Title

Quantifying the Utility--Privacy Tradeoff in the Internet of Things

Journal

ACM Transactions on Cyber-Physical Systems, 2(2)

ISSN

2378-962X

Authors

Dong, Roy
Ratliff, Lillian J
Cárdenas, Alvaro A
et al.

Publication Date

2018-04-30

DOI

10.1145/3185511

Peer reviewed

Quantifying the Utility–Privacy Tradeoff in the Internet of Things

ROY DONG, University of California, Berkeley
LILLIAN J. RATLIFF, University of Washington
ALVARO A. CÁRDENAS, University of Texas at Dallas
HENRIK OHLSSON and S. SHANKAR SASTRY, University of California, Berkeley

The Internet of Things (IoT) promises many advantages in the control and monitoring of physical systems from both efficacy and efficiency perspectives. However, in the wrong hands, the data might pose a privacy threat. In this article, we consider the tradeoff between the operational value of data collected in the IoT and the privacy of consumers. We present a general framework for quantifying this tradeoff in the IoT, and focus on a smart grid application for a proof of concept. In particular, we analyze the tradeoff between smart grid operations and how often data are collected by considering a realistic direct-load control example using thermostatically controlled loads, and we give simulation results to show how its performance degrades as the sampling frequency decreases. Additionally, we introduce a new privacy metric, which we call inferential privacy. This privacy metric assumes a strong adversary model and provides an upper bound on the adversary's ability to infer a private parameter, independent of the algorithm he uses. Combining these two results allows us to directly consider the tradeoff between better operational performance and consumer privacy.

CCS Concepts: • **Security and privacy → Economics of security and privacy**; *Information-theoretic techniques*; *Privacy-preserving protocols*;

Additional Key Words and Phrases: Privacy, smart grid

ACM Reference format:
Roy Dong, Lillian J. Ratliff, Alvaro A. Cárdenas, Henrik Ohlsson, and S. Shankar Sastry. 2018. Quantifying the Utility–Privacy Tradeoff in the Internet of Things. *ACM Trans. Cyber-Phys. Syst.* 2, 2, Article 8 (May 2018), 28 pages.
https://doi.org/10.1145/3185511

1 INTRODUCTION

The Internet of Things (IoT) collects and transmits a large amount of data. These data enable a multitude of advantages to all parties, but also presents a privacy risk to consumers who are now sharing data that may contain private information or may be statistically correlated with infor-

This work was supported in part by Foundations Of Resilient CybEr-Physical Systems (FORCES), which receives support from the National Science Foundation (NSF award numbers CNS-1238959, CNS-1238962, CNS-1239054, CNS-1239166). This work was also supported by NSF CNS-1553683.
Authors' addresses: R. Dong and S. S. Sastry, 337 Cory Hall, Berkeley, CA, 94720; emails: {roydong, sastry}@eecs.berkeley.edu; L. J. Ratliff, 185 Stevens Way Campus Box 352500, Seattle, WA 98195-2500; email: ratliffl@uw.edu; A. A. Cárdenas, 800 West Campbell Rd, EC31, Richardson, TX 75080-3021; email: alvaro.cardenas@utdallas.edu; H. Ohlsson, 1300 Seaport Blvd #500, Redwood City, CA 94063; email: henrik.ohlsson@c3iot.ai.

https://doi.org/10.1145/3185511

mation considered private. For example, in the smart grid, IoT devices promise better efficiency in energy distribution, more reliability, and transparency to electric utility customers in their energy consumption. On the other hand, monitoring energy consumption at high granularity can allow the inference of detailed information about consumers' lives such as the times they eat, when they watch TV, and when they take a shower (Lisovich et al. 2010). Such information is highly valuable and will be sought by many parties, including advertising companies (Anderson and Fuloria 2010), law enforcement (Smith 2012), and criminals (Thompson 2011).

As the IoT modernizes our lives and infrastructures, privacy emerges as a major concern. In this article, we hope to address privacy issues in the IoT by presenting a formal framework by which privacy-aware IoT service models can be designed. Currently, legislators are investigating technology-aware policies to ensure customer privacy; the final form of these policies is yet to be determined but will likely have a large impact on the shape of the IoT in the future. Additionally, researchers are looking at ways to design these systems in a fashion that acknowledges privacy.

Again referring to the smart grid example for concreteness: governments, researchers, and organizations are working on privacy standards and policies to guide advanced metering infrastructure (AMI) deployments. Researchers have considered the issue of data privacy in smart grid infrastructures and have proposed novel mechanisms for protecting the collected data (encryption, access control, and cryptographic commitments) (Kursawe et al. 2011; Rial and Danezis 2011), by anonymization and aggregation (Taban and Gligor 2009; Li et al. 2010), and by preventing inferences and re-identification from databases that allow queries from untrusted third parties (via differential privacy) (Acs and Castelluccia 2011).

To successfully understand the utility–privacy tradeoff in these smart grid operations, we must quantify two things. First, we must model the tradeoff between the quality of collected data and performance of smart grid operations. Second, we must understand how data quality affects an adversary's ability to infer a consumer's private information. More generally, in the IoT, we will have to quantify this utility–privacy tradeoff for any data collected to design privacy-aware IoT service models.

The underlying philosophy of our work is that these data transmission policies often unintentionally transmit information about private parameters unrelated to the original control goal: We separate operational parameters from parameters users may consider "private." Furthermore, the operational goals of a systems operator are different from the inferential goals of a privacy-breaching adversary. Thus, different types of analyses are needed to understand the tradeoff between data collection and smart grid performance versus the tradeoff between data collection and user privacy.

For example, to quantify how much data are needed for smart grid operations, we consider how the performance of proposed direct load control (DLC) mechanisms change as fewer and fewer measurements are received by the controller. To quantify the privacy risk in these mechanisms, we use recent results in nonintrusive load monitoring (NILM) to give theoretical guarantees on when NILM algorithms will fail: Adversaries will not be able to infer the device usage of a consumer from observing the aggregate power consumption of a building. Additionally, we model the private parameters of a consumer and the inferences that can be made about private parameters from device usage patterns.

Once this analysis is done, we can apply our framework for understanding the utility–privacy tradeoff in the IoT. We note that quantifying the operational value of data, picking a useful privacy metric, and applying the chosen metric to the appropriate models to provide a meaningful guarantee of privacy are all highly context-dependent actions. The hopes of a one-size-fits-all privacy solution is likely impossible due not only to formal properties of different IoT technologies but also social aspects of the system and what is considered "private" in the technology's application

domain. However, we maintain hope for a general framework by which to begin to devise and analyze privacy solutions.

The rest of the article is organized as follows. In Section 2, we discuss the literature in applied privacy. In Section 3, we explicate a general framework for modeling the utility–privacy tradeoff in IoT applications. In Section 4, we introduce a new privacy metric, *inferential privacy*, and specify contexts in which this privacy metric is applicable. Then, we apply this utility–privacy framework in a smart grid example in Section 5. Finally, we conclude in Section 6.

1.1 Notation

Throughout this article, we will use the following notation.

$\mathbb{R}$ will refer to the set of real numbers, and $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ will refer to the set of nonnegative real numbers. $\mathbb{N}$ will refer to the set of natural numbers.

We will consider random variables defined on some probability space $(\Omega, \mathcal{A}, \Pr)$ taking values in the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-field on $\mathbb{R}$. A random variable $X : \Omega \to \mathbb{R}$ has a probability distribution $P_X : \mathcal{B}(\mathbb{R}) \to [0, 1]$ if $P_X(B) = \Pr(X^{-1}(B))$ for any $B \in \mathcal{B}(\mathbb{R})$. In this article, we will suppose all random variables have probability distributions that are absolutely continuous with respect to the Lebesgue measure λ, so we will let p_X denote the probability density function such that $P_X(B) = \int_B p_X(x)\lambda(dx)$. Similarly, we will write $X \sim P_X$ to denote the fact that X is distributed according to probability distribution P_X.

For a proposition p, we let $1\{p\}$ denote the indicator function of that proposition, i.e., $1\{p\} = 1$ if p is true, and $1\{p\} = 0$ otherwise.

2 BACKGROUND AND RELATED WORK

Some of the earliest literature in applied privacy was ensuring that surveys could be conducted in a privacy-preserving fashion. These methods were called *randomized response* methods (Warner 1965; Greenberg et al. 1969). These researchers noticed there was structural bias when surveys requested sensitive information, such as whether or not a subject was HIV-positive. The key component for guaranteeing privacy was to given individual subjects *deniability*: a positive answer could either be a true response or due to the randomness in the survey procedure.

The next advances in the applied privacy literature was in statistical databases. In Dalenius (1977), the author argues that any definition of privacy should satisfy the following desideratum: Nothing can be learned about a user with the database that could not be learned without the database. One attempt to satisfy this desideratum was k-anonymity (Sweeney 2002), which provides methods to ensure that for any one user, there are at least $k - 1$ users who appear indistinguishable from said user.

More recently, the advent of Big Data has introduced many databases with potentially sensitive data that could be utilized by an adversary as side information to infer private facts. For example, in Narayanan and Shmatikov (2006), the authors are able to take anonymized Netflix data and, using publicly available information from IMDB, recover the identities of individual users. These results pushed researchers to no longer consider privacy of a database in isolation but in the larger context of widely available side information.

Recent research has been focused on designing privacy metrics to quantify privacy risks. We will outline two major examples briefly and introduce a third privacy metric in Section 4.

Arguably the most popular privacy metric, *differential privacy*, was introduced in Dwork (2006). Differential privacy requires an exogenous adjacency relationship, which specifies pairs of potential values for private parameters that we hope to keep indistinguishable. With this adjacency relationship, differential privacy is a bound on the change in the distribution of the observables between any two adjacent private parameters.

Differential privacy is attack-agnostic, in the sense that as a metric it does not suppose the adversary launches a particular type of inference attack. Furthermore, differential privacy is also agnostic to the amount of side information an adversary has, since it simply captures how much the distribution of the observable changes for small perturbations to private parameters. Since our work was first submitted, the authors of Yang et al. (2017) provide a survey of results using differential privacy to publish data and its effect on utility, as measured by statistical properties of the privatized data.

An alternative definition uses an information-theoretic metric to quantify privacy loss. In particular, *mutual information* between a private parameter and a public observable has recently become a popular metric (du Pin Calmon and Fawaz 2012; Sankar et al. 2013). One interpretation of the mutual information is the difference between the entropy of the prior distribution and the entropy of the posterior distribution (Cover and Thomas 1991); from that perspective, this metric has an intuitive interpretation as quantifying the reduction in the uncertainty of an adversary due to a public observable. This metric is attack-agnostic, since it simply quantifies a statistical relationship between private and public variables. However, this requires a specification of the available side information to an adversary, as reflected in the prior distribution.

It is our belief that, similarly to previous technological changes, the Internet of Things will motivate a sea change in how privacy is perceived, defined, quantified, and treated. All the previously state references consider privacy in the context of databases, but a nascent area of research is the investigation of how privacy can be understood in the context of systems with dynamics. In Stankovic (2014), the author cites privacy as one of the key research problems in IoT, and in Lin et al. (2017), the authors note that privacy issues are one of the open problems in IoT, with as most perturbation-based methods cause decreases in the utility of data.

Recent work in this regard includes the extension of differential privacy to Kalman filtering (Le Ny and Pappas 2014), constrained optimization (Han et al. 2014) and convex optimization (Hsu et al. 2014), distributed control (Huang et al. 2014), and online learning (Dong et al. 2015). Similarly, there have been efforts to consider information-theoretic metrics in the context of dynamic systems such as the smart grid (Rajagopalan et al. 2011; Jia et al. 2016). There have also been game-theoretic approaches to modeling how participants in IoT systems may selectively hide private information (Jin et al. 2017).

There is also recent work studying the behavioral aspects of people's relationship to their data and privacy decisions. For example, there are methods in what is often referred to as *usable privacy*. These methods outline a particular situation in which there is a privacy problematic and design solutions for these situations. For example, Shay et al. (2016) works on designing password policies that are reliable, easy to remember, and still secure against most attacks. In Roesner et al. (2012), the authors re-design smartphone applications to limit their access to private materials in a convenient, easy-to-implement and user-intuitive fashion. For example, rather than granting the Facebook app constant access to one's camera, a user can instead only allow Facebook to access the camera when he or she clicks a camera button, which will be chosen from a pre-determined set of icons that serve as clear indicators that grant camera access permission.

