Published in final edited form as:
IEEE Signal Process Mag. 2013 September 1; 30(5): 86–94. doi:10.1109/MSP.2013.2259911.

Signal Processing and Machine Learning with Differential Privacy:

Algorithms and challenges for continuous data

Anand D. Sarwate [research assistant professor] and
Toyota Technological Institute at Chicago. He received B.S. degrees in electrical engineering and mathematics from the Massachusetts Institute of Technology in 2002 and a Ph.D. degree in electrical engineering from the University of California at Berkeley in 2008. His research is on distributed signal processing, optimization, machine learning, information theory, and statistics. (asarwate@ttic.edu)

Kamalika Chaudhuri [assistant professor]
Department of Computer Science and Engineering, University of California, San Diego. She received a bachelor of technology degree in computer science and engineering from the Indian Institute of Technology, Kanpur, in 2002, and a Ph.D. degree in computer science from the University of California at Berkeley in 2007. Her research focuses on the design and analysis of machine-learning algorithms and their applications. In particular, she is interested in privacy-preserving machine learning, where the goal is to develop machine-learning methods for sensitive data while still preserving the privacy of the individuals in the data set. (kchaudhuri@ucsd.edu)

Private companies, government entities, and institutions such as hospitals routinely gather vast amounts of digitized personal information about the individuals who are their customers, clients, or patients. Much of this information is private or sensitive, and a key technological challenge for the future is how to design systems and processing techniques

for drawing inferences from this large-scale data while maintaining the privacy and security of the data and individual identities. Individuals are often willing to share data, especially for purposes such as public health, but they expect that their identity or the fact of their participation will not be disclosed. In recent years, there have been a number of privacy models and privacy-preserving data analysis algorithms to answer these challenges. In this article, we will describe the progress made on differentially private machine learning and signal processing.

INTRODUCTION

There are many definitions and models for privacy-preserving computation, and a recent survey by Fung et al. compares several different approaches [1]. Many of these models have been shown to be susceptible to composition attacks, in which an adversary observing the output of the algorithm exploits prior knowledge to reidentify individuals [2]. For example, the adversary could use publicly available records such as voting polls [3]. Defining privacy is not simple, and the words privacy, confidentiality, and security have many different meanings across different communities. It has become increasingly clear that there is no real separation between individuals' identity and their data—the pattern of data associated with an individual is itself uniquely identifying.

Differential privacy is a cryptographically motivated definition of privacy [4] that has gained significant attention over the past few years in the machine-learning and data-mining communities. There are a few variant definitions [5]–[7], but for the purposes of this survey, differential privacy measures privacy risk by a parameter ε that bounds the log-likelihood ratio of the output of a (private) algorithm under two databases differing in a single individual's data. When ε is small, the inferences that an adversary can make observing the output of the algorithm will be similar regardless of whether that individual is in the data set or not. There have been other surveys of differential privacy literature; in particular, Dwork and Smith's survey [8] covers much of the earlier theoretical work. The privacy guarantees made in differential privacy are statistical in nature and are different than those based on cryptography [9] or information theory [10].

Initial work on differential privacy was motivated by problems in official statistics such as publishing "sanitized" data tables. A different approach is the interactive query model: a user poses queries to a curator of the database who then provides approximate answers. The approximation is designed to protect the privacy of individual data entries. From these two settings, the literature has spread to cover more complex data processing algorithms such as real-time signal processing [11]–[13], classification [14]–[16], dimensionality reduction [17], [18], and auction design [19].

> DIFFERENTIAL PRIVACY IS A CRYPTOGRAPHICALLY MOTIVATED DEFINITION OF PRIVACY THAT HAS GAINED SIGNIFICANT ATTENTION OVER THE PAST FEW YEARS IN THE MACHINE-LEARNING AND DATA-MINING COMMUNITIES.

In these applications, the key challenge is evaluating the impact of the privacy constraint on the performance or utility of the algorithm. Privacy is in tension with utility; a completely private algorithm releases nothing. However, if the available data set contains many individuals, there is a tradeoff between the privacy guarantee ε, utility, and the number of data points (or sample size) n. This tradeoff will, in general, depend on properties of the data, such as its dimension, range, or sparsity. The choice of how to measure utility differs across application areas. For example, for statistical estimation, we may measure the quality of the estimate by mean-squared error (MSE), whereas for classification, we may measure the expected loss. Calculating the achievable privacy and accuracy levels for a given amount

of data provides a way of comparing different differentially private algorithms for the same task.

While the theory of differential privacy has undergone significant development, there is substantial work left to be done to extend the framework to practical applications. In particular, much of the theory has been developed for data taking discrete values, and there are many challenges raised by continuous data, ranging from the implementation of differentially private algorithms [20] to theoretical foundations [21]. In this tutorial, we will focus on differentially private statistical methods and algorithms that operate on continuous data. We will describe statistical estimators, classification procedures, dimensionality reduction techniques, and signal processing techniques.

The theory for differential privacy using continuous data is different than for discrete data. For example, learning classifiers is easier with discrete data. If the number of possible classifiers, or hypotheses, is finite or the data is discrete, learning the best classifier is possible if the number of data points n grows logarithmically with the size of the hypothesis set or the data domain [22], [23]: for data in $\{0, 1\}^d$, the sample size n must grow linearly with d. On the other hand, when data is allowed to be continuous and the hypothesis class is allowed to be infinite, distribution-free learning is impossible [24]: either we need prior knowledge about the data distribution, or n will depend on the data distribution. Thus there is no uniform upper bound on the sample requirement. This holds even for simple classes such as learning thresholds and linear classifiers: in the absence of a privacy constraint, we can pick an n such that we learn the true hypothesis for any data distribution, but to learn the true hypothesis with differential privacy we must choose n as a function of the data distribution.

Techniques from signal processing have the potential to greatly expand differentially private algorithms for continuous data. Our focus on continuous data means we will not discuss the many active research topics in differential privacy for discrete data—in particular, we will not discuss some of the progress made in software systems engineering for differential privacy [25]–[27], algorithms for computing histograms and contingency tables [28], [29], or the large body of work on privacy-preserving data release (references can be found in recent works [18], [30]).

LEARNING FROM SENSITIVE DATA

There are n records in the database $\mathscr{D} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, where each $\mathbf{x}_i$ is a vector in $\mathbb{R}^d$ and corresponds to the data of an individual i. The d elements of a vector $\mathbf{x}$ correspond to different numerical features. We will assume that the ranges of these features are normalized such that $\|x\| \leq 1$, where $\|\cdot\|$ is the Euclidean norm. Although we are focusing on continuous data in this survey, there is extensive literature on differentially private methods for discrete data.

