Annual Reviews in Control 47 (2019) 412–422

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Information-theoretic approaches to privacy in estimation and control☆

Ehsan Nekouei[a,*], Takashi Tanaka[b], Mikael Skoglund[a], Karl H. Johansson[a]

[a] KTH Royal Institute of Technology, Stockholm, Sweden
[b] University of Texas at Austin, USA

ARTICLE INFO

Article history:
Received 5 February 2019
Revised 10 April 2019
Accepted 11 April 2019
Available online 23 April 2019

Keywords:
Privacy
Information theory
Networked control systems

ABSTRACT

Network control systems (NCSs) heavily rely on information and communication technologies for sharing information between sensors and controllers as well as controllers and actuators. When estimation, control or actuation tasks in a NCS are performed by an untrusted party, sharing information might result in the leakage of private information. The current paper reviews some of the recent results on the privacy-aware decision-making problems in NCSs. In particular, we focus on static and dynamic decision-making problems wherein privacy is measured using information-theoretic notions. We also review the applications of these problems in smart buildings and smart grids.

Contents

1. Introduction

Advanced Information and Communication Technologies (ICTs) have significantly facilitated the exchange of information between sensors, controllers and actuators within a control system. Enabled by ICT, we can respond to global challenges, such as carbon emission and energy consumption reductions, by designing

☆ This work is supported by the Knut and Alice Wallenberg Foundation, the Swedish Foundation for Strategic Research and the Swedish Research Council.

* Corresponding author.
E-mail addresses: nekouei@kth.se (E. Nekouei), ttanaka@utexas.edu (T. Tanaka), skoglund@kth.se (M. Skoglund), kallej@kth.se (K.H. Johansson).

https://doi.org/10.1016/j.arcontrol.2019.04.006
1367-5788/

high-performance networked control systems (NCSs) such as smart grid and smart buildings. For example, the efficiency of heating, ventilation and air-conditioning (HVAC) systems can be improved, up to 40%, by providing occupancy information of the building to the HVAC control unit, *e.g.,* see (Balaji, Xu, Nwokafor, Gupta, & Agarwal, 2013; Erickson & Cerpa, 2010) and (Kleiminger, Santini, & Mattern, 2014).

Moreover, enabled by ICT, the system designer can offload the computational burden of control, optimization and estimation algorithms to remote and powerful computation units, *e.g.,* cloud computing centers. Hence, substantial performance improvements can be gained by implementing complex and computationally demanding algorithms. As a result, NCSs are able to provide invaluable services such as intelligent transportation and smart energy management systems, thanks to the information exchange and computation capacities supplied by ICT.

Due to the distributed architecture of NCSs, the control, estimation and actuation tasks might be performed by different and possibly untrusted entities. For example, a cloud-based controller might operate a building's HVAC system based on its occupancy information. However, the occupancy information can be used to infer the location traces of individual occupants of the building. Thus, sharing the occupancy information with an untrusted controller might result in the leakage of private information, *i.e.,* location traces of occupants.

Although the design of NCSs is a well-studied research area, the privacy aspect of these systems is not well-understood and is a relatively new research avenue. The current paper reviews the recent results on privacy-aware decision-making problems in NCSs. In each problem, a NCS shares certain information with an untrusted party which is responsible for performing a certain task, *e.g.,* estimation or control tasks. The shared information is correlated with private variables which carry private information, *e.g.,* the state of a plant. The objective is to perform the desired task reliably while the leakage of private information is kept below a certain level. In particular, we focus on the scenarios wherein information-theoretic notions are used to capture the leakage of private information due to the sharing information. These problems can be broadly divided into two categories:

1. *Static setting:* This setting is comprised of decision-making scenarios wherein the private information is modeled as a sequence of independent and identically distributed random variables. Mutual information and conditional entropy have been used as privacy metrics in this setting. In particular, the design of privacy filters for hypothesis testing and the design of privacy-aware estimators will be discussed in this setting.
2. *Dynamic setting:* This setting includes the decision-making problems wherein the private information evolves in time according to a certain evolution law. For example, the electricity demand of a household, which carries private information, is usually modeled as a stochastic process with a certain dynamic. Besides mutual information, directed information has been used as the privacy metric in this setting. Examples of privacy-aware decision-making in the dynamic setting include the state privacy for a plant controlled remotely by a cloud-based controller, privacy of location traces in occupancy-based HVAC control and privacy-aware operation of household electricity storage device for ensuring tenants' privacy.

The properties of the decision-making problems and their solutions are reviewed in each setting. We note that the decision-making problems with information-theoretic privacy measures typically have desirable properties. For example, the privacy filter design problem with information-theoretic privacy measures can be typically cast as convex optimization problems. Thus, the design problem becomes tractable and systematic which allows one to easily study the privacy-performance trade-offs. Moreover, the information-theoretic notions provide strong guarantees for privacy by imposing universal bounds on the performance of untrusted parties in recovering private information, *e.g.,* see the discussion in Subsection 3.2.

1.1. Organization of the paper

The rest of this paper is organized as follows. Next section introduces the information-theoretic privacy metrics and other common privacy notions in the literature. Section 3 presents the privacy-aware decision-making problems in the static setting. In this section, we first review the privacy filter design framework of (Liao, Sankar, Tan, & du Pin Calmon, 2018) for hypothesis testing problems. Next, the results of (Nekouei, Sandberg, Skoglund, & K.H., 2018a) on the design of optimal privacy-aware estimators are reviewed. Finally, the leakage level of private information in a multi-sensor estimation problem is reviewed which was studied in (Nekouei, Skoglund, & Johansson, 2018b).

The privacy in dynamic decision-making problems is discussed in Section 4. In this section, we first review the optimal privacy-aware operation of a household's electricity storage device for ensuring the privacy of the electricity demand. This result is based on the paper (Li, Khisti, & Mahajan, 2018). Next, the privacy-aware disturbance attenuation framework of (Jia, Dong, Sastry, & Spanos, 2017) is reviewed. Finally, the results of (Tanaka, Skoglund, Sandberg, & Johansson, 2017) on the state privacy of a linear dynamical system in a cloud-based control setting are reviewed. Section 5 presents a smart building application of the privacy-aware disturbance attenuation problem which appeared in (Jia et al., 2017).

2. Privacy notions

In this section, we first describe the information-theoretic privacy metrics used in this paper. Then, we briefly review other privacy notions, such as differential privacy, and their applications in estimation and control problems.

2.1. Information-theoretic privacy metrics

Conditional entropy is a common privacy metric in the literature. Let X and Y denote two discrete random variables with the support sets $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$, respectively. Then, the conditional entropy of X given Y is defined as

$$\mathsf{H}[X|Y] = -\sum_{i,j} \mathsf{Pr}\big(X = x_i, Y = y_j\big) \log \mathsf{Pr}\big(X = x_i | Y = y_j\big).$$

The conditional discrete entropy of X given Y is interpreted as the ambiguity level regarding X after observing Y. As an example, assume that X is a private random variable and Y is the output of a privacy filter. Then, a relatively large value of $\mathsf{H}[X|Y]$ indicates a large ambiguity level regarding X after observing Y which implies a high privacy level.

Using the fact that conditioning reduces entropy (Cover & Thomas, 2006), the conditional entropy of X given Y can be upper bounded as

$$0 \leq \mathsf{H}[X|Y] \leq \mathsf{H}[X]$$

where $\mathsf{H}[X]$ is the entropy of X defined as

$$\mathsf{H}[X] = -\sum_{i} \mathsf{Pr}(X = x_i) \log \mathsf{Pr}(X = x_i).$$

If X and Y are independent, *i.e.,* Y contains no information about X, we have $\mathsf{H}[X|Y] = \mathsf{H}[X]$. Using Fano's inequality

(Cover & Thomas, 2006), the error probability of estimating X from Y can be lower bounded by

$$\Pr\big(X \neq \hat{X}(Y)\big) \geq \frac{\mathrm{H}[X|Y] - 1}{\log |\mathcal{X}|} \tag{1}$$

where $\hat{X}(Y)$ is an arbitrary estimator of X using Y and $|\mathcal{X}|$ is the cardinality of the support set of X. Fano's inequality is not necessarily tight when the marginal distribution of X is fixed (Ho & Verdú, 2010). The reader is referred to (Feder & Merhav, 1994) and (Ho & Verdú, 2010) for tighter, but more sophisticated, bounds on the error probability in terms of entropy.

