University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Farokhi, F;Sandberg, H

Title:
Fisher information privacy with application to smart meter privacy using HVAC units

Date:
2020

Citation:
Farokhi, F. & Sandberg, H. (2020). Fisher information privacy with application to smart meter privacy using HVAC units. Farokhi, F (Ed.). Privacy in Dynamical Systems, Privacy in Dynamical Systems, (1), pp.3-17. Springer.

Persistent Link:
https://hdl.handle.net/11343/285055

Chapter 1
Fisher Information Privacy with Application to Smart Meter Privacy Using HVAC Units

Farhad Farokhi and Henrik Sandberg

DRAFT

Abstract In this chapter, we use Heating, Ventilation, and Air Conditioning (HVAC) units to preserve the privacy of households with smart meters in addition to regulating indoor temperature. We model the effect of the HVAC unit as an additive noise in the household consumption. The Cramér-Rao bound is used to relate the inverse of the trace of the Fisher information matrix to the quality of an adversary's estimation error of the household private consumption from the aggregate consumption of the household with the HVAC unit. This establishes the Fisher information as the measure of privacy leakage. We compute the optimal privacy-preserving policy for controlling the HVAC unit through minimizing a weighted sum of the Fisher information and the cost operating the HVAC unit. The optimization problem also contains the constraints on the temperatures of the house.

1.1 Introduction

Smart meters are commonly used in modern electricity networks for metering consumption and sensing variables-of-interest for feedback control. The data gathered by the smart meters might also be used by adversaries to infringe on the privacy of households [14, 26, 19]. Therefore, appropriate privacy-preserving mechanisms must be developed and implemented to ensure the privacy of households with smart meters.

A common tool in privacy analysis is differential privacy [3, 4, 17, 13, 15]. Differential privacy requires that outputs, such as smart meter readings, do not noticeably change if an individual's data, such as its habits and appliance usage, changes.

Farhad Farokhi
The University of Melbourne and CSIRO's Data61 e-mail: farhad.farokhi@unimelb.edu.au
e-mail: farhad.farokhi@data61.csiro.au

Henrik Sandberg
KTH Royal Institute of Technology e-mail: hsan@kth.se

Privacy-preserving mechanisms that meet the standards of differential privacy often rely on additive noises with slow-decaying distributions, such as the Laplace noise. Previously, differentially-private mechanisms have been previously used in private smart metering [21, 1].

Although potentially powerful when dealing with consumption aggregates across multiple households, implementing differentially-private mechanisms at household level remains impractical due to multiple reasons. First, in addition sensing and control purposes by the grid operator, smart meters are used for metering and billing purposes and thus adding noise might not be desirable from the perspective of the retailer and the household. Second, an additive Laplace noise might result in an inferior utility [21, 2, 20, 10] and increase the computational complexity of optimal filtering [7]. Differential privacy is also susceptible to adversarial attacks [12]. Finally, and perhaps possibly most importantly, because of the lack of a systematic way for setting the so-called privacy parameter or budget, differential privacy has shown to be sometimes ineffective in practice [11, 23].

Fueled by these criticisms, information-theoretic privacy has gathered momentum as a rival approach for measuring private information leakage and designing privacy-preserving policies [18]. Information-theoretic privacy dates back to wiretap channels [24] and their extensions [25]. Information-theoretic measures of privacy leakage rely on mutual information (or relative entropy) [18], Fisher information [9, 8], maximin information [6] for measuring private information leakage and cast the privacy problem as a generalized rate-distortion problem. The use of entropy as a measure of privacy forces the private dataset to be statistically distributed with known distributions. This motivates the extension of information-theoretic privacy to use the Fisher information and maximin information as measures of private information leakage. This is because these information measures does not impose statistical assumptions on the private dataset.

In this paper, we use Heating, Ventilation, and Air Conditioning (HVAC) units to preserve the privacy of households. The effect of the HVAC unit can be modeled as an additive noise in the household consumption but there are several important distinctions with the case of additive measurement noise. Firstly, the HVAC units cannot be run arbitrarily as their main goal is to keep the temperature of the house within a comfortable zone and thus noises with unbounded support, such as the Gaussian and Laplace noises, cannot be implemented. Secondly, the measurements of the voltage and the current provided by the smart meter in this case are exact and thus estimation complexities and metering issues are entirely avoided. We consider an adversary that aims to estimate the household's consumption, as closely as possible, from the aggregate consumption, i.e., the combined private consumption of the household and the consumption of the HVAC unit. We use the Cramér-Rao bound [22, p. 169] to relate the variance of the estimation error of unbiased estimators of the household consumption from the aggregate consumption by the adversary to the inverse of the trace of the Fisher information matrix. This enables the use of the Fisher information as a measure of privacy. We can then find the optimal policies for controlling the HVAC unit through minimizing the Fisher information. We also consider the cost operating the HVAC unit and add that to the optimization problem

DRAFT

in order to capture the trade-off between privacy and utility/comfort. We do this while enforcing the constraints on the indoor temperature of the house.