Alternatively, researchers quantify privacy preferences using behavioral studies. For example, in Acquisti et al. (2015), the authors experiment to see how much information users will reveal about themselves in different contexts. One interesting result from this area is known as the "paradox of control": When users are given more control over their privacy settings, they tend to reveal more about themselves. It is not entirely known why people do this, but one hypothesis that is common in the community is that control over privacy settings creates a false sense of security: Users naturally believe they have the ability to recall private information once it is "out there."

In Butler et al. (2015), researchers consider the efficacy of robot teleoperation under different privacy-preserving filters. Robot operators are given tasks to perform by remotely controlling a robot, with the visual feedback obscured by different image filters. The privacy levels are evaluated by user survey questions.

A lot of behavioral methods are very compelling, because it is difficult to turn a person's emotions and decision making around their privacy into a mathematical object. Researchers using these methods focus on ensuring that, when they evaluate privacy levels, it closely aligns with either a person's privacy assessment or a person's decision to reveal information. Grounding this with user studies ensures that this research maintains applicability to real-life applications.

However, these studies are often done on a case-by-case basis and it is difficult to generalize across multiple domains and studies to develop a unifying framework of privacy design. In this article, we will focus on formulations of privacy that attempt to mathematically quantify the level of privacy of data, independent of a user's perception of their own privacy risks.

The contributions of this article are as follows.

(1) We introduce a framework for quantifying the utility–privacy tradeoff in the Internet of Things. That is, we consider how modifying the *quality* of data, e.g., data with different levels of additive noise or sampled at lower frequencies, affects the operational value and the privacy levels of data.
(2) We instantiate this framework in the context of the smart grid by analyzing the operational value of data as well as the privacy risks inherent in data for different sampling rates in the operation of a recently proposed direct load control method. We are able to calculate the utility–privacy tradeoff for these programs.
(3) In the process of applying this framework, we extended a state-of-the-art direct load control mechanism to handle missing measurements.
(4) We introduce a new privacy metric, *inferential privacy*, broadly applicable in IoT settings. This privacy metric is a bound on the probability that an adversary will correctly infer a private parameter based on observations; its formulation allows us to leverage many classical results in optimal estimator design and minimax risk.

3 UTILITY-PRIVACY FRAMEWORK

In this section, we introduce a framework for quantifying the tradeoff between the operational utility of data and the privacy levels of consumers.

Privacy-preserving mechanisms can be divided into two categories: mechanisms that control *access* to data or mechanisms that vary the *quality* of data.

Access control methods have been researched primarily by the cryptography community, with very strong results (Diffie and Hellman 1979). The former can provide strong guarantees of privacy against outside adversaries but does not protect users from privacy breaches by those who have access to the data. For example, your utility company should have access to your energy consumption, but they may be able to infer aspects of your lifestyle from these patterns (Lisovich et al. 2010; Dong et al. 2013). In the Internet of Things, Google can receive your Nest sensor readings, and Fitbit may receive your GPS and step data, but unless there is an operational justification for the collection of data, it is likely that users will find this data collection invasive and unnecessary.

In contrast, quality-based methods have been researched by several communities. For example, most differential privacy mechanisms add noise to the data (Dwork 2006; Dwork and Roth 2014): As the noise levels increase, the quality of the data decreases and privacy levels increase as well. As another example, systems can sample real-time data less frequently to increase the privacy levels of consumers; these mechanisms are considered in Cárdenas et al. (2012) and Giraldo et al. (2014).

By modifying the quality of the data prior to its transmission, these methods guarantee privacy against both outside adversaries and insiders. However, the modifications to the data's quality must be carefully designed to not erode its original utility; if the data are no longer useful for its intended purposes, then the efficiency and comfort benefits of these novel technologies will be lost.

In this article, we will focus on privacy-preserving mechanisms that vary the quality of data to achieve different levels of privacy. By using this framework, system designers can intentionally set the level of privacy that achieves a desirable balance between collecting a quality of data that is sufficiently useful as well as privacy preserving. Currently, in practice, when design parameters relevant to privacy are considered, they are often chosen in an ad hoc fashion, without a principled way to analyze how variations in design parameters affect privacy and utility. However, we note that this framework does not provide a clear means to analyze the privacy tradeoff for different levels of access control for different users.

3.1 IoT Architecture

One of the defining aspects of the Internet of Things is a wide variety of architectures, network configurations, and protocols simultaneously co-existing across multiple application domains. We generally model IoT systems as having three components:

(1) *Data sensors and transmitters:* These are devices located at the physical systems that sense and transmit data. For example, this includes advanced metering infrastructures and Fitbit devices. In our framework, we assume the adversary does not have the ability to tamper with these devices. In other words, design parameters such as the sampling rate of sensors cannot be modified. However, the adversary may be able to listen to any data recorded by these devices.
(2) *Network:* Data transmitters upload the sensor data onto a network, which transmits the data to intended recipients, e.g., Fitbit data are transmitted to the phone application and Fitbit's servers, advanced metering infrastructures transmit energy consumption data to the utility company. In our framework, we assume the adversary may have access to any packets transmitted on the network.
(3) *Network listeners:* The third component of the system is any entity that listens to network transmissions. This covers both legitimate listeners and illegitimate eavesdroppers. We consider the privacy concerns of consumers against both legitimate and illegitimate listeners. This is a distinct facet of privacy we consider important. For example, if an advanced metering infrastructure records high-frequency data of a user's energy consumption, encrypts the data, and sends it to the utility company, then there may be no breach by an adversary, and, therefore, no security issue. However, if the utility company has no reason to collect such high-frequency data of users, there is a potential privacy breach of the consumer. Even legitimate listeners should have to justify the collection of potentially private data by the operational utility of said data.

Essentially, we abstract out the problems of the adversary's identity by assuming that the adversary has access to any data that is recorded by data sensors. In particular, we analyze the privacy of consumers with regards to both legitimate and illegitimate listeners that view the data.

We note that in most IoT applications, energy consumption is a fundamental concern, as many IoT devices are meant to function for relatively long periods of time without recharging. For some design parameters, the energy usage is not significantly affected. However, some forms of privacy-preserving mechanisms do affect the power usage of devices, e.g., encryption techniques require a fair amount of computational power on the part of small microprocessors. Insofar as this energy usage can be modeled as an operational cost, our framework can capture this problem.

If the design parameter is the location of computational load of data processing, then this becomes an access control problem that is outside the scope of our framework. For example, if Fitbit were to implement additive noise to the user's data for privacy, should this process happen at the physical Fitbit device or in the smartphone application? Our framework is meant to handle variations in data quality, not that entities have access to un-noised data.

3.2 The Utility of Data

In the Internet of Things, the utility of a particular set of data comes from the improvement in the performance of some service due to said data. To model these systems, we follow a control-theoretic framework.

We are interested in the performance of our system at some set of times $T \subset \mathbb{R}_+$. This includes the discrete time cases $T = \{0, 1, \ldots, N\}$ for some $N \in \mathbb{N}$ or $T = \{0, 1, \ldots\}$, as well as the continuous time cases $T = [0, T_f]$ for some $T_f \in \mathbb{R}$ or $T = [0, \infty)$. For simplicity, we will assume operation of the system begins at time $t = 0$ and $0 \in T$.

Our system has some state space $\mathcal{X}$, which represents all possible configurations of the system at one point in time. Usually, we will take $\mathcal{X} = \mathbb{R}^n$ for some $n \in \mathbb{N}$. We will denote the state at time $t \in T$ as $x(t)$. Similarly, the control actions we can take on the system live in some input space, $\mathcal{U}$, with the input at time t denoted $u(t)$. The dynamics of the system are captured in a function $\phi : \mathcal{X} \times \mathcal{U}^T \to \mathcal{X}^T$, which takes an initial condition and an input signal across all T and specifies which trajectory in $\mathcal{X}^T$ the system will follow. For example, in the context of linear time-invariant systems, if $\phi(x_0, u) = x$, then x is the unique solution to differential equation $\dot{x}(t) = Ax(t) + Bu(t)$ with initial condition $x(0) = x_0$.

The performance of the system is evaluated with respect to a cost function $J : \mathcal{X}^T \times \mathcal{U}^T \to \mathbb{R}$. The system has an initial condition $x_0 \in \mathcal{X}$ and obeys the system dynamics ϕ. The system operator wants to pick a $u \in \mathcal{U}^T$ such that $J(\phi(x_0, u), u)$ is kept low. Ideally, the optimal control problem would be solved: $\min_u J(\phi(x_0, u), u)$. However, this often is difficult, and, in practice, we will use controllers that will approximate the optimal control strategy subject to information and tractability constraints.

To attempt to minimize this cost, the system operators will design a controller. This controller will determine the input $u \in \mathcal{U}^T$ that will be given to the system. However, this controller will have a limited amount of data about the system. In our framework, we will consider how variations in the quality of the data affect the system operator's control decisions and therefore affect the realized cost of the system. For some quality level q and time $t \in T$, we will let $Y(q, t)$ denote the data available to the controller at time t.[1] With these data, the controller will pick a control input $u(t) \in \mathcal{U}$. We let this process be denoted $u_c(Y(q, t), t) \in \mathcal{U}$.

With this controller specified, we can consider the mapping from quality level q to realized cost J. That is, for a particular quality q, the controller will use the controller and issue control command $u_c(Y(q, t), t) \in \mathcal{U}$ at each time $t \in T$. This will cause the realized cost to be $J(\phi(x_0, u_q), u_q)$, where $u_q \in \mathcal{U}^T$ is defined as $u_q(t) = u_c(Y(q, t), t) \in \mathcal{U}$ for every $t \in T$.

Abstractly, this allows us to quantify the utility of data by showing how the control performance of the cyber-physical system erodes for different quality levels of data. As previously mentioned, we will instantiate this in a concrete example in Section 5.

[1]For generality, we have not included details of what space these objects q and $Y(q, t)$ live in. Formally, q can live in a general space, but we will often think of $q \in \mathbb{R}$. For example, q can denote the sampling period of our system, as we will explore in Section 5. Similarly, $Y(q, t)$ can live in some arbitrary space for each q and t. In the example in Section 5, $Y(q, t)$ will be a collection of random variables that the controller can observe at time t.

3.3 The Privacy of Data

Data are collected from consumers with the intent of improving IoT operations. However, these data also allow the inference of private information about consumers, unrelated to IoT operations. This section quantifies how much information about the private lives of consumers is contained in data.

In the previous section, we fixed a set of time indices $T \subset \mathbb{R}_+$, and defined a data mechanism $Y(q,t)$ for each quality level q and time t. This data mechanism defined what information is collected and transmitted, and we quantified how a controller's performance changes as the quality level q is varied. In this section, we will consider how variations in q affect the privacy levels of consumers in the data mechanism $Y(q,t)$. We take a statistical perspective on privacy: What is the inferential power of these new observations relative to some private parameter? Our model is as follows.

Users have a private parameter $\theta \in \Theta$, which they wish to protect. These private parameters θ live in a space Θ with some particular structure, which depends on the privacy metric in use. In differential privacy, the private parameter space Θ is equipped with an "adjacency" relationship specifies which pairs $(\theta, \theta') \in \Theta \times \Theta$, which should be indistinguishable. For information-theoretic metrics and the inferential privacy metric used in Section 4, θ is seen as a random variable taking finitely many values, i.e., Θ has finitely many elements and there exist a prior distribution P_θ for the random variable θ.

These privacy metrics should be general enough in definition to allow evaluation for any data mechanism Y under consideration. Additionally, it will depend on the quality q: So our privacy valuations will be a function of the structure of our data mechanism, as well as the quality level. This will be denoted $m(Y,q)$. This framework is general enough to capture any quality-varying privacy-preserving mechanisms, and this generality is needed to be able to encompass the spectrum of possible privacy risks and information structures in IoT. In Section 4, we will discuss privacy metrics in more detail. In particular, the majority of Section 4 will cover theoretical results that bound the probability an adversary will correctly infer an underlying private parameter from observable data.

4 PRIVACY METRICS AND INFERENTIAL PRIVACY

In Section 3, we outlined a general framework for quantifying the utility–privacy tradeoff in Internet of Things applications. Before covering an concrete example in the smart grid, we will discuss a new privacy metric, *inferential privacy*. Informally, inferential privacy is a guaranteed lower bound on the probability an adversary will correctly infer θ from the observations $(Y(q,t))_{t\in T}$.