Mutual information is another commonly used information-theoretic notion of privacy in the literature. Using the concepts of entropy and conditional entropy, the mutual information between X and Y is defined as

$$\mathrm{I}[X;Y] = \mathrm{H}[X] - \mathrm{H}[X|Y] \tag{2}$$

When X carries private information, a small value of $\mathrm{I}[X;Y]$ indicates a low level of private information leakage via the random variable Y. The equality above expresses the connection between mutual information and conditional entropy.

We next define the notion of directed information which is used in the literature as a privacy metric. Let $X^n = (X_1, \ldots, X_n)$ and $Y^n = (Y_1, \ldots, Y_n)$. Then, the directed information from X^n to Y^n is defined as

$$\mathrm{I}[X^n \to Y^n] = \sum_{i=1}^{n} \mathrm{I}\big[X^i; Y_i \big| Y^{i-1}\big]$$

where $\mathrm{I}[X^i; Y_i|Y^{i-1}]$ is the mutual information between X^i and Y_i given Y^{i-1} defined as

$$\mathrm{I}\big[X^i; Y_i \big| Y^{i-1}\big] = \mathrm{H}\big[X^i \big| Y^{i-1}\big] - \mathrm{H}\big[X^i \big| Y^i\big]$$

These privacy metrics can be defined for continuous random variables in a similar fashion, *e.g.,* see (Cover & Thomas, 2006).

Finally, we define the Kullback-Leibler divergence between two probability distributions. Let P and Q be two probability distributions defined over the same discrete probability space. Then, the Kullback-Leibler divergence between P and Q is defined as

$$\mathrm{D}[P\|Q] = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

where $Q(i) \neq 0$ for all i.

Information-theoretic notions, such as conditional entropy, mutual information and Kullback-Leibler divergence, have been widely used in the literature to characterize the performance of various decision-making problems. For example, Fano's inequality in (1) provides a universal lower bound on the error probability of all estimators. Similarly, the asymptotic behaviors of the Bayes rule and the Neyman-Pearson test in hypothesis testing problems are expressed in terms of the Kullback-Leibler divergence. The interested reader is referred to (Cover & Thomas, 2006; Csiszár & Shields, 2004; Zhang, 2006) and references therein for more details on the connection between information theory, statistics and decision-making theories.

2.2. Other privacy notions

Besides information-theoretic notions, differential privacy and homomorphic encryption have been used in the literature for ensuring privacy. The notion of differential privacy was proposed in (Dwork, McSherry, Nissim, & Smith, 2006) to address the database privacy problem. Given two databases which only differ in one row, a differentially private mechanism generates randomized outputs which are almost statistically identical over two databased. Thus, if a participant adds her data to a database or removes it, no output will become significantly more or less informative than other outputs. The interested reader is referred to (Dwork & Roth, 2014) for more details on privacy-preserving data analysis using differential privacy.

Differential privacy has been used to design privacy-aware estimation, filtering and average consensus algorithms. The authors in (Ny & Pappas, 2014) designed various filtering schemes, which ensure the privacy of states or measurements of dynamical systems, based on differential privacy. The authors in (Sandberg, Dán, & Thobaben, 2015) studied the state estimation problem in a power distribution network subject to differential privacy constraints for the consumers. A differential privacy mechanism was proposed in (Wang, Huang, Mitra, & Dullerud, 2017) for a distributed linear multi-agent control problem which guarantees the privacy of agents' preferences, *e.g.,* their way-points in a congestion-aware navigation application. The authors in (Nozari, Tallapragada, & Cortés, 2017) and (Mo & Murray, 2017) developed algorithms which ensure the privacy of the initial states of different agents in an average consensus problem. The notion of differential privacy was used in (Nozari, Tallapragada, & Cortés, 2018) to design distributed convex optimization algorithms which preserve the privacy of objective functions.

Finally, we note that homomorphic encryption has been used in the literature to improve the privacy of networked control systems against eavesdropping attacks. In (Farokhi, Shames, & Batterham, 2017), the Paillier encryption method was used to ensure the privacy of sensors measurements in networked control systems as well as distributed systems such as distributed formation seeking algorithms. The authors in (Kogiso & Fujita, 2015) proposed controller encryption schemes, based on public-key RSA, which conceal controller parameters as well as the information available at the controller, *e.g.,* measurements.

3. Information-theoretic privacy in static settings

In this section, we first review the design of privacy filters for hypothesis testing problems which appeared in (Liao et al., 2018). Next, the framework of (Nekouei et al., 2018a) on the design of optimal privacy-aware estimators is reviewed. Finally, we review the results of (Nekouei et al., 2018b) on the leakage level of private information in a multi-sensor estimation problem under two information sharing schemes.

3.1. Privacy-aware hypothesis testing

Hypothesis testing methods are widely employed for fault detection in networked control systems, *e.g.,* see (Mehra & Peschon, 1971) and (Chen & Patton, 1999) and references therein. The authors in (Liao et al., 2018) studied the design of optimal privacy filters for a hypothesis testing problem with mutual information as the privacy metric. In their set-up, a sensor observes a sequence of independent and identically distributed random variables, denoted by $X^n = \{X_1, \cdots, X_n\}$, where each X_i takes values in the finite set $\mathcal{X} = \{x_1, \ldots, x_l\}$. The random variables X_is are drawn randomly according to an unknown probability distribution $P = [p_1, \ldots, p_l]^\top$. However, it is a priori known that the probability distribution P belongs to the finite set of distributions $\{P_1, \ldots, P_m\}$.

The objective of the hypothesis testing task is to determine the underlying distribution of the measurements. Typically, the underlying distribution of the observations is estimated by a test rule which is obtained by minimizing a certain loss function (Poor, 2013). For example, in a binary hypothesis testing problem with two hypotheses $H_1: \quad P = P_1$ and $H_2: \quad P = P_2$, the optimal test minimizing the miss detection probability subject to an upper bound on the false alarm rate is the Neyman-Pearson test (Poor, 2013). Moreover, according to the Chernoff-Stein Lemma, the

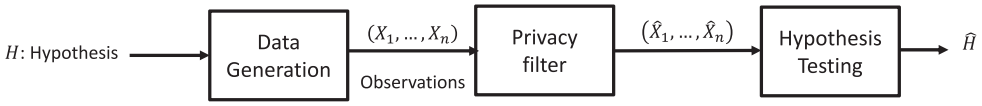

Fig. 1. Privacy-aware hypothesis testing set-up.

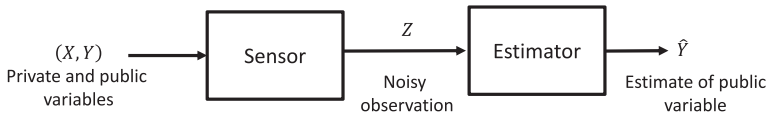

Fig. 2. A single sensor estimation set-up.

asymptotic exponent of the miss detection probability under the Neyman-Pearson test is given by $\mathrm{D}[P_1 \| P_2]$ (Cover & Thomas, 1991).

In many hypothesis testing problems, the observation signals contain private information, *e.g.,* in a medical diagnostic problem, one might be able to identify a patient using her medical test results. Thus, the privacy of the observation signals might be compromised when the hypothesis testing task is performed by an untrusted party. The work (Liao et al., 2018) proposed an information-theoretic framework for ensuring privacy in hypothesis testing problems. In this framework, a privacy filter generates a distorted version of the measurements and hypothesis testing is performed on the distorted data as shown in Fig. 1.

The privacy filter is designed such that the leakage of private information is kept below a certain level. More precisely, the privacy filter takes the observed signals as input and outputs an element from $\hat{\mathcal{X}} = \{\hat{x}_1, \cdots, \hat{x}_{\hat{l}}\}$. The output of the privacy mechanism is revealed to a possibly untrusted party for hypothesis testing.

The privacy mechanism is represented by an $l \times \hat{l}$ conditional probability matrix $W = [W_{ij}]$ where W_{ij} denotes the probability that the privacy mechanism selects $\hat{x}_j$ as its output when its input is x_i. The privacy mechanism W is selected according to the solution of the following optimization problem

$$\max_{W} \min_{k=1,2,\cdots,m} \quad \mathrm{D}[P_k W \| P_1 W] \\ \mathrm{I}[P_k; P_k W] \leq \epsilon_k, k = 1, \cdots, m \tag{3}$$

where the objective is to maximize the smallest error exponent such that the leakage of private information is kept below a desired level. Here, the information leakage is captured by the mutual information between the input and output of the privacy filter for all possible distributions of the measurements. Moreover, it was shown that the design of the privacy mechanism W in the high privacy regime, *i.e.,* when ϵ_ks are small, can be cast as a semi-definite optimization problem, see (Liao et al., 2018) for more details.

