Privacy Preserving Machine Learning, A Survey

Kritika Prakash
Machine Learning Lab
International Institute of Information Technology,
Hyderabad, India
kritika.prakash@research.iiit.ac.in

Sujit Gujar
Machine Learning Lab
International Institute of Information Technology,
Hyderabad, India
sujit.gujar@iiit.ac.in

***Abstract*—This survey focuses on understanding various solutions and mechanisms of preserving privacy and their application in Machine Learning algorithms. It gives a concise summary of the privacy-utility trade-offs achieved by various methods of preserving privacy in different settings of querying, involvement of data sources and distribution. It pays special attention to the technique - differential privacy.**

In general, we can assume that data sources value their privacy and a participant in a survey would participate only if he/she has the guarantee that his/her data is at least as private as it would be if he/she chose not to participate in the survey at all.

CONTENTS

I. INTRODUCTION

Anonymity is not enough. In 2006, Netflix announced a $1 million prize challenge to improve their movie recommendation service. They publicly released a large dataset containing approximately 100 million movie ratings created by approximately 480 thousand Netflix subscribers over six years. Netflix claimed to have removed all customer identifying information. But this definition of anonymity is not absolute - they didn't take into account how much a potential adversary (trying to breach the privacy of a user) needs to know about a Netflix subscriber in order to identify his/her record in the dataset. Thus, the question

of compromising privacy becomes existential in order to prove that a public release of a dataset does not breach it's participants' privacy and maintains their anonymity.

Shortly after Netflix released it's prize dataset, a linkage attack was created and published which linked the Netflix database with the publicly available view of the IMDB database, resulting in a privacy breach of many of the Netflix subscribers.

Some conventional approaches to privacy are anonymization, sanitization (sampled subset) and controlling access and flow of information. But these do not provide the privacy guarantee.

Privacy of individuals has become a need with rising popularity in today's age of information. This survey of research work over the past few years on differential privacy and privacy preserving machine learning aims to understand the current state of theoretically guaranteed privacy bounds, privacy mechanisms in various scenarios, their implications on the effectiveness of machine learning algorithms and adversarial attacks trying to breach the privacy of participants in a survey or a dataset.

We first provide an introduction to the field of differential privacy and privacy preserving machine learning, followed by a classification of papers explored in a hierarchical manner, and then we present a summary of each paper in the subsequent section.

II. Differential Privacy Basics

Differentially private database mechanisms can make confidential data widely available for accurate data analysis, without resorting to data clean rooms, data usage agreements, data protection plans, or restricted views. Differential privacy addresses the paradox of learning nothing about an individual while learning useful information about a population.

Differential privacy guarantees that the impact on the participant of a study is the same independent of whether or not he was in the study. It is the conclusions reached in the study that affect the participant, not his presence or absence in the data set.

Differential privacy is a definition, not an algorithm. For a given computational task T and a given value of ϵ there will be many differentially private algorithms for achieving T in an ϵ-differentially private manner. Some will have better accuracy than others. When ϵ is small, finding a highly accurate ϵ-differentially private algorithm for T can be difficult. Therefore, differential privacy makes it impossible to guess whether one participated in a database with large probability.

A. Randomized Algorithm

A randomized algorithm M with domain A and discrete range B is associated with a mapping $M : A \to \Delta(B)$. On input $a \in A$, the algorithm M outputs $M(a) = b$ with probability $(M(a))_b$ for each $b \in B$. The probability space is over the coin flips of the algorithm M.

B. Distance between Databases

The l_1 norm of a database x is denoted $||x||_1$ (number of records in the database) and is defined to be:

$$||x||_1 = \sum_{i=1}^{|\chi|} |x_i| \tag{1}$$

The l_1 distance between two databases x and y is $||x - y||_1$.