The design of privacy-preserving policies based on the Fisher information and its interpretation through the Cramér-Rao bound are have been previously studied in [9, 8]. However, using this metric for designing a privacy-preserving control unit for HVAC units, in this chapter, is entirely novel.

The rest of the chapter is organized as follows. In Section 1.2, we present the problem formulation by discussing the merits of using the Fisher information as a measure of privacy and modeling HVAC units. We present optimal privacy-preserving policies in Section 1.3. Finally, we present a numerical example in Section 1.4 and conclude the paper in Section 1.5.

1.2 Problem Formulation

Let $x[k] \in \mathbb{R}^n$, for all $k \in \mathfrak{T} := \{0, \dots, T\}$ with T denoting the privacy-preservation horizon, be an arbitrary deterministic sequence which is desired to be kept private. At every time instant k, the following report is made:

$$y[k] = x[k] + w[k], \tag{1.1}$$

where $w[k]$ is an additive noise for privacy preservation. We assume that the additive noise over the horizon $\mathfrak{T}$ must satisfy $Cw[0:T] \leq d$ with probability one, where $w[0:T] = [w[0]^\top \ \cdots w[T]^\top]^\top$. We assume that the probability measure of the additive noise $w[0:T]$ is absolutely continuous with respect to the Lebesgue measure μ. Therefore, there exists probability density function γ such that

$$\mathbb{P}\{w[0:T] \in \mathfrak{W}\} = \int_{\mathfrak{W}} \gamma(w)\mu(\mathrm{d}w),$$

where $\mathfrak{W} \subseteq \mathbb{R}^{nT}$ is any Borel-measurable set. Further, we may define $\Omega := \{w[0:T] \,|\, Cw[0:T] \leq d\}$. In this chapter, we restrict ourselves to additive noise $w[k]$ that is independent of the data $x[k]$. One could imagine that a causal stochastic map from $x[k]$ to $w[k]$ may yield an even higher level of privacy in some cases but that is out of our scope. We restrict our attention for selecting the design variable γ to the set of density functions that fulfill the following standing assumption.

Assumption 1 *(i)* $\mathbb{P}\{w[0:T] \in \Omega\} = 1$, *(ii)* $\gamma(w[0:T]) > 0$ for all $w[0:T] \in \operatorname{int}\Omega$, and *(iii)* $\gamma(w[0:T]) = 0$ for all $w[0:T] \in \partial\Omega$.

The first part of the assumption is required for the feasibility of the additive noise. This will be discussed thoroughly in the case of HVAC systems at the end of this section. The second and the third parts of the assumption ensure the validity of the Cramér-Rao bound for relating the estimation error an adversary to the Fisher

DRAFT

information. In what follows, Γ denotes the set of all density functions γ that meet Assumption 1.

In order to keep $x[k]$, $\forall k \in \mathfrak{T}$, private, we opt to make the estimation error $\sum_{k=0}^{T} \mathbb{E}\{\|x[k] - \hat{x}[k]\|_2^2\}$ as large as possible, where $\hat{x}[0:T]$ is an estimate of $x[0:T]$ based on the transmitted measurements $y[0:T]$. This is however a function of the adversary's estimation policy, which might not be available at the time of computing the privacy-preserving additive noise. Therefore, we use the Cramér-Rao bound [22, p. 169] to relate the estimation error to a fundamental property of the additive noise $w[0:T]$, known as the Fisher information. Under Assumption 1, for any unbiased estimate of $x[0:k]$ given measurements $y[0:T]$ denoted by $\hat{x}[0:T]$, we have

$$\begin{aligned}\frac{1}{T+1}\sum_{k=0}^{T}\mathbb{E}\{\|x[k]-\hat{x}[k]\|_2^2\} &= \frac{1}{T}\mathbb{E}\{\|x[0:k]-\hat{x}[0:T]\|_2^2\}\\ &\geq \frac{n}{\text{trace}(\mathfrak{I}(\gamma))},\end{aligned}$$

where

$$\begin{aligned}\mathfrak{I}(\gamma) &:= \int_{\Omega}\left\|\frac{1}{\gamma(w[0:T])}\nabla\gamma(w[0:T])\right\|_2^2\gamma(w[0:T])\mu(\mathrm{d}w[0:T]))\\ &= \int_{\Omega}\frac{1}{\gamma(w[0:T])}\sum_{k=0}^{T}\sum_{i=1}^{n}\left(\frac{\partial}{\partial w_i[k]}\gamma(w[0:T])\right)^2\mu(\mathrm{d}w[0:T])).\end{aligned}\tag{1.2}$$

We consider unbiased estimators; however, biased estimators can be treated similarly to [9]. We can also define a measure of cost for implementing the additive noise $w[0:T]$ by the privacy-preserving agent for balancing against privacy as

$$\mathfrak{Q}(\gamma) := \frac{1}{T+1}\int_{\Omega} f(w[0:T])\gamma(w[0:T])\mu(\mathrm{d}w[0:T]),$$

where $f : \Omega \to \mathbb{R}$ is a function determining the cost of the realization $w[0:T]$. To extract the optimal privacy-preserving policy, we can solve the optimization problem

$$\min_{\gamma\in\Gamma} \quad \mathfrak{I}(\gamma) + \varrho\mathfrak{Q}(\gamma),\tag{1.3}$$

where $\varrho > 0$ determines the balance between privacy and quality. Note, here, ϱ can be seen as a Lagrange multiplier for the following optimization problem:

$$\min_{\gamma\in\Gamma} \quad \mathfrak{I}(\gamma),\tag{1.4a}$$

$$\text{s.t.} \quad \mathfrak{Q}(\gamma) \leq \rho.\tag{1.4b}$$

DRAFT

In fact, it can be shown that, for any $\varrho > 0$, there exists $\rho > 0$ such that the solutions of Problems 1.3 and 1.4 are equal due to the convexity of the optimization problem and linearity of the constraints [16].