Depending on the context, different privacy metrics may be more applicable than others. As argued in the philosophy of privacy (Nissenbaum 2004), we believe that a plurality of privacy metrics and definitions is required to capture the essence of a concept as context dependent and essentially contested as privacy (Solove 2002).

Differential privacy gives a very powerful guarantee that is both agnostic to attacks and adversary's available side information; however, many practical applications require a particular structure of the uncertainty, such as additive independent Laplacian or Gaussian noise. For example, it is not clear how to consider how the level of differential privacy varies in a dynamical system when the sampling rate is adjusted.

In contrast, information-theoretic metrics lend themselves nicely to the design of noise in a fashion that is often optimal with respect to some criterion. In du Pin Calmon and Fawaz (2012), the authors are able to design an optimal noising scheme subject to a performance constraint in database estimation, and in Rajagopalan et al. (2011), the authors consider compression schemes in

the context of the smart grid, and provide a theoretical bounds on the information leakage subject to a distortion constraint.

Our work on privacy builds on a hypothesis testing framework, which has been well studied in the information theory (Cover and Thomas 1991) and statistics (Keener 2010) communities. Variational calculus methods for statistics were first introduced by Neyman and Pearson (1933) and have been a fruitful way to find optimal estimators. Additionally, a popular metric for critiquing the performance of an estimator is known as the minimax risk, which measures an estimator's expected loss against a worst-case distribution (Le Cam 1973); the minimax risk can act as a measure of the difficulty of a hypothesis testing problem. Alternatively, Fano was able to analyze the difficulty of the hypothesis testing problem by considering the entropy and mutual information between the parameter of interest and the observables (Cover and Thomas 1991; Yu 1997); these results were extended to observations on the continuum in Han and Verdú (1994). Each of these methods can provide a measure of the hypothesis testing problem's difficulty, which we use as a guarantee for privacy.

Throughout this section, we will be analyzing the privacy level for a fixed quality q. Naturally, one can vary q afterwards to see the effect of quality on privacy levels.

4.1 User and Data Mechanism Model

First, we introduce a model for how the private parameter θ influences the observed data $(Y(q,t))_{t\in T}$. We will allow $\mathcal{Y}$ to denote the possible values of $(Y(q,t))_{t\in T}$.

Assumption 1. *The private parameter θ follows a distribution P_θ. Similarly, $(Y(q,t))_{t\in T}$ given θ has a conditional distribution $P_{y|\theta}$.*

We note that, formally, this assumption is quite succinct, but, in practice, determining these distributions is rarely trivial.

4.2 Adversary Model

Next, we introduce our adversary model.

Assumption 2. *Our adversary is able to observe the transmitted data $(Y(q,t))_{t\in T}$ and has knowledge of P_θ and $P_{y|\theta}$. Additionally, this adversary has an arbitrary amount of computational power.*

This adversary has access to the measured data signal and also holds priors on the consumer's private information θ. He also knows how this private information affects the consumer's usage of IoT devices, $P_{y|\theta}$. Although this adversary has quite a bit of knowledge about the consumers, he does not hold arbitrary side information.

We note that it may not be realistic to suppose the adversary has access to P_θ and $P_{y|\theta}$. However, any adversary who tries to infer θ from y with less information will only do worse than our adversary model. Thus, this model provides a conservative estimate against all weaker adversary models.

In the development of the theory in this section, we will use the convention that a citation listed inside the headers of propositions and corollaries will be the source of various theoretical claims; we refer the reader to these citations for proofs, as well as illuminating discussion of the surrounding theory.

4.3 Inferential Privacy Metric

Our privacy metric is the probability of error if an adversary tries to infer the private variable θ.

Definition 4.1. Under the usage model outlined in Assumption 1, a system is *"α inferentially private"* if, for any estimator $\widehat{\theta} : \mathcal{Y} \to \Theta$, we have

$$\Pr(\widehat{\theta}[(Y(q,t))_{t\in T}] \neq \theta) \geq \alpha. \tag{1}$$

This estimator can be based on information in P_θ and $P_{y|\theta}$.

Here we note that this is in essence an ex-ante privacy metric, i.e., the privacy is spread across Θ according to P_θ. As often arises in many statistical estimation problems, an ex-post privacy metric, i.e., a privacy metric that guarantees privacy for every type, is not a well-posed problem.

For example, suppose $\Theta = \{0, 1\}$, and consider the estimator $\widehat{\theta} \equiv 0$. For any consumer of type $\theta = 0$, the adversary will correctly infer their type with this estimator. In other words, an adversary can always violate the privacy of one type of consumer by making the blanket assumption that everyone is a fixed type. In a sense, we gain privacy by noting that the adversary has to be successful across the different types Θ (weighted according to P_θ).

Regardless of the algorithm the adversary uses, we can bound the probability it will successfully breach a consumer's privacy. Furthermore, this formula allows us to vary the quality level q, such as how often data are collected and transmitted. We will examine this on a concrete example in Section 5. This guarantee is also simple for consumers to interpret and can be used in the design of privacy contracts between the IoT service provider and consumers (Ratliff et al. 2015).

We will derive results that allow us to calculate values of α that satisfy the condition given in Definition 4.1. This section is a generalization of some of our previous work (Dong et al. 2014), which considered this definition in the context of aggregate energy observations.

There are three methods by which we can derive lower bounds. Depending on the particular form of P_θ and $P_{y|\theta}$, some forms of the lower bound may be easier to calculate than others.

In the sequel, we will refer to $(Y(q,t))_{t\in T}$ as Y, leaving the dependence on q implicit. Recall that we will use P_X to refer to the probability distribution of a random variable X and p_X to refer to its probability density function or probability mass function, as appropriate.

4.4 Likelihood-Based Methods

Let $\Theta = \{1, \ldots r\}$. We can define the maximum a posteriori (MAP) estimator $\widehat{\theta}_{MAP}$, which maximizes $\Pr(\widehat{\theta}(Y) = \theta)$.

Proposition 4.2 (Neyman and Pearson 1933; Cover and Thomas 1991). *Under Assumption 1,* $\Pr(\widehat{\theta}(Y) = \theta)$ *is maximized by*

$$\widehat{\theta}_{MAP}(Y) = \arg\max_{i\in\Theta}(p_\theta(i) \cdot p_{y|\theta}(Y|i)). \tag{2}$$

The proof of this proposition follows a variational calculus approach that was pioneered in Neyman and Pearson (1933).

The optimality of the MAP estimator with respect to the prior P_θ immediately leads to a guarantee of privacy.

Corollary 4.3. *Under the model outlined in Assumption 1, the system is α inferentially private, where $\alpha = \Pr(\widehat{\theta}_{MAP}(Y) \neq \theta)$. Furthermore, the system is* not *α' inferentially private for any $\alpha' > \alpha$.*

Proof. This follows immediately by noting that $\widehat{\theta}_{MAP}$ maximizes $\Pr(\widehat{\theta}(Y) = \theta)$; no other estimator $\widehat{\theta}(Y)$ can do better with respect to the adversary's goal of inferring the true value of θ. □

Although this bound is optimal, it is often difficult to calculate. In these instances, some of the latter bounds may be used as a surrogate.

4.5 Le Cam's Method

Le Cam's method is typically used in assessing minimax risk. More specifically, it is used to find a lower bound on the worst-case loss for an estimator. Here, "worst case" means that the performance of the estimator is evaluated on the distribution for which the loss is maximized. For more details, we refer readers to Le Cam (1973) and Yu (1997).

We present Le Cam's lemma in the context of our usage model. Again, let $\Theta = \{1, 2, \ldots, r\}$. First, we offer two definitions of distances between probability distributions.

Definition 4.4. The *total variation distance* between two densities p and q on a measure space $(X, \mathcal{A}, \mu)$ is given by

$$\begin{aligned} \|p - q\|_{TV} &= \sup_{A \in \mathcal{A}} \left| \int_A p(x) - q(x) \mu(dx) \right| \\ &= \frac{1}{2} \int_X |p(x) - q(x)| \mu(dx). \end{aligned} \tag{3}$$

Definition 4.5. The *Kullback-Leibler (KL) divergence* between two densities p and q on a measure space $(X, \mathcal{A}, \mu)$ is given by

$$D_{KL}(p\|q) = \int p(x) \log \frac{p(x)}{q(x)} \mu(dx). \tag{4}$$

Similarly, we will define the KL divergence between two random variables X and Y to be the KL divergence between their densities, and it will be denoted $D_{KL}(X\|Y)$.

Using Definitions 2 and 3, we now state Le Cam's Lemma.

Proposition 4.6 (Le Cam's lemma (Le Cam 1973; Yu 1997)). *Assume the usage model outlined in Assumption 1. Then, for any estimator $\widehat{\theta} : \mathcal{Y} \to \Theta$ and any distinct $j, j' \in \Theta$, we have*

$$\Pr(\widehat{\theta}(Y) \neq \theta \mid \theta = j) + \Pr(\widehat{\theta}(Y) \neq \theta \mid \theta = j') \geq \quad 1 - \|p_{y|\theta}(\cdot|j) - p_{y|\theta}(\cdot|j')\|_{TV}. \tag{5}$$

A quick corollary is a lower bound on the probability of error:

Corollary 4.7. *Under the assumptions of Proposition 4.6, $\Pr(\widehat{\theta} \neq \theta)$ is bounded below by*

$$\max_{j \neq j'} \left[\min(p_\theta(j), p_\theta(j')) \cdot (1 - \|p_{y|\theta}(\cdot|j) - p_{y|\theta}(\cdot|j')\|_{TV}) \right]. \tag{6}$$

Proof. Invoking Proposition 4.6, we have that, for any estimator $\widehat{\theta} : \mathcal{Y} \to \Theta$

$$\begin{aligned} \Pr(\widehat{\theta}(Y) \neq \theta) &= \sum_{i \in \Theta} p_\theta(i) \Pr(\widehat{\theta}(Y) \neq \theta | \theta = i) \\ &\geq p_\theta(j) \Pr(\widehat{\theta}(Y) \neq \theta | \theta = j) + p_\theta(j') \Pr(\widehat{\theta}(Y) \neq \theta | \theta = j') \\ &\geq \min(p_\theta(j), p_\theta(j')) \cdot \left[\Pr(\widehat{\theta}(Y) \neq \theta | \theta = j) \Pr(\widehat{\theta}(Y) \neq \theta | \theta = j') \right]. \end{aligned}$$

This held for any $j \neq j'$, so we have our desired result. □

In practice, it will suffice to find an over-approximation of the total variation distance. For example, we have Pinsker's inequality:

Proposition 4.8 (Pinsker's inequality (Tsybakov 2009)). *For any densities p and q:*

$$\|p - q\|_{TV} \leq \sqrt{\frac{1}{2} D_{KL}(p\|q)}. \tag{7}$$

Thus, we can provide a guarantee of inferential privacy.

Proposition 4.9. *Under the model outlined in Assumption 1, the system is α inferentially private, where*

$$\alpha = \max_{j \neq j'} \left[\min(p_\theta(j), p_\theta(j')) \cdot (1 - \|p_{y|\theta}(\cdot|j) - p_{y|\theta}(\cdot|j')\|_{TV}) \right]. \tag{8}$$

Proof. This follows immediately from the lower bound across all estimators provided by Corollary 4.7. □

4.6 Fano's Method

Fano's inequality relates the probability of error for a hypothesis test to the entropy and mutual information between a parameter and its estimator. Traditionally a concept defined on discrete random variables, it has also been extended to handle continuous distributions (Cover and Thomas 1991; Ibragimov and Has'minskii 1991; Han and Verdú 1994; Yu 1997). Here we will state Fano's inequality in the context of our usage model, where $\Theta = \{1, \ldots, r\}$.

Proposition 4.10 (Fano's inequality (Yu 1997)). *In the model of Assumption 1, for any estimator $\widehat{\theta} : \mathcal{Y} \to \Theta$, the probability of error $P(\widehat{\theta}(Y) \neq \theta)$ is bounded below by*

$$\frac{1}{\log(r-1)} \left[\log r - \frac{1}{r^2} \sum_{i,j} D_{KL} \left[p_{y|\theta}(\cdot|i) \middle\| p_{y|\theta}(\cdot|j) \right] - \log 2 \right]. \tag{9}$$

Thus, we have the quick corollary:

Corollary 4.11. *Under the usage model outlined in Assumption 1, the system is α inferentially private, where α is given by Equation (9).*

Proof. This follows immediately from the lower bound provided by Proposition 4.10. □

5 THE UTILITY–PRIVACY TRADEOFF IN DIRECT LOAD CONTROL PROGRAMS

In this section, we instantiate our utility–privacy framework in a concrete context. Specifically, we consider the privacy of DLC programs in the smart grid. By focusing on one application, we can show how these design principles can be practically applied. This allows us to explicitly analyze one design parameter, sampling frequency, and its effect on the utility of data as well as the privacy of users.

