PRIVACY-COST TRADE-OFF IN A SMART METER SYSTEM WITH A RENEWABLE ENERGY SOURCE AND A RECHARGEABLE BATTERY

Ecenaz Erdemir, Pier Luigi Dragotti and Deniz Gündüz

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

ABSTRACT

We study the privacy-cost trade-off in a smart meter (SM) system with a renewable energy source (RES) and a finite-capacity rechargeable battery (RB). Privacy is measured by the mutual information rate between the energy demand and the energy received from the grid, where the latter also determines the cost, and hence, reported by the SM to the utility provider (UP). We consider a renewable energy generation process that fully charges the RB at random time instants, and its realization is assumed to be known also by the UP. We reformulate the problem as a Markov decision process (MDP), and solve it by dynamic programming (DP) to design battery charging and discharging policies that minimize a linear combination of the privacy leakage and energy cost. We also propose a lower bound and two alternative low-complexity energy management policies, one of which is shown numerically to perform close to the MDP solution.

Index Terms— Home energy management, Markov decision processes, privacy, smart meter.

1. INTRODUCTION

Smart meters (SMs) are essential components of smart grids: they collect real-time consumption data of a household, and report it to the utility provider (UP). SM measurements can be used for time-of-use pricing, trading user-generated energy, and mitigating load variations [1]. However, SM readings can also reveal details about consumer's private activities, which they may not want to share with the UP. Various techniques have been proposed in the literature to enable SM privacy [2–13], which can be categorized as those based on SM data manipulation, and those based on demand shaping. While the former focuses on modifying SM measurements [2, 3], the latter directly manipulates energy consumption exploiting physical resources, such as a rechargeable battery (RB) [4–8] or a renewable energy source (RES) [9–13]. Manipulating the SM readings reduces the relevance of the reported vaues for grid management and load prediction, limiting the benefits of SMs. Moreover, the grid operator can place sensors outside a household and obtain the real consumption data, since they own and control the infrastructure.

On the other hand, demand shaping tackles these issues by manipulating the real consumption. In demand shaping, instantaneous demand of a user can be supplied partially by the power grid, as the rest can be provided by the RB or RES. This effectively filters the energy consumption time series, reducing the information leakage to the UP.

In [7], information theoretic privacy with a RB is formulated as a Markov decision process (MDP). Markovian energy demand is considered, and the minimum leakage is obtained numerically through dynamic programming (DP), while a single-letter expression is obtained for an independent and identically distributed (i.i.d) demand. This approach is extended to the scenario with a RES in [9]. In [8], privacy-cost trade-off is examined with a RB. Due to Markovian demand and price processes, the problem is formulated as a partially observable MDP with belief-dependent rewards (ρ-POMDP), and solved by DP for infinite-horizon.

We consider the privacy-cost trade-off of a SM system with both a RB and a RES [8, 14]. We measure privacy as the mutual information rate between the user demand and the energy received from the grid. We define the cost as the average amount of energy received from the grid, and study the trade-off between the cost and privacy by setting their weighted sum as the objective function. We formulate the problem as a MDP with a continuous belief state, and solve it numerically by DP by quantizing the belief state. Our contribution with respect to [8] is the inclusion of a RES into the system, which provides additional privacy. While [9] also considers RB and RES jointly, here we study the privacy-cost trade-off, and present numerical solutions focusing on a particular renewable energy arrival process that fully recharges the RB at random time instances. We also provide two low-complexity policies and a lower bound, which exploit the special structure of the renewable energy process. We show numerically that the policy targeting a fixed recharge period performs very close to the infinite-horizon MDP solution, providing a low-complexity alternative for practical systems.