The interested reader is referred to (He & Tay, 2017; He, Tay, & Sun, 2016; Li, 2017; Li & Oechtering, 2015, 2017; Sun & Tay, 2016) and references therein for the privacy aspect of hypothesis testing problems under other privacy metrics.

3.2. Optimal privacy-aware estimator design problem

The optimal design of privacy-aware estimators of static random variables was studied in (Nekouei et al., 2018a). In their set-up, a sensor collects noisy information about a private random variable and a public random variable, as shown in Fig. 2. Let Z denote the sensor measurement and X and Y denote the private and public random variables, respectively. The random variable X contains private information which should be kept hidden from any untrusted party. The value of the public random variable is estimated using Z and the estimate is revealed to an untrusted party. The estimate of Y based on Z is denoted by $\hat{Y}(Z)$.

Due to the dependency of $\hat{Y}(Z)$ on X, an untrusted user can infer about the value of X by observing the output of the estimator. The objective of the privacy-aware estimator design problem is to obtain the optimal randomized estimator of the public random variable while a certain privacy level of the private random variable is guaranteed. The random variables X and Y take values in the finite sets $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, \ldots, y_m\}$, respectively, and Z takes value in $\mathbb{R}$. The privacy level of X is defined as the conditional discrete entropy of X given the output of the estimator, *i.e.,* $\mathrm{H}[X|\hat{Y}(Z)]$ which captures the ambiguity level regarding X after observing $\hat{Y}(Z)$. Thus, the privacy increases as $\mathrm{H}[X|\hat{Y}(Z)]$ becomes large. Moreover, according to Fano's inequality, the performance of any estimator of X using $\hat{Y}(Z)$ is limited by this privacy metric (see Section 2 for more details).

To define the randomized estimator of Y, let $\{B_i\}_{i=1}^N$ denote a partition of $\mathbb{R}$ where B_1 and B_N are semi-infinite intervals and where B_i, $2 \leq i \leq N-1$, are of the form $B_i = [a_{i-1}, a_i)$, $a_i > a_{i-1}$. We assume that the estimator has only access to the discretized version of Z. That is, the estimator knows the index of the bin which contains Z. Then, a randomized estimator of Y is defined as

$$\hat{Y}_P(Z) = \begin{cases} y_1 & \text{w.p.} \quad P_{1l}, \\ \vdots & \vdots \\ y_m & \text{w.p.} \quad P_{ml}, \end{cases} \text{, if} \quad Z \in B_l \tag{4}$$

where $\sum_i P_{il} = 1$ for all $l \in \{1, \cdots, N\}$. Thus, the estimator selects y_i as its output with probability P_{il} when the measurement belongs to B_l.

Let $Y_P(Z)$ denote an estimator of Y based on Z with $\mathrm{H}[X|Y_P(Z)] = \mathrm{H}_0$. Then, we say this estimator achieves the privacy level of H_0. An estimator of Y with $\mathrm{H}_0 = 0$ does not guarantee any privacy level and an estimator with $\mathrm{H}_0 = \mathrm{H}[X]$ achieves the maximum privacy level. Notice that in the latter, the output of the estimator is independent of X. The optimal privacy-aware estimator of the public random variable Y is the solution of the following optimization problem

$$\begin{aligned} \underset{\{P_{il}\}_{i,l}}{\text{minimize}} \quad & \mathrm{E}\big[L\big(Y, \hat{Y}_P(Z)\big)\big] \\ & P_{il} \geq 0, \forall i, l \\ & \sum_i P_{il} = 1, \quad \forall l \\ & \mathrm{H}\big[X \big| \hat{Y}_P(Z)\big] \geq \mathrm{H}_0 \end{aligned} \tag{5}$$

where $L(\cdot, \cdot)$ is a loss function quantifying the estimation loss. In the estimator design problem above, the randomization probabilities of bins are the optimization variables and the last constraint ensures the privacy level of X stays above the desired level H_0. Note that the optimal privacy-aware estimator is an estimator, with the privacy level of at least H_0, which minimizes the estimation loss. The objective function and the privacy constraint in the optimization problem above can be written as

$$\mathrm{E}\big[L\big(Y, \hat{Y}_P(Z)\big)\big] = \sum_{i,k} L(y_i, y_k) \Pr\big(Y = y_i, \hat{Y}_P(Z) = y_k\big)$$

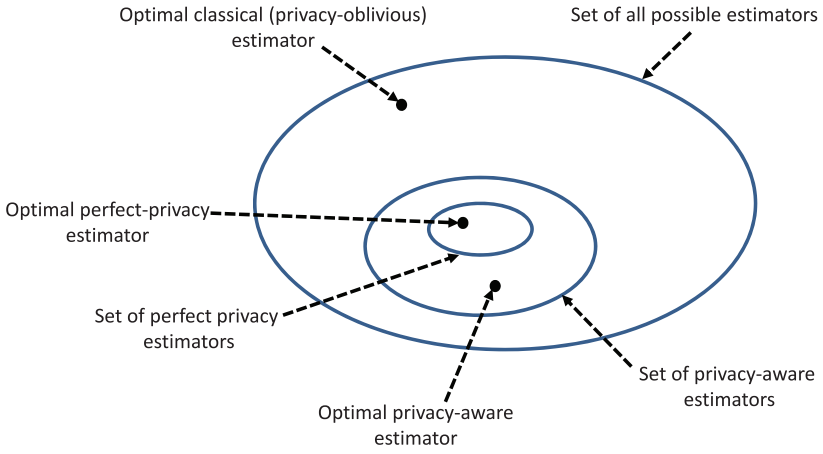

Fig. 3. A schematic representation of the sets of all possible estimators, all privacy-aware estimators with the privacy level at least $\mathsf{H}_0 > 0$ and all perfect-privacy estimators.

and

$$\mathsf{H}\big[X\big|\hat{Y}_P(Z)\big] = -\sum_{y,x} \mathsf{Pr}\big(\hat{Y}_P(Z)=y, X=x\big)\log \mathsf{Pr}\big(X=x|\hat{Y}_P(Z)=y\big)$$

respectively.

Fig. 3 shows the basic idea behind the privacy-aware estimator design problem. Note that, for $\mathsf{H}_0 = 0$, the feasible set of the optimization problem above becomes the set of all possible estimators, *i.e.,* the privacy constraint becomes inactive. In the absence of the privacy constraint, the solution of (5) may not satisfy the privacy constraint. The privacy constraint limits the feasible set of this optimization problem to the set of estimators with a privacy level at least equal to H_0.

Next theorem shows that the set of privacy-aware estimators is a convex set and the optimization problem above is a convex problem.

Theorem 1 *(Nekouei et al., 2018a). The objective function in the optimization problem (5) is linear and the privacy constraint is convex in the optimization variables. Thus, the optimal privacy-aware estimator design problem in (5) is a convex optimization problem.*

Thus, the optimal privacy-aware estimator can be designed by solving a convex optimization problem.

Next, the notion of perfect-privacy is defined.

Definition 1. An estimator of the public random variable satisfies the perfect-privacy condition if its output is independent of the private random variable.

Next theorem presents the necessary and sufficient condition for an estimator to satisfy the perfect-privacy condition.

Theorem 2 *(Nekouei et al., 2018a). Let $\hat{Y}_P(Z)$ denote a randomized estimator of Y and define the matrix $\Phi = [\phi_{jl}]_{jl}$ with $\phi_{jl} = \mathsf{Pr}(Z \in B_l|X = x_j) - \mathsf{Pr}(Z \in B_l)$. Then, $\hat{Y}_P(Z)$ satisfies the perfect-privacy condition if and only if $\boldsymbol{P}_i \in \mathsf{Null}(\Phi)$ for all $i \in \{1, \cdots, m\}$ where $\boldsymbol{P}_i = [P_{i1}, \cdots, P_{iN}]^\top$ and $\mathsf{Null}(\Phi)$ is the null space of the matrix Φ.*

Note that the set of perfect-privacy estimators is a subset of the set of privacy-aware estimators as shown in Fig. 3. Moreover, it can be shown that the set of perfect-privacy estimators is a convex ploytope, and the optimal perfect-privacy estimator can be obtained by solving a linear optimization problem (Nekouei et al., 2018a). The interested reader is referred to (d. P. Calmon et al., 2017) and (Rassouli & Gündüz, 2017) for more details on the perfect-privacy condition and its relation to the notion of maximal correlation.