DLC has been a promising future direction for the smart grid for a variety of reasons. By controlling loads that can be modified without much impact on consumer satisfaction, we can allay many costs by shifting loads from peak demand and compensating for real-time load imbalances. Additionally, as renewable energy penetration increases, the generation side of power is growing more uncertain and will require demand flexibility. In this section, we will consider the load imbalance signal as exogenous, and use a DLC scheme to try and compensate the imbalance.

Additionally, such DLC policies are being deployed today. For example, Pacific Gas & Electric deployed the SmartAC program in spring 2007 (Alexander et al. 2008). Another provider of demand response (DR) services has recruited over 1.25 million residential customers in DLC programs and has deployed over 5 million DLC devices in the United States. In California, they have successfully curtailed over 25MW of power consumption since 2007 (Meehan 2013). As these programs are being deployed on a large scale, it is important to consider the privacy aspects of these programs (Lisovich et al. 2010).

In this article, we consider different sampling rates as a method of varying the quality of data q. Our motivations for this are twofold.

First, there are many cases where noise-free data are required, for practical, regulatory, performance, or economic reasons. For example, suppose random noise is added to your energy consumption signal before being transmitted to the utility company. A consequence of this mechanism is that the energy bill you receive will not be a deterministic function of your energy usage but rather a random variable with a conditional dependence on your energy usage. Many consumers may be unhappy with this mechanism in which they may be billed for more energy than they used, and a lot of regulatory overhead would be necessary for a utility company to roll out such a mechanism, even in the face of statistical arguments that the effect of such a random mechanism is negligible in the long run.

Second, an analysis of the effect of sampling rates on operational performance is the first step in enacting the *data minimization* principle for dynamical systems. In the United States, the Obama Administration examined privacy issues in its June 2011 smart grid policy framework report (Chopra et al. 2011). The report recommends that State and Federal regulators should consider, as a starting point, methods to ensure that consumers' detailed energy usage data are protected in a manner consistent with federal Fair Information Practice (FIP) Principles. One of the key principles is data minimization. This principle is consistent with the notion of privacy by design (Cavoukian 2011).

Similarly, the FIP principle of data minimization appear in smart grid privacy recommendations by the National Institute of Standards and Technology (The Smart Grid Interoperability Panel - Cyber Security Working Group 2014), the North American Energy Standards Board (North American Energy Standards Board 2015), the Department of Energy (Department of Energy 2010), the Texas Legislature and Public Utility Commission (Public Utility Commission of Texas 2014), and the California Public Utilities Commission (CPUC) (Peevey 2011).

The NISTIR 7628 (Locke and Gallagher 2014) expresses the data minimization principle in the smart grid context as follows:

> Limit the collection of data to only that necessary for Smart Grid operations, including planning and management, improving energy use and efficiency, account management, and billing.

All these recommendations and policy proposals have been broad in coverage by necessity, as regulators do not want to burden electric utilities with specific limits on what they can collect. However, electric utilities who want to follow these privacy recommendations do not have a sound reasoning principle to help them decide how much data is too little or too much. Our goal in this section is to start discussing scientifically sound principles that can help determine how much data to collect to achieve a certain level of functionality of the grid and how much privacy is granted to consumers under this data collection policy.

By analyzing the effect that sampling rate has on Smart Grid operations, we can begin to quantify the utility of data, a necessary first step to enacting data minimization. Intuitively, there should be a sampling rate where higher sampling frequencies have a negligible effect on the system's performance. For example, this could be due to the ability of the controller to leverage this high-frequency data or the timescales of the system itself. Conversely, there should intuitively be a sampling rate that is so low that the system's performance is comparable to the performance should the controller receive no measurements at all. Finding these regimes of operation is the goal of the first half of our framework.

As mentioned previously, there are several approaches to preserve the privacy of a consumer participating in an advanced metering infrastructure (AMI), including adding noise to data, modifying how data are aggregated, and the duration of data retention (Kursawe et al. 2011; Rial and Danezis 2011; Taban and Gligor 2009; Li et al. 2010; Acs and Castelluccia 2011; Sankar et al. 2013).

These quality-varying mechanisms are currently an active topic of research. We note that our work is complementary to these other privacy policies. Our analysis is meant to assist electric utilities in following privacy recommendations: We seek to determine how much data to collect and how often it should be collected. Once this is in place, encryption, anonymization, and aggregation techniques can be employed in tandem.

We evaluate the performance of a widely studied DLC scheme as a function of the sampling rate. As we will later show, increasing the sampling period is a means of improving the privacy of consumers. In particular, we focus on a DLC application using thermostatically controlled loads (TCLs) to manage load imbalances.

Our framework provides a means for grid operators to pick the sampling frequency of data in a structured way. This allows system designers to adhere to the data minimization principle, as well as show the design procedure to verify that the data minimization principle was implemented.

We note that this is an application in which the theory of differential privacy cannot be directly applied. The theoretical development of differential privacy often maintains bounds on the changes in the distribution of the output; most of the theory proves how additive noise can guarantee differential privacy. In contrast, varying the sampling rate can be seen as changing the codomain of the observables, and it is not obvious how to apply the theory of differential privacy in this context. This is another motivation for our arguments that a plurality of privacy metrics is needed, as some metrics may not be applicable in the analysis of some design parameters, e.g., sampling frequency.

This section is organized as follows. First, we outline the DLC model. Then, we will define a controller for this DLC model that corrects for load imbalances and formally specify how its control actions vary with different data sampling frequencies. Next, we estimate the performance of this controller with simulations. Afterward, we consider how privacy levels vary with different data sampling frequencies. Finally, we invoke the framework outlined in Section 3 to quantify the utility–privacy tradeoff in this application.

5.1 DLC Model

In this section, we consider one recently proposed DLC program for concreteness. We note that our contribution is a general framework for numerically analyzing the sensitivity of these DLC programs to different information collection policies. We consider this research to be complementary to other research in how parameters affect system performance (Lu 2012; Lu and Zhang 2013).

TCLs, which are often heating, ventilation, and air conditioning (HVAC) systems for buildings, are a promising avenue for the implementation of DLC policies (Callaway 2009; Perfumo et al. 2012). This is due to the fact that buildings have a thermal inertia and can, in essence, store energy. Moreover, power consumption can be deferred and shifted while resulting in an imperceptible change in temperature.

For easy reference, we include a table of the interpretation of various variables in Figure 1.

5.1.1 Thermostatically Controlled Load Model. There are several TCL and DLC models in the literature, e.g., Ruiz et al. (2009), Moura et al. (2013), and Mathieu et al. (2013), and our analysis can easily be applied to any of these models. For concreteness, we consider the model presented in Mathieu et al. (2013).

Let $\mathcal{I}$ denote the set of TCLs participating in a DLC program. We model the temperature evolution of each TCL $i \in \mathcal{I}$ as a discrete-time difference equation:

$$x_i(k+1) = a_i x_i(k) + (1-a_i)[T_{a,i}(k) - m_i(k)T_{g,i}] + \epsilon_i(k). \tag{10}$$

Variable	Interpretation
$\mathcal{I}$	The set of TCLs participating in a DLC program.
$x_i(k)$	The internal temperature of TCL $i \in \mathcal{I}$ at time k.
R_i	The thermal resistance of TCL i.
C_i	The thermal capacitance of TCL i.
h_B	The 'base sampling period', e.g. the time step of the system dynamics.
a_i	$a_i = \exp(-h_B/(R_iC_i))$, a term which appears in the dynamics of TCL i.
$T_{a,i}$	The ambient temperature around TCL i.
$P_{\text{trans},i}$	The energy transfer rate of TCL i.
P_i	The power consumed by TCL i when in the ON state.
$T_{g,i}$	The temperature gain for TCL i in the ON state, given by $T_{g,i} = R_iP_{\text{trans},i}$.
$m_i(k)$	The state of TCL i at time k; $m_i(k) = 1$ means that the TCL is in the ON state.
$\epsilon_i(k)$	The noise process of TCL i at time k.
$T_{\text{set},i}$	The temperature setpoint of TCL i.
δ_i	The deadband of TCL i.
P_{des}	The desired power signal to track.
β	A placeholder variable capturing all the parameters known by the DLC program.
h	The subsampling period, i.e. data is transmitted every h time steps of the underlying dynamics.
T_k	The time indices in which measurements are available at time k.
$\hat{x}_i(k), \hat{m}_i(k)$	The estimates of $x_i(k)$ and $m_i(k)$ maintained by the DLC program, respectively.
N_{bin}	The number of bins used by the centralized DLC controller.
φ_i	A function that normalizes the deadband of TCL i to $[0, 1]$.
ψ_i	A function that identifies the bin TCL i is currently in.

Fig. 1. A table of the variable names and their interpretation for our DLC model.

In the above equation, $x_i(k)$ is the internal temperature of TCL i at time k, $T_{a,i}$ is the ambient temperature around TCL i, m_i is the state of TCL i (either 1 or 0), and ϵ_i is a noise process.[2] The term $a_i = \exp(-h_B/(R_iC_i))$, where h_B is the base sampling period,[3] R_i is the thermal resistance of TCL i, and C_i is the thermal capacitance of TCL i. The T_g term represents the temperature gain when a TCL is in the ON state, and $T_g = R_iP_{\text{trans},i}$, where $P_{\text{trans},i}$ is the energy transfer rate of TCL i. Let P_i denote the power consumed by TCL i when it is in the ON state.

The local control for TCL i is modeled by the variable m_i. We assume the local controller performs an ON/OFF hysteresis control based on its setpoint and deadband. For a cooling TCL, this is defined as

$$m_i(k+1) = \begin{cases} 0 & \text{if } x_i(k+1) < T_{\text{set},i} - \delta_i/2 \\ 1 & \text{if } x_i(k+1) > T_{\text{set},i} + \delta_i/2 \\ m_i(k) & \text{otherwise} \end{cases}. \tag{11}$$

[2]Our development focuses on air conditioning for notational simplicity, but similar statements can be made for heaters.
[3]Here, h_B denotes the timescale of the dynamics. Later, we will introduce how often the direct load controller may receive fewer measurements to preserve privacy, and this subsampling period will be denoted h.

In these equations, $T_{\text{set},i}$ and δ_i are the temperature setpoint and deadband of TCL i, respectively. If $m_i(k) = 1$, then we say that TCL i is in the ON state at time k, and, similarly, $m_i(k) = 0$ means that i is in the OFF state at k.

In the next few sections, we will assume that these local control signals can also be overridden by the direct load controller, replacing Equation (11). The controller will only intermittently has access to observations $(x_i(k), m_i(k))$, due to a privacy-aware sampling policy.

5.1.2 Direct Load Control Objective. We consider DLC policies that attempt to compensate for load imbalances and defer demands from peak times by switching TCLs between the ON state and the OFF state. The marginal cost of peak loads and unexpected load imbalances is responsible for a large portion of the preventable costs in the electricity grid; for a more detailed treatment of the benefits and impact of a DLC policy that can shave demand, we refer the reader to Callaway and Hiskens (2011).

Formally, we consider the load imbalance as an exogenous variable. In particular, the centralized DLC controller is given some desired power trajectory P_{des} for the TCLs.[4] The goal of the controller is to minimize the error between the actual power consumed by the TCLs and the signal P_{des}, i.e., it wishes to minimize $\sum_k \left|\sum_{i\in\mathcal{I}} P_i m_i(k) - P_{\text{des}}(k)\right|$.

5.1.3 Direct Load Control Capabilities. To achieve the DLC objective, we assume the centralized DLC controller has the capability of telling TCLs to switch modes between ON and OFF when the temperature $x(k)$ is between $T_{\text{set},i} - \delta_i/2$ and $T_{\text{set},i} + \delta_i/2$. More explicitly, if the centralized DLC controller issues a command to a TCL to switch from OFF to ON, then the TCL turns on its air conditioner earlier than it would have in the absence of a control command. This DLC command will override the local controller. We assume that the centralized DLC controller has no control authority when the temperature is outside of the deadband, with the local controller deterministically in the OFF state when $x(k) < T_{\text{set},i} - \delta_i/2$ and in the ON state when $x(k) > T_{\text{set},i} + \delta_i/2$.