2. SYSTEM MODEL

We consider a discrete time system model, illustrated in Fig. 1, in which the energy demand of the user and energy requested from the grid at time slot t are denoted by $X_t \in \mathcal{X}$ and $Y_t \in \mathcal{Y}$, respectively, where $(|\mathcal{X}|, |\mathcal{Y}| < \infty)$. The RB state of charge at the beginning of time slot t is denoted by
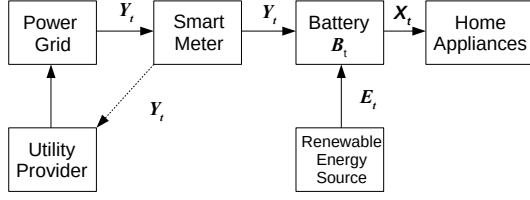
Fig. 1. Illustration of the SM system model.

$B_t \in \mathcal{B}:=\{0, \ldots, B_{max}\}$. The RB charging and discharging process is assumed ideal without any losses (see [6] for a model with energy losses). $E_t \in \mathcal{E}:=\{0, B_{max}\}$ units of energy is generated by the RES at the beginning of each time slot t; that is, when the renewable energy arrives, it completely recharges the RB, and it can be used by the appliances only through the RB. The E_t process is assumed to be independent of X_t, and known by the UP. We assume that X_t and E_t are i.i.d. with distributions P_X and P_E, respectively.

The appliances' energy demand is always satisfied; that is, $E_t + B_t + Y_t \geq X_t$, $\forall t$. In addition, intentional energy waste to provide privacy, or selling energy to the grid are not allowed. Therefore, the battery state of charge is updated as

$$B_{t+1} = \min(E_t + B_t, B_{max}) + Y_t - X_t, \qquad \forall t, \quad (1)$$

where Y_t is chosen such that $B_{t+1} \leq B_{max}$.

The amount of energy requested from the grid is determined by a randomized battery charging policy $\boldsymbol{q} = \{q_t\}_{t=1}^{\infty}$, where q_t is a conditional probability distribution $q_t(Y_t|X^t, B^t, E^t, Y^{t-1})$, which randomly decides on the amount of energy received from the grid at time t, given the histories of demand $X^t:=\{X_1, \ldots, X_t\}$, battery charge B^t, energy generation E^t, and grid energy Y^{t-1}. Our goal is to find an energy management policy, $\{q_t^*\}_{t=1}^{\infty}$, which provides the best trade-off between the privacy and cost.

2.1. Privacy Measure
We measure the privacy of a policy $\boldsymbol{q} = \{q_t\}_{t=1}^{T}$ over T time slots by the *information leakage rate*, $L_{\boldsymbol{q}}(T)$, defined as the average mutual information between the demand side load X^T and initial RB charge B_1, and the SM readings Y^T:

$$L_{\boldsymbol{q}}(T) := \frac{1}{T} I(X^T, B_1; Y^T|E^T), \quad (2)$$

where E^T is known by the UP. It can be shown, similarly to [7], that there is no loss of optimality in considering policies of the form $q_t(Y_t|X_t, B_t, E^t, Y^{t-1})$; that is, it is sufficient to consider only the current demand and battery state. Hence, (2) can be rewritten in an additive form

$$L_{\boldsymbol{q}}(T) = \frac{1}{T} \sum_{t=1}^{T} I(X_t, B_t; E_t, Y_t|Y^{t-1}, E^{t-1}). \quad (3)$$

The Markovity of optimal actions and the additive objective function in (3) allow us to represent the privacy component of our problem as a MDP with state $S_t=\{B_t, X_t\} \in \mathcal{S}$. However, the leakage at time t depends on Y^{t-1} and E^{t-1},

resulting in a growing state space in time. Therefore, a belief state $\beta_t(s_t)$ is defined as the causal posterior probability distribution over the state space given Y^{t-1} and E^{t-1}:

$$\beta_t(s_t) = P^{\boldsymbol{q}}(S_t = s_t|Y^{t-1} = y^{t-1}, E^{t-1} = e^{t-1}). \quad (4)$$

The control actions chosen by randomized policies are the conditional probabilities of energy received from the grid given the state and belief, and denoted by $a_t(y_t|s_t, e_t) = P^{\boldsymbol{q}}(Y_t = y_t|S_t=s_t, E_t=e_t, \beta_t)$ [9].