C. (ϵ, δ)-Differential Privacy

A randomized algorithm M with domain $\mathbb{N}^{|\chi|}$ is $(\epsilon, \delta)-$ differentially private if for all $S \subseteq$ Range (M) and for all $x, y \in N^{|\chi|}$ such that $||x - y||_1 \leq 1$:

$$Pr[M(x) \in S] \leq exp(\epsilon) Pr[M(y) \in S] + \delta \tag{2}$$

We are interested in values of δ in the order of $1/||x||_1$.

(ϵ, δ) - differential privacy ensures that for all adjacent x, y, the absolute value of the privacy loss will be bounded by ϵ with probability at least $1 - \delta$. Differential Privacy is immune to post processing.

D. l_1-sensitivity

The l_1-sensitivity of a function $f : \mathbb{N}^{|\chi|} \to R^k$ is:

$$\Delta f = \max ||f(x) - f(y)||_1 \tag{3}$$

where $||x - y||_1 = 1$ and $x, y \in \mathbb{N}^{|\chi|}$

The l_1-sensitivity of a function f captures the magnitude by which a single individual's data can change the function f in the worst case. This gives an upper bound on how much we need to perturb the data to preserve privacy. The l_1 sensitivity of counting queries is 1.

E. The Laplace Mechanism

Given any function $f : \mathbb{N}^{|\chi|} \to R^k$, the Laplace Mechanism is defined as:

$$M_L(x, f(.), \epsilon) = f(x) + (Y_1, Y_2, ..., Y_k) \tag{4}$$

where Y_i's are independently identically distributed (i.i.d) random variables drawn from the Laplacian distribution $Lap(\Delta f/\epsilon)$. The Laplace Mechanism preserves $(\epsilon, 0)$-differential privacy.

F. The Exponential Mechanism

The exponential mechanism was designed for situations in which we wish to choose the best response but adding noise directly to the computed quantity can completely destroy its value. For example, setting a price of an auction, where the goal is to maximize the revenue is a highly sensitive value.
The exponential mechanism $M_E(x, u, R)$ selects and outputs an element $r \in R$ with probability proportional to $exp(\epsilon u(x, r)/2\Delta u)$. The exponential mechanism preserves $(\epsilon, 0)$-differential privacy.

G. Query Types

- Large Sensitivity Queries
- Numerical Queries (Counting)
- Set-Based Queries
- Graph-Based Queries

H. Composition of Queries

What happens when the same private database is queried multiple times? Are the privacy bounds compromised? Can we prove different theoretical privacy for this setting? Does the new bound depend on the number of queries? Can it be made independent of the number of queries?

Let $M_i : \mathbb{N}^{|\chi|} \rightarrow R_i$ be an (ϵ_i, δ_i)-differentially private algorithm for $i \in [k]$. Then if $M_{[k]} : \mathbb{N}^{|\chi|} \rightarrow \prod_{i=1}^{k} R_i$ is defined to be $M_{[k]}(x) = (M_1(x), M_2(x), ..., M_k(x))$, then $M_{[k]}$ is $(\sum_{i=1}^{k} \epsilon_i, \sum_{i=1}^{k} \delta_i)$-differentially private.

I. Optimal bound on Composition of Queries

(Advanced Composition)
For all $\epsilon, \delta, \delta^{'} \geq 0$, the class of (ϵ, δ)-differentially private mechanisms satisfy $(\epsilon^{'}, k\delta + \delta^{'})$-differential privacy under k-fold adaptive composition for:

$$\epsilon^{,} = \sqrt{2kln(1/\delta')}\epsilon + k\epsilon(e^{\epsilon} - 1). \quad (5)$$

III. Differential Privacy in Machine Learning

Many machine learning tasks can be performed under the constraint of differential privacy. Contrary to intuition, privacy constraints need not be at odds with the goals of machine learning. Some of the Machine Learning algorithms that have been modified in literature to suit privacy needs are:

- Logistic Regression
- Low Rank Approximation
- Principal Component Analysis
- Support Vector Machines
- Deep Learning
- K-Means Clustering
- Linear Classification

IV. Privacy Preserving Machine Learning Approaches

A. Data Perturbation

Data perturbation approaches can be grouped into two main categories: the probability distribution approach and the value distortion approach. The probability distribution approach replaces the data with another sample from the same (or estimated) distribution or by the distribution itself, and the value distortion approach perturbs data elements or attributes directly by either additive noise, multiplicative noise, or some other randomization procedures.