1.2.1 Motivating Example

Assume that the private energy consumption of a household at time instant $k \in \mathfrak{T}$ by $x[k]$. The house has a heating, ventilation, and air conditioning (HVAC) unit that the occupants can use for providing a comfortable leaving place and masking their private energy consumption profile and thus preserving their privacy. The total energy consumption of the house is given by

$$y[k] = x[k] + w[k],$$

where $w[k]$ denotes electricity consumption of the unit. Privacy can be achieved by ensuring that the total energy consumption of the house (i.e., the private profile plus the HVAC's consumption) $y[k]$ cannot be used to construct an accurate estimate of the household's private consumption $x[k]$. Assume that

$$z[k+1] = Az[k] + Bw[k], \quad z[0] = z_0,$$

where $z[k] \in \mathbb{R}^m$ denotes the states of the HVAC unit and the temperature in the house. An example of a simple model for the HVAC unit is

$$\begin{aligned}\tau[k+1] &= z[k] - a(\tau[k] - T_a) + bw[k] \\ &= (1-a)\tau[k] + aT_a + bw[k],\end{aligned}$$

where $\tau[k]$ is the average temperature of the house and T_a is the ambient temperature. Defining $z[k] = \tau[k] - T_a$ results in

$$z[k+1] = Az[k] + Bw[k], \tag{1.5}$$

where $A = (1-a)$. The primary use of an air conditioning system is to provide a comfortable environment for the occupants. This can be guaranteed by ensuring that

$$\bar{C}z[k] \leq \bar{d}, \quad \forall k \in \mathfrak{T}.$$

An example of constraints on the temperature is that $T_{\min} \leq \tau[k] \leq T_{\max}$, which can be captured by

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} z[k] \leq \begin{bmatrix} T_{\max} - T_a \\ T_a - T_{\min} \end{bmatrix}. \tag{1.6}$$

Note that

DRAFT

$$\begin{bmatrix} z[1] \\ z[2] \\ \vdots \\ z[T+1] \end{bmatrix} = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^{T+1} \end{bmatrix} z_0 + \begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^T B & A^{T-1}B & \cdots & B \end{bmatrix} \begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[T] \end{bmatrix}.$$

Therefore, the constraints on the state of the HVAC unit can be rewritten as

$$\underbrace{\begin{bmatrix} \bar{C}B & 0 & \cdots & 0 \\ \bar{C}AB & \bar{C}B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{C}A^T B & \bar{C}A^{T-1}B & \cdots & \bar{C}B \end{bmatrix}}_{=:C} \begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[T] \end{bmatrix} \leq \underbrace{\begin{bmatrix} \bar{d} \\ \bar{d} \\ \vdots \\ \bar{d} \end{bmatrix} - \begin{bmatrix} \bar{C}A \\ \bar{C}A^2 \\ \vdots \\ \bar{C}A^{T+1} \end{bmatrix} z_0}_{=:d}. \tag{1.7}$$

The problem of finding the optimal privacy-preserving control input for an HVAC unit can then be cast in the form of (1.3).

1.3 Privacy-Preserving Policies

We slightly restrict the set of policies over which we search for the solution of (1.3) to be able to use tools from calculus of variations. To do so, we define the complete normed space $\mathfrak{C}^1(\Omega)$ as the set of continuously differentiable functions with the norm $\sup_{w\in\Omega} |\gamma(w)| + \sup_{w\in\Omega} \|\nabla\gamma(w)\|_\infty$. For this set, we can prove the following result.

Theorem 1 *Let $\beta : \mathbb{R}^{nT} \to \mathbb{R}$ and $\lambda \in \mathbb{R}$ be solutions of the following partial differential equation:*

$$\nabla^2\beta(w[0:T]) + \frac{1}{4}\left(\lambda - \frac{\varrho}{T+1} f(w[0:T])\right)\beta(w[0:T]) = 0, w[0:T] \in \Omega, \tag{1.8a}$$

$$\beta(w[0:T]) = 0, w[0:T] \in \partial\Omega, \tag{1.8b}$$

$$\beta(w[0:T]) \neq 0, w[0:T] \in \operatorname{int}\Omega, \tag{1.8c}$$

$$\int_\Omega \beta(w[0:T])^2 \mu(\mathrm{d}w[0:T]) = 1. \tag{1.8d}$$

Then, $\gamma(w[0:T]) = \beta(w[0:T])^2$ is a solution of (1.3) *over $\Gamma \cap \mathfrak{C}^1(\Omega)$. If there are multiple solutions over $\Gamma \cap \mathfrak{C}^1(\Omega)$, all exhibit the same cost in the sense of* (1.3)*.*