We follow the approach in [9] for updating the belief state, and define the per-step leakage of taking action $a_t(y_t|s_t, e_t)$ which is incurred by the policy $\boldsymbol{q}$ at each step as,

$$l_t(s_t, e^t, a_t, y^t; \boldsymbol{q}) := \log \frac{a_t(y_t|s_t, e_t)P_E}{P^{\boldsymbol{q}}(y_t, e_t|y^{t-1}, e^{t-1})}. \quad (5)$$

The average leakage rate over a finite-horizon T, $\frac{1}{T}\mathbb{E}_{\boldsymbol{q}}[\sum_{t=1}^{T} l_t(s_t, e^t, a_t, y^t)]$, is equal to the original formulation in (3). Given belief and action probabilities, average information leakage rate at time t is formulated as,

$$\mathbb{E}_{\boldsymbol{q}}[l_t(s_t, e^t, a_t, y^t)] = \sum_{\substack{s_t \in \boldsymbol{S}, e_t \in \boldsymbol{E} \\ y_t \in \boldsymbol{Y}}} \beta_t(s_t)a_t(y_t|s_t, e_t)P_E$$

$$\times \log \frac{a_t(y_t|s_t, e_t)P_E}{\sum_{\hat{s}_t \in \boldsymbol{S}} \beta_t(\hat{s}_t)a_t(y_t|\hat{s}_t, \hat{e}_t)P_E}$$

$$:= \mathcal{L}(\beta_t, a_t). \quad (6)$$

2.2. Energy Cost
Energy cost is defined as the average amount of energy received from the grid over T time slots,

$$C_{\boldsymbol{q}}(T) := \frac{1}{T} \sum_{t=1}^{T} Y_t. \quad (7)$$

We remark that, differently from [8], we do not consider time-varying energy unit cost, although our model can easily be extended in this direction. In the context of [8], i.e., in the absence of a RES, our cost model would result in a deterministic energy cost, independent of policy q_t. However, in the presence of a RES, our cost model follows [12], and incentivizes the maximum exploitation of locally generated renewable energy. For example, when privacy is not a concern, cost minimizing policy would use battery energy first, to be able to store the arriving renewable energy as much as possible. The energy cost averaged over a finite-horizon T is simply $\frac{1}{T}\mathbb{E}_{\boldsymbol{q}}[\sum_{t=1}^{T} y_t]$. The average per-step cost can be represented in terms of belief and action probabilities as follows:

$$\mathcal{C}(\beta_t, a_t) := \mathbb{E}_{\boldsymbol{q}}[y_t] = \sum_{\substack{s_t \in \boldsymbol{S}, e_t \in \boldsymbol{E}, \\ y_t \in \boldsymbol{Y}}} \beta_t(s_t)P_E a_t(y_t|s_t, e_t)y_t$$

2.3. Weighted Total Privacy Leakage and Energy Cost
We have two distinct performance measures, which are not necessarily aligned. Therefore, we define the objective function as the weighted sum of the information leakage rate and the average cost over all the feasible policies, as $T \to \infty$,

$$U^*(\gamma) = \lim_{T \to \infty} \inf_{\boldsymbol{q}} [\gamma L_{\boldsymbol{q}}(T) + (1 - \gamma)C_{\boldsymbol{q}}(T)], \quad (8)$$

where $\gamma \in [0, 1]$ is determined by the user according to her preference regarding privacy and cost. Our goal is to design

a policy $q^*(\gamma)$, which is the argmin of the right hand side of (8), satisfying the energy management rules. This problem can be modeled as an MDP with state $\beta_t(s_t)$ and action $a_t(y_t|s_t, e_t)$. The corresponding Bellman equations can be written similarly to [7] and [9]. We include the instantaneous weighted objective function, $\mathcal{U}_\gamma(\beta_t, a_t) = [\gamma\mathcal{L}(\beta_t, a_t) + (1 - \gamma)\mathcal{C}(\beta_t, a_t)]$, into the Bellman operator,