Remark 1. Information-theoretic methods for improving data privacy have been investigated in the literature, *e.g.,* see (Asoodeh, Alajaji, & Linder, 2016; Asoodeh, Diaz, Alajaji, & Linder, 2017; Basciftci, Wang, & Ishwar, 2016; Kalantari, Sankar, & Kosut, 2017; Moraffah & Sankar, 2015; du Pin Calmon & Fawaz, 2012) and references therein. In this line of research, the objective is to design privacy preserving filters which operate on a (directly observable) public random variable which is correlated with a private random variable.

3.3. Information leakage of estimators

The leakage level of private information in a multi-sensor estimation problem was studied in (Nekouei et al., 2018b). Consider a multi-sensor estimation problem with M sensors in which the measurement of sensor $i \in \{1, \ldots, M\}$ at time $k \in \mathbb{N}$ can be written as

$$Z_k^i = Y_k + X_k^i + N_k^i \tag{6}$$

where $\{Y_k\}_k$ is a common process observed by all sensors, $\{X_k^i\}_k$ is a local process only observed by sensor i and N_k^i denotes the measurement noise of sensor i at time k. The local process $\{X_k^i\}_{i,k}$s are assumed to be private as they contain information about the local environments of sensors.

For each i, the sequence of random variables $\{N_k^i\}_k$ is assumed to be a set of i.i.d. random variables and the random variables $\{Y_k, X_k^i, N_k^i, i \in \{1, \ldots, M\}\}_k$ are assumed to be mutually independent.

The objective of the estimation problem is to obtain a reliable estimate of Y_k using an untrusted entity named the "cloud" which receives a function of the sensors' measurements at each time instance. Since the sensors' measurements are correlated with the local processes, the cloud can infer about the local process of each sensor based on the received information. The privacy level of the local processes is studied under two schemes for sharing the sensors' measurements with the cloud: a local scheme, and a global scheme.

3.3.1. Privacy of the local scheme

Under the local scheme, each sensor first estimates Y_k using the maximum a posteriori probability (MAP) estimator, and then transmits its estimate to the cloud. Then, the cloud obtains an estimate of Y using the local estimates of sensors. A pictorial representation of the local information sharing scheme is shown in Fig. 4(a). Let $\{\hat{Y}_k^i\}_{i=1}^M$ denote the collection of received information from the sensors at time k where $\hat{Y}_k^i$ is the estimate of Y_k by sensor i. In the

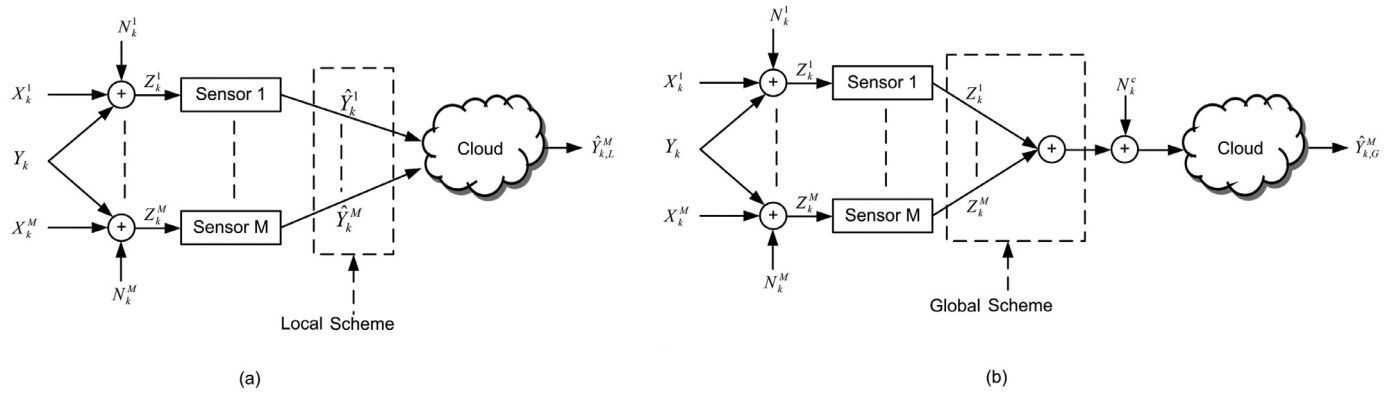

Fig. 4. Cloud-based multi-sensor estimation with local (a) and global (b) information sharing schemes.

local scheme, the privacy level of the local process of sensor i, $\{X_k^i\}_k$, is defined as the conditional entropy of X_k^i given the estimates of Y_k by all the sensors, *i.e.,* $\mathsf{H}[X_k^i|\hat{Y}_k^1,\ldots,\hat{Y}_k^M]$.

Next theorem derives a lower bound on the privacy level of local processes under the local information sharing scheme.

Theorem 3 *(Nekouei et al., 2018b). The privacy level of X_k^i under the local scheme can be lower bounded as*

$$\mathsf{H}\big[X_k^i\big|\hat{Y}_k^1,\ldots,\hat{Y}_k^M\big] \geq \mathsf{H}\big[X_k^i\big] - \mathsf{I}\big[X_k^i; Y_k, \hat{Y}_k^i\big]. \tag{7}$$

According to Theorem 3, the privacy level of the local process of sensor i is lower bounded by the difference between the discrete entropy of X_k^i and the mutual information between X_k^i and $(Y, \hat{Y}_k^i)$. Using the fact that conditioning reduces entropy and the lower bound in Theorem 3, we have

$$\mathsf{H}\big[X_k^i\big] - \mathsf{I}\big[X_k^i; Y_k, \hat{Y}_k^i\big] \leq \mathsf{H}\big[X_k^i\big|\hat{Y}_k^1,\ldots,\hat{Y}_k^M\big] \leq \mathsf{H}\big[X_k^i\big]$$

Hence, under the local scheme, the *privacy loss* of each sensor i is at most be equal to $\mathsf{I}[X_k^i; Y_k, \hat{Y}_k^i]$.

3.3.2. Privacy of the global scheme

Under the global scheme, sensors simultaneously transmit their measurements to the cloud. The received signal at the cloud at time k under the global scheme can be expressed as

$$Z_k^{c,M} = \left(\sum_{i=1}^{M} Z_k^i\right) + N_k^c$$

where N_k^c is the additive noise at time k. Then, the cloud obtains an estimate of Y_k using $Z_k^{c,M}$. In the global scheme, the privacy level of sensor i is defined as the discrete conditional entropy of X_k^i given the received signal by the cloud, *i.e.,* $\mathsf{H}[X_k^i|Z_k^{c,M}]$.

To study the privacy level in the global scheme, the following assumptions on the noise distributions are imposed.

- B1. The measurement noise of each sensor i at each time instance is modeled as a Gaussian random variable with zero mean and variance σ_i^2.
- B2. The received noise in the cloud at each time instance is modeled as a Gaussian random variable with zero mean and variance σ_c^2.
- B3. We have $0 < \sigma_{\min}^2 = \min(\sigma_c^2, \inf_i \sigma_i^2)$.
- B4. The sequence of random variables $\{N_k^c\}_k$ is assumed to be i.i.d. and independent of other random variables.

Next theorem establishes a lower bound on the privacy level of the global information sharing scheme.

Theorem 4 *(Nekouei et al., 2018b). Under Assumptions **B1-B4**, the privacy level of sensor i can be lower bounded as*

$$\mathsf{H}\big[X_k^i\big|Z_k^{c,M}\big] \geq \mathsf{H}\big[X_k^i\big] - \frac{\max_{x,x'\in\mathcal{X}^i}|x-x'|^2}{2(M+1)\sigma_{\min}^2} \tag{8}$$

where $\mathcal{X}^i$ is the support set of X_k^i.

The lower bound in Theorem 4 is a function of the number of sensors, $\sigma_{\min}^2$ and the "width" of the support set of X_k^i which is defined as $\max_{x,x'\in\mathcal{X}^i}|x-x'|$. This result implies that the privacy level of X_k^i converges to its maximum value $\mathsf{H}[X_k^i]$ at the rate of $O(1/M)$ when the number of sensors increases. Thus, the global information sharing scheme is asymptotically perfectly private as the number of sensors becomes large.