Proof. First, we can see that the cost function and the constraint sets are convex [9]. Now, following the result of [16], the Lagrangian can be constructed as

DRAFT

$$\mathcal{L} := \int_\Omega \frac{1}{\gamma(w[0:T])} \sum_{k=0}^{T} \sum_{i=1}^{n} \left(\frac{\partial}{\partial w_i[k]} \gamma(\mathrm{d}w[0:T]) \right)^2 \mu(\mathrm{d}w[0:T]))$$
$$+ \frac{\varrho}{T+1} \int_\Omega f(w[0:T]) \gamma(w[0:T]) \mu(\mathrm{d}w[0:T])$$
$$- \lambda \left(\int_\Omega \gamma(w[0:T]) \mu(\mathrm{d}w[0:T]) - 1 \right)$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier corresponding to the equality constraint that

$$\int_\Omega \gamma(w[0:T]) \mu(\mathrm{d}w[0:T]) = 1.$$

Using Theorem 5.3 in [5, p. 440], it can be seen that the extrema must satisfy

$$-\lambda + \frac{\varrho}{T+1} f(w[0:T]) - \frac{1}{\gamma(w[0:T])^2} \sum_{k=0}^{T} \sum_{i=1}^{n} \left(\frac{\partial \gamma(w[0:T])}{\partial w_i[k]} \right)^2$$
$$- 2 \sum_{k=0}^{T} \sum_{i=1}^{n} \frac{\partial}{\partial w_i[k]} \left[\frac{1}{\gamma(w[0:T])} \frac{\partial \gamma(w[0:T])}{\partial w_i[k]} \right] = 0. \quad (1.9)$$

Using the change of variable $\gamma(w[0:T]) = \beta(w[0:T])^2$, we can rewrite the optimality condition in (1.9) as

$$-\lambda + \frac{\varrho}{T+1} f(w[0:T]) - \frac{4}{\beta(w[0:T])^2} \sum_{k=0}^{T} \sum_{i=1}^{n} \left(\frac{\partial \beta(w[0:T])}{\partial w_i[k]} \right)^2$$
$$- 2 \sum_{k=0}^{T} \sum_{i=1}^{n} \frac{\partial}{\partial w_i[k]} \left[\frac{2}{\beta(w[0:T])} \frac{\partial \beta(w[0:T])}{\partial w_i[k]} \right] = 0,$$

which is equivalent to

$$-\lambda + \frac{\varrho}{T+1} f(w[0:T]) - \frac{4}{\beta(w[0:T])} \sum_{k=0}^{T} \sum_{i=1}^{n} \frac{\partial^2 \beta(w[0:T])}{\partial w_i[k]^2} = 0.$$

This concludes the proof. □

The partial differential equations in Theorem 1 can be solved to find the optimal privacy-preserving policy. However, simultaneously solving the partial differential equation with boundary conditions in (1.8a)-(1.8c) and the non-linear equation in (1.8d) is complex task in general, except for specific costs and constraints studied in [9, 8]. In what follows, a more tractable algorithm for finding the optimal privacy-preserving policy is proposed for the case where $f(w[0:T]) = 0$. Note that the inclusion of the cost function f is not the only way for trading off privacy and performance since the constraint $Cw[0:T] \leq d$ can also be used to ensure the additive noise is "small".

DRAFT

Theorem 2 *Assume* $f(w[0:T]) = 0$. *Let* $A_{\tilde{\xi}} \geq 0$, $\phi_{\tilde{\xi}} \in (-\pi, +\pi]$, *and* $\lambda \geq 0$ *be solutions of the following nonlinear complex equations:*

$$\sum_{\tilde{\xi} \in \aleph} A_{\tilde{\xi}} \cos\left(2\pi \sum_{k=0}^{T} \sum_{i=1}^{n} \tilde{\xi}_i[k] w_i[k] + \phi_{\tilde{\xi}}\right) = 0, \quad w[0:T] \in \partial\Omega, \tag{1.10a}$$

$$\sum_{\tilde{\xi} \in \aleph} A_{\tilde{\xi}} \cos\left(2\pi \sum_{k=0}^{T} \sum_{i=1}^{n} \tilde{\xi}_i[k] w_i[k] + \phi_{\tilde{\xi}}\right) \neq 0, \quad w[0:T] \in \operatorname{int}\Omega, \tag{1.10b}$$

$$\sum_{\tilde{\xi} \in \aleph} A_{\tilde{\xi}}^2 = 1, \tag{1.10c}$$

with

$$\aleph := \left\{ \tilde{\xi} \in \mathbb{R}^{nT} \;\middle|\; \sum_{k=0}^{T} \sum_{i=1}^{n} \tilde{\xi}_j[\ell]^2 = \frac{\lambda}{16\pi^2} \right\}. \tag{1.10d}$$

Then,

$$\gamma(w[0:T]) = \left(\sum_{\tilde{\xi} \in \aleph} A_{\tilde{\xi}} \cos\left(2\pi \sum_{k=0}^{T} \sum_{i=1}^{n} \tilde{\xi}_i[k] w_i[k] + \phi_{\tilde{\xi}}\right) \right)^2$$

is a solution of (1.3) *over* $\Gamma \cap \mathfrak{C}^1(\Omega)$.

