t)$ as

$$g_t(x_t, e^t, b_t, u_t, y^t) \triangleq \log \frac{q_E(e_t|e_{t-1})u_t(y_t|x_t, e_t, b_t)}{p(y_t, e_t|y^{t-1}, e^{t-1})}. \quad (21)$$

By considering the average n-horizon cost, it is possible to show that this formulation is equivalent to the original problem

$$\mathcal{I}(n, \pi) = \frac{1}{n}\mathbb{E}^\pi\left[\sum_{t=1}^n g_t(x_t, e^t, b_t, u_t, y^t)\right],$$

$$= \frac{1}{n}\sum_{t=1}^n \sum_{\substack{s_t \in \mathcal{S} \\ y^t \in \mathcal{Y}^t, e^t \in \mathcal{E}^t}} p(s_t, e_t, y_t|y^{t-1}, e^{t-1})g_t(s_t, e^t, u_t, y^t),$$

$$= \mathcal{I}(n, f'), \quad (22)$$

where u_t is also a function of (y^{t-1}, e^{t-1}) since $u_t = \pi_t(\beta_t)$. Given a policy π, β_t and u_t are determined by (y^{t-1}, e^{t-1}). Then, it is possible to write

$$\mathbb{E}^\pi\left[g_t(X_t, E^t, B_t, U_t, Y^t)\right]$$

$$= I(X_t, B_t; E_t, Y_t|Y^{t-1} = y^{t-1}, E^{t-1} = e^{t-1}),$$

$$= \sum_{\substack{x_t \in \mathcal{X}, e^t \in \mathcal{E}^t \\ b_t \in \mathcal{B}, y^t \in \mathcal{Y}^t}} p(x_t, e_t, b_t|y^{t-1}, e^{t-1})p(y_t|x_t, b_t, e^t, y^{t-1})$$

$$\times \log \frac{q_E(e_t|e_{t-1})u_t(y_t|x_t, e_t, b_t)}{p(y_t, e_t|y^{t-1}, e^{t-1})},$$

$$= \sum_{\substack{s_t \in \mathcal{S}, y_t \in \mathcal{Y} \\ e_t \in \mathcal{E}}} \beta_t(s_t)q_E(e_t|e_{t-1})u_t(y_t|s_t, e_t)$$

$$\times \log \frac{q_E(e_t|e_{t-1})u_t(y_t|s_t, e_t)}{\sum_{\tilde{s}_t \in \mathcal{S}} \beta_t(\tilde{s}_t)q_E(e_t|e_{t-1})u_t(y_t|\tilde{s}_t, e_t)},$$

$$= I(X_t, B_t; E_t, Y_t|\beta_t, q_E, u_t), \quad (23)$$

where the last step confirms that $(Y^{t-1}, E^{t-1}) \to (\beta_t, q_E, u_t) \to (X_t, E_t, B_t, Y_t)$ is a Markov chain.

The following lemma summarizes the results of this section.

Lemma 2. *Without loss of optimality, the SM privacy problem (5) for the scenario in which the UP knows the realizations of the process E^n can be modeled as an MDP, such that*

1) *the state at time t is given by (20),*
2) *the action at time t is specified by $u_t(y_t|x_t, e_t, b_t)$,*
3) *and the instantaneous cost is given by (23).*

Bellman equations for the finite and infinite horizon problems can be obtained as in Section III, with the consequent changes in the formulations of β, u and π.

A final corollary, counterpart of Corollary 2, holds.

Corollary 4. *Let X and E be Markov chains of order m and l, respectively, $S_t \triangleq (X_{t-m+1}^t, B_t)$, and q_E^l the l-th order*

transition probability. The previous passages follow also for this scenario, where

1) *the state is $\beta_t(s_t) \triangleq p(s_t|y^{t-1}, e^{t-1})$,*
2) *the action is $u_t(y_t|s_t, e_t)$,*
3) *the cost is $g_t(s_t, u_t, y^t, e^t) \triangleq \log \frac{q_E^l(e_t|e_{t-l}^{t-1})u_t(y_t|s_t, e_t)}{p(y_t, e_t|y^{t-1}, e^{t-1})}$.*

V. CONCLUSIONS

We have studied the information leakage rate in an SM system by considering the availability of a RES and a finite capacity battery at the consumer side. The minimum information leakage rate has been characterized for both the scenario in which the UP does not know the realizations of the renewable energy process, and the scenario in which the UP knows them. For both scenarios, we have formulated the minimum information leakage rate as an additive cost function, and cast the problem as an MDP, thereby finding the expressions for the corresponding Bellman equations. The optimal leakage rate for a given scenario can be obtained by discretizing the continuous belief state and applying dynamic programming techniques.

REFERENCES

[1] U. Greveler, P. Glösekötter, B. Justus, and D. Loehr, "Multimedia content identification through smart meter power usage profiles," in *Proc. Int. Conf. on Inform. and Knowledge Eng.*, Las Vegas, NV, USA, Jul. 2012.

[2] Y. Kim, E. Ngai, and M. Srivastava, "Cooperative state estimation for preserving privacy of user behaviors in smart grid," in *Proc. IEEE Int. Conf. on Smart Grid Commun.*, Brussels, Belgium, Oct. 2011, pp. 178–183.

[3] C. Efthymiou and G. Kalogridis, "Smart grid privacy via anonymization of smart metering data," in *Proc. IEEE Int. Conf. on Smart Grid Commun.*, Gaithersburg, MD, USA, Oct. 2010, pp. 238–243.

[4] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Proc. IEEE Int. Conf. on Commun.*, Cape Town, South Africa, May 2010, pp. 1–5.

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

[6] J. Koo, X. Lin, and S. Bagchi, "Privatus: Wallet-friendly privacy protection for smart meters," in *Proc. European Symp. on Research in Comput. Security*, Pisa, Italy, Sep. 2012, pp. 343–360.

[7] D. Varodayan and A. Khisti, "Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Prague, Czech Republic, May 2011, pp. 1932–1935.

[8] L. Zuxing, T. J. Oechtering, and M. Skoglund, "Privacy-preserving energy flow control in smart grids," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Shanghai, China, Mar. 2016.

[9] S. Li, A. Khisti, and A. Mahajan, "Privacy-optimal strategies for smart metering systems with a rechargeable battery," *arXiv:1510.07170 [cs.IT]*, Oct. 2015.

[10] O. Tan, D. Gündüz, and H. V. Poor, "Increasing smart meter privacy through energy harvesting and storage devices," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1331–1341, Jul. 2013.

[11] J. Gomez-Vilardebo and D. Gündüz, "Privacy of smart meter systems with an alternative energy source," in *Proc. IEEE Int. Symp. on Inform. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2572–2576.

[12] G. Giaconi, D. Gündüz, and H. V. Poor, "Smart meter privacy with an energy harvesting device and instantaneous power constraints," in *Proc. IEEE Int. Conf. on Commun.*, London, UK, Jun. 2015, pp. 7216–7221.

[13] H. Sandberg, G. Dáan, and R. Thobaben, "Differentially private state estimation in distribution networks with smart meters," *arXiv:1503.08490 [math.OC]*, Mar. 2015.

[14] S. Tatikonda and S. Mitter, "The capacity of channels with feedback," *IEEE Trans. on Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009.

[15] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. II*, 3rd ed. Belmont, MA, USA: Athena Scientific, 2007.