$$[T_a v](\beta) = \mathcal{U}_\gamma(\beta, a) + \sum_{\substack{s\in S, e\in E \\ y\in Y}} \beta(s)a(y|s, e)P_E v(\phi(\beta, y, a)), \quad (9)$$

where v is the value function and the updated belief state is represented by $\beta_{t+1} = \phi(\beta_t, y_t, a_t)$. The mplementation of DP for infinite-horizon is as follows:

- For λ constant [15], the value function v is time-homogeneous and defined iteratively:

$$\lambda + v(\beta) = \min_a [T_a v](\beta). \quad (10)$$

- Time-homogeneous optimal policy, $\mathbf{q}^* = (q^*, q^*, \dots)$,

$$q^*(y_t|s_t, e_t, \beta_t) = a(y_t|s_t, e_t). \quad (11)$$

While an exact DP solution cannot be achieved due to the continuous belief, we provide an approximate numerical solution. To be able to solve the problem numerically by DP, we discretize the belief $\beta_t(s_t)$. At each value iteration, we quantize the updated belief, $\beta_{t+1}(s_{t+1})$, by rounding it to the closest discrete belief value.

3. LOW-COMPLEXITY POLICIES

Due to the special renewable energy generation process we consider here, the problem is an episodic MDP, which resets to an initial state of full RB at every renewable energy instant. Between two consecutive energy arrivals, energy transitions occur only between the grid, RB and home appliances. Hence, for each time period between two charging instances, the system can be modeled as a SM with only a RB and no RES. Accordingly, we formulate a finite-horizon privacy-cost trade-off problem for a SM system with an initially full RB, which will be used to propose a low-complexity policy as well as a lower bound for the original problem.

In the finite-horizon problem with no RES, as before, the user demand is always satisfied by imposing $Y_t + B_t \geq X_t$, $\forall t$, and RB charge is updated by $B_{t+1} = B_t + Y_t - X_t, \forall t$. Randomized battery charging policies, q_t, are of the form $q_t(Y_t|X_t, B_t, Y^{t-1})$. The information leakage rate induced by the policy $\mathbf{q}$ over a finite-horizon between two consecutive energy arrivals is given by,

$$\bar{L}_{\mathbf{q}}(T) := \frac{1}{T}I(X^T, B_1; Y^T). \quad (12)$$

As in Section 2, (12) can be written in an additive form:

$$\bar{L}_{\mathbf{q}}(T) = \frac{1}{T}\sum_{t=1}^{T} I(X_t, B_t; Y_t|Y^{t-1}). \quad (13)$$

Similarly to the original problem, the finite-horizon problem with no RES can also be formulated as a MDP with belief $\beta_t(b_t, x_t)$. Control actions used to determine energy received from the grid are defined as $a_t(y_t|b_t, x_t) = P^{\mathbf{q}}(Y_t = y_t|B_t = b_t, X_t = x_t, \beta_t)$. To solve the finite-horizon

problem by DP, we use the method in [7] for belief updates, and express the average information leakage rate at time t in terms of belief and action probabilities as follows:

$$\mathbb{E}_{\mathbf{q}}[\bar{l}_t(b_t, x_t, a_t, y^t; \mathbf{q})] = \sum_{\substack{b_t\in\mathbf{B}, x_t\in\mathbf{X} \\ y_t\in\mathbf{Y}}} \beta_t(b_t, x_t)a_t(y_t|b_t, x_t)$$

$$\times \log \frac{a_t(y_t|b_t, x_t)}{\sum_{\hat{b}_t\in\mathbf{B}, \hat{x}_t\in\mathbf{X}} \beta_t(\hat{b}_t, \hat{x}_t)a_t(y_t|\hat{b}_t, \hat{x}_t)}$$