AN EXAMPLE

Suppose that each record $\mathbf{x}(i)$ represents the numerical readings from d different sensors that are monitoring different quantities (temperature, heart rate) related to the health of a patient. For simplicity, we will assume that each of the measurements has been normalized so that $\mathbf{x}_i \in [0,1]^d$. Given readings from these sensors across a large group of n patients, we can ask many statistical and signal processing questions. What is the average reading across the population of a given feature? How are two of the features correlated with each other? Can we predict one of the features from another? Do the data points lie (approximately) on a k-dimensional subspace with $k < d$? We would like to answer these questions while satisfying a quantifiable notion of privacy.

DEFINING PRIVACY

Differential privacy seeks to provide guarantees about the process of computing functions on sensitive data and has a number of features that make it an attractive approach to quantifying privacy. Privacy is guaranteed by ensuring that the process is randomized with the following promise: an algorithm is differentially private if the participation of any record (corresponding to a single individual) in the database does not alter the probability of any outcome by very much. This definition has many features: it is resistant to attacks to which other privacy models are susceptible [2], it bounds the privacy risk to each individual, and it degrades gracefully as an individual's data is used in multiple computations.

DEFINITION 1

An algorithm $\mathscr{A}_{priv}(\cdot)$ taking values in a set $\mathscr{T}$ provides ε-differential privacy if

$$\mathbb{P}\left(\mathscr{A}_{priv}\left(\mathscr{D}\right) \in \mathscr{S}\right) \leq e^{\epsilon} \cdot \mathbb{P}\left(\mathscr{A}_{priv}\left(\mathscr{D}'\right) \in \mathscr{S}\right) \quad (1)$$

for all measurable $\mathscr{S} \subseteq \mathscr{T}$ and all data sets $\mathscr{D}$ and $\mathscr{D}'$ differing in a single entry. It provides (ε, δ)-differential privacy if

$$\mathbb{P}\left(\mathscr{A}_{priv}\left(\mathscr{D}\right) \in \mathscr{S}\right) \leq e^{\epsilon}\mathbb{P}\left(\mathscr{A}_{priv}\left(\mathscr{D}'\right) \in \mathscr{S}\right) + \delta \quad (2)$$

for all $\mathscr{S} \subseteq \mathscr{T}$ and all data sets $\mathscr{D}$ and $\mathscr{D}'$ differing in a single entry.

Here we assume that each entry in the database $\mathscr{D}$ corresponds to a single individual. Privacy parameters are ε and δ, and low ε and δ ensure more privacy [4], [21]. The second privacy guarantee [31] is weaker, and reduces to the first one when $\delta = 0$. Variants of (ε, δ)-differential privacy such as (1, ε, δ)-indistinguishability [7] and δ-probabilistic privacy [32] have also been considered in the literature; we focus on the most popular variant for our purpose.

There are two important features of differentially private algorithms. First, if $\boldsymbol{v}$ is the output of an ε-differentially private algorithm $\mathscr{A}_{priv}$, then any function $\boldsymbol{g}(\boldsymbol{v})$ of the output also guarantees ε-differential privacy. That is, postprocessing of the output does not change the privacy guarantee, as long as that postprocessing does not use the original data. The second key feature is how the privacy guarantees are affected by multiple computations on the data. If we run algorithms $\mathscr{A}_{priv}^{(1)}$ and $\mathscr{A}_{priv}^{(2)}$ on the data with privacy guarantees ε_1 and e_2, then the pair ($\mathscr{A}_{priv}^{(1)}$, $\mathscr{A}_{priv}^{(2)}$) guarantees differential privacy with privacy risk at most $\varepsilon_1 + \varepsilon_2$. Somewhat better guarantees may be obtained if we are allowed (ε, δ)-differential privacy [33].

GENERIC METHODS FOR DIFFERENTIAL PRIVACY

For a given algorithm or function $\mathscr{A}_{nonpriv}$, there are many general methods for generating an approximation $\mathscr{A}_{priv}$ of the algorithm that satisfies one of these privacy definitions. These approaches are illustrated in Figure 1. The methods introduce the privacy-preserving randomness in different ways, but most involve adding noise during some step of the original algorithm $\mathscr{A}_{nonpriv}$. We describe below four key approaches for obtaining differential privacy.

INPUT PERTURBATION

Suppose we would like to provide the data from our body-network sensors to a third party. The easiest method for guaranteeing differential privacy is to add noise to the data itself. If **x** is a real *d*-dimensional vector, then a differentially private version of **x** is

$$\hat{\mathbf{x}}=\mathbf{x}+\boldsymbol{Z}, \quad (3)$$

where **Z** is a random *d*-dimensional vector with density

$$pz\,(\mathbf{z}) \propto \quad exp\left(-\frac{\epsilon}{2}\|\mathbf{z}\|\right). \quad (4)$$

By adding this noise to each individual data vector $\mathbf{x}i$ in $\mathscr{D}$, we can guarantee that the resulting database $D \hat{=} (\hat{\mathbf{x}_1}, \hat{\mathbf{x}_2}, \ldots, \hat{\mathbf{x}_n})$ is an ε-differentially private approximation to $\mathscr{D}$. In the scalar case this corresponds to adding noise with a Laplace distribution. This is not the only distribution that can guarantee differential privacy—in particular, for a given utility on the output the noise distribution that maximizes utility while providing differential privacy may have a different shape.

OUTPUT PERTURBATION

Suppose now that we wish to calculate the average of each of the sensor readings across the population. In this situation, our desired algorithm $\mathscr{A}_{nonpriv}$ simply computes a function $f\,(\mathscr{D})$ of the data, and we can obtain differential privacy by adding noise to $f\,(\mathscr{D})$. The amount of noise we need to add depends on the sensitivity of the function *f* to changes in its input. The global sensitivity is the maximum difference of the function over all pairs of databases $\mathscr{D}$ and $\mathscr{D}'$ differing in a sin gle individual

$$S\,(f) = \max_{\mathscr{D}\sim\mathscr{D}'} \|f\,(\mathscr{D}) - f\left(\mathscr{D}'\right)\|, \quad (5)$$

where $\|\cdot\|$ is the Euclidean norm. We can then compute an ε-differentially private approximation of *f*:

$$\hat{f}\,(\mathscr{D}) = f\,(\mathscr{D}) + \boldsymbol{Z}, \quad (6)$$

where **Z** is a random *d*-dimensional vector with density

$$pz\,(\mathbf{z}) \propto \; exp\left(-\frac{\epsilon}{S\,(f)}\|\mathbf{z}\|\right). \quad (7)$$

For example, to compute the average vector $f\,(\mathscr{D}) = (1/2)\sum_{i=1}^{n}\mathbf{x}_i$, the sensitivity S(f)=2/n. This is the (global) sensitivity method [4], and there are many variants to handle other more relaxed notions of sensitivity. For example, the smoothed sensitivity method [34] tries to approximate a function *f* which has large *S*(*f*) only in the "worst case" by adding noise as a function of a "smoothed" version of the sensitivity at the given $\mathscr{D}$.