Note that the control policy effectively tightens the deadband. In particular, this control policy maintains customer satisfaction in the sense that the effective deadband is never larger than the user-specified deadband.

Our model of a direct load controller is as follows. We assume the centralized DLC controller has access to the parameters $\beta = (a_i, T_{a,i}, T_{g,i}, T_{\text{set},i}, \delta_i, P_i)$ for each TCL $i \in \mathcal{I}$. In other words, the controller knows the dynamics of each TCL. However, it is only able to observe the signals $(x_i(k), m_i(k))$ for certain values of k, determined by the privacy-aware sampling policy. One of the contributions of this article is the extension of a DLC controller to situations where measurements are intermittent.

For the rest of this section, we will assume a privacy-preserving sampling policy that considers subsampling rates. In other words, our sampling policy is parameterized by a subsampling period $h \in \mathbb{N}$, and at time k, the centralized controller has access to the measurements $(x(k), m(k))_{k\in T_k}$, where the set $T_k = \{hl : l \in \mathbb{N}, hl \le k\}$ denotes the time indices in which measurements are available.[5]

5.2 Direct Load Controller

In this section, we outline a DLC policy inspired by work in the recent literature (Callaway 2009; Mathieu et al. 2013). Our model of a direct load controller is as follows. First, the controller

[4]We consider this load imbalance signal exogenous. In future work, we hope to examine elements of generation, such as scheduling, and how it is influenced by these programs.

[5]For simplicity, we assume that either all the TCLs transmit their state information at time k or none of them do. More asynchronous transmissions can be handled with some additional notational baggage.

maintains an estimate of the thermal state of each TCL. Let $\hat{x}_i(k)$ and $\hat{m}_i(k)$ denote the estimates of $x_i(k)$ and $m_i(k)$, respectively.

The estimator acts as follows:

$$\hat{x}_i(k) = \begin{cases} x_i(k) & \text{if } k \in T_k \\ a_i \hat{x}_i(k-1) + (1-a_i)[T_{a,i}(k-1) - \hat{m}_i(k-1) T_{g,i}] & \text{if } k \notin T_k \end{cases}, \tag{12}$$

$$\hat{m}_i(k) = \begin{cases} m_i(k) & \text{if } k \in T_k \\ 0 & \text{if } k \notin T_k \text{ and } \hat{x}_i(k) < T_{\text{set},i} - \delta_i/2 \\ 1 & \text{if } k \notin T_k \text{ and } \hat{x}_i(k) > T_{\text{set},i} + \delta_i/2 \\ \hat{m}_i(k-1) & \text{otherwise} \end{cases}. \tag{13}$$

At time k, the estimator uses the observation if it is available. If no measurement is available, then it evolves the estimates according to the dynamics with known parameters β, under the assumption that $\epsilon_i(k) = 0$. Similarly, it supposes that a TCL does not switch states under the local controller, unless the estimate of the thermal state of the TCL leaves the deadband.

These estimates are used to issue control commands. Our controller takes a binning approach, as seen in recent research (Callaway 2009; Mathieu et al. 2013). Each TCL is assigned to a bin based on its thermal state relative to its deadband and whether or not it is in the ON or OFF state.

More formally, let N_{bin} be a parameter of our centralized DLC controller. N_{bin} is an even number denoting the number of bins our controller uses. For the ON states, we assign $N_{\text{bin}}/2$ bins, and for the OFF states, we assign $N_{\text{bin}}/2$ bins. Then, for each $i \in \mathcal{I}$, we define the following functions. First, we define a normalizing function for each TCL $\varphi_i : \mathbb{R}_+ \to \mathbb{R}$ as

$$\varphi_i(x) = [x - (T_{\text{set},i} - \delta_i/2)]/\delta_i. \tag{14}$$

This function normalizes x so, if x is in the deadband, then $\varphi_i(x)$ lies in the interval $[0, 1]$, e.g., $\varphi_i(T_{\text{set},i} - \delta_i/2) = 0$ and $\varphi_i(T_{\text{set},i} + \delta_i/2) = 1$.

Next, define the function $\psi_i : \mathbb{R}_+ \to \{0, 1, 2, \ldots, N_{\text{bin}}/2\}$ as

$$\psi_i(x) = \begin{cases} 1 & \text{if } 0 \le \varphi_i(x) < 1/(N_{\text{bin}}/2) \\ 2 & \text{if } 1/(N_{\text{bin}}/2) \le \varphi_i(x) < 2/(N_{\text{bin}}/2) \\ \vdots & \\ N_{\text{bin}}/2 & \text{if } 1 - 1/(N_{\text{bin}}/2) \le \varphi_i(x) < 1 \\ 0 & \text{otherwise} \end{cases}. \tag{15}$$

This function evenly partitions the interval $[T_{\text{set},i} - \delta_i/2, T_{\text{set},i} + \delta_i/2)$ into $N_{\text{bin}}/2$ bins of length $\delta_i/(N_{\text{bin}}/2)$, and assigns 0 if x lies outside this interval. Bins are indexed by an integer in $\{1, 2, \ldots, N_{\text{bin}}/2\}$ and a state in $\{\text{ON}, \text{OFF}\}$. Thus, if the state estimate of TCL i at time k is $(\hat{x}_i(k), \hat{m}_i(k))$, then it will be assigned to bin $(\psi_i(\hat{x}_i(k)), \hat{m}_i(k))$ at time k. The number of TCLs in bin (n, m) at time k is $\sum_{i \in \mathcal{I}} 1\{\psi_i(x_i(k)) = n \text{ and } m_i(k) = m\}$. The estimated number of TCLs in bin (n, m) at time k is $\sum_{i \in \mathcal{I}} 1\{\psi_i(\hat{x}_i(k)) = n \text{ and } \hat{m}_i(k) = m\}$.

Also, note that a TCL may not fall into any bin; this corresponds to when the TCL's thermal state is out of its deadband. Since we are considering deadband tightening strategies to maintain customer satisfaction, if a TCL is outside its deadband, then we cannot issue control commands to it.

Based on its estimate of how many TCLs are in each bin, the controller issues a command to each bin, stating what fraction of the TCLs in each bin should switch states. Here, for simplicity, we assume that every TCL consumes the same amount of power when on, i.e., $P_i = P$ for all $i \in \mathcal{I}$.

More concretely, the controller switches TCLs at time k based on the mismatch between the estimated power consumed $(\sum_{i \in \mathcal{I}} \hat{m}_i(k))\, P$ at time k and the desired power consumption $P_{\text{des}}(k)$ at time k. For example, suppose that it is time k. The estimated number of TCLs in the ON state is

$\sum_{i\in\mathcal{I}} \hat{m}_i(k)$. If $(\sum_{i\in\mathcal{I}} \hat{m}_i(k))\, P > P_{\text{des}}(k)$, then too many TCLs are on, and our controller will issue a command to switch from ON to OFF to some TCLs. It will try to turn off $\lfloor P_{\text{des}}(k)/P \rfloor - \sum_{i\in\mathcal{I}} \hat{m}_i(k)$ TCLs.

To do so, it will issue a probability to each bin, based on how many TCLs are estimated to be in each bin. Since we prefer to switch TCLs that are likely to switch to an OFF state soon, we start by turning off items in bin (1, ON). If there are more than enough TCLs in bin (1, ON), then we issue a fraction based on how many TCLs we wish to turn off and the estimated number in a bin. If there are not enough, then we command every TCL in the bin to turn off and move on to the next bin (2, ON). The algorithm is described in more detail in Algorithm 1.

ALGORITHM 1: The centralized DLC controller's algorithm at time k for issuing commands to bins to reduce power consumption.

Input: The estimated states of each TCL: $(\hat{x}_i(k), \hat{m}_i(k))$, the desired power signal $P_{\text{des}}(k)$, and the power of individual TCLs P.
Output: Updated mode estimates $\hat{m}'_i(k)$. (Commands for time k are also issued to each bin.)

```
Initialize the number of TCLs to switch, N, and the bin number b.
N = ⌊P_des(k)/P⌋ − Σ_{i∈I} m̂_i(k)
b = 1
while N > 0 and b ≤ N_bin/2 do
    Calculate the estimated number of TCLs in bin (b, ON).
    n = Σ_{i∈I} 1{ψ_i(x̂_i(k)) = b and m̂_i(k) = ON}
    if n ≥ N then
        There are enough TCLs. Switch as many as are needed.
        c = N/n
        N = 0
    else
        There are not enough TCLs. Switch all of them.
        c = 1
        N = N − n
    end
    Issue the calculated command to the bin (b, ON).
    issueCommand(c, (b, ON))
    Update the estimate by having TCL mode estimates switch as necessary.
    for i ∈ I do
        if ψ_i(x̂_i(k)) = b and m̂_i(k) = ON then
            flip ~ Bernoulli(c)
            if flip = 1 then
                m̂'_i(k) = 1 − m̂_i(k)
            else
                m̂'_i(k) = m̂_i(k)
            end
        end
    end
    b = b + 1
end
return m̂'_i(k)
```

An analogous process takes place if $(\sum_{i\in\mathcal{I}} \hat{m}_i(k))\, P < P_{\text{des}}(k)$ and the controller must turn TCLs on. This algorithm would be the same, only the variable b in Algorithm 1 would be initialized with $N_{\text{bin}}/2$ and would decrement across iterations, and ON would be replaced with OFF.

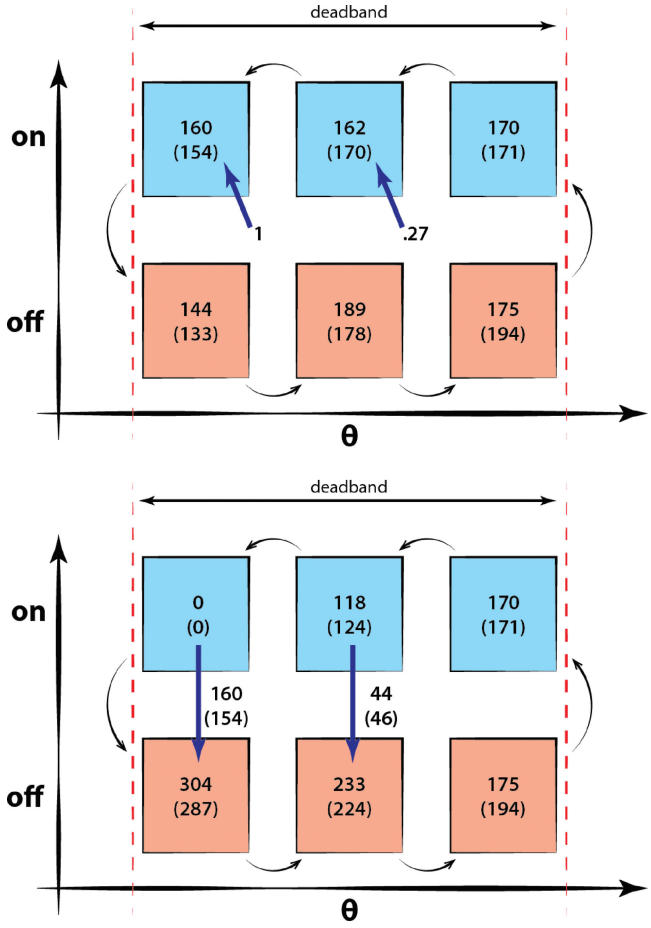

Fig. 2. An example execution of Algorithm 1.

At the level of an individual TCL, the TCL can calculate which bin it is in, based on its true state $(x_i(k), m_i(k))$ and its deadband. When its been receives a command c, it will switch states with probability c. Using a probability allows the centralized controller to issue commands without broadcasting individual TCL identities and without explicit knowledge of which TCLs will switch. Additionally, a TCL can decide whether or not to switch entirely on its own, without coordination or communication with other members of its bin.