B. Differential Privacy

Modifying the data in a manner such that upon querying the database, no more information about a participant in revealed than the amount revealed if he/she didn't participate in the data survey.

C. Distributed ML

Extraction of patterns at a given node in a distributed network by exchanging only the minimal necessary information among the participating nodes.

D. Secure Multi-Party Computation

Evaluate a function of the secret inputs coming from two or more input sources, such that no party learns anything more than the expected output of the function.

E. Homomorphic Encryption

Homomorphic Encryption is a special case of a Secure Multi-Party Computation, where an encryption transformation is applied to the data, such that certain distances and operations on the data are preserved in the encrypted space.

F. Data Swapping

Transform the database by switching a subset of attributes between selected pairs of records such that lower order frequency counts are preserved and data confidentiality is not compromised.

G. k-Anonymity

The information of each person contained in the publicly released dataset cannot be distinguished from at least $k - 1$ other distinct people.

H. Rule Hiding

Transform the database such that sensitive rules are masked, and at the same time, underlying patterns can still be discovered. Decrease the support of sensitive rules using a user-specified threshold of hiding rules.

V. Research Paper Index

This section provides a high-level overview of the research papers surveyed.

Table Index			
Paper	**Appeared in**	**Problem Addressed**	**Novel Approach & Contributions**
Foundation Level: Differential Privacy			
The Algorithmic Foundations of Differential Privacy [1]	Foundations and Trends in Theoretical Computer Science, 2014	Strengthening the security of existing systems, surveys and algorithms in a cost-effective manner, such that the privacy of the data sources is not compromised by their decision to participate.	Cynthia Dwork co-founded the field of Differential Privacy. This book provides a detailed understanding of differential privacy as a solution framework to the existing theoretical privacy breaches over surveys and datasets.
The Promise of Differential Privacy: A Tutorial on Algorithmic Techniques [2]	Foundations of Computer Science (FOCS), IEEE, 2011	Understanding differential privacy.	A rich explanatory tutorial on the topic of differential privacy.
Differential Privacy: A Short Tutorial	Presented by Wang Yuxiang	Differential privacy along with its applications.	An intuitive introduction to differential privacy, its mechanisms, advanced techniques, and its applications in Machine Learning.
Mechanism Level: Papers Explored in Depth			
The Johnson-Lindenstrauss Transform Itself Preserves Differential Privacy [3]	Foundations of Computer Science (FOCS), IEEE, 2012	Approximating cut-queries in a graph such that edge-differential privacy is preserved. Estimating the variance of a matrix in a given direction in a differentially-private manner.	Used the existing Johnson-Lindenstrauss Transform to preserve differential privacy without compromising on utility bounds to approximate cut-queries in graphs and estimate directional variance in a matrix. Error introduced is data independent (Additive Gaussian noise).
Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining [4]	IEEE Transactions on knowledge and Data Engineering, 2006	Privacy preserving distributed data mining, specifically computing statistical aggregates, clustering, PCA & classification.	Used multiplicative random projection matrices to preserve privacy while preserving certain statistical characteristics' invariance. Explores Independent Component Analysis as a privacy breach.
Interactive Privacy via the Median Mechanism [5]	Proceedings of the forty-second ACM symposium on Theory of computing, ACM, 2010	Answering a large number of arbitrary predicate queries arriving online in a privacy-preserving manner.	Mechanism introduced which answers exponentially more number of interactive (online) queries - the Median mechanism, by exploiting correlations across queries.
Analyze Gauss: Optimal Bounds for Privacy-Preserving Principal Component Analysis [6]	Proceedings of the forty-sixth annual ACM symposium on Theory of computing. ACM, 2014	Releasing a private low dimensional approximation of a database represented as a database.	Compute PCA with the optimal privacy-utility trade-off. Used randomized response algorithm. Created a private online algorithm with nearly optimal regret.
Papers Explored Briefly			
Deep Learning with Differential Privacy [7]	ACM, 2016	Crowdsourced datasets containing sensitive information can expose private information when used for training deep learning models.	Explore new algorithmic techniques for learning specific to deep-learning.
GELU-Net: A Globally Encrypted, Locally Unencrypted Deep Neural Network for Privacy-Preserved Learning [8]	IJCAI, 2018	Collaborative training of deep-learning models in a privacy-preserving manner per client.	Encrypt the data using Partial Homomorphic Encryption scheme and perform linear computations. Perform non-linear computations on unencrypted data. Presented a novel privacy-preserving deep learning architecture.
Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netflix Prize Dataset) [9]	Proceedings 29^{th} IEEE Symposium on Security and Privacy, 2008	Privacy breach attacks on "anonymized" publicly released datasets of real-world users.	Presents a new class of statistical de-anonymization attacks against high-dimensional micro-data, such as individual preferences, recommendations, transaction records. Apply the de-anonymization methodology to Netflix dataset and create an adversary (privacy breach) attack.
Secure Face Matching Using Fully Homomorphic Encryption [10]	2018	Making face representation systems & databases more secure, as face recognition is strongly linked to an individual's identity.	Fully homomorphic encryption based framework to secure a database of face templates. It preserves the privacy of the users and prevents information leakage, while maintaining their utility through operatibility in the encrypted domain.