EXPONENTIAL MECHANISM

Suppose we would like to publish a predictor of a patient's heart rate after an activity using *k* readings of the heart rate during the activity. Given a set of linear predictors $\{P_k\}$, which are publicly known, we would select one of them in a differentially private way. We can

measure the quality of a linear predictor P_k of order k by the MSE $M(P)$ of the predictions. Using these measurements, we can determine k^*, the k that maximizes $M(P_k^*)$. In this setting, adding noise to the optimal k may not make sense, but the exponential mechanism [35] gives a way of choosing an output biased toward having higher utility. Let $q(\mathscr{D}, k) = -M(P_k^*)$ measure the utility of the order-k predictor and define its sensitivity as

$$S(q) = \max_{k, \mathscr{D} \sim \mathscr{D}'} |q(\mathscr{D}, k) - q\left(\mathscr{D}', k\right)|. \quad (8)$$

This is the maximum change in the quality for any output k and any database $\mathscr{D}$. The exponential mechanism picks a random value of k with distribution

$$p(k) \propto exp\left(-\frac{\epsilon}{2S(q)} q(\mathscr{D}, k)\right). \quad (9)$$

FOR DATA THAT LIES IN A BOUNDED DOMAIN, MANY BASIC STATISTICS CAN BE EASILY COMPUTED WITH DIFFERENTIAL PRIVACY AND RELATIVELY HIGH ACCURACY.

This approach, due to McSherry and Talwar [35], is very general and is not restricted to selecting from discrete sets; it can be used whenever a natural performance measure $q(\mathscr{D}, \cdot)$ exists for the algorithm $\mathscr{A}_{nonpriv}$. In many cases, sampling from the distribution in (9) is easy, but for some $q(\mathscr{D}, \cdot)$ we do not know how to sample from the corresponding distribution in polynomial time.

OBJECTIVE PERTURBATION

Suppose in our example that some of the patients we are monitoring had heart attacks. We would like to classify future patients into high or low risk for heart attacks using the same monitoring data. We can learn such a classifier using regularized convex optimization. Chaudhuri et al. [14] introduced an approach that adds noise to the objective function of the optimization to obtain a differentially private approximation. That is, given an algorithm $\mathscr{A}_{nonpriv}$, which computes an output $\mathbf{f}$ via a minimization of a (strongly) convex function $J(\mathrm{g}, \mathscr{D})$, we can get a differentially private algorithm $\mathscr{A}_{priv}$ by adding noise prior to minimization

$$\mathrm{f} = \underset{\mathrm{g}}{argmin}\left(J(\mathbf{u}, \mathscr{D}) + \mathrm{g}^T \boldsymbol{Z}\right), \quad (10)$$

where the distribution of $\mathbf{Z}$ has the same shape as (4) in the previous examples, but the coefficient in the exponent must be chosen as a function of the sensitivity of the optimization [14].

If we use Gaussian noise for input, output, and objective perturbation, we can obtain algorithms that will guarantee (ε, δ)-differential privacy—the parameters of the Gaussian noise will depend on ε, δ, and the specific target function $\mathscr{A}_{nonpriv}$. In general, the sensitivity parameters depend on the $\mathscr{A}_{nonpriv}$ that we want to approximate but not on the actual data $\mathscr{D}$ that is given. The sample-and-aggregate framework [34] tries to relax this condition by approximating the function value on subsets of the actual data; this may result in less noise for many data sets. More recent work has focused on how to exploit properties of the data (for example, incoherence [36], [37]) to develop algorithms that add less noise and have better performance. Notable among these methods is the propose-test-release framework

[38], which uses a differentially private test on the data to check if a property holds and then picks an algorithm tuned to exploit this property.

DIFFERENTIAL PRIVACY IN STATISTICS

One of the most basic tasks in sensitive data analysis is the computation of basic descriptive statistics, such as means, variances, and other parameters of the data distribution. In our patient-monitoring example, we may wish to know the average resting heart rate of patients or how heart rate correlates with activity level. Publishing the exact value does not preserve differential privacy. For example, two data sets $\mathscr{D}$ and $\mathscr{D}^{'}$ differing in a single entry will have different means, so the inequality (1) will not hold when $\mathscr{S}$ contains $\mathscr{A}_{priv}\left(\mathscr{D}\right)$ but not $\mathscr{A}_{priv}\left(\mathscr{D}^{'}\right)$. To prevent such privacy violation, we can compute these statistics in a differentially private way. We can often use standard methods such as those in Figure 1 to guarantee differential privacy. For data that lies in a bounded domain, many basic statistics can be easily computed with differential privacy and relatively high accuracy. When each individual's data is a scalar $x_i \in [0, 1]$ and this interval is known in advance, many statistical estimates can be made private and consistent [39]. Starting from the first works on differential privacy, estimators have been proposed for statistics such as the mean [4], median [34], covariance matrices [40], [41], and a wide range of nonparametric problems [21], including density estimation [42].

EXAMPLE 1: SAMPLE MEAN

Suppose we wanted to compute the average heart rate across the patient population. For bounded data, the global sensitivity method of [4] gives us a very simple differentially private approximation to sample mean. If $(x_1, \ldots, x_n)$ is the input data set, then the estimate is

$$A\left(x\right) = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{\epsilon n} Z,$$

where ε is the privacy parameter and Z is random noise drawn from a Laplace distribution with unit variance. If n and ε are large, this provides a fairly accurate additive approximation to the sample mean. Figure 2(a) shows a histogram of outputs of this procedure for a data set of size $n = 1{,}000$ and for $\varepsilon = 0.1$. The same technique can be used to develop differentially private approximations to variance and higher moments, that is, to all linear statistical functionals.

EXAMPLE 2: SAMPLE MEDIAN

Suppose instead that we want to compute the median heart rate. The global sensitivity approach, however, does not apply to the sample median because the global sensitivity of sample median is high: in a data set with m zeros and $m + 1$ ones, switching a single element can move the sample median from one to zero. Here we can use the exponential mechanism to compute a differentially private approximation to the sample median for data drawn from a bounded domain. For any $y \in [0, 1]$, let $F_n(y)$ be the empirical cumulative distribution function of the input data $(x_1, \ldots, x_n)$. That is, $F_n(y)$ is the fraction of data points x_i for which $x_i \leq y$. By choosing the quality function $q\left(\mathscr{D}, y\right) = \left| \left(1/2\right) - F_n\left(y\right) \right|$, we have $S(q) = 1/n$. This quality function is maximized at the true median, and the variance of a sample drawn from the exponential mechanism decreases with n. Sampling an estimate from the distribution in (9) guarantees ε-differential privacy. Figure 2(b) illustrates the distribution of outputs for this procedure for a data set of size $n = 1{,}000$ and $\varepsilon = 0.1$. A different algorithm

for computing a differentially private approximation to the sample median that adds noise proportional to the smoothed sensitivity was provided by Nissim et al. [34].