$$:= \bar{\mathcal{L}}(\beta_t, a_t). \quad (14)$$

The average per-step energy cost for the finite-horizon problem is determined similarly to Section 2.2, and represented in terms of the belief and action probabilities as:

$$\bar{\mathcal{C}}(\beta_t, a_t) := \mathbb{E}_{\mathbf{q}}[y_t] = \sum_{\substack{b_t\in\mathbf{B}, x_t\in\mathbf{X} \\ y_t\in\mathbf{Y}}} \beta_t(b_t, x_t)a_t(y_t|b_t, x_t)y_t. \quad (15)$$

The weighted objective function to be minimized for the finite-horizon privacy-cost trade-off is denoted by $\bar{\mathcal{U}}_\gamma(\beta_t, a_t) = [\gamma\bar{\mathcal{L}}(\beta_t, a_t) + (1 - \gamma)\bar{\mathcal{C}}(\beta_t, a_t)]$. We can first quantize the belief state, and solve the resulting MDP with a finite state space by DP recursively using the Bellman operator in (9) with the corresponding changes for finite-horizon.

In the next subsection, using the solution of the finite-horizon problem above, we will propose a low-complexity solution for the original infinite-horizon problem with an RB and RES, and a resetting energy generation process of the form $E_t = \{0, B_{max}\}$.

3.1. Threshold Policy (TP)

In TP, we fix a target horizon n, and after each RB recharge instance, start employing the optimal energy management policy for this finite-horizon, derived in the previous section. We follow the optimal policy for horizon n until either the battery is recharged again, in which case we restart with the same policy, or we reach the time horizon n. If the RB is not recharged at time $(n + 1)$, we assume that we simply provide all the energy demand directly from the grid, resulting in full information leakage. The intuition behind this scheme follows from the law of large numbers, which suggests that the RB will be charged after $n = \frac{1}{P_E}$ time slots with high probability. We will consider policies with a fixed time horizon of $n = \frac{1}{P_E}$, as well as those with an optimized time horizon. Our numerical results in Section 5 show that the performance with optimized but fixed time horizon closely follows that of the infinite-horizon solution.

3.2. Battery Conditioned Policy (BCP)

We propose another low-complexity policy, which depends only on the current input load. In BCP, when there is no demand, we allow the RB to be recharged by the grid with a probability P_{C_i} for each battery state $B_t = i$, for $i = \{0, \dots, B_{max}\}$. On the other hand, when there is energy demand, the RB is discharged with a probability P_{D_i} for each battery state. As before, intentional energy waste is not allowed. When there is demand in the case of an empty RB, it is entirely supplied from the grid. We choose (P_{C_i}, P_{D_i}) values that minimize (3) by an exhaustive grid search on $[0, 1]^2$.

4. LOWER BOUND

Next, we provide a lower bound on the privacy-cost trade-off by assuming that the user non-causally knows the times at which the RES recharges the RB. In Fig. 2, these time instances are represented by consecutive arrows. The weighted sum of finite-horizon leakage rate and average energy cost, minimized over policy $\boldsymbol{q}$, is denoted by $\bar{U}^*(\gamma, T_k)$ in Fig. 2. Given i.i.d. P_E, the probability that the RB is recharged after T_k time slots is given by

$$f(T_k; P_E) = P_E(1 - P_E)^{T_k}. \quad (16)$$

If the RB recharge instances are known in advance, the problem reduces to the finite-horizon MDP for each inter-arrival period, and can be solved as outlined in Section 3.
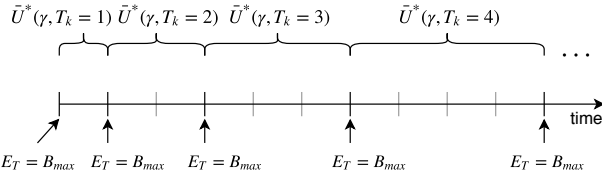

Fig. 2. Renewable energy generation instances and privacy-cost rate for the corresponding intervals.