Proof. Let the Fourier transform of β be given by

$$\hat{\beta}(\xi) := \int_{\Omega} \exp(-2\pi\imath \xi^\top w[0:T]) \beta(w[0:T]) \mu(\mathrm{d}w[0:T]), \quad \forall \xi \in \mathbb{R}^{nT}.$$

We can decompose ξ as $(\xi_i[k])_{i,k}$ according to the structure of $w[0:T]$. Hence, we can rewrite the Fourier transform as

$$\hat{\beta}(\xi) = \int_{\Omega} \exp\left(-2\pi\imath \sum_{k=0}^{T} \sum_{i=1}^{n} \xi_i[k] w_i[k]\right) \beta(w[0:T]) \mu(\mathrm{d}w[0:T]).$$

Similarly, the inverse Fourier transform is given by

$$\beta(w[0:T]) := \int_{\Omega} \exp\left(2\pi\imath \sum_{k=0}^{T} \sum_{i=1}^{n} \xi_i[k] w_i[k]\right) \hat{\beta}(\xi) \mu(\mathrm{d}w[0:T]).$$

Therefore,

$$\frac{\partial^2 \beta(w[0:T])}{\partial w_j[\ell]^2} = \int_{\Omega} \exp\left(2\pi\imath \sum_{k=0}^{T} \sum_{i=1}^{n} \xi_i[k] w_i[k]\right) (-4\pi^2 \xi_j[\ell]^2 \hat{\beta}(\xi)) \mu(\mathrm{d}w[0:T]).$$

DRAFT

This implies that the Fourier transform of $\partial^2\beta(w[0:T])/\partial w_j[\ell]^2$ is given by $-4\pi^2\xi_j[\ell]^2\hat{\beta}(\xi)$. Taking Fourier transform from the partial differential equation in Theorem 1 results in

$$\left(\frac{\lambda}{4} - 4\pi^2 \sum_{k=0}^{T}\sum_{i=1}^{n} \xi_j[\ell]^2\right)\hat{\beta}(\xi) = 0.$$

This implies that

$$\hat{\beta}(\xi) = \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)} c_{\tilde{\xi}}\delta(\xi-\tilde{\xi}).$$

Taking inverse Fourier transform from this solution results in

$$\beta(w[0:T]) = \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)} c_{\tilde{\xi}} \exp\left(2\pi\imath \sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_i[k]w_i[k]\right).$$

Furthermore, the Parseval's Theorem shows that

$$\begin{aligned}\int_{\Omega}\beta(w[0:T])^2\mu(\mathrm{d}w[0:T]) &= \int_{\mathbb{R}^{nT}}|\hat{\beta}(\xi)|^2\mu(\mathrm{d}\xi)\\ &= \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)}|c_{\tilde{\xi}}|^2.\end{aligned}$$

Noting that $\beta(w[0:T]) = \beta(w[0:T])^*$ (because $\beta(w[0:T])$ is real), we get $c^*_{\tilde{\xi}} = c_{-\tilde{\xi}}$, which implies that $\angle c_{\tilde{\xi}} = -\angle c_{-\tilde{\xi}}$. Also,

$$\begin{aligned}\beta(w[0:T]) &= \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)}|c_{\tilde{\xi}}|\exp(\imath\angle c_{\tilde{\xi}})\exp\left(2\pi\imath\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_i[k]w_i[k]\right)\\ &= \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)}|c_{\tilde{\xi}}|\left(\cos\left(2\pi\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_i[k]w_i[k]+\angle c_{\tilde{\xi}}\right)\right.\\ &\qquad\left.+\imath\sin\left(2\pi\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_i[k]w_i[k]+\angle c_{\tilde{\xi}}\right)\right)\\ &= \sum_{\tilde{\xi}:\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_j[\ell]^2=\lambda/(16\pi^2)}|c_{\tilde{\xi}}|\cos\left(2\pi\sum_{k=0}^{T}\sum_{i=1}^{n}\tilde{\xi}_i[k]w_i[k]+\angle c_{\tilde{\xi}}\right),\end{aligned}$$

where the last equality follows from that $\angle c_{\tilde{\xi}} = -\angle c_{-\tilde{\xi}}$. □

We can develop a numerical method for approximately solving (1.3) by discretizing the frequency space and enforcing the conditions in Theorem 2. The results of Theorem 2 can be seen as the extension of the results of [8] on optimality of square

DRAFT

cosine additive noise for privacy preservation to more general support sets for the additive noise.

Finally, we find the optimal privacy-preserving HVAC control in the next corollary.