VI. Classification of Papers

A. Foundation Level: Differential Privacy

Papers:

- The Algorithmic Foundations of Differential Privacy
- The Promise of Differential Privacy: A Tutorial on Algorithmic Techniques
- Differential Privacy: A Tutorial

B. Mechanism Level: Papers Explored in Depth

Papers:

- The Johnson Lindenstrauss Transform Itself Preserves Differential Privacy
- Interactive Privacy via the Median Mechanism
- Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining
- Analyze Gauss: Optimal Bounds for Privacy-Preserving Principal Component Analysis

C. Papers Explored Briefly

Papers:

- Deep Learning with Differential Privacy
- GELU-Net: A Globally Encrypted, Locally Unencrypted Deep Neural Network for Privacy-Preserved Learning
- Secure Face Matching Using Fully Homomorphic Encryption

VII. Foundation Level: Differential Privacy

A. The Algorithmic Foundations of Differential Privacy

Key Points

- Privacy-preserving data analysis
- Model of computation
- Differential privacy
- Randomized response
- Laplace Mechanism
- Exponential Mechanism
- Composition theorems
- Sparse vector techniques

B. The Promise of Differential Privacy: A Tutorial on Algorithmic Techniques

Key Points

- Explaining the need for privacy and differential privacy
- Differential privacy definition
- Privacy loss - utility trade-off
- Differential Privacy Overview

C. Differential Privacy: A Short Tutorial

Key Points

- Intuition behind how anonymization is not enough to preserve privacy
- What differential privacy can and cannot do (scope).
- Global Sensitivity and Laplace Mechanism
- Composition of queries
- Sparse-vector technique
- Exponential mechanism and net-mechanism
- DP Logistic Regression
- DP Low-Rank Approximation
- DP Principal Component Analysis
- DP Support Vector Machine

VIII. Mechanism Level: Papers Explored in Depth

A. The Johnson-Lindenstrauss Transform Itself Preserves Differential Privacy

Key Points

- Novel approach of preserving differential privacy.
- If we take two neighbouring databases $D, D^{'}$ then multiplying D or $D^{'}$ with a vector of i.i.d. normal Gaussians yields two statistically close distributions from the DP perspective.
- JL Transform used to approximate cut-queries in graph databases: the number of edges crossing a $(S, S^{'})$-cut in a graph, such that edge differential privacy is preserved, using only additive noise.
- Different from many other methods as sanitization of the database is performed instead of sanitization of a query result. This allows publishing of the sanitized database publicly.
- JL Transform applied to the task of estimating the variance of a given matrix in any direction. A sanitized covariance matrix is published in a DP manner, such that the noise introduced is additive and independent of the matrix dimensions.