Once the optimal performance is evaluated for all $T_k > 0$, the lower bound can be derived by taking their average using the probability mass function in (16):

$$F_\gamma(P_E) = \sum_{k=1}^{\infty} f(T_k; P_E)\bar{U}^*(\gamma, T_k), \quad (17)$$

where the coefficient $f(T_k; P_E)$ approaches zero as $T \to \infty$, while $\bar{U}^*(\gamma, T_k)$ approaches the infinite-horizon privacy-cost trade-off. For the numerical solution of the infinite-sum indicated in (17), we perform the summation for finite $k=\{1, \ldots, K\}$ such that $\sum_{k=K+1}^{\infty}\{f(T_k; P_E)\bar{U}^*(\gamma, T_k)\} < \epsilon$. To obtain the minimum K satisfying this inequality, we first consider the worst case information leakage rate and average energy cost, where all the demand is supplied by the grid, $Y_t = X_t$, and denote the lower bound by

$$F_\gamma(P_E) \leq \sum_{k=1}^{K} f(T_k; P_E)\bar{U}^*(\gamma, T_k) + \sum_{k=K+1}^{\infty} f(T_k; P_E)\bar{U}_w(\gamma),$$

where $\bar{U}_w(\gamma) := [\gamma H(X) + (1 - \gamma)\mathbb{E}[X]]$ represents the worst case privacy-cost trade-off, in which $H(X)$ and $\mathbb{E}[X]$ are the entropy and expected value of the demand, respectively. Hence, we choose the minimum K value that satisfies $\sum_{k=K+1}^{\infty} f(T_k; P_E)\bar{U}_w(\gamma) = (1 - P_E)^{T_K}\bar{U}_w(\gamma) < \epsilon$. We can find a finite T_K satisfying this inequality for any $\epsilon > 0$.

5. A SIMPLE BINARY EXAMPLE

We consider a simple scenario with $(\mathcal{X}, \mathcal{Y}) = \{0, 1\}$, $\mathcal{E} = \{0, 2\}$ and $\mathcal{B} = \{0, 1, 2\}$. We emphasize that obtaining numerical results for larger alphabets is challenging as the belief grows with the state space, and so does the computational complexity, also due to the quantization of the belief. For simplicity, demand and energy generation processes are assumed to be i.i.d. with Bernoulli $P_X = 0.5$ and $P_E \in [0, 1]$, respectively.

Extensions to Markovian E_t process is straightforward for TP and BCP; however, the MDP formulation requires including E_t in the state, and updating the belief accordingly. We consider a privacy-cost trade-off weight of $\gamma = 0.5$.

The weighted total privacy leakage and energy cost for TP, BCP and infinite-horizon MDP are depicted in Fig. 3, together with the lower bound. The average weighted cost decreases with P_E, since the demand can be mostly supplied by the RES, decreasing both the cost and leakage. The lower bound is obtained from (17) evaluated over a sufficiently long T. While the lower bound is not tight in general, it also shows us the value of predicting the energy generation instances for optimizing the privacy and cost. Two plots of TP are obtained corresponding to different horizons. For the first TP plot, the finite-horizon is set to be $n = \frac{1}{P_E}$. Since TP leads to full information leakage when energy arrives later than the set horizon, this approach has a higher privacy-cost trade-off compared to the infinite-horizon DP solution of the original problem. For the second TP plot, for each P_E value, the best horizon value is selected by searching over the set $n = [1 : 15]$. We observed that, the optimal fixed horizon is typically longer than $\frac{1}{P_E}$, which reduces the probability of full leakage. Interestingly, the performance of TP with optimized yet fixed horizon follows that of the infinite-horizon MDP solution very closely. We remark here that the curve obtained for the infinite-horizon MDP solution is an approximation as well, due to the quantization of the belief. Finally, we observe that the performance of the BCP scheme can outperform that of fixed horizon TP policy for high P_E values.
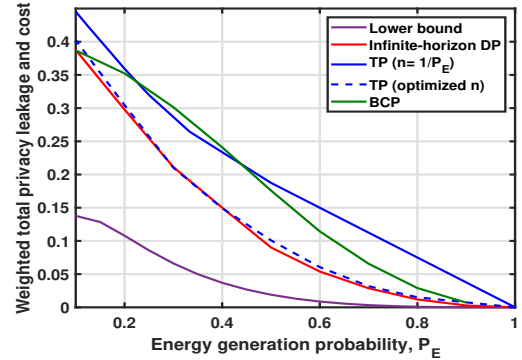