Corollary 1 *Assume $f = 0$. For the HVAC unit in* (1.5) *with the temperature constraint in* (1.6)*, the optimal privacy-preserving policy is given by*

$$w[k] = \begin{cases} \frac{1}{b} v[k], & k = 0 \\ -\frac{1-a}{b} v[k-1] + \frac{1}{b} v[k], & k > 0, \end{cases} \tag{1.11}$$

where $v[k]$, for all $k \in \mathfrak{T}$, are i.i.d. distributed according to the probability density function:

$$p(v[k]) = \frac{2}{T_{\max} - T_{\min}} \cos^2 \left(\frac{\pi}{T_{\max} - T_{\min}} \left[v[k] - \frac{T_{\max} + T_{\min}}{2} + c_k \right] \right), \tag{1.12}$$

with $c_k = T_a + (1-a)^{k+1}(\tau[0] - T_a)$.

Proof. Define the change of variable

$$v[0:T] := \begin{bmatrix} v[0] \\ v[1] \\ \vdots \\ v[T] \end{bmatrix} = \underbrace{\begin{bmatrix} B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^T B & A^{T-1} B & \cdots & B \end{bmatrix}}_{=:\Psi} \begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[T] \end{bmatrix}$$

Clearly, the constraints in (1.7) can be rewritten as

$$(I \otimes C) v[0:k] \leq d,$$

where $\otimes$ denotes the Kronecker product. The constraints are separable. Let $\beta(w[0:T]) = \beta'(\Psi^{-1} w[0:T])$. In what follows, we show that

$$\beta'(v[0:T]) = \beta'_0(v[0]) \ldots \beta'_T(v[T])$$

is a solution of (1.3). With this change of variable, we get

$$\begin{aligned} \nabla^2 \beta(w[0:T]) \Big|_{w[0:T] = \Psi^{-1} v[0:T]} &= \operatorname{trace}(\Psi^{-\top} D^2 \beta'(\Psi^{-1} v[0:k]) \Psi^{-1}) \\ &= \operatorname{trace}(D^2 \beta'(\Psi^{-1} v[0:k])(\Psi^\top \Psi)^{-1}) \\ &= \sum_{k=0}^{T} \frac{\mathrm{d}^2 \beta'_k(v[k])}{\mathrm{d} v[k]^2} \left(\prod_{\ell \neq k} \beta'_\ell(v[\ell]) \right) [(\Psi^\top \Psi)^{-1}]_{kk}. \end{aligned}$$

DRAFT

Therefore, the optimality condition in Theorem 1 can be decomposed into

$$\frac{\mathrm{d}^2\beta_k'(v[k])}{\mathrm{d}v[k]^2} + \frac{1}{4}\lambda_k\beta_k'(v[k]) = 0,$$

where λ_k, $\forall k$, are constants such that $\sum_{k=0}^{T}\lambda_k[(\Psi^\top\Psi)^{-1}]_{kk} = \lambda$. The rest follows from the same line of reasoning as in [8]. □

Note that separation of the variables in the proof of Corollary 1 is only possible if there are constraint on either $z[k]$ or $w[k]$. This method cannot be perused if there are constraints on both $z[k]$ and $w[k]$. In this case, the solution must be approximated by implementing a discretized version of Theorem 2. Let $f(w) = w^\top w$. For the privacy-preserving policy in Corollary 1, we have

$$\begin{aligned}
\mathfrak{Q}(\gamma) =& \frac{1}{T+1}\int w[0:T]^\top w[0:T]\gamma(w[0:k])\mu(\mathrm{d}w[0:T]) \\
=& \frac{1}{T+1}\int v[0:T]^\top\Psi^{-\top}\Psi^{-1}v[0:T]\gamma(v[0:k])\mu(\mathrm{d}w[0:T]) \\
=& \frac{1}{T+1}\left(\frac{1}{b^2}\left[\left(\frac{1}{12}-\frac{1}{2\pi^2}\right)(T_{\max}-T_{\min})^2 + \left(\frac{T_{\max}+T_{\min}}{2}-c_T\right)^2\right]\right. \\
&\left. + \frac{(1-a)^2+1}{b^2}\sum_{k=0}^{T-1}\left[\left(\frac{1}{12}-\frac{1}{2\pi^2}\right)(T_{\max}-T_{\min})^2 + \left(\frac{T_{\max}+T_{\min}}{2}-c_k\right)^2\right]\right).
\end{aligned}$$

Therefore,

$$\lim_{T\to\infty}\mathfrak{Q}(\gamma) = \frac{(1-a)^2+1}{b^2}\left[\left(\frac{1}{12}-\frac{1}{2\pi^2}\right)(T_{\max}-T_{\min})^2 + \left(\frac{T_{\max}+T_{\min}}{2}-T_a\right)^2\right]. \tag{1.13}$$

If privacy was not an issue, any sequence of control actions for the HVAC unit $w[0:T]$ could be selected so long as the temperature in the house is kept between $T_{\min}$ and $T_{\max}$. An example of such control actions is to ensure that the temperature is constantly kept at $(T_{\min}+T_{\max})/2$. In this case, we get

$$\lim_{T\to\infty}\mathfrak{Q}(\gamma) = \frac{(1-a)^2+1}{b^2}\left(\frac{T_{\max}+T_{\min}}{2}-T_a\right)^2. \tag{1.14}$$