CONNECTION TO ROBUST STATISTICS

The success of individual statistical estimators raises the question of whether we can find properties that make a statistical estimator easier to approximate under differential privacy. It turns out that a key property is robustness. Robust statistics is a subfield of statistics that studies the effect of contaminations and changes in the data on the performance of estimators. Robust estimators are insensitive to changes in the data. For example, for data drawn from an unbounded domain, the sample mean is not robust because a single outlier can arbitrarily perturb the mean. On the other hand, the median is robust for distributions where the density at the median is positive. There are several measures of robustness, and an extensive literature on robust statistical estimation [43].

Dwork and Lei [38] identified a connection between robust statistics and differential privacy, and introduced differentially private approximations to several robust statistical estimators, including trimmed mean, interquartile range and regression. This connection was made concrete by Chaudhuri and Hsu [44], who showed that the gross error sensitivity (GES), a measure of robustness, dictates the finite sample convergence rate of a differentially private approximation to any estimator T on a distribution F.

Given an estimator T and a distribution F, the influence function of T at F along x at scale ρ is defined as

$$IF_{\rho}(T, F, x) = \frac{T((1-\rho)F + \rho\delta_x) - T(F)}{\rho},$$

where δ_x is a point mass at x. The influence function can be intuitively thought of as a directional derivative of T at F along the point mass at x at a step size of ρ. The GES of T at F at scale ρ is defined to be $\mathrm{GES}_{\rho}(T, F) = \sup_x |\mathrm{IF}_{\rho}(T, F, x)|$; thus the GES is the absolute value of the maximum directional derivative. Chaudhuri and Hsu [44] prove two results. First, they give a differentially private approximation to the plug-in estimator $T(F_n)$ when T has a bounded range—the additional error due to privacy grows as $O(\mathrm{GES}_{\rho}(T, F)/\varepsilon n)$. Second, they show that the convergence rate of any differentially private approximation to $T(F)$ has to grow as $\Omega(\mathrm{GES}_{\rho}(T, F)/\varepsilon n)$ either for F or for some F' in a small neighborhood around F. In both cases, the scale parameter ρ is $O(1/\varepsilon n)$. These results show that GES characterizes how amenable an estimator is to differentially private approximation.

Lei [45] provided differentially private approximations of M-estimators, a class of robust estimators, by quantizing the data and then building an estimator on a perturbed histogram. Suppose, in our example, that all of the features have been normalized to lie in [0, 1] so the data lie in $[0, 1]^d$. The algorithm chooses a parameter h_n, partitions the space into cubes of side-length h_n, computes an estimate of the data density by counting the fraction of points lying in each cube, and adds Laplace noise to these counts to guarantee differential privacy. Computing an M-estimator using this density estimate preserves differential privacy. Lei shows h_n appropriately the error of the estimator can be driven to 0 as $n \to \infty$.

SIGNAL PROCESSING AND MACHINE LEARNING WITH PRIVACY

There is a growing body of research on privacy-preserving algorithms for machine-learning and signal processing tasks. For example, there are algorithms for privacy-preserving classification [14], [15], [46], [47], regression [16], [45], principal components analysis

(PCA) [17], [37], [40], [48], boosting [33], and online learning [49]. A different framework was proposed by Duchi et al. [50], who analyze statistical risk minimization via a noisy (privacy-preserving) gradient descent procedure. There has been much work on the theory of learning with differential privacy; in this section we instead focus on recent applied work and open practical challenges in differentially private machine learning.

CLASSIFICATION AND REGRESSION

In our example, suppose that we would like to learn a rule for classifying patients into high- or low-risk categories for a heart attack. Classification is a simple and fundamental machine-learning task and, for discrete data, researchers have developed algorithms to compute differentially private decision trees [51]–[53]. For continuous data, the most common approach to classification is empirical risk minimization (ERM). For example, for logistic regression, a regularized ERM procedure takes labeled data $\{(\mathbf{x}_i, y_i): i = 1, 2, \ldots, n\}$ with features $\mathbf{x}_i \in \mathbb{R}^d$ and labels $y_i \in \{-1, +1\}$ and finds vector $\mathbf{f}$ such that new points can be labeled by $\text{sgn}(\mathbf{f}^T\mathbf{x})$. This is done by solving the following minimization:

$$\mathrm{f}=\underset{\mathrm{g}\in\ ^d}{argmin}\left(\frac{1}{n}\sum_{i=1}^{n} log\left(1+e^{yi\cdot \mathrm{g}^T\mathbf{x}i}\right)+\frac{\Lambda}{2}\|\mathrm{g}\|^2\right), \quad (11)$$

where $\|\mathbf{g}\|^2$ is a regularizer to prevent overfitting and Λ is a tradeoff parameter. There have been several approaches to differentially private classification. Output perturbation computes the ERM solution in (11) and adds noise. Objective perturbation [14] solves a modified version of the program

$$\mathrm{f}_{priv}=\underset{\mathrm{g}\in\ ^d}{argmin}\left(\frac{1}{n}\sum_{i=1}^{n} log\left(1+e^{yi\cdot \mathrm{g}^T\mathbf{x}i}\right)+\frac{\Lambda}{2}\|\mathrm{g}\|^2+\boldsymbol{Z}^T\mathrm{g}\right). \quad (12)$$

The noise $\mathbf{Z}$ guarantees differential privacy. To measure utility for classification we can calculate the expected loss of the differentially private classifier. The theoretical guarantee on the loss for objective perturbation is lower than that for output perturbation, which adds noise to $\mathbf{f}$ in (11). Objective perturbation also has an empirical performance closer to the non-private classifier $\mathbf{f}$ in (11). Follow-up work has expanded the class of functions for which the classifier works [46], and the initial empirical evidence is promising [54], [55]. Another method for that is based on perturbing the objective function, the functional mechanism, was recently proposed by Zhang et al. [16]. They claim, incorrectly, that Chaudhuri et al. [14] solve a nonstandard form of logistic regression; however, their method, based on adding noise to a Taylor-series approximation of (11), can also achieve lower classification error than output perturbation. In general, differentially private approximations (both output and objective perturbation) to the optimization in (11) guarantee differential privacy for the exact minimizer. The effect of approximate computation from numerical methods on the privacy guarantee is an open question.