4. Information-theoretic privacy in dynamic settings

This section discusses the privacy filter design problem in the dynamic setting. We first review the privacy-aware operation of a storage device for ensuring the privacy of electricity demand of a household. In this problem, the demand is modeled as an exogenous private process and its privacy is ensured by the optimal operation of the storage device. Then, we study the design of privacy filters for two closed-loop control problems. The first problem is the privacy-aware disturbance attenuation for a linear dynamical system. Here, the disturbance is modeled as an exogenous private process and the objective is to design privacy filters for sharing disturbance information with an untrusted controller. The second problem is the privacy-aware control of linear dynamical systems. In this problem, the state trajectory is considered as private information and the objective is to design privacy filters for sharing the sensors measurements with an untrusted controller.

4.1. Privacy-aware operation of electricity storages in households

Information-theoretic privacy of the smart metering system of a household equipped with an electricity storage device was studied in (Li et al., 2018) and (Li, Oechtering, & Skoglund, 2016). In this subsection, the results of (Li et al., 2018) are reviewed. Consider the smart metering system of a household as illustrated in Fig. 5. The electricity demand of the household contains private information such as the absence or presence of the tenants and it should be kept hidden from an untrusted party. However, a utility company (or an eavesdropper), with access to the electricity consumption level of the household, can make inference about the household's electricity demand.

Let X_t $(t \geq 1)$ denote the household's electricity demand at time t, excluding the electricity storage unit. The demand process is

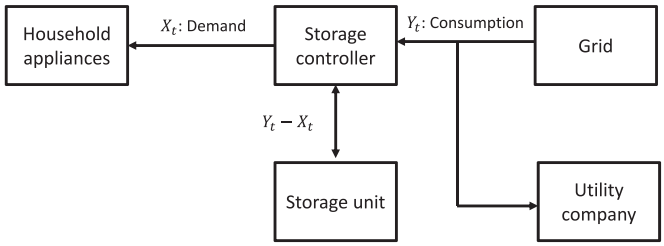

Fig. 5. The smart metering system of a household equipped with an electricity storage.

modeled as a first-order time-homogeneous irreducible and aperiodic Markov chain. Let Y_t denote the total electricity consumption of the household, including that of the storage unit, at time t. Assuming that the storage unit is lossless, we have

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

where S_t is the state of charge of the storage unit at time t. We assume that the random variables X_t, Y_t and S_t take values in $\mathcal{X} = \{0, 1, \dots, m_x\}$, $\mathcal{Y} = \{0, 1, \dots, m_y\}$, and $\mathcal{S} = \{0, 1, \dots, m_s\}$, respectively, for all $t \geq 1$. We also assume that X_1 is independent of S_1.

Given (x_t, s_t), *i.e.,* the realizations of the demand and the state of charge at time t, the feasible support set of Y_t is given by

$$\mathcal{Y}(s_t, y_t) = \{y \in \mathcal{Y} : s_t - x_t + y \in \mathcal{S}\}$$

Let $q_t(y|x^t, s^t, y^{t-1})$ denote a randomized charging policy of the storage at time t which, in general, might depend on all the available information at time t. Given (x_t, s_t), Y_t takes values in $\mathcal{Y}(s_t, y_t)$, thus, we have

$$\sum_{y \in \mathcal{Y}(s_t, y_t)} q_t\left(y \,|\, x^t, s^t, y^{t-1}\right) = 1$$

Let $\boldsymbol{q} = (q_1, q_2, \cdots, q_T)$ denote a randomized charging policy for the storage unit up to time T. The set of all feasible policies is denoted by $\mathcal{Q}_T$. Then, the optimal privacy-aware charging policy is defined as the solution of the following optimization problem

$$\underset{\boldsymbol{q} \in \mathcal{Q}_T}{\text{minimize}} \quad \frac{1}{T} \mathsf{I}\left[X^T, S_1; Y^T\right] \tag{9}$$

where $X^T = [X_1, \dots, X_T]$ and $Y^T = [Y_1 \dots, Y_T]$ are the histories of the household's electricity demand and consumption levels up to time T, respectively. Thus, the objective is to minimize the leakage of information about the demand and the initial state of charge of the storage by the optimal operation of the storage unit.

It has been shown in (Li et al., 2018) that it is optimal to restrict the set of feasible policies to the charing policies of the form $q_t(y_t|x_t, s_t, y^{t-1})$ which only depend on the current demand, the current state of charge and the history of consumption up to time t. To formalize this idea, we define the belief state as

$$\pi_t(x, s) = \Pr\left(X_t = x, S_t = s | Y^{t-1} = y^{t-1}\right)$$

Let $\mathcal{P}_{X,S}$ denote the space of joint probability distributions of the demand and the state of the charge random variables. Then, for any $V : \mathcal{P}_{X,S} \to \mathbb{R}$ and any $\pi \in \mathcal{P}_{X,S}$, the Bellman operator $\mathcal{B}_a$ is defined as

$$[\mathcal{B}_a V](\pi) = \mathsf{I}[a; \pi] + \sum_{x,y,s} \pi(x, s) a(y|x, s) V(\phi(\pi, y, a))$$

where $a(y|x, s.)$ represents a charging policy and ϕ is the nonlinear function which determines the evolution of the belief, *i.e.,* $\pi_{t+1} = \phi(\pi_t, y_t, a_t)$ and a_t is the charging policy at time t.

The following theorem characterizes the optimal privacy-aware operation of the storage device.

Theorem 5 *(Li et al., 2018). Consider the optimization problem (9). Then, we have*

1. *For any $\pi \in \mathcal{P}_{X,S}$, the value functions are iteratively defined as*
$$V_t(\pi) = \min_a [\mathcal{B}_a V_{t+1}](\pi) \tag{10}$$
for $t \in \{1, \cdots, T\}$ and $V_{T+1}(\pi) = 0$.
2. *The optimal policy $\boldsymbol{q}^\star = (q_1^\star, \cdots, q_T^\star)$ is given by*
$$q_t^\star\left(y_t \,|\, x_t, s_t, y^{t-1}\right) = f_t^\star(\pi_t)$$
where $f_t^\star(\pi_t)$ is the minimizer of the right hand side of (10).
3. *The minimum leakage of the private information is given by $\frac{1}{T} V_1(P_{X_1}(x) P_{S_1}(s))$ where $P_{X_1}(\cdot)$ and $P_{S_1}(\cdot)$ denote the distributions of X_1 and S_1, respectively.*

The second item in the theorem above indicates the optimality of the policies of the form $q_t(y_t|x_t, s_t, y^{t-1})$.

4.2. Privacy-aware disturbance attenuation using an untrusted controller

Consider the fully observable plant

$$X_{t+1} = AX_t + BU_t + W_t$$

where X_t denotes the state of the plant at time t, W_t denotes the disturbance signal at time t and the pair (A, B) is assumed to be controllable. The disturbance signal, which carries private information, is modeled as a stochastic process taking values in the finite set $\{w_0, \dots, w_M\}$. Moreover, we assume that it can be observed using a sensor. The control objective is to steer the states of the system to a desired region using an untrusted controller. Since the disturbance signal can be measured, the control objective can be easily achieved if the controller has access to the disturbance. The authors in (Jia et al., 2017) studied the design of privacy filters for sharing the disturbance information with the controller in the context of the occupancy-based HVAC control. Here, we present the privacy filter design framework of (Jia et al., 2017) in a more abstract framework and its application to the HVAC control problem will be discussed in the next section.

At time t, a privacy filter takes W_t as input and outputs a distorted version of it, denoted by V_t which takes values in $\{w_0, \dots, w_M\}$. Then, the controller receives V_t. Given $W_t = w_i$, the random variable V_t takes w_j as its value with probability P_{ji}. Thus, the controller's information regarding the disturbance is not necessarily the same as the true disturbance signal. The randomization probabilities are designed such that the leakage of private information is minimized while a certain performance level for the closed-loop control system is ensured.