The ratio of costs in (1.13) and (1.14) captures the price of privacy (PoP), which is given by

$$\begin{aligned}
\mathrm{PoP} :=& 1 + \left(\frac{1}{3}-\frac{2}{\pi^2}\right)\left(\frac{T_{\max}-T_{\min}}{T_{\max}+T_{\min}-2T_a}\right)^2 \\
\approx& 1 + 0.1307\left(\frac{T_{\max}-T_{\min}}{T_{\max}+T_{\min}-2T_a}\right)^2.
\end{aligned}$$

DRAFT

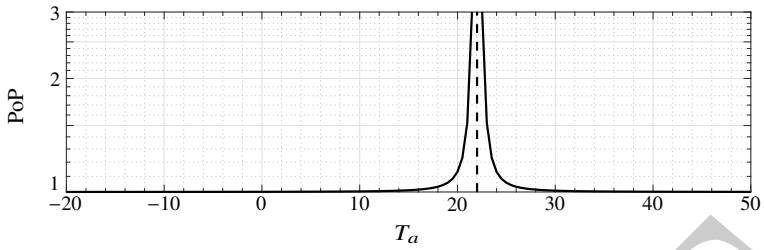


Fig. 1.1 The temperature $\tau[k]$ versus time k for the HVAC system with the optimal privacy-preserving controller in Corollary 1. The dashed lines show the constraints on the temperature.

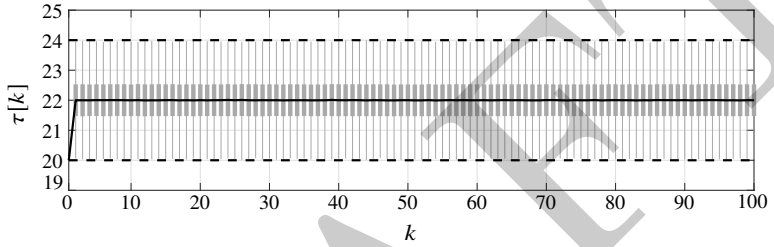


Fig. 1.2 Statistics of the temperature $\tau[k]$ versus time k extracted from hundred-thousands runs of the HVAC system with the optimal privacy-preserving controller in Corollary 1. The boxes, i.e., the vertical lines at each iterations, illustrate the range of 25% to 75% percentiles of the temperature and the thin gray lines illustrate the maximum and minimum temperatures in all the runs. The black line show the median temperature. The dashed lines show the constraints on the temperature.

1.4 Numerical Example

Consider the motivating example in Subsection 1.2.1 with horizon $T = 100$. We set $a = 0.1$ and $b = 0.5$. Further, we assume that constraints such as (1.6) must be enforced with $T_{\min} = 20$ and $T_{\max} = 24$, capturing the comfort zone of the occupants. We also assume that the starting temperature is given by $\tau[0] = 20$, which is within the comfort zone. Figure 1.1 shows the temperature $\tau[k]$ versus time k for the HVAC system with the optimal privacy-preserving controller in Corollary 1 with the dashed lines capturing the constraints on the temperature. Noting that the optimal privacy-preserving controller in Corollary 1 is stochastic, it makes sense to also observe the statistics of the temperature. Figure 1.2 illustrates the statistics of the temperature $\tau[k]$ versus time k extracted from one hundred-thousand runs of the HVAC system with the optimal privacy-preserving controller in Corollary 1. The boxes, i.e., the vertical lines at each iterations, illustrate the range of 25% to 75% percentiles of the temperature and the thin vertical gray lines illustrate the maximum and minimum temperatures in all the runs. The black line show the median temperature. The dashed lines illustrate the boundary of the comfort zone. Evidently, the temperature always stays within the comfort zone and hovers around 22 degrees. For this example,

DRAFT

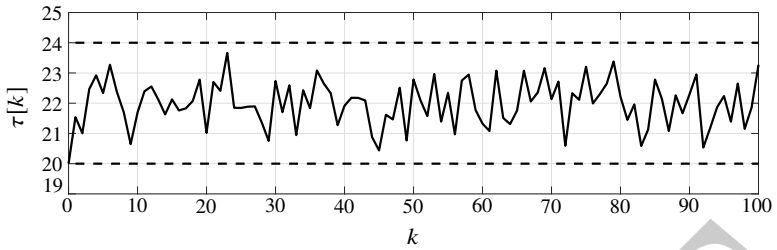


Fig. 1.3 Price of privacy versus the ambient temperature.

PoP = 1.0082 (i.e., 0.82% increase in consumption due to privacy), which is a negligible price to pay in average for privacy.

Figure 1.3 shows the price of privacy PoP versus the ambient temperature T_a. Note that as the ambient temperature moves towards 22 degrees, the price of privacy goes to infinity. This is because at 22 degrees, there is no need to use the HVAC unit for staying comfortable and all the energy consumption is essentially used for preserving privacy.