DIMENSIONALITY REDUCTION

Another fundamental building block of machine-learning and signal processing systems is dimensionality reduction. Data may be presented in high dimension, but the underlying phenomenon may be fundamentally low dimensional. The simplest example of this is when the data all lie on or close to a low-dimensional subspace of the original space. In this setting, the singular value decomposition (SVD) of the data covariance matrix computes this low-dimensional subspace—this is also known as the PCA algorithm. Given a set of n vectors $\mathscr{D} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, where each $\mathbf{x}_i \in \mathbb{R}^d$ corresponds to the private data of one

individual, let $X = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^T$ be the matrix whose rows are the data vectors $\{\mathbf{x}_i\}$, and $A = (1/n)X^TX$ denote the $d \times d$ second moment matrix of the data. The SVD gives $A = V^T \Lambda V$, where Λ is a $d \times d$ diagonal matrix with diagonal elements $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_d(A) \geq 0$ and V is orthonormal. The top-k subspace of A is the first k rows of V, which we denote by $V_k(A)$.

There have been several proposed approaches to approximating the top-k PCA subspace while preserving differential privacy. The sublinear queries (SULQ) method [40] adds noise to the matrix A and then computes the SVD of the noisy matrix. Chaudhuri et al. [17] propose using the exponential mechanism [35] to sample a random k-dimensional subspace that approximates the top-k PCA subspace. This corresponds to sampling from the matrix Bingham distribution, which has the density

$$f(U) \propto exp\left(n\frac{\epsilon}{2}tr\left(UAU^T\right)\right), \quad (13)$$

where U is a $k \times d$ matrix whose rows are orthonormal. This distribution has maximal density at $U = V_k(A)$, and samples a random subspace which is close to the true subspace [17], [48].

A major difficulty is sampling from the Bingham distribution. Because differential privacy is a property of the output distribution, the privacy guarantees are contingent on accurately sampling from the distribution. Kapralov and Talwar [48] propose an intricate procedure for drawing samples according to (13) when $k = 1$, but the running time can become prohibitive in the data dimension. Chaudhuri et al. propose using a Gibbs sampler [56], which is simple to implement; unfortunately, there is no rigorous analysis of the convergence time of the sampler. Developing a practical and exact sampler for this distribution is an open question.

> ONE OF THE GOALS OF THIS ARTICLE IS TO INSPIRE ENGINEERS TO TAKE SOME OF THE IDEAS FROM DIFFERENTIAL PRIVACY AND APPLY IT TO THEIR SIGNAL PROCESSING PROBLEMS.

TIME SERIES AND FILTERING

One of the goals of this article is to inspire engineers to take some of the ideas from differential privacy and apply it to their signal processing problems. There has been some recent work connecting problems in signal processing and information theory to issues in differential privacy. Rastogi and Nath [57] proposed a method for dealing with queries on data sets where each individual's data is a time-series data, such as body weight. Their approach performs differentially private perturbation of a query sequence in the Fourier domain and uses homomorphic encryption to enable distributed noise addition. Fan and Xiong [13] look at how to publish a differentially private version of a single time series by learning a linear predictor and using Kalman filtering. To control the amount of privacy lost, they adaptively choose whether to release the output of the differentially private predictor or add Laplace noise to the true sample. This approach improves over the discrete Fourier transform approach [57] in many cases.

Le Ny and Pappas [11], [12] recently studied differential privacy in a signal processing framework. They studied the difference between input and output perturbation in the context of aggregating signals and using Kalman filter estimation and show that in some cases noise addition at the input is better due to the benefits of filtering. This stands in contrast to many machine-learning examples in which noise addition at the input may incur too much perturbation for learning to be possible.

PRACTICAL ISSUES AND LIMITATIONS

The literature on differentially private algorithms is growing rapidly, but there are many open questions that remain. While many of the theoretical results imply that estimating statistics or learning while preserving differential privacy is possible [22], [39], some of these results depend on technical assumptions [24], [58], such as discrete data, finite hypothesis sets, or bounded range, which may not hold in all settings. Understanding the fundamental limits for continuous data may shed some light on which signal processing tasks are possible under differential privacy.

A more immediate issue is how to choose ε and and δ in the first place. It is clear that smaller ε and δ guarantee more privacy [4], and while there are heuristics [8] for choosing ε, interpreting the privacy risk for practitioners is challenging. Because a single data set may be used in multiple computations, the composition rule for privacy implies that we should choose a total ε for all computations on the data and "budget" privacy for each computation. There is little consensus on how to choose δ for (ε, δ) differential privacy: experiments often use small but constant δ but Ganta et al. [2] suggest δ much less than $1/n^2$ is more appropriate.

For a given privacy level ε, we need a larger sample size n to achieve the same level of utility or approximation error. For smaller sample sizes, the randomization for differential privacy can some times be prohibitive [29]. In such settings it may not be possible to provide a meaningful level of differential privacy. In some applications, such as medical data mining, the amount of data n is fixed, and the question becomes one of finding the lowest ε such that the sacrifice in utility is acceptable.

The privacy definitions rely on an idealized model of computation. Recent work has shown that standard implementations of floating point arithmetic may be problematic from a privacy perspective [20]. Since every computation has to be made differentially private, more complex systems such as PINQ [25], AIRAVAT [26], and GUPT [27] may only work with a large value of ε. Even so, there are privacy risks arising from how these systems are implemented, in particular, the time it takes to respond to a query can disclosed information [59].

FUTURE CHALLENGES

Ideas from differential privacy are already beginning to influence some systems, but many theoretical and practical challenges remain. Some core topics in signal processing are being explored now, and the rich body of expertise in the signal processing community can help spur the development of new privacy-preserving data processing algorithms and systems. The literature on differential privacy is growing rapidly, and we were only able to touch on a few topics here. We hope that interested readers will investigate the wide range of topics that have been studied through the lens of differential privacy.

From a signal processing perspective, there are several directions that should be explored in future research. First, in many signal processing applications, signal acquisition is part of the design; an open question is how to best integrate privacy considerations while measuring the signal. For example, how should we represent the signal if it is later going to be used in a differentially private system? Can we design signal acquisition methods which themselves guarantee privacy?

Second, the signals associated with an individual may be more complex than the d-dimensional vectors we considered in this survey. Although some work has been done with unidimensional time series, there are many interesting open questions for prediction and

forecasting methods, transforms, and other core signal processing tasks. Image processing is another important topic that received little attention in the existing privacy literature. Images are very high-dimensional signals, and the data requirements of many differentially private machine-learning methods scale poorly with the data dimension. However, images are also very structured signals, and this structure could potentially be used to develop algorithms with better theoretical guarantees and practical performance.

Networked information systems are another emerging application for differential privacy. Large-scale data mining often involves parties who wish to collaborate but do not wish to divulge their data. While there have been cryptographic approaches to this problem, differentially private distributed algorithms are still in their infancy [60], [61]. Social networks and other distributed collection and measurement systems also provide a rich source of applications for privacy-preserving algorithms.