The control inputs are designed using a model predictive controller (MPC). The controller at each time-step receives the distorted disturbance value and assumes that the disturbance remains constant during the control horizon. Let $J_{\text{MPC}}(W_t, V_t, X_t)$ denote the optimal control cost which depends on the current state of the system as well as true and distorted disturbance values at time t. Then, the optimal privacy filter is the solution of the following optimization problem:

$$\begin{aligned}
&\underset{\{P_{ji}\}_{i,j}}{\text{minimize}} \quad \mathsf{I}[W_t; V_t] \\
&P_{ji} \geq 0, \forall i, j \\
&\sum_j P_{ji} = 1, \quad \forall i \\
&\mathsf{E}\left[J_{\text{MPC}}(W_t, V_t, X_t) - J_{\text{MPC}}\left(W_t, \tilde{X}_t\right) \middle| W = w\right] \leq \delta_1, \\
&\mathsf{E}\left[\left\| X_{\text{MPC}}(W_t, V_t, X_t) - X_{\text{MPC}}\left(W_t, \tilde{X}_t\right) \right\| \middle| W = w\right] \leq \delta_2 \\
&\forall w, \forall \left\| X_t - \tilde{X}_t \right\| \leq \delta_3
\end{aligned} \tag{11}$$

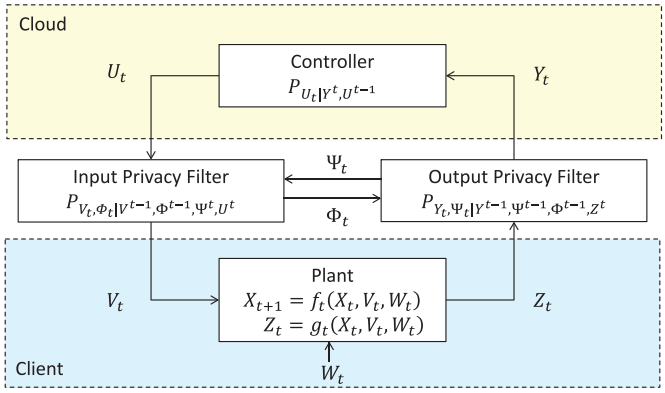

Fig. 6. Privacy filter: General model.

where $J_{\mathrm{MPC}}(W_t, \tilde{X}_t)$ is the optimal cost when the controller has access to the true disturbance signal and the state of the system is $\tilde{X}_t$, $X_{\mathrm{MPC}}(V_t, Y_t, X_t)$ and $X_{\mathrm{MPC}}(W_t, \tilde{X}_t)$ are the resulting states when the controller has access to the distorted and true disturbance values, respectively.

The objective of the optimization problem above is to minimize the leakage of private information. The third constraint ensures that for two sufficiently close states of the system X_t and $\tilde{X}_t$, the difference between the control costs of the system using the true and distorted disturbance information is small. The fourth constraint also guarantees that, for sufficiently close initial states X_t and $\tilde{X}_t$, the resulting states after one-step MPC iteration will be close.

The next theorem shows that the privacy filter design problem above can be solved efficiently.

Theorem 6 *(Jia et al., 2017). The optimal privacy filter design problem (11) is a convex optimization problem.*

4.3. Privacy in cloud-based control systems

Consider the regulation problem of a dynamical system using an untrusted cloud-based controller. Let $X^t \triangleq (X_1, \ldots, X_t)$ denote the state sequence of the local plant up to time t which is considered to be private. To perform the control task, (possibly, a function of) sensors' measurements are shared with the cloud. The objective is to preserve the privacy of the state trajectory while a certain regulation performance is ensured. The information-theoretic privacy of cloud-based control of dynamical system was studied in (Tanaka et al., 2017).

The general structure of the class of privacy filters is illustrated in Fig. 6. An output filter prevents raw sensor data to be disclosed to the cloud. An input privacy filter replaces the control input U_t with a different value V_t to enhance privacy. In general, the input and output filters can communicate with each other via messages Ψ_t and Φ_t. Privacy filters and controller algorithms are in general randomized policies and have memories of the past observations. Thus, we model them as stochastic kernels of the forms specified in Fig. 6. Fig. 7 shows a simpler form of a privacy filter in which the control input commanded by the cloud is directly applied to the plant. Since there is no input filter, this architecture is easier to implement. For the rest of this section, we focus on this simple architecture in Fig. 7, and discuss privacy notions and privacy filter design problems exclusively for this architecture.

Since the control is a multi-stage problem, we consider the following privacy metric which captures the total privacy loss over the considered time-horizon

$$\sum_{t=1}^{T} I(X^t; Y_t | Y^{t-1}, U^{t-1}) =: I(X^T \to Y^T \| U^{T-1}). \tag{12}$$

The notation on the right hand side of (12) is introduced in (Kramer, 2003). We refer to this quantity as Kramer's causally conditioned directed information. Note that under appropriate postulates regarding a privacy metric, it can be shown that the causally conditioned directed information is a proper privacy metric for multi-stage decision-making problems (Tanaka et al., 2017).

Suppose that the performance of the cloud-based control system is measured by a stage-wise additive cost function $\sum_{t=1}^{T} \mathbb{E}c(X_{t+1}, U_t)$. Then, privacy loss in cloud-based control with a given control performance requirement δ is minimized by solving

$$\min \; I(X^T \to Y^T \| U^{T-1}) \tag{13a}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} \mathbb{E}c(X_{t+1}, U_t) \le \delta. \tag{13b}$$

Likewise, the best achievable control performance under the privacy constraint is characterized by flipping the constraint and objective functions in (13). In both cases, the optimization domain is the space of the sequence of Borel measurable stochastic kernels

$$\mathcal{D} = \{P_{U_t|Y^t,U^{t-1}} \;,\; P_{Y_t|Y^{t-1},U^{t-1},Z^t}\}_{t=1}^{T} \tag{14}$$

characterizing joint controller and output privacy filter policies.[1] Since (13) is an infinite dimensional optimization problem, it is in general difficult to obtain an explicit form of an optimal solution. Thus, we consider a special case in which (13) becomes a tractable optimization problem. Suppose the plant in Fig. 7 is a fully observable linear dynamical system

$$X_{t+1} = A_t X_t + B_t U_t + W_t, \quad Z_t = X_t$$

where $W_t \sim \mathcal{N}(0, \Sigma_t^W)$ is a sequence of independent Gaussian random variables. We assume $\Sigma_t^W \succ 0$ for $t = 1, \ldots, T$. Assume also that $c(\cdot, \cdot)$ in (13) is a convex quadratic function, and that the problem (13) can be written as

$$\min \; I(X^T \to Y^T \| U^{T-1}) \tag{15a}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} \mathbb{E}(\|X_{t+1}\|_{Q_t}^2 + \|U_t\|_{R_t}^2) \le \delta. \tag{15b}$$

The domain of optimization is (14).[2]

It can be shown that the minimum privacy leakage characterized by (15) is lower bounded by the optimal value of

$$\min \; I(X^T \to U^T) \tag{16a}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} \mathbb{E}(\|X_{t+1}\|_{Q_t}^2 + \|U_t\|_{R_t}^2) \le \delta \tag{16b}$$

where again the domain of optimization is $\mathcal{D}$ given by (14). In what follows, we provide an optimal joint controller and output privacy filter policy that solves (15).

Theorem 7 *(Tanaka et al., 2017). The policy shown in Fig. 8 is an optimal solution to (15).*

[1] We consider $P_{U_1|Y^1,U^0} = P_{U_1|Y^1}$ and $P_{Y_1|Y^0,U^0,Z^1} = P_{Y_1|Z^1}$.

[2] Problem (15) is identical to the problem considered in (Tanaka & Sandberg, 2015), except that in (Tanaka & Sandberg, 2015), an optimal solution is provided under the restriction that the stochastic kernels in (14) are Linear-Gaussian.

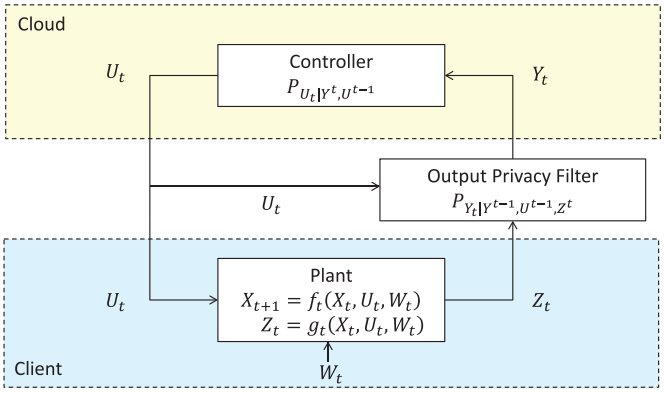

Fig. 7. Privacy filter: Output filter only.