The JL Transform

Fix any $0 < \eta < 1/2$. Let M be a $r * m$ matrix whose entries are i.i.d. samples from $\mathcal{N}(0, 1)$. Then $\forall x \in \mathbb{R}^m$:

$$Pr_M[1/r||Mx||^2 \notin (1 \pm \eta)||x||^2] \leq 2exp(-\eta^2 r/8) \quad (6)$$

Model

Consider an undirected weighted graph $G = (V(G), E(G))$ represented as a $\binom{n}{2} * n$ edge matrix, E_G. The vertices are ordered arbitrarily, and for each pair of vertices u, v where $u < v$, there exists a row in E_G.

$$E_{G(u,v,x)} = \begin{cases} \sqrt{w_{u,v}}, & \text{if } u \sim_G v \text{ and } x = u \\ -\sqrt{w_{u,v}}, & \text{if } u \sim_G v \text{ and } x = v \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where $u \sim_G v$ denotes that (u, v) is an edge in G.

Let $L_G = E_G^T E_G$ be the Laplacian of the graph.

Publishing a Perturbed Laplacian

Algorithm: Publishing a perturbed Laplacian of a graph while preserving differential privacy using the JL Transform. This algorithm preserves (ϵ, δ)-differential privacy.

Algorithm 1: Outputting the Laplacian of a Graph while Preserving Differential Privacy

Input: A n-node graph G, parameters: $\epsilon, \delta, \eta, \nu > 0$
Output: A Laplacian of a graph $\tilde{L}$

Set $r = \frac{8\ln(2/\nu)}{\eta^2}$, and $w = \frac{\sqrt{32r\ln(2/\delta)}}{\epsilon}\ln(4r/\delta)$
For every $u \neq v$, set $w_{u,v} \leftarrow \frac{w}{n} + \left(1 - \frac{w}{n}\right) w_{u,v}$.
Pick a matrix M of size $r \times \binom{n}{2}$, whose entries are iid samples of $\mathcal{N}(0,1)$.
return $\tilde{L} = \frac{1}{r} E_G^{\intercal} M^{\intercal} M E_G$

Publishing a Covariance Matrix

Algorithm: Publishing a perturbed Covariance Matrix by adding random Gaussian noise while maintaining its utility - privacy trade-off. This algorithm preserves (ϵ, δ)-differential privacy.

Algorithm 3: Outputting a Covariance Matrix while Preserving Differential Privacy

Input: A $n \times d$ matrix A. Parameters $\epsilon, \delta, \eta, \nu > 0$.

Set $r = \frac{8\ln(2/\nu)}{\eta^2}$ and $w = \frac{16\sqrt{r\ln(2/\delta)}}{\epsilon}\ln(16r/\delta)$.
Subtract the mean from A by computing $A \leftarrow A - \frac{1}{n}\mathbf{1}\mathbf{1}^T A$.
Compute the SVD of $A = U\Sigma V^{\intercal}$.
Set $A \leftarrow U(\sqrt{\Sigma^2 + w^2 I_{n\times d}})V^{\intercal}$.
Pick a matrix M of size $r \times n$ whose entries are iid samples of $\mathcal{N}(0,1)$.
return $\tilde{C} = \frac{1}{r} A^{\intercal} M^{\intercal} M A$.