In this article, we were only able to give an introduction to the extensive literature on differential privacy. Differentially private algorithms for continuous data are the most relevant for signal processing. Privacy impacts time series and real-time processing differently than offline algorithms such as parameter estimation. Through application of domain-specific metrics and signal assumptions, we believe that it will be possible to achieve meaningful privacy-utility tradeoff curves for many signal processing applications. However, more work is needed to explore the potential of differential privacy and related ideas in signal processing systems; we hope that this article will help motivate that work.

Acknowledgments

The work of the authors was supported in part by the National Institutes of Health under award U54-HL108460.

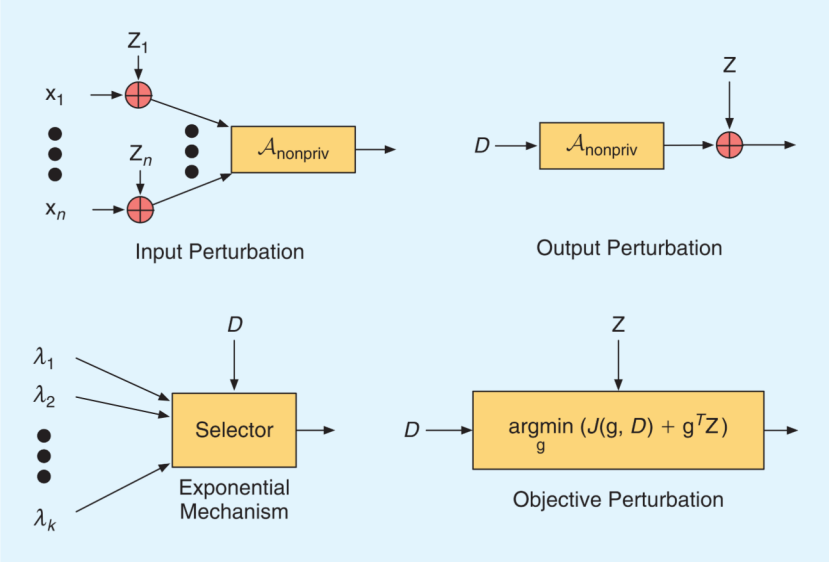
REFERENCES

1. Fung BCM, Wang K, Chen R, Yu PS. Privacy-preserving data publishing: A survey of recent developments. ACM Comput. Surv. Jun; 2010 42(4):14, 1–14, 53. [Online] Available: http://dx.doi.org/10.1145/1749603.1749605.
2. Ganta, SR.; Kasiviswanathan, SP.; Smith, A. Composition attacks and auxiliary information in data privacy.. presented at the 14th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '08); [Online]. Available: http://dx.doi.org/10.1145/1401890.1401926
3. Sweeney L. k-Anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzz. Knowl.-Based Syst. Oct; 2002 10(5):557–570. [Online] Available: http://dx.doi.org/10.1142/S0218488502001648.
4. Dwork C, McSherry F, Nissim K, Smith A. Theory of Cryptography (Lecture Notes in Computer Science Series. Mar 4–7.2006 3876 [Online]. Available: http://dx.doi.org/10.1007/11681878_14.
5. Rastogi, V.; Hay, M.; Miklau, G.; Suciu, D. Relationship privacy: Output perturbation for queries with joins.. presented at 28th ACM SIGMOD-SIGACTSIGART Symp. Principles Database Systems (PODS '09); [Online]. Available: http://dx.doi.org/10.1145/1559795.1559812
6. Kifer, D.; Machanavajjhala, A. No free lunch in data privacy.. presented at 2011 ACM SIGMOD Int. Conf. Management Data; [Online]. Available: http://dx.doi.org/10.1145/1989323.1989345
7. Chaudhuri, K.; Mishra, N. Lecture Notes in Computer Science Series, vol. 4117. Advances in Cryptology— CRYPTO 2006. Aug. 2006 [Online]. Available: http://dx.doi.org/10.1007/11818175_12
8. Dwork C, Smith A. Differential privacy for statistics: What we know and what we want to learn. J. Privacy Confident. 2009; 1(2):135–154. [Online] [Online]. Available: http://repository.cmu.edu/jpc/vol1/iss2/2.
9. Vaidya J, Clifton CW, Zhu YM. Privacy Preserving Data Mining (Advances in Information Security Series. 2006; 19 [Online]. Available: http://dx.doi.org/10.1007/978-0-387-29489-6.