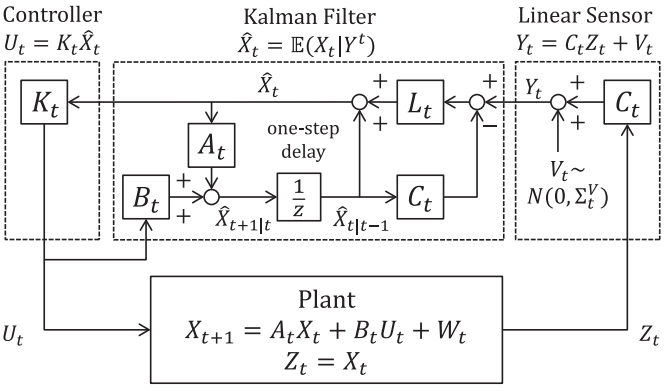

Fig. 8. Structure of optimal policy for problem (16).

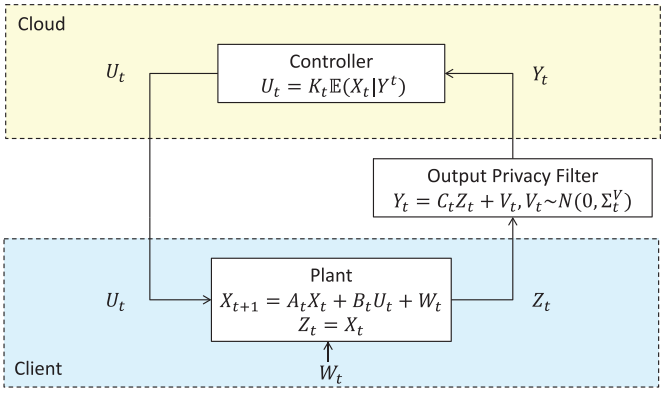

Fig. 9. Structure of the optimal joint controller and output privacy filter for the cloud-based LQG control problem (15). Although the output privacy filter is allowed to utilize public random variable U^{t-1} (as shown in Fig. 7), it turns out that this information need not be used.

Next theorem provides an explicit form of the joint control and output privacy filter policy solving (15).

Theorem 8 *(Tanaka et al., 2017). An optimal joint controller and output privacy filter characterized by an optimal solution to (15) is in the form shown in Fig. 9. An optimal choice of matrices C_t, Σ_t^V, L_t (Kalman gains) and K_t (feedback control gains) are obtained by Algorithm 1. Moreover, the optimal value of (15) is equal to the optimal value of the determinant maximization problem in Algorithm 1.*

5. A smart building application of privacy-aware control

The design of privacy filters for occupancy-based HVAC control problem was studied in (Jia et al., 2017). Consider the occupancy-based control of the heating, ventilation and air conditioning (HVAC) system of a smart building as shown in Fig. 10. The energy consumption of the HVAC system can be substantially reduced by utilizing the occupancy information of different building zones. However, sharing this information with an untrusted party might result in the loss of residents' privacy. That is, the location traces of the individual residents can be inferred from the occupancy data.

To formally define the privacy filter design problem, let the set $Z = \{z_1, \ldots, z_N\}$ denote the zones inside the building and the set $O = \{o_1, \ldots, o_M\}$ denote the building's occupants. The location of the mth occupant at time t is denoted by the random variable X_t^m which takes values in Z. The occupancy level of zone n at time t can be expressed as

$$Y_t^n = \sum_{m=1}^{M} 1_{\{X_t^m = z_n\}}$$

where $1_{\{\}}$ is an indicator function. The location trace of each occupant is modeled as a first-order Markov chain and the location traces of different occupants are assumed to be mutually independent.

Algorithm 1 Joint controller and privacy filter design for cloud-based LQG control.

(Tanaka et~al., 2017)

1. Determine feedback control gains K_t via the backward Riccati recursion:

$$S_t = \begin{cases} Q_t & \text{if } t = T \\ Q_t + \Phi_{t+1} & \text{if } t = 1, \cdots, T-1 \end{cases}$$
$$\Phi_t = A_t^\top (S_t - S_tB_t(B_t^\top S_tB_t + R_t)^{-1}B_t^\top S_t)A_t$$
$$K_t = -(B_t^\top S_tB_t + R_t)^{-1}B_t^\top S_tA_t$$

2. Solve a determinant maximization problem with respect to $P_{t|t} \succ 0$, $\Pi_t \succ 0$, $t = 1, \ldots, T$ subject to LMI constraints:

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{t=1}^{T} \log\det \Pi_t^{-1} + c_1 \\ \text{s.t.} \quad & \sum_{t=1}^{T} \mathrm{Tr}(\Theta_t P_{t|t}) + c_2 \le D, \\ & P_{1|1} \preceq P_{1|0}, P_{T|T} = \Pi_T, \\ & P_{t+1|t+1} \preceq A_tP_{t|t}A_t^\top + \Sigma_t^W, \quad t = 1, \ldots, T-1 \\ & \begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t}A_t^\top \\ A_tP_{t|t} & A_tP_{t|t}A_t^\top + \Sigma_t^W \end{bmatrix} \succeq 0, \quad t = 1, \ldots, T-1 \end{aligned}$$

where $\Theta_t = K_t^\top (B_t^\top S_tB_t + R_t)K_t$, $t = 1, \ldots, T$ and

$$c_1 = \tfrac{1}{2}\log\det P_{1|0} + \tfrac{1}{2}\sum_{t=1}^{T-1} \log\det \Sigma_t^W$$
$$c_2 = \mathrm{Tr}(N_1P_{1|0}) + \sum_{t=1}^{T} \mathrm{Tr}(\Sigma_t^W S_t).$$

3. For each $t = 1, \ldots, T$, choose (e.g., by the singular value decomposition) a full row rank matrix C_t and a positive definite matrix Σ_t^V such that

$$C_t^\top \Sigma_t^{V^{-1}} C_t = P_{t|t}^{-1} - (A_{t-1}P_{t-1|t-1}A_{t-1}^\top + \Sigma_{t-1}^W)^{-1}.$$

4. Determine the Kalman gains by

$$L_t = P_{t|t-1}C_t^\top (C_tP_{t|t-1}C_t^\top + \Sigma_t^V)^{-1}$$

where $P_{t+1|t} = A_tP_{t|t}A_t^\top + \Sigma_t^W$.

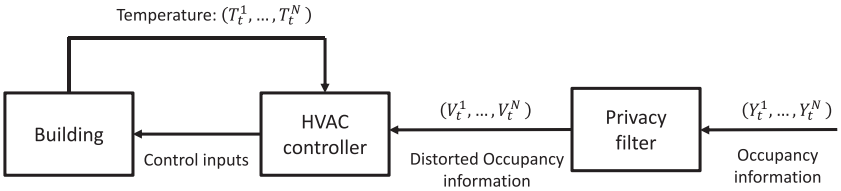

Fig. 10. An occupancy-based HVAC control system.

To improve the efficiency of the HVAC system, the occupancy data needs to be shared with the HVAC controller. However, sharing the occupancy data might result in the privacy loss of occupants due to the correlation between the occupancy data and the location traces of occupants. The objective of a privacy filter is to ensure the occupants' privacy while a certain efficiency level for the HVAC system is guaranteed.

A privacy filter takes the occupancy level of different zones as input and outputs a distorted version of the occupancy level. Then, the controller receives the distorted occupancy levels. Let V_t^n denote the distorted occupancy level of zone n at time t which takes value in $\{0, \cdots, M\}$. Given $Y_t^n = i$, the random variable V_t^n takes j as its value with probability P_{ji}^n. Thus, the controller's information regarding the occupancy level of each zone is not necessarily the same as the true occupancy level of that zone. The randomization probabilities are designed such that the leakage of private information is minimized while a certain performance level for the closed-loop control system is preserved.

The discretized dynamics of the temperature in zone n can be expressed as

$$C^n \frac{T_{t+1}^n - T_t^n}{\Delta} = R^n T_t + c_0 Y_t^n + m_{s,t}^n c_p \left(T_{s,t}^n - \frac{T_{t+1}^n + T_t^n}{2} \right)$$

where c_0 is the thermal load per person, C^n is the thermal capacity of the zone n, Δ is the discretization step, R^n denotes the heat transfer between zone n and other zones, c_p is the thermal capacity of the air. The control inputs are the supplied air mass flow rate and the temperature, denoted by $m_{s,t}^n$ and $T_{s,t}^n$, respectively.

The authors in (Jia et al., 2017) showed that under mild assumptions the privacy filter for the HVAC system can be designed using the framework in Subsection 4.2.