B. Random Projection-Based Multiplicative Data Perturbation for Privacy Preserving Distributed Data Mining

Suppose there are N organizations $O_1, O_2 ... O_N$. Each organization O_i has a private transaction database DB_i. A third party data miner wants to learn certain statistical properties of the union of these databases $N_i = 1DB_i$. These organizations are comfortable with this, but they are reluctant to disclose their raw data. How could the data miner perform data analysis without compromising the privacy of the data? This is generally referred to as the census problem. In this scenario, the data is usually distorted and its new representation is released, and anybody has arbitrary access to the published data.

Key Points

- Use multiplicative random projection matrices for privacy preserving data mining.
- Computing statistical aggregates like inner product matrix, correlation coefficient matrix, and Euclidean distance matrix from distributed privacy sensitive data owned by multiple parties. This has direct applications in PCA, clustering and classification.
- Explores Independent Component Analysis as a possible tool for breaching privacy in deterministic multiplicative perturbation-based models such as random orthogonal transformation and random rotation.
- Proposes an approximate random projection-based technique to improve the level of privacy protection while still preserving certain statistical characteristics of the data.

In this paper, we mainly focus on the value distortion approach.

Random Orthogonal Transformation

An orthogonal transformation [28] is a linear transformation $R : \mathbb{R}^n \to \mathbb{R}^n$, which preserves the length of vectors as well as the angles between them. Let R be an nxn random orthogonal matrix. If U, V are linear transformations of the datasets X, Y respectively,
$U = XR$, and $V = YR$, then we have

$$UU^T = XX^T, VV^T = YY^T, UV^T = XRR^TY^T = XY^T \tag{8}$$

This allows a third party to analyze the data (distance and correlation) without accessing the raw data. But this transformation is just a pure rotation or a rotoinversion, which is very easy to decipher, and not secure enough.

Independent Component Analysis

Independent Component Analysis (ICA) is a technique for discovering independent hidden factors that are underlying a set of linear or nonlinear mixtures of some unknown variables, where the mixing system is also unknown.

$$u(t) = Rx(t) \tag{9}$$

where $x(t) = (x_1(t), x_2(t), ..., x_m(t))^T$ is an $m-$dimensional vector collecting m independent source signals which vary with time. R is a constant kxm unknown mixing matrix. $u(t) = (u_1(t), u_2(t), ..., u_k(t))^T$ is the observed matrix. Uniqueness of recovered signals found by ICA will have scaling and permutation ambiguities.

Decomposability

A linear filter is designed to get the recovered signal $y(t) = (y_1(t), y_2(t), ..., y_l(t))^T$.

$$y(t) = Bu(t) = Zx(t) \tag{10}$$

where B is an $l * k$-dimensional separating matrix, and $Z = BR$ is an lxm matrix. When $k \geq m$ and R has full column rank, there always exists a matrix B such that $Z = BR = I$, the identity matrix, and all signals can be recovered with permutation and scaling ambiguities. At least all but one signals need to be non-Gaussian.

When $l \leq k < m$ (fewer receivers than sources), the source signals can at most be separated into k disjoint groups from the mixtures, and at most $k - 1$ original signals can be separated out.

When the mixing matrix R is not two-row decomposable ($m \geq 2k - 1, m \geq 2$, with i.i.d. entries chosen from a continuous distribution), there is no linear method that can find a matrix B to separate out any of the source signals.

Random Projection-based Multiplicative Perturbation

The Johnson-Lindenstrauss Lemma introduces random projections from a high-dimensional space to a randomly chosen low-dimensional space. Both row-wise and column-wise projections are performed. This is done in such a manner that statistical dependencies among the observations are maintained, along with preserving privacy.

Applications

This work can be applied to machine learning algorithms like K-Means Clustering and Linear Classification.