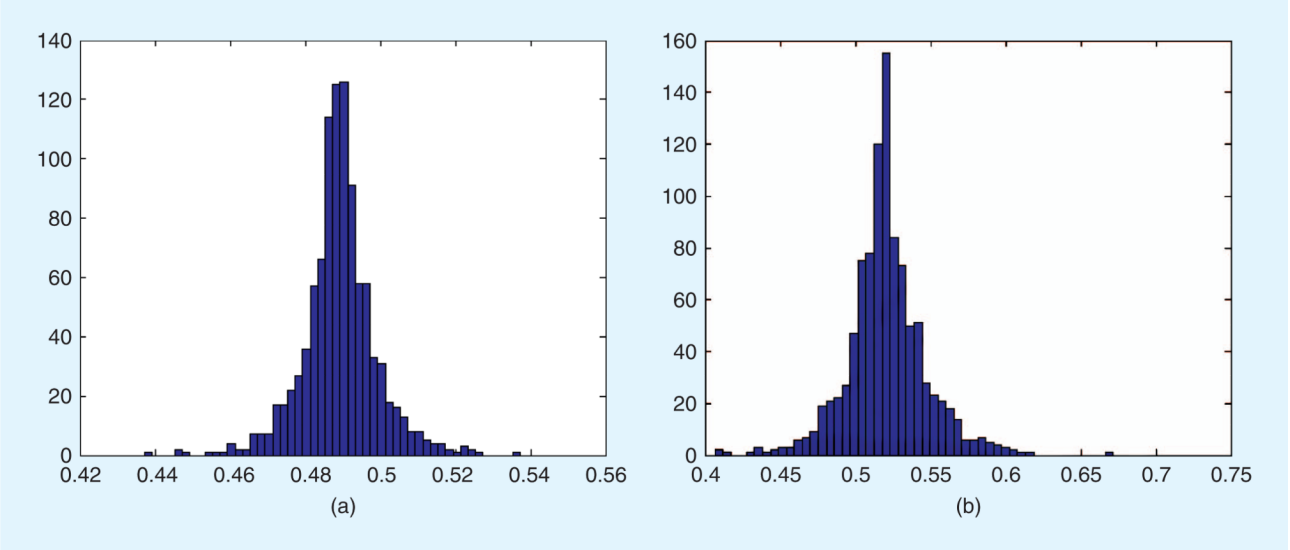
10. Sankar, L.; Rajagopalan, SR.; Poor, HV. Utility-privacy tradeoff in databases: An information-theoretic approach.. IEEE Trans. Inform. Forensics Sec. [Online]. to be published. Available: http://dx.doi.org/10.1109/TIFS.2013.2253320
11. Le Ny, J.; Pappas, GJ. Differentially private filtering.. presented at 51st Conf. Decision and Control (CDC); Dec. 2012 [Online]. Available: http://dx.doi.org/10.1109/CDC.2012.6426355
12. Le Ny, J.; Pappas, GJ. Differentially private Kalman filtering.. presented at 50th Annu. Allerton Conf. Communications, Control and Computing; Oct. 2012 [Online]. Available: http://dx.doi.org/10.1109/Allerton.2012.6483414
13. Fan, L.; Xiong, L. Real-time aggregate monitoring with differential privacy.. presented at 21st ACM Int. Conf. Information and Knowledge Management (CIKM '12); [Online]. Available: http://dx.doi.org/10.1145/2396761.2398595
14. Chaudhuri K, Monteleoni C, Sarwate AD. Differentially private empirical risk minimization. J. Mach. Learn. Res. Mar.2011 12:1069–1109. [Online]. Available: http://jmlr.csail.mit.edu/papers/v12/chaudhuri11a.html. [PubMed: 21892342]
15. Rubinstein BIP, Bartlett PL, Huang L, Taft N. Learning in a large function space: Privacy-preserving mechanisms for SVM learning. J. Privacy Confident. 2012; 4(1):65–100. [Online] Available: http://repository.cmu.edu/jpc/vol4/iss1/4/.
16. Zhang J, Zhang Z, Xiao X, Yang Y, Winslett M. Functional mechanism: Regression analysis under differential privacy. Proc. VLDB Endowment. Jul; 2012 5(11):1364–1375. [Online] Available: http://vldb.org/pvldb/vol5/p1364_junzhang_vldb2012.pdf.
17. Chaudhuri K, Sarwate A, Sinha K. Near-optimal algorithms for differentially-private principal components. J. Mach. Learn. Res. to be published.
18. Hardt M, Ligett K, McSherry F. Advances in Neural Information Processing Systems. 2012; 25 [Online]. Available: http://books.nips.cc/papers/files/nips25/NIPS2012_1143.pdf.
19. Ghosh, A.; Roth, A. Selling privacy at auction.. presented at 12th ACM Conf. Electronic Commerce (EC '11); [Online]. Available: http://dx.doi.org/10.1145/1993574.1993605
20. Mironov, I. On significance of the least significant bits for differential privacy.. presented at ACM Conf. Computer and Communications Security (CCS '12); [Online]. Available: http://research.microsoft.com/apps/pubs/?id=173034
21. Wasserman L, Zhou S. A statistical framework for differential privacy. J. Amer. Stat. Assoc. 2010; 105(489):375–389. [Online] Available: http://dx.doi.org/10.1198/jasa.2009.tm08651.
22. Kasiviswanathan, SA.; Lee, HK.; Nissim, K.; Raskhodnikova, S.; Smith, A. What can we learn privately?. presented at IEEE 49th Annu. IEEE Symp. Foundations Computer Science (FOCS '08); [Online]. Available: http://dx.doi.org/10.1109/FOCS.2008.27
23. Blum, A.; Ligett, K.; Roth, A. A learning theory approach to non-interactive database privacy.. presented at 40th Annu. ACM Symp. Theory Computing (STOC '08); [Online]. Available: http://dx.doi.org/10.1145/1374376.1374464
24. Chaudhuri, K.; Hsu, D. JMLR Workshop and Conference Proceedings Series, vol. 19. Proceedings of the 24th Annual Conference on Learning Theory (COLT `11); Jun. 2011 [Online]. Available: http://www.jmlr.org/ proceedings/papers/v19/chaudhuri11a/chaudhuri11a.pdf
25. McSherry F. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. Commun. ACM. Sep; 2010 53(9):89–97. [Online] Available: http://dx.doi.org/10.1145/1810891.1810916.
26. Roy, I.; Setty, STV.; Kilzer, A.; Shmatikov, V.; Witchel, E. Airavat: Security and privacy for mapreduce. Proc. 7th USENIX Conf. Networked Systems Design and Implementation (NSDI '10); Berkeley, CA.
27. Mohan, P.; Thakurta, A.; Shi, E.; Song, D.; Culler, D. GUPT: Privacy preserving data analysis made easy. Proc. 2012 ACM SIGMOD Int. Conf. Management Data; p. 349-360.
28. Barak, B.; Chaudhuri, K.; Dwork, C.; Kale, S.; McSherry, F.; Talwar, K. Privacy, accuracy, and consistency too: A holistic solution to contingency table release.. presented at 26th ACM SIGMOD-SIGACT-SIGART Symp. Principles Database Systems (PODS '07); [Online]. Available: http://dx.doi. org/10.1145/1265530.1265569