6. Conclusions

In this paper, we reviewed privacy-aware decision-making problems wherein information-theoretic privacy metrics were used to capture the leakage of private information. In particular, we focused on the recent results on the design of optimal privacy filters in both static and dynamic settings. In the static setting, the recent work on privacy-aware hypothesis testing and estimation were reviewed. In the dynamic setting, we reviewed the optimal privacy-aware operation of a household's electricity storage, the optimal privacy-aware disturbance attenuation problem and the design of privacy filter for controlling a linear dynamical system using an untrusted controller.

Despite strong privacy guarantees, the research on information-theoretic privacy has been limited to centralized settings. The design of *distributed* privacy filters, under information-theoretic privacy metrics, for distributed settings *e.g.,* average consensus, distributed control and distributed optimization, is a promising research avenue from both theoretical and practical perspectives.

Conflict of Interest

None.

References

Asoodeh, S., Alajaji, F., & Linder, T. (2016). Privacy-aware MMSE estimation. In *2016 ieee international symposium on information theory (isit)* (pp. 1989–1993).

Asoodeh, S., Diaz, M., Alajaji, F., & Linder, T. (2017). Privacy-aware guessing efficiency. In *2017 ieee international symposium on information theory (isit)* (pp. 754–758).

Balaji, B., Xu, J., Nwokafor, A., Gupta, R., & Agarwal, Y. (2013). Sentinel: Occupancy based hvac actuation using existing wifi infrastructure within commercial buildings. In *Proceedings of the 11th acm conference on embedded networked sensor systems* (pp. 17:1–17:14).

Basciftci, Y. O., Wang, Y., & Ishwar, P. (2016). On privacy-utility tradeoffs for constrained data release mechanisms. In *2016 information theory and applications workshop (ita)* (pp. 1–6).

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Norwell, MA, USA: Kluwer Academic Publishers.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York, NY, USA: Wiley-Interscience.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. Wiley-Interscience.

Csiszár, I., & Shields, P. (2004). *Information theory and statistics: A tutorial*. Foundations and Trends®in Communications and Information Theory.

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography* (pp. 265–284). Berlin, Heidelberg: Springer Berlin Heidelberg.

Dwork, C., & Roth, A. (2014). The Algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science, 9*(3-4), 211–407.

Erickson, V. L., & Cerpa, A. E. (2010). Occupancy based demand response hvac control strategy. In *Proceedings of the 2nd acm workshop on embedded sensing systems for energy-efficiency in building* (pp. 7–12).

Farokhi, F., Shames, I., & Batterham, N. (2017). Secure and private control using semi-homomorphic encryption. *Control Engineering Practice, 67*, 13–20.

Feder, M., & Merhav, N. (1994). Relations between entropy and error probability. *IEEE Transactions on Information Theory, 40*(1), 259–266.

He, X., & Tay, W. P. (2017). Multilayer sensor network for information privacy. In *2017 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 6005–6009).

He, X., Tay, W. P., & Sun, M. (2016). Privacy-aware decentralized detection using linear precoding. In *2016 ieee sensor array and multichannel signal processing workshop (sam)* (pp. 1–5).

Ho, S., & Verdú, S. (2010). On the interplay between conditional entropy and error probability. *IEEE Transactions on Information Theory, 56*(12), 5930–5942.

Jia, R., Dong, R., Sastry, S. S., & Spanos, C. J. (2017). Privacy-enhanced architecture for occupancy-based hvac control. In *Proceedings of the 8th international conference on cyber-physical systems, ICCPS '17* (pp. 177–186). New York, NY, USA: ACM.

Kalantari, K., Sankar, L., & Kosut, O. (2017). On information-theoretic privacy with general distortion cost functions. In *2017 ieee international symposium on information theory (isit)* (pp. 2865–2869).

Kleiminger, W., Santini, S., & Mattern, F. (2014). Smart heating control with occupancy prediction: how much can one save? In *Proceedings of the 2014 acm international joint conference on pervasive and ubiquitous computing: Adjunct publication* (pp. 947–954).

Kogiso, K., & Fujita, T. (2015). Cyber-security enhancement of networked control systems using homomorphic encryption. In *2015 54th ieee conference on decision and control (cdc)* (pp. 6836–6843).

Kramer, G. (2003). Capacity results for the discrete memoryless network. *IEEE Transactions on Information Theory, 49*(1), 4–21.

Li, S., Khisti, A., & Mahajan, A. (2018). Information-theoretic privacy for smart metering systems with a rechargeable battery. *IEEE Transactions on Information Theory, 64*(5), 3679–3695.

Li, Z. (2017). *Privacy-by-Design for Cyber-Physical Systems* Ph.D. thesis..

Li, Z., & Oechtering, T. J. (2015). Privacy-aware distributed bayesian detection. *IEEE Journal of Selected Topics in Signal Processing, 9*(7), 1345–1357.

Li, Z., & Oechtering, T. J. (2017). Privacy-constrained parallel distributed neyman–pearson test. *IEEE Transactions on Signal and Information Processing over Networks, 3*(1), 77–90.

Li, Z., Oechtering, T. J., & Skoglund, M. (2016). Privacy-preserving energy flow control in smart grids. In *2016 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 2194–2198).

Liao, J., Sankar, L., Tan, V. Y. F., & du Pin Calmon, F. (2018). Hypothesis testing under mutual information privacy constraints in the high privacy regime. *IEEE Transactions on Information Forensics and Security, 13*(4), 1058–1071.

Mehra, R., & Peschon, J. (1971). An innovations approach to fault detection and diagnosis in dynamic systems. *Automatica, 7*(5), 637–640.
Mo, Y., & Murray, R. M. (2017). Privacy preserving average consensus. *IEEE Transactions on Automatic Control, 62*(2), 753–765.
Moraffah, B., & Sankar, L. (2015). Information-theoretic private interactive mechanism. In *2015 53rd annual allerton conference on communication, control, and computing* (pp. 911–918).
Nekouei, E., Sandberg, H., Skoglund, M., & K. H. , J. (2018a). Privacy-aware Minimum Error Probability Estimation: An Entropy Constrained Approach. *Technical Report*.
Nekouei, E., Skoglund, M., & Johansson, K. H. (2018b). Privacy of information sharing schemes in a cloud-based multi-sensor estimation problem. In *2018 annual american control conference (acc)* (pp. 998–1002).
Nozari, E., Tallapragada, P., & Cortés, J. (2017). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica, 81*, 221–231.
Nozari, E., Tallapragada, P., & Cortés, J. (2018). Differentially private distributed convex optimization via functional perturbation. *IEEE Transactions on Control of Network Systems, 5*(1), 395–408.
Ny, J. L., & Pappas, G. J. (2014). Differentially private filtering. *IEEE Transactions on Automatic Control, 59*(2), 341–354.
du Pin Calmon, F., Makhdoumi, A., Médard, M., Varia, M., Christiansen, M., & Duffy, K. R. (2017). Principal inertia components and applications. *IEEE Transactions on Information Theory, 63*(8), 5011–5038.
du Pin Calmon, F., & Fawaz, N. (2012). Privacy against statistical inference. In *2012 50th annual allerton conference on communication, control, and computing* (pp. 1401–1408).
Poor, H. V. (2013). *An introduction to signal detection and estimation*. Springer Science & Business Media.
Rassouli, B., & Gündüz, D. (2017). *On perfect privacy and maximal correlation*.
Sandberg, H., Dán, G., & Thobaben, R. (2015). Differentially private state estimation in distribution networks with smart meters. In *2015 54th ieee conference on decision and control (cdc)* (pp. 4492–4498).
Sun, M., & Tay, W. P. (2016). Privacy-preserving nonparametric decentralized detection. In *2016 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 6270–6274).
Tanaka, T., & Sandberg, H. (2015). SDP-Based joint sensor and controller design for information-regularized optimal LQG control. In *The 54th IEEE Conference on Decision and Control (CDC)*.
Tanaka, T., Skoglund, M., Sandberg, H., & Johansson, K. (2017). Directed Information as Privacy Measure in Cloud-based Control. *Technical Report*. KTH Royal Institute of Technology, Sweden.
Wang, Y., Huang, Z., Mitra, S., & Dullerud, G. E. (2017). Differential privacy in linear distributed control systems: entropy minimizing mechanisms and performance tradeoffs. *IEEE Transactions on Control of Network Systems, 4*(1), 118–130.
Zhang, T. (2006). Information-theoretic upper and lower bounds for statistical estimation. *IEEE Transactions on Information Theory, 52*(4), 1307–1321.