C. Interactive Privacy via the Median Mechanism

Key Points

- Interactive Differentially Private mechanism via the Median mechanism for answering arbitrary predicate queries that arrive online.
- Given fixed accuracy and privacy constraints, this mechanism can answer exponentially more queries than the previously best known interactive privacy mechanism (the Laplace mechanism, which independently perturbs each query result).
- The median mechanism is the first privacy mechanism capable of identifying and exploiting correlations among queries in an interactive setting.
- Efficient implementation of the median mechanism with running time polynomial in the number of queries, database size and domain size.
- Guarantees on privacy for all input databases and accurate query results for all input distributions.

The interactive or online model permits the database to be queried adaptively by the data analyst. This way, the data analyst can decide which query to pose next based on the observed responses to previous queries. The key challenge is to design an interactive mechanism that outperforms the Laplace mechanism. For this, we need to determine correlations between different output perturbations on the fly, independent of future queries.

Model

Privacy parameter: α
Number of queries: k
Accuracy parameter: ϵ
The above three are hard constraints on the performance of the mechanism. The mechanism obeys these constraints with a value of δ inversely polynomial in k, n and a value of τ negligible in k, n.
Database size: n

The Median Mechanism

- Among any set of k queries, there are $O(\log k \log|X|)$ hard queries, which act as a basis for all other queries, as the answers to these hard queries can completely determine the answer to all other queries (up to $\pm\epsilon$).
- They design a method to privately release an indicator vector which distinguishes between hard and easy queries online. A query is easy if s majority of the databases that are consistent (up to $\pm\epsilon$) with the previous answers of the mechanism would answer the current query accurately.
- The Median Mechanism answers the small number of hard queries using independent Laplace perturbations.
- If a user knows that query i is easy, then it can generate the mechanisms answer on its own.
- They show how to release the classification of queries as easy and hard at a low cost of privacy, as there can only be $O(\log k \log|X|)$ hard queries.

---

1. Initialize $C_0 = \{$ databases of size m over X $\}$.
2. For each query $f_1, f_2, \ldots, f_k$ in turn:
   (a) Define r_i as in (4) and let $\hat{r}_i = r_i + \mathrm{Lap}(\frac{2}{\epsilon n \alpha'})$.
   (b) Let $t_i = \frac{3}{4} + j \cdot \gamma$, where $j \in \{0, 1, \ldots, \frac{1}{\gamma}\frac{3}{20}\}$ is chosen with probability proportional to 2^{-j}.
   (c) If $\hat{r}_i \geq t_i$, set a_i to be the median value of f_i on C_{i-1}.
   (d) If $\hat{r}_i < t_i$, set a_i to be $f_i(D) + \mathrm{Lap}(\frac{1}{n\alpha'})$.
   (e) If $\hat{r}_i < t_i$, set C_i to the databases S of C_{i-1} with $|f_i(S) - a_i| \leq \epsilon/50$; otherwise $C_i = C_{i-1}$.
   (f) If $\hat{r}_j < t_j$ for more than $20m \log|X|$ values of $j \leq i$, then halt and report failure.

Figure 1: The Median Mechanism.

D. Analyze Gauss: Optimal Bounds for Privacy-Preserving Principal Component Analysis

Key Points

- Problem: Privately releasing a low dimensional approximation to a set of data records represented as a matrix A in which each row corresponds to an individual and each column to an attribute.
- Goal: Compute a subspace that captures the covariance of A as much as possible (Principal Component Analysis).
- Constraint: Each row of A has l_2 norm bounded by 1. Privacy guarantee is defined w.r.t. addition or removal of a single row.
- Result: Randomized response algorithm provides a nearly-optimal additive quality gap compared to the best possible singular subspace of A.
- When $A^T A$ has a large eigenvalue gap, the quality improves significantly.
- Result: Upon combining randomized response mechanism along with the *following the perturbed leader* algorithm, they obtain a private online algorithm with nearly optimal regret.