29. Yang X, Fienberg SE, Rinaldo A. Differential privacy for protecting multi-dimensional contingency table data: Extensions and applications. J. Privacy Confident. 2012; 4(1):101–125. [Online] Available: http://repository.cmu.edu/jpc/vol4/iss1/5.
30. Ding, B.; Winslett, M.; Han, J.; Li, Z. Differentially private data cubes: Optimizing noise sources and consistency.. presented at 2011 ACM SIGMOD Int. Conf. Management Data; [Online] Available: http://dx.doi.org/10.1145/1989323.1989347
31. Dwork, C.; Kenthapadi, K.; McSherry, F.; Mironov, I.; Naor, M. Lecture Notes in Computer Science Series, vol. 4004. Advances in Cryptology—EUROCRYPT 2006. [Online]. Available: http://dx.doi.org/10.1007/11761679_29
32. Machanavajjhala, A.; Kifer, D.; Abowd, JM.; Gehrke, J.; Vilhuber, L. Privacy: Theory meets practice on the map.. presented at IEEE 24th Int. Conf. Data Engineering (ICDE); Jun. 2008 [Online]. Available: http://dx.doi.org/10.1109/ICDE.2008.4497436
33. Dwork, C.; Rothblum, G.; Vadhan, S. Boosting and differential privacy.. presented at 51st Annu. IEEE Symp. Foundations Computer Science (FOCS '10); Oct. 2010 [Online]. Available: http://dx.doi.org/10.1109/FOCS.2010.12
34. Nissim, K.; Raskhodnikova, S.; Smith, A. Smooth sensitivity and sampling in private data analysis.. presented at 39th Annu. ACM Symp. Theory Computing (STOC '07); [Online]. Available: http://dx.doi.org/10.1145/1250790.1250803
35. McSherry, F.; Talwar, K. Mechanism design via differential privacy.. presented at 48th Annu. IEEE Symp. Foundations Computer Science (FOCS '07); [Online]. Available: http://dx.doi.org/10.1109/FOCS.2007.41
36. Hardt, M.; Roth, A. Beating randomized response on incoherent matrices.. presented at 44th Annu. ACM Symp. Theory Computing (STOC '12); [Online]. Available: http://dx.doi.org/10.1145/2213977.2214088
37. Hardt, M.; Roth, A. Beyond worst-case analysis in private singular vector computation. Proc. 45th Annu. ACM Symp. Theory Computing (STOC '13); New York. June 2013;
38. Dwork, C.; Lei, J. Differential privacy and robust statistics.. presented at 41st Ann. ACM Symp. Theory Computing (STOC '09); [Online]. Available: http://dx.doi.org/10.1145/1536414.1536466
39. Smith, A. Privacy-preserving statistical estimation with optimal convergence rates.. presented at 43rd Annu. ACM Symp. Theory Computing (STOC '11); [Online]. Available: http://dx.doi.org/10.1145/1993636.1993743
40. Blum, A.; Dwork, C.; McSherry, F.; Nissim, K. Practical privacy: The SuLQ framework.. presented at 24th ACM SIGMOD-SIGACT-SIGART Symp. Principles Database Systems (PODS '05); [Online]. Available: http://dx.doi.org/10.1145/1065167.1065184
41. Blocki, J.; Blum, A.; Datta, A.; Sheffet, O. The Johnson– Lindenstrauss Transform itself preserves differential privacy.. presented at IEEE 53rd Annu. Symp. Foundations Computer Science (FOCS); Oct. 2012 [Online]. Available: http://dx.doi.org/10.1109/FOCS.2012.67
42. Hall R, Rinaldo A, Wasserman L. Differential privacy for functions and functional data. J. Mach. Learn. Res. 2013; 14:703–727. [Online] Available: http://jmlr.csail.mit.edu/papers/v14/hall13a.html.
43. Huber, PJ. CBMS-NSF Regional Conference Series in Applied Mathematics. Robust Statistical Procedures. 2nd ed.1996. [Online]. Available: http://dx.doi.org/10.1137/1.9781611970036
44. Chaudhuri, K.; Hsu, D. Convergence rates for differentially private statistical estimation.. presented at 29th Int. Conf. Mach. Learn. (ICML-12); [Online]. Available: http://icml.cc/2012/papers/663.pdf
45. Lei J. Differentially private M-estimators. Advances in Neural Information Processing Systems. 2011; 24 [Online]. Available: http://books.nips.cc/papers/files/nips24/NIPS2011_0256.pdf.
46. Kifer, D.; Smith, A.; Thakurta, A. JMLR Workshop and Conference Proceedings Series, vol. 23. Proceedings of the 25th Annual Conference on Learning Theory (COLT '12); Jun. 2012 [Online]. Available: http://jmlr.csail.mit.edu/proceedings/papers/v23/kifer12/kifer12.pdf
47. Cormode, G. Personal privacy vs population privacy: Learning to attack anonymization.. presented at 17th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '11); [Online]. Available: http://dx.doi.org/10.1145/2020408.2020598

48. Kapralov, M.; Talwar, K. On differentially private low rank approximation. Proc. 24th Annu. ACM–SIAM Symp. Discrete Algorithms (SODA `13); New Orleans, LA. p. 1395-1414.
49. Jain, P.; Kothari, P.; Thakurta, A. JMLR Workshop and Conference Proceedings Series, vol. 23. Proceedings of the 25th Annual Conference on Learning Theory (COLT '12); Jun. 2012 [Online]. Available: http://www.jmlr.org/proceedings/papers/v23/jain12/jain12.pdf
50. Duchi J, Jordan M, Wainwright M. Advances in Neural Information Processing Systems. 2012; 25 [Online]. Available: http://books.nips.cc/papers/files/nips25/NIPS2012_0682.pdf.
51. Friedman, A.; Schuster, A. Data mining with differential privacy.. presented at 16th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining (KDD '10); [Online]. Available: http://dx.doi.org/10.1145/1835804.1835868
52. Mohammed, N.; Chen, R.; Fung, BCM.; Yu, PS. Differentially private data release for data mining.. presented at 17th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '11); [Online]. Available: http://dx.doi.org/10.1145/2020408.2020487
53. Jagannathan G, Pillaipakkamnatt K, Wright RN. A practical differentially private random decision tree classifier. Trans. Data Privacy. 2012; 5(1):273–295.
54. Williams O, McSherry F. Advances in Neural Information Processing Systems. 2010; 23 [Online]. Available: http://books.nips.cc/papers/files/nips23/NIPS2010_1276.pdf.
55. Pathak MA, Raj B. Large margin Gaussian mixture models with differential privacy. IEEE Trans. Dependable Secure Comput. Jul-Aug;2012 9(4):463–469. [Online] Available: http://dx.doi.org/10.1109/TDSC.2012.27.
56. Hoff PD. Simulation of the matrix Bingham–von Mises–Fisher distribution, with applications to multivariate and relational data. J. Comput. Graph. Statist. 2009; 18(2):438–456.
57. Rastogi, V.; Nath, S. Differentially private aggregation of distributed time-series with transformation and encryption.. presented at 2010 ACM SIGMOD Int. Conf. Management Data; [Online]. Available: http://dx.doi.org/10.1145/1807167.1807247
58. Beimel, A.; Kasiviswanathan, SP.; Nissim, K. Lecture Notes in Computer Science Series, vol. 5978. Theory of Cryptography. Feb 9–11. 2010 [Online]. Available: http://dx.doi.org/10.1007/978-3-642-11799-2_26
59. Haeberlen, A.; Pierce, BC.; Narayan, A. Proc. 20th USENIX Conf. Security. Berkeley, CA: 2011. Differential privacy under fire.
60. Proserpio, D.; Goldberg, S.; McSherry, F. A workflow for differentially-private graph synthesis.. presented at 2012 ACM Workshop Online Social Networks (WOSN '12); [Online]. Available: http://dx.doi.org/10.1145/2342549.2342553
61. McSherry, F.; Mahajan, R. Differentially-private network trace analysis.. presented at ACM SIGCOMM 2010 Conf.; [Online]. Available: http://dx.doi.org/10.1145/1851182.1851199



[FIG1].
An illustration of different approaches for guaranteeing differential privacy.



[FIG2].
a comparison of computing the mean and the median. (a) outputs of 1,000 runs of the differentially private sample mean algorithm. (b) outputs of 1,000 runs of the differentially private sample median algorithm.