1.5 Conclusions

We used HVAC units to preserve the privacy of households with smart meters by modelling the effect of the HVAC unit as an additive noise in the household consumption. We computed the optimal privacy-preserving policy for controlling the HVAC unit by minimizing a weighted sum of the measure of privacy, based on the Fisher information, and the cost operating the HVAC unit.

References

1. Ács, G., Castelluccia, C.: I have a DREAM! (DiffeRentially privatE smArt Metering). In: T. Filler, T. Pevný, S. Craver, A. Ker (eds.) Information Hiding: 13th International Conference, IH 2011, Prague, Czech Republic, May 18-20, 2011, Revised Selected Papers, pp. 118–132. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
2. Bambauer, J., Muralidhar, K., Sarathy, R.: Fool's gold: an illustrated critique of differential privacy. Vanderbilt Journal of Entertainment & Technology Law **16**, 701 (2013)
3. Dwork, C.: Differential privacy: A survey of results. In: M. Agrawal, D. Du, Z. Duan, A. Li (eds.) Theory and Applications of Models of Computation: 5th International Conference, TAMC 2008, Xi'an, China, April 25-29, 2008. Proceedings, pp. 1–19. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
4. Dwork, C.: Differential privacy. In: H.C.A. van Tilborg, S. Jajodia (eds.) Encyclopedia of Cryptography and Security. Springer US, Boston, MA (2011)
5. Edwards, C.H.: Advanced Calculus of Several Variables. Academic Press (1973)

DRAFT

6. Farokhi, F.: Development and analysis of deterministic privacy-preserving policies using non-stochastic information theory. arXiv preprint arXiv:1810.11153v4 (2018)
7. Farokhi, F., Milosevic, J., Sandberg, H.: Optimal state estimation with measurements corrupted by Laplace noise. In: Proceedings of the 55th IEEE Conference on Decision and Control (2016)
8. Farokhi, F., Sandberg, H.: Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries. IEEE Transactions on Smart Grid **9**(5), 4726–4734 (2018)
9. Farokhi, F., Sandberg, H.: Ensuring privacy with constrained additive noise by minimizing Fisher information. Automatica **99**, 275–288 (2019)
10. Garfinkel, S.L., Abowd, J.M., Powazek, S.: Issues encountered deploying differential privacy. In: Proceedings of the 2018 Workshop on Privacy in the Electronic Society, pp. 133–137 (2018)
11. Greenberg, A.: How one of apple's key privacy safeguards falls short (2017). `https://www.wired.com/story/apple-differential-privacy-shortcomings/`
12. Haeberlen, A., Pierce, B.C., Narayan, A.: Differential privacy under fire. In: USENIX Security Symposium (2011)
13. Han, S., Topcu, U., Pappas, G.J.: Differentially private convex optimization with piecewise affine objectives. In: Proceedings of the 53rd IEEE Conference on Decision and Control, pp. 2160–2166 (2014)
14. Hart, G.W.: Residential energy monitoring and computerized surveillance via utility power flows. IEEE Technology and Society Magazine **8**(2), 12–16 (1989)
15. Huang, Z., Wang, Y., Mitra, S., Dullerud, G.E.: On the cost of differential privacy in distributed control systems. In: Proceedings of the 3rd International Conference on High Confidence Networked Systems, pp. 105–114 (2014)
16. Jeyakumar, V., Wolkowicz, H.: Zero duality gaps in infinite-dimensional programming. Journal of Optimization Theory and Applications **67**(1), 87–108 (1990)
17. Le Ny, J., Pappas, G.J.: Differentially private filtering. IEEE Transactions on Automatic Control **59**(2), 341–354 (2014)
18. Liang, Y., Poor, H.V., Shamai, S.: Information theoretic security. Foundations and Trends® in Communications and Information Theory **5**(4–5), 355–580 (2009)
19. McDaniel, P., McLaughlin, S.: Security and privacy challenges in the smart grid. IEEE Security and Privacy **7**(3), 75–77 (2009)
20. Muralidhar, K., Sarathy, R.: Does differential privacy protect terry gross' privacy? In: J. Domingo-Ferrer, E. Magkos (eds.) Privacy in Statistical Databases, pp. 200–209. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
21. Sandberg, H., Dán, G., Thobaben, R.: Differentially private state estimation in distribution networks with smart meters. In: Proceedings of the 54th IEEE Conference on Decision and Control, pp. 4492–4498 (2015)
22. Shao, J.: Mathematical Statistics. Springer Texts in Statistics. Springer-Verlag New York (2003)
23. Tang, J., Korolova, A., Bai, X., Wang, X., Wang, X.: Privacy loss in Apple's implementation of differential privacy on macOS 10.12. arXiv preprint arXiv:1709.02753 (2017)
24. Wyner, A.D.: The wire-tap channel. Bell System Technical Journal, The **54**(8), 1355–1387 (1975)
25. Yamamoto, H.: A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers. IEEE Transactions on Information Theory **29**(6), 918–923 (1983)
26. Zoha, A., Gluhak, A., Imran, M.A., Rajasegarar, S.: Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. Sensors **12**(12), 16838–16866 (2012)

DRAFT