IX. Papers Explored Briefly

A. Deep Learning with Differential Privacy

Key Points

- By tracking detailed information about privacy loss, much tighter estimates on the overall privacy loss are obtained.
- Improve computational efficiency of differentially private training by introducing new techniques such as efficient algorithms for computing gradients for individual training examples, sub-dividing tasks into smaller batches to reduce memory footprint, and applying differentially private principal projection on the input layer.

B. GELU-Net: A Globally Encrypted, Locally Unencrypted Deep Neural Network for Privacy-Preserved Learning

Key Points

- Previously used Fully Homomorphic Encryption mechanisms are very costly to compute and unsuitable for large-scale neural networks.
- Activation functions are not cryptographically computable, and have previously been approximated by polynomials. This results in unstable training.
- Solution: Globally Encrypted, Locally Unencrypted Deep Neural Network (GELU-Net)
  Partition a deep neural network to two non-colluding parties. The first party performs linear computations on encrypted data utilizing a partially homomorphic cryptosystem of encryption. The second party executes non-polynomial computations in the original domain in a privacy preserving manner.
- GELU-Net is more efficient than fully homomorphic encryption, and retains the original accuracy of the algorithm, along with its stability.
- Algorithms for privacy preserving back propagation and forward propagation.
- Extensive analysis of security and complexity. Provide a 14 to 35 times speed-up w.r.t state-of-the-art solution.

C. Robust de-anonymization of large datasets (how to break anonymity of the Netflix prize dataset)

Key Points

- This example of privacy leak is via de-anonymization of Netflix data.
- Sparsity of data: With large probability, no two profiles are similar up to ϵ. In Netflix data, no two records are similar more than 50%.
- The adversary knows with high probability the true identity of the Netflix profile if it can be matched up to 50% similarity to a profile in the IMDB dataset.
- This paper proposes an efficient random algorithm adversarial attack to break privacy via de-anonymization.

D. Secure Face Matching Using Fully Homomorphic Encryption

Key Points

- Explore practicality and scope of using a fully homomorphic encryption based framework to secure a database of face templates.
- Face template matching is performed directly in the encrypted domain. The aim is to neither compromise on accuracy, nor on privacy.
- Explore a batching and dimensionality reduction scheme to improve the trade-off between computational complexity and accuracy of face matching. Multiple homomorphic encryptions are performed in a single operation.
- Algorithm for homomorphic inner product.

References

[1] Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." Foundations and Trends in Theoretical Computer Science 9.34 (2014): 211-407.

[2] Dwork, Cynthia. "The promise of differential privacy: A tutorial on algorithmic techniques." Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on. IEEE, 2011.

[3] Blocki, Jeremiah, et al. "The johnson-lindenstrauss transform itself preserves differential privacy." Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on. IEEE, 2012.

[4] Liu, Kun, Hillol Kargupta, and Jessica Ryan. "Random projection-based multiplicative data perturbation for privacy preserving distributed data mining." IEEE Transactions on knowledge and Data Engineering 18.1 (2006): 92-106.

[5] Roth, Aaron, and Tim Roughgarden. "Interactive privacy via the median mechanism." Proceedings of the forty-second ACM symposium on Theory of computing. ACM, 2010.

[6] Dwork, Cynthia, et al. "Analyze gauss: optimal bounds for privacy-preserving principal component analysis." Proceedings of the forty-sixth annual ACM symposium on Theory of computing. ACM, 2014.

[7] Abadi, Martin, et al. "Deep learning with differential privacy." Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2016.

[8] Zhang, Qiao, et al. "GELU-Net: A Globally Encrypted, Locally Unencrypted Deep Neural Network for Privacy-Preserved Learning." IJCAI. 2018.

[9] Narayanan, Arvind, and Vitaly Shmatikov. "Robust de-anonymization of large datasets (how to break anonymity of the Netflix prize dataset). 2008." University of Texas at Austin (2008).

[10] Boddeti, Vishnu Naresh. "Secure Face Matching Using Fully Homomorphic Encryption." arXiv preprint arXiv:1805.00577 (2018).