Fig. 3. Privacy-cost trade-off of the lower bound, TP, BCP and infinite-horizon MDP w.r.t. P_E for $\gamma = 0.5$ and $P_X = 0.5$.

6. CONCLUSIONS

We have studied the privacy-cost trade-off in a SM system equipped with both RB and RES. Motivated by the episodic nature of the problem, we proposed a low-complexity TP policy to solve this infinite-horizon problem by solving simplified finite-horizon problems with only RB and no RES. We also proposed the BCP policy, whose actions depend only on the demand. We numerically showed for a binary example that the fixed-horizon policy that ignores the RES process can achieve a near-optimal performance. As a future work, we will try to quantify/bound the gap between the two policies.

7. REFERENCES

[1] G. Giaconi, D. Gündüz, and H. V. Poor, "Privacy-aware smart metering: Progress and challenges," *IEEE Signal Processing Magazine*, vol. 35, no. 6, pp. 59–78, Nov 2018.

[2] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," *in Proc. 6th Workshop Security and Trust Management*, vol. 6710, pp. 226238, 2017.

[3] Y. Kim, E. C. H. Ngai, and M. B. Srivastava, "Cooperative state estimation for preserving privacy of user behaviors in smart grid," *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 178–183, 2011.

[4] M. Arrieta and I. Esnaola, "Smart meter privacy via the trapdoor channel," *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 277–282, 2017.

[5] S. Han, U. Topcu, and G. J. Pappas, "Event-based information-theoretic privacy: A case study of smart meters," *American Control Conference (ACC)*, pp. 2074–2079, July 2016.

[6] R. R. Avula, T. J. Oechtering, and D. Månsson, "Privacy-preserving smart meter control strategy including energy storage losses," *arXiv e-prints*, Mar. 2018.

[7] S. Li, A. Khisti, and A. Mahajan, "Information-theoretic privacy for smart metering systems with a rechargeable battery," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3679 – 3695, 2018.

[8] J. Yao and P. Venkitasubramaniam, "On the privacy-cost tradeoff of an in-home power storage mechanism," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 115–122.

[9] G. Giaconi and D. Gündüz, "Smart meter privacy with renewable energy and a finite capacity battery," in *IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2016, pp. 1–5.

[10] G. Giaconi, D. Gündüz, and H. V. Poor, "Smart meter privacy with renewable energy and an energy storage device," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 129–142, Jan 2018.

[11] J. Gomez-Vilardebo and D. Gündüz, "Smart meter privacy for multiple users in the presence of an alternative energy source," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 132–141, Jan 2015.

[12] O. Tan, D. Gündüz, and H. V. Poor, "Increasing smart meter privacy through energy harvesting and storage devices," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1331–1341, July 2013.

[13] J. Chin, T. Tinoco De Rubira, and G. Hug, "Privacy-protecting energy management unit through model-distribution predictive control," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 3084–3093, Nov 2017.

[14] O. Tan, J. Gomez-Vilardebo, and D. Gündüz, "Privacy-cost trade-offs in demand-side management with storage," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1458–1469, June 2017.

[15] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1994.