An example of this control algorithm is depicted in Figure 2. In the top figure, we see how the TCLs are divided into bins, with $N_{\text{bin}} = 6$. The number in each bin denotes how many TCLs are actually in the bin, the number in parentheses denotes the estimated number of TCLs in the bin. There are an estimated 495 TCLs on, so the estimated total power consumption of the TCLs is 1.2375MW. Suppose, in an extreme case, we wish to decrease power consumption by 500kW. Thus, we would have to turn off 200 TCLs. According to the estimate, if we tell every TCL in the (1, ON) bin (the top-left bin), then 154 TCLs will turn off. Therefore, we must tell 46 TCLs in the bin (2, ON) to turn off as well, where there is estimated to be 170 TCLs. Thus, the control command issued to the bin (1, ON) is 1, to bin (2, ON) is $46/170 = 0.27$, and to all other bins is 0. In the bottom figure, the TCLs actually in each bin switch from the ON state to the OFF state according to a Bernoulli

coin flip, with probability equal to the command issued, and the estimates are updated based on the expected number of TCL switches. The numbers inside the bin represent the actual number of TCLs in each bin after the switching is completed, and the estimated number of TCLs in each bin after the switching is completed.

Closing the loop on this model development, we have the following model of the TCL with the control actuations by the centralized DLC controller. The closed-loop dynamics are given by the following equation:

$$x_i(k+1) = a_i x_i(k) + (1-a_i)[T_{a,i}(k) - \tilde{m}_i(k) T_{g,i}] + \epsilon_i(k). \tag{16}$$

Here, the parameters are the same as in Equation (16). Note that the only difference in these dynamics and the open-loop dynamics without DLC is the modification of the $\tilde{m}_i(k)$ term. Furthermore, the mode of the TCL with DLC, $\tilde{m}_i(k)$ is given by

$$\tilde{m}_i(k) = \begin{cases} 1 - m_i(k) & \text{with probability } c \\ m_i(k) & \text{with probability } 1-c \end{cases}. \tag{17}$$

$\tilde{m}_i(k)$ will depend on the local control law and the centralized DLC law described in Algorithm 1, with preference given to the centralized command. Here, $m_i(k)$ is the local control law as defined in Equation (11).

5.3 DLC Model Simulations

For simulations, we assume each TCL consumes $P_i = 2.5$kW when in the ON state, and we consider a DLC controller in control of 1000 TCLs. Parameters for each TCL i are drawn independently, from distributions based on recent studies of a 250m^2 home (Callaway 2009; Mathieu et al. 2013). The time step h_B was chosen to be $h_B = 1$ minute, and the number of bins $N_{\text{bin}} = 10$.

The ambient temperature $T_a = 32^\circ$C for all TCLs,[6] and the noise process $\epsilon_i(k)$ is independent across k and distributed according to a $N(0, 0.0005)$ distribution[7] for each k.

California Independent System Operator (CAISO) market signals are given in 5 minute intervals (Mathieu et al. 2013; Delparte 2018), so for simulations, the signal P_{des} is independently drawn from a $U(875\text{kW}, 1.35\text{MW})$ distribution[8]. That is, $P_{\text{des}}(k)$ is uniformly drawn for $k \in \{0, 5, 10, \dots\}$. For other values of k, we take the linear interpolation.

Simulations of the aggregate power consumption of all the TCLs is shown in Figure 3 for the uncontrolled case, the case where $h = 1$ minute, and the case where $h = 30$ minutes. Comparing the top plot with the middle and bottom plots, we can see that a DLC policy can reduce the load imbalance even when the controller does not always receive measurements. However, small unforeseen temperature deviations can cause the controller's performance to degrade if enough measurements are not provided, as seen by comparing the middle and bottom plots.

Additionally, the thermal state of one TCL is shown in Figure 4. We can see that the temperature inside the TCL remains inside the deadband, resulting in no loss of comfort to the consumer, in all three cases.

In Figure 5, we plot the error between the actual power consumption and the desired load imbalance compensation signal. First, we randomly drew a P_{des} signal and TCL parameters. Then, for this fixed P_{des} signal and TCL parameters, we ran 500 trials for each sampling period h, and we

[6]For these simulations, we assumed that the ambient temperature is constant across 1 hour, which can be reasonable for this short timeframe.

[7]This is the variance of the noise for one time step, so 0.0005 models the variance of temperature across $h_B = 1$ minute.

[8]This framework can handle other distributions for the load imbalance signal, but a uniform distribution was chosen as a noninformative prior (Keener 2010). The parameters of the distribution were chosen as reasonable values for which energy consumption could be compensated. From simulations, we find that a larger interval is more difficult to track, as expected.

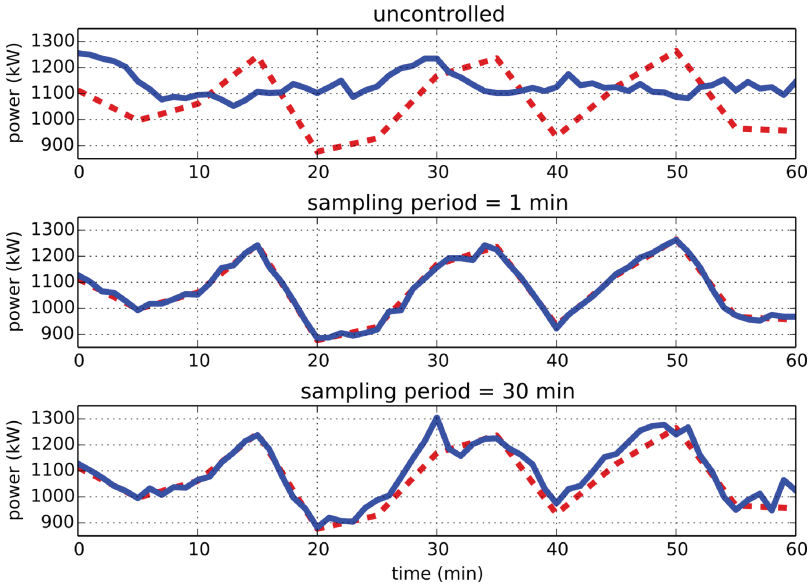

Fig. 3. A sample simulation of the aggregate power consumption of 1000 TCLs. The solid blue line represents the actual power consumption, and the dotted red line represents the desired power consumption. The top figure shows the power consumption in the absence of any control commands, the middle figure shows the power consumption with a sampling period of $h = 1$ minute, and the bottom figure shows the power consumption with a sampling period of $h = 30$ minutes.

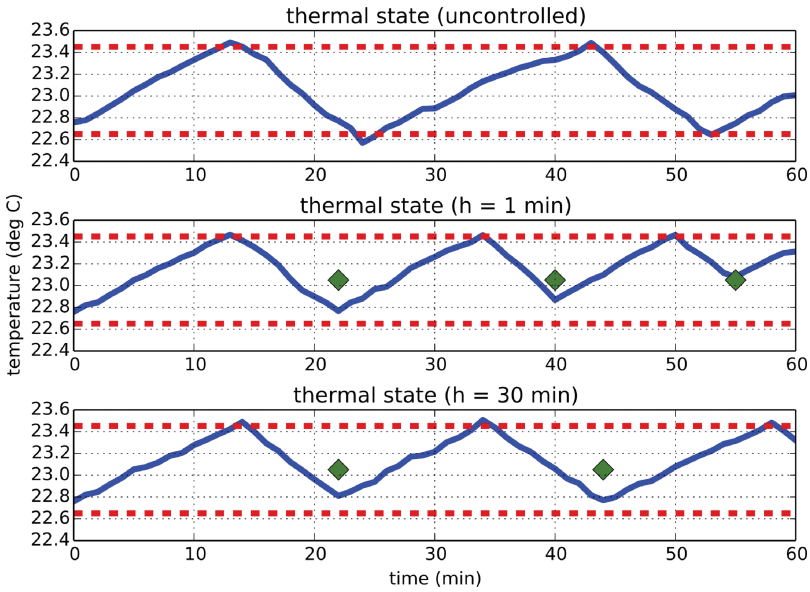

Fig. 4. The thermal state of one sample TCL. The top graph shows the thermal states of the TCL when there is no control. The middle and bottom graphs show the thermal states based on a controller that receives observations every $h = 1$ minute and $h = 30$ minutes, respectively. The dotted red lines indicate the deadband limits. The diamonds indicate when the DLC policy issued control commands to the TCL.

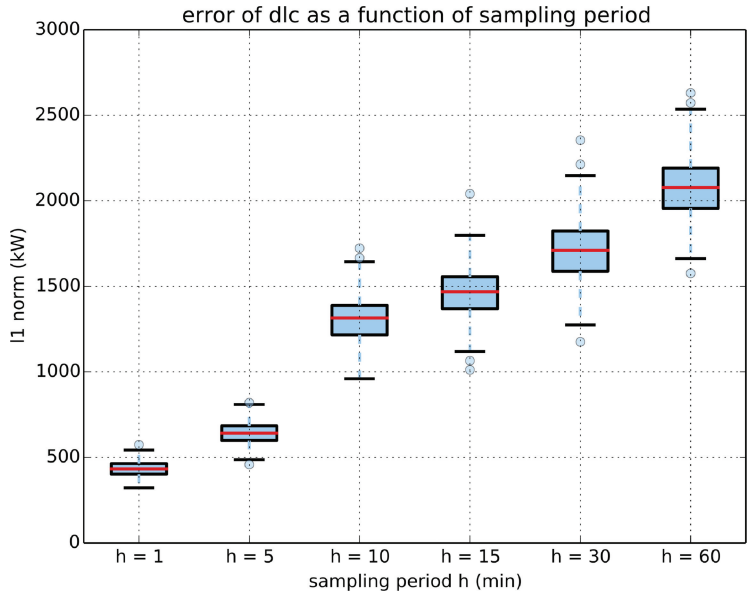

Fig. 5. A plot of how the error between the actual power consumed by the TCLs and the desired power consumption signal empirically varies with the sampling period h. The value we are plotting is $\| \sum_{i \in \mathcal{I}} P_i m_i - P_{\text{des}} \|_1$. The whiskers indicate all data points within 1.5 times the interquartile range. For reference, the error after 500 simulations of uncontrolled TCLs has an empirical mean of 5.39MW with a standard error of 302kW.

consider the empirical distribution of the difference between the actual power consumed by all the TCLs and the desired power signal: $\sum_{i \in \mathcal{I}} P_i m_i - P_{\text{des}}$. We used the ℓ_1 norm on the error signal, so, if we assume a fixed price for spot market electricity purchases/sales throughout the hour interval, then this is directly proportional to the cost the utility company must pay.

5.4 DLC Privacy Analysis

In this example, we suppose households consider their income private. However, their income levels will affect their behaviors at home; in this article, we focus on how their cooking behaviors change. To model this, we use data from the U.S. Energy Information Administration's 2009 Residential Energy Consumption Survey (RECS) (Berry 2009). By observing these different cooking behaviors through a household's energy consumption, an adversary may infer the income of the household.

Formally, let $\Theta = \{\theta_L, \theta_M, \theta_U\}$ denote the private parameter corresponding to lower- (less than \$20,000), middle- (\$20,000 to \$59,999), and upper- (\$60,000 or more) class incomes. Across 113.6 million U.S. homes, 23.7 million households are θ_L, 48.7 million are θ_M, and 41.2 million are θ_U (Berry 2009). This will be our prior, P_θ.

Furthermore, we look at the overall energy consumption of each consumer type. These data are shown in Figure 6. For each type, we fit a log-normal distribution to the overall energy consumption. To estimate the location parameter μ and scale parameter σ, we used the unbiased, minimum variance estimators (Keener 2010, Chapter 4) on the log of the data.[9] We assume the scale parameter is the same for all three private parameters, and we can see that these distributions

[9]Recall that the log-normal distribution, denoted $\ln N(\mu, \sigma)$, is defined by a location parameter μ and scale parameter σ, with density $x \mapsto \frac{1}{x\sigma\sqrt{2\pi}} \exp(\frac{(\ln x - \mu)^2}{2\sigma^2})$ for $x > 0$.

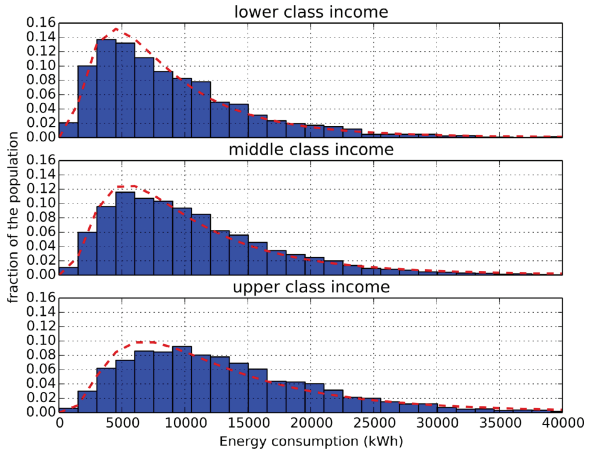

Fig. 6. Histograms of the United States household total annual energy consumptions in each income level in 2009 (Berry 2009), corresponding to private parameters θ_L, θ_M, and θ_U. The data roughly follow a log-normal distribution. The location parameters are $\mu_L = 8.88$, $\mu_M = 9.06$, and $\mu_H = 9.31$, and we assumed all three distributions had the sample scale parameter, $\sigma = 0.49$. To model sampling, we assume that this data is representative of energy consumption on smaller timescales as well.

approximate the data quite well. We can see that a household's income level is correlated with the energy consumption.

Thus in this framework, θ determines the distribution across the observable energy consumption y, i.e., $P_{y|\theta}(\cdot \mid \theta)$ is a log-normal distribution. For example, $P_{y|\theta}(\cdot \mid \theta_L)$ is the $\ln N(\mu_L, \sigma^2)$ distribution. We assume that power consumption on smaller timescales is distributed similarly to this annual data, and these distributions are independent across time.[10] In other words, if a household consumed P kWh in a year, then we assume they consumed roughly $P/365$kWh a day.

With this assumption, we can consider the distribution of energy consumption at different sampling rates. Note that, if we sample at high frequencies, then we receive more measurements than in the low-frequency case, but each measurement is less informative with regards to the consumer's income level.[11]

Since the scale parameters are the same for all three distributions, we can explicitly calculate the MAP using the theory of exponential families (Keener 2010, Chapter 2). Then, using Proposition 4.3, we can calculate the probability an adversary can infer the private parameters, i.e., income level, from the AMI signals. This is represented in Figure 7.

We can see that very high frequency data provide few guarantees of privacy of income level, but this privacy level, α, quickly increases as the sampling period h increases. Furthermore, we can note relationships between time horizons, sampling rates, and privacy. For example, 1 hour of data sampled every 3 minutes is as informative as 6 hours of data sampled every 15 minutes.

Although we focus on a particular example here, this framework can be applied to more detailed models, i.e., more informed adversaries, and other private parameters. For example, we consider

[10]This assumption is likely valid for certain timescales, but will not hold in general. In future work, we hope to analyze the distributions of energy consumption data at different sampling rates.

[11]Here, we scale the data according to the timescale, and, as before, we used the uniform, minimum variance estimators on the log of the data (Keener 2010, Chapter 4). For example, if we receive measurements every minute, then the location parameters for each measurement are $\mu_L = 0.014$, $\mu_M = 0.016$, and $\mu_H = 0.017$, whereas if we receive measurements hourly, then the location parameters are $\mu_L = 0.82$, $\mu_M = 0.99$, and $\mu_H = 1.26$.

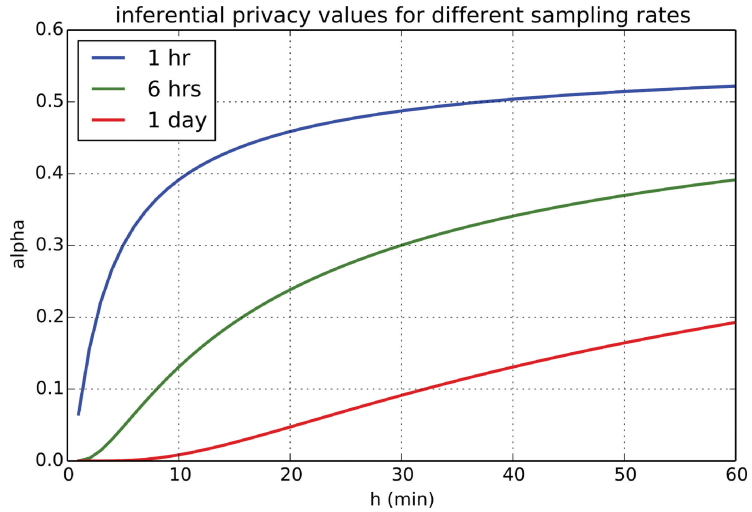

Fig. 7. A plot of the inferential privacy value α as a function of the sampling period h. Each line corresponds to a time horizon in which an adversary can receive samples. Intuitively, we would expect the privacy value α to decrease for longer horizons. Note that our framework accounts for the fact that with longer sampling periods, we receive fewer measurements, but each individual measurement is more informative.

the case where an adversary has knowledge of correlation across time and high-frequency dynamics across time in Dong et al. (2014). We are also currently examining the effects of sampling for longer time horizons (Faisal and Cárdenas 2015).

5.5 Framework Application

In Section 5.1, we outlined a formal model of using DLC of TCLs to correct for load imbalances. In Section 5.2, we defined the control policy of a direct load controller. Importantly, in our model we accounted for how varying sampling rates affect the performance. This required adapted a proposed control law for various sampling rates. In Section 5.3, we quantified the effect of various sampling rates on the operational efficiency of this load imbalance correction program. Additionally, in Section 5.4, we consider how energy consumption data can reveal income levels and quantify this private information leakage using an inferential privacy metric.

Now, we place this in the context of the framework outlined in Section 3.

The set of time indices we care about is $T = \{0, 1, \ldots, N\} \subset \mathbb{N}$. For the utility of data, the state space $\mathcal{X} = (\mathbb{R} \times \{OFF, ON\})^{|\mathcal{I}|}$, where $|\mathcal{I}|$ denotes the number of TCLs participating in the DLC program. At each time step, the controller issues a command between $[0, 1]$ to each bin. The input space is given by $\mathcal{U} = [0, 1]^N_{\text{bin}}$, and the dynamics ϕ are given by Equations (16) and (17). The data sampling policy determining the observables are modeled by $Y(h, t) = (x(k), m(k))_{k \in T_k}$ with $T_k = \{hl : l \in \mathbb{N}, hl \leq k\}$. Note here that the quality parameter is h, the subsampling period. Thus, $Y(h, t)$ takes values in $\cup_{n=1}^{N} \mathbb{R}^n$. This allows us to define the controller u_c as in Algorithm 1. Additionally, the cost is given by $J(x, u) = \sum_{k \in T} \| \sum_{i \in \mathcal{I}} P_i m_i(k) - P_{\text{des}} \|_1$. The privacy metric used is inferential privacy. Thus, the choice of subsampling period h affects the structure of the observation function Y and affects the privacy levels $m(Y, q)$.

Using this framework, we can combine the results in Figures 5 and 7. This is presented in Figure 8. As expected, we can see that lower levels of privacy for consumers result in better load imbalance correction for the direct load controller. This framework allows us to quantify

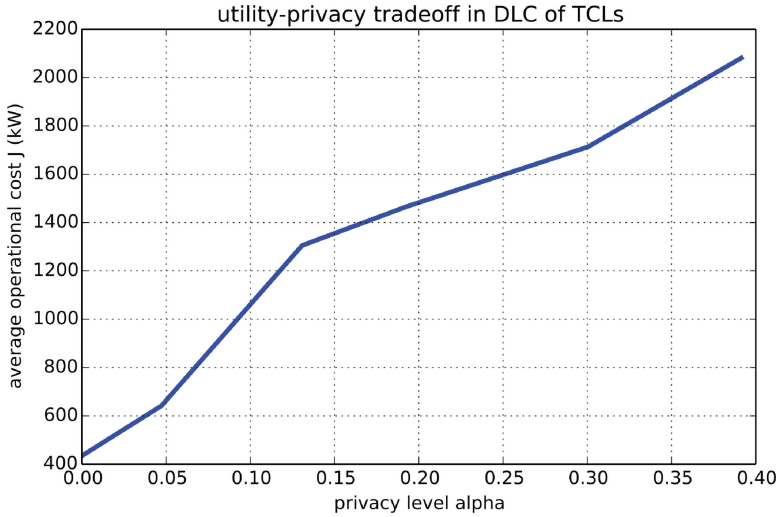

Fig. 8. The utility–privacy tradeoff in a smart grid application. This depicts a direct load controller's ability to compensate for load imbalances, as a function of the inferential privacy levels of consumers. We chose the time-horizon here to be $N = 6$ hours.

this tradeoff: Modeling this tradeoff is the first step towards designing systems that account for privacy and is essential for formulating economic models of data exchange.

We note that the underlying principles of design and analysis throughout this section are quite portable: They can be applied in many IoT domains. What is important is to identify a design parameter that has an effect on the quality of data and a mathematical fashion with which we can quantify the utility of data as well as the privacy of data; our framework provides a means to calculate these quantitative metrics.

6 CONCLUSIONS

In this article, we introduce a framework for quantifying the tradeoff between the utility of data and the privacy loss due to data. Specifically, we consider how variations in the *quality* of data can improve or degrade the operational performance of controllers that utilize these data, and how these variations in quality can change the privacy of users, where privacy is quantified by an appropriately chosen privacy metric.

Additionally, we introduced a new privacy metric, *inferential privacy*, which bounds the probability an adversary can correctly infer private parameters of a consumer based on public observables. The formulation is attack-agnostic and relies on statistical properties of the system under consideration. We are able to leverage the existing literature in statistics and information theory to provide several bounds on the level of inferential privacy, as well.

Finally, we apply our utility–privacy tradeoff framework in a smart grid application. We consider the direct load control of thermostatically controlled loads and analyze how its performance degrades as it receives samples less and less frequently—a privacy preserving metering policy. One of our contributions is a framework for understanding the utility of data in DLC programs, as well as understanding the private information about consumers contained in the data.

As the Internet of Things grows, the potential for privacy breaches is only going to increase. Already, consumers are beginning to become more privacy-aware and sensitive about data transmission policies, and legislative processes are looking into technology-aware policies to handle

privacy issues. Moving forward, these technologies need to evolve with a carefully considered privacy component: The utility of collected data must justify the privacy risks involved.

REFERENCES

Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (2015), 509–514. DOI: http://dx.doi.org/10.1126/science.aaa1465 arXiv:http://science.sciencemag.org/content/347/6221/509.full.pdf.

Gergely Acs and Claude Castelluccia. 2011. I have a DREAM! (DiffeRentially privatE smArt metering). In *Information Hiding*. Lecture Notes in Computer Science, Vol. 6958. Springer, Berlin, 118–132. DOI: http://dx.doi.org/10.1007/978-3-642-24178-9_9

Michael Alexander, Ken Agnew, and Miriam Goldberg. 2008. New approaches to residential direct load control in california. In *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*.

Ross Anderson and Shailendra Fuloria. 2010. On the security economics of electricity metering. In *Proceedings of the9th Workshop on the Economics of Information*.

Chip Berry. 2009. *Residential Energy Consumption Survey*. Technical Report. U.S. Energy Information Administration.

Daniel J. Butler, Justin Huang, Franziska Roesner, and Maya Cakmak. 2015. The privacy-utility tradeoff for remotely teleoperated robots. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI'15)*. ACM, 27–34.

Colin Meehan. 2013. Increasing Demand Response Capabilities in California. Docket No. 13-IEP-1F. California Energy Commission, Sacramento, CA.

David Delparte. 2018. Business Practice Manual for Market Operations version 56. California Independent System Operators, Folsom, CA.

Michael R. Peevey. 2011. Decision Adopting Rules to Protect the Privacy and Security of the Electricity Usage Data of the Customers of Pacific Gas and Electric Company, Southern California Edison Company, and San Diego Gas & Electric Company. California Public Utilities Commission, San Francisco, CA.

D. S. Callaway and I. A. Hiskens. 2011. Achieving controllability of electric loads. *Proc. IEEE* 99, 1 (2011), 184–199. DOI: http://dx.doi.org/10.1109/JPROC.2010.2081652

Duncan S. Callaway. 2009. Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy. *Energ. Conv. Manage.* 50, 5 (2009), 1389–1400. DOI: http://dx.doi.org/10.1016/j.enconman.2008.12.012

Alvaro A. Cárdenas, Saurabh Amin, Galina Schwartz, Roy Dong, and S. Shankar Sastry. 2012. A game theory model for electricity theft detection and privacy-aware control in AMI systems. In *Proceedings of the 50th Allerton Conference on Communication, Control, and Computing*. 1830–1837. DOI: http://dx.doi.org/10.1109/Allerton.2012.6483444

Ann Cavoukian. 2011. Privacy by Design: Strong Privacy Protection – Now, and Well into the Future. A Report on the State of PbD to the 33rd International Conference of Data Protection and Privacy Commissioners. Information & Privacy Commissioner, Ontario, Canada.

Thomas M. Cover and Joy A. Thomas. 1991. *Elements of Information Theory*. Wiley-Interscience.

T. Dalenius. 1977. Towards a methodology for statistical disclosure control. *Statistik Tidskrift* 15 (1977), 429–444.

Department of Energy. 2010. Data Access and Privacy Issues Related to Smart Grid Technologies. Department of Energy, Washington, D.C.

W. Diffie and M. E. Hellman. 1979. Privacy and authentication: An introduction to cryptography. *Proc. IEEE* 67, 3 (Mar. 1979), 397–427. DOI: http://dx.doi.org/10.1109/PROC.1979.11256

R. Dong, W. Krichene, A. M. Bayen, and S. S. Sastry. 2015. Differential privacy of populations in routing games. In *Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC'15)*. 2798–2803. DOI: http://dx.doi.org/10.1109/CDC.2015.7402640

Roy Dong, Lillian Ratliff, Henrik Ohlsson, and S. Shankar Sastry. 2014. Fundamental limits of nonintrusive load monitoring. In *Proceedings of the 3rd International Conference on High Confidence Networked Systems (HiCoNS'14)*. ACM, 11–18. DOI: http://dx.doi.org/10.1145/2566468.2566471

Roy Dong, Lillian J. Ratliff, Henrik Ohlsson, and S. Shankar Sastry. 2013. Energy disaggregation via adaptive filtering. In *Proceedings of the 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton'13)*. 173–180. DOI: http://dx.doi.org/10.1109/Allerton.2013.6736521

F. du Pin Calmon and N. Fawaz. 2012. Privacy against statistical inference. In *Proceedings of the 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton'12)*. 1401–1408. DOI: http://dx.doi.org/10.1109/Allerton.2012.6483382

Cynthia Dwork. 2006. Differential privacy. In *Proceedings of the International Colloquium on Automata, Languages and Programming*. Springer, 1–12.

Cynthia Dwork and Aaron Roth. 2014. *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends in Theoretical Computer Science.

M. Faisal and A. A. Cárdenas. 2015. How the quantity and quality of training data impacts re-identification of smart meter users. In *Proceedings of the IEEE Smart Grid Communications Conference*.

J. Giraldo, A. Cárdenas, E. Mojica-Nava, N. Quijano, and R. Dong. 2014. Delay and sampling independence of a consensus algorithm and its application to smart grid privacy. In *Proceedings of the IEEE 53rd Annual Conference on Decision and Control*. 1389–1394. DOI: http://dx.doi.org/10.1109/CDC.2014.7039596

Bennie G. Thompson. 2011. Electricity Grid Modernization: Progress Being Made on Cybersecurity Guidelines, But Key Challenges Remain to Be Addressed. Report to Congressional Requesters. United States Government Accountability Office, Washington, D.C.

Bernard G. Greenberg, Abdel-Latif A. Abul-Ela, Walt R. Simmons, and Daniel G. Horvitz. 1969. The unrelated question randomized response model: Theoretical framework. *J. Am. Statist. Assoc.* 64, 326 (1969), 520–539. DOI: http://dx.doi.org/10.1080/01621459.1969.10500991 arXiv:http://www.tandfonline.com/doi/pdf/10.1080/01621459.1969.10500991

Shuo Han, Ufuk Topcu, and George J. Pappas. 2014. Differentially private distributed constrained optimization. *arXiv* (2014).

Te Han and S. Verdú. 1994. Generalizing the fano inequality. *IEEE Trans. Inf. Theory* 40, 4 (1994), 1247–1251. DOI: http://dx.doi.org/10.1109/18.335943

Justin Hsu, Zhiyi Huang, Aaron Roth, and Zhiwei Steven Wu. 2014. Jointly private convex programming. *arXiv* (2014).

Zhenqi Huang, Yu Wang, Sayan Mitra, and Geir E. Dullerud. 2014. On the cost of differential privacy in distributed control systems. In *Proceedings of the 3rd International Conference on High Confidence Networked Systems (HiCoNS'14)*. ACM, New York, NY, 105–114. DOI: http://dx.doi.org/10.1145/2566468.2566474

I. A. Ibragimov and R. Z. Has'minskii. 1991. *Statistical Estimation—Asymptotic Theory*. Springer-Verlag New York.

Ruoxi Jia, Roy Dong, S. Shankar Sastry, and Costas Spanos. 2016. Privacy-enhanced architecture for occupancy-based HVAC control (submitted).

Richeng Jin, Xiaofan He, and Huaiyu Dai. 2017. On the tradeoff between privacy and utility in collaborative intrusion detection systems-a game theoretical approach. In *Proceedings of the Hot Topics in Science of Security: Symposium and Bootcamp (HoTSoS'17)*. ACM, New York, NY, 45–51. DOI: http://dx.doi.org/10.1145/3055305.3055311

Robert W. Keener. 2010. *Theoretical Statistics: Topics for a Core Course*. Springer.

Klaus Kursawe, George Danezis, and Markulf Kohlweiss. 2011. Privacy-friendly aggregation for the smart-grid. In *Proceedings of the 11th International Conference on Privacy Enhancing Technologies (PETS'11)*. 175–191.

L. Le Cam. 1973. Convergence of estimates under dimensionality restrictions. *Ann. Stat.* 1, 1 (1973), 38–53.

J. Le Ny and G. J. Pappas. 2014. Differentially private filtering. *IEEE Trans. Autom. Contr.* 59, 2 (2014), 341–354. DOI: http://dx.doi.org/10.1109/TAC.2013.2283096

Fenjun Li, Bo Luo, and Peng Liu. 2010. Secure information aggregation for smart grids using homomorphic encryption. In *Proceedings of the 1st IEEE International Conference on Smart Grid Communications (SmartGridComm'10)*. 327–332. DOI: http://dx.doi.org/10.1109/SMARTGRID.2010.5622064

J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao. 2017. A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. *IEEE IoT J.* 4, 5 (Oct 2017), 1125–1142. DOI: http://dx.doi.org/10.1109/JIOT.2017.2683200

M. A. Lisovich, D. K. Mulligan, and S. B. Wicker. 2010. Inferring personal information from demand-response systems. *IEEE Secur. Priv.* 8, 1 (2010), 11–20. DOI: http://dx.doi.org/10.1109/MSP.2010.40

Ning Lu. 2012. An evaluation of the HVAC load potential for providing load balancing service. *IEEE Trans. Smart Grid* 3, 3 (2012), 1263–1270. DOI: http://dx.doi.org/10.1109/TSG.2012.2183649

Ning Lu and Yu Zhang. 2013. Design considerations of a centralized load controller using thermostatically controlled appliances for continuous regulation reserves. *IEEE Trans. Smart Grid* 4, 2 (2013), 914–921. DOI: http://dx.doi.org/10.1109/TSG.2012.2222944

J. L. Mathieu, S. Koch, and D. S. Callaway. 2013. State estimation and control of electric loads to manage real-time energy imbalance. *IEEE Trans. Power Syst.* 28, 1 (2013), 430–440. DOI: http://dx.doi.org/10.1109/TPWRS.2012.2204074

S. Moura, J. Bendtsen, and V. Ruiz. 2013. Observer design for boundary coupled PDEs: Application to thermostatically controlled loads in smart grids. In *Proceedings of the IEEE 52nd Annual Conference on Decision and Control*. 6286–6291. DOI: http://dx.doi.org/10.1109/CDC.2013.6760883

Arvind Narayanan and Vitaly Shmatikov. 2006. How to break anonymity of the netflix prize dataset. arXiv:cs/0610105. Retrieved from https://arxiv.org/abs/cs/0610105

J. Neyman and E. S. Pearson. 1933. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. Roy. Soc. Lond. A* 231, 1 (1933), 289–337.

Helen Nissenbaum. 2004. Privacy as contextual integrity. *Washington Law Review* 79, 1 (2004), 119–158.

North American Energy Standards Board. 2015. NAESB Privacy Policy. Retrieved April 25, 2018 from https://www.naesb.org/privacy.asp.

Aneesh Chopra, Vivek Kundra, and Phil Weiser. 2011. A Policy Framework for the 21st Century Grid: Enabling Our Secure Energy Future. National Science and Technology Council (NSTC) Subcommittee on Smart Grid, Washington, D.C.

Cristian Perfumo, Ernesto Kofman, Julio H. Braslavsky, and John K. Ward. 2012. Load management: Model-based control of aggregate power for populations of thermostatically controlled loads. *Energ. Conv. Manage.* 55, 1 (2012), 36–48. DOI: http://dx.doi.org/10.1016/j.enconman.2011.10.019

Public Utility Commission of Texas. 2014. Electric Substantive Rules – Chapter 25. (2014).

S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor. 2011. Smart meter privacy: A utility-privacy framework. In *Proceedings of the 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm'11)*. 190–195. DOI: http://dx.doi.org/10.1109/SmartGridComm.2011.6102315

Lillian J. Ratliff, Carlos Barreto, Roy Dong, Henrik Ohlsson, Alvaro Cárdenas, and S. Shankar Sastry. 2015. Effects of risk on privacy contracts for demand-side management. *arXiv:1409.7926v3* (2015).

Alfredo Rial and George Danezis. 2011. Privacy-preserving smart metering. In *Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society (WPES'11)*. ACM, 49–60. DOI: http://dx.doi.org/10.1145/2046556.2046564

Franziska Roesner, James Fogarty, and Tadayoshi Kohno. 2012. User interface toolkit mechanisms for securing interface elements. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST'12)*. ACM, New York, NY, 239–250. DOI: http://dx.doi.org/10.1145/2380116.2380147

N. Ruiz, I. Cobelo, and J. Oyarzabal. 2009. A direct load control model for virtual power plant management. *IEEE Trans. Power Syst.* 24, 2 (2009), 959–966. DOI: http://dx.doi.org/10.1109/TPWRS.2009.2016607

L. Sankar, S. R. Rajagopalan, and H. V. Poor. 2013. Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Trans. Inf. Forens. Secur.* 8, 6 (2013), 838–852. DOI: http://dx.doi.org/10.1109/TIFS.2013.2253320

Richard Shay, Saranga Komanduri, Adam L. Durity, Phillip (Seyoung) Huh, Michelle L. Mazurek, Sean M. Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2016. Designing password policies for strength and usability. *ACM Trans. Inf. Syst. Secur.* 18, 4, Article 13 (May 2016), 34 pages. DOI: http://dx.doi.org/10.1145/2891411

Glenn Smith. 2012. Marijuana bust shines light on utilities. Retrieved April 30, 2018 from https://www.postandcourier.com/news/marijuana-bust-shines-light-on-utilities/article_f63a8bed-9a43-5429-aaef-99f7eb0f71f0.html.

Daniel J. Solove. 2002. Conceptualizing privacy. *Cali. Law Rev.* 90, 4 (2002), 1087.

J. A. Stankovic. 2014. Research directions for the internet of things. *IEEE IoT J.* 1, 1 (Feb. 2014), 3–9. DOI: http://dx.doi.org/10.1109/JIOT.2014.2312291

L. Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 5 (2002), 557–570.

G. Taban and V. D. Gligor. 2009. Privacy-preserving integrity-assured data aggregation in sensor networks. In *Proceedings of the International Conference on Computational Science and Engineering*, Vol. 3. 168–175. DOI: http://dx.doi.org/10.1109/CSE.2009.389

Gary Locke and Patrick D. Gallagher. 2014. NISTIR 7628 – Guidelines for Smart Grid Cyber Security: Vol. 2, Privacy and the Smart Grid. The Smart Grid Interoperability Panel—Cyber Security Working Group, Washington, D.C.

Alexandre B. Tsybakov. 2009. *Introduction to Nonparametric Estimation*. Springer, New York.

Stanley L. Warner. 1965. Randomized response: A survey technique for eliminating evasive answer bias. *J. Am. Stat. Assoc.* 60, 309 (1965), 63–69. DOI: http://dx.doi.org/10.1080/01621459.1965.10480775 arXiv:http://www.tandfonline.com/doi/pdf/10.1080/01621459.1965.10480775 PMID: 12261830.

X. Yang, T. Wang, X. Ren, and W. Yu. 2017. Survey on improving data utility in differentially private sequential data publishing. *IEEE Trans. Big Data* 1, 1 (2017), 1–1. DOI: http://dx.doi.org/10.1109/TBDATA.2017.2715334

Bin Yu. 1997. Assouad, fano, and le cam. In *Festschrift for Lucien Le Cam*. Springer, Berlin, 423–435.

Received July 2016; revised June 2017; accepted January 2018