Statistic Maximal Leakage

Shuaiqi Wang
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Email: shuaiqiw@andrew.cmu.edu

Zinan Lin
Microsoft Research
Redmond, WA, 98052, USA
Email: zinanlin@microsoft.com

Giulia Fanti
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Email: gfanti@andrew.cmu.edu

2024 IEEE International Symposium on Information Theory (ISIT) | 979-8-3503-8284-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISIT57864.2024.10619258

***Abstract*—We introduce a privacy metric called *statistic maximal leakage* that quantifies how much a privacy mechanism leaks about a specific secret, relative to the adversary's prior information about that secret. Statistic maximal leakage is an extension of the well-known *maximal leakage*. Unlike maximal leakage, it protects a single, known secret. We show that statistic maximal leakage satisfies composition and post-processing properties. Additionally, we show how to efficiently compute it in the special case of deterministic data release mechanisms. We analyze two important mechanisms under statistic maximal leakage: the quantization mechanism and randomized response. We show theoretically and empirically that the quantization mechanism achieves better privacy-utility tradeoffs in the settings we study.**

I. Introduction

A common barrier to data sharing is the risk of leaking private information contained in, or correlated with, the released data [1]. Since the seminal work of Yamamoto [2], many papers have studied how to release a random variable without leaking correlated secret information [3]–[9]. In this work, we consider a data holder that possesses data drawn from a distribution parameterized by a random variable Θ, and knows a secret, which is represented as a discrete random variable G computed as a function $\mathfrak{g}$ of Θ. The data holder's goal is to release Θ', a perturbed version of Θ, while optimizing the tradeoff between the *leakage* about G and the *utility* of Θ'.

We study privacy metrics that satisfy three properties:
(1) Prior-independence: The metric should not depend on any party's prior over the input data, up to determining the support of the distribution. This arises because priors may be difficult to obtain in practice; metrics that require such knowledge may admit mechanisms that are fragile to prior mis-specification [10].
(2) Secret-specific: The metric should use knowledge of the secret (distribution). In other words, we assume the secret function $\mathfrak{g}$ is known, which describes how to obtain the secret G from the input data. We want a metric that depends explicitly on $\mathfrak{g}$. This is primarily for efficiency reasons; by utilizing known information, we may be able to add less noise or perturbation to our data.
(3) Composition and Post-processing: Informally, composition describes what happens to our privacy metric when one or more mechanisms are applied sequentially, i.e., a bound on how the privacy metric degrades. For examples, in differential privacy (DP), the privacy parameter ϵ degrades additively when a mechanism is applied multiple times to the same dataset. Post-processing states that if one applies an arbitrary (possibly random) function to the output of a privacy mechanism, as long as the function does not depend on the original data, the privacy metric in question does not degrade. Composition and post-processing are two very useful properties exhibited by DP, which have contributed to its widespread usage in practical settings such as machine learning pipelines [11].

Today there exist metrics that satisfy all three of these properties. To the best of our knowledge, these metrics are all inspired by differential privacy. Examples include attribute privacy [12], distribution privacy [13], and distribution inference [14]. Due to their conservative assumptions and formulation, they require large amounts of noise in practice [10], [15].[1]

In this work, we study an information-theoretic privacy metric that satisfies the above three properties. For a special class of data with distributions drawn from a parametric family, and parameter vectors drawn from a finite set, we propose a privacy metric inspired by maximal leakage [8], [9], which we call *statistic maximal leakage*. It trivially satisfies the first two properties, and we show that it also satisfies composition and post-processing. Note that although composition and post-processing have been previously proved for an extension of maximal leakage called *pointwise maximal leakage* [16], their result and proofs do not apply to statistic maximal leakage.

Given a definition and properties for statistic maximal leakage, we next study how to compute it. In general, computing statistic maximal leakage is intractable. However, we show that for the class of deterministic mechanisms, statistic maximal leakage can be computed in polynomial time by solving a maximum flow problem.

We next analyze two natural mechanisms that have been studied widely in the privacy literature: the quantization mechanism [10], [17] and randomized response [18], [19]. Quantization-based mechanisms have been shown to achieve (near)-optimal privacy-utility tradeoffs in several privacy frameworks, including summary statistic privacy [10] and non-randomized privacy [17]. Randomized response is a widely-adopted mechanism [18] that achieves optimal privacy-utility tradeoffs for differentially-private data collection [20]. We show that both mechanisms satisfy non-trivial statistic maximal leakage guarantees. Further, under general tabular datasets, we analyze their privacy-utility tradeoffs. Our results show that under most cases, the quantization mechanism achieves better tradeoffs.

[1]Note that for vanilla differential privacy, the secret it considers is sample-level, and thus does not represent general secret functions of the underlying distribution of the input data without modification. Further discussion of the limitations of differential privacy in our setting can be found in [10].

Finally, we apply the quantization mechanism to a real tabular dataset. We show that when instantiated with an appropriate statistic maximal leakage parameter, the quantization mechanism effectively protects the secret while still ensuring high utility of the released data.

II. Related Work

While there are many ways of categorizing existing privacy metrics, we consider our three desired properties:

(1) Prior-independence: Many existing information-theoretic metrics for quantifying privacy leakage require knowledge of a prior distribution of the data and possibly the secret G. These include metrics based on mutual information [6], min entropy [3], [10], [15], [21], maximal leakage [9] and other f-divergences [22]. This is a strong assumption in practice, and hence several metrics have instead aimed to remove dependence on the data prior, up to determining the support of the distribution. Perhaps the most well-known example of a prior-independent metric is differential privacy [23] and its variants (e.g., attribute privacy [12], distribution privacy [13], distribution inference [24]). Another example is maximal (α, β)-leakage [25]. In this work, we are interested in a practically-motivated class of *prior-invariant* metrics.

(2) Secret-specific: Not all privacy metrics assume prior knowledge of the secret that needs to be hidden. For example, maximal leakage and its current variants [9], [16], [25]–[27] assume the secret is unknown *a priori*, and hence is a worst-case metric over all secrets. In contrast, many other metrics assume prior knowledge of the secret [10], [12], [15], [28]; this can allow for better utility, as it does not require protecting against arbitrary secrets. We assume the secret, represented as a function mapping the input data distribution to the secret quantity, is known. Using this information, we aim to achieve better privacy-utility tradeoffs than secret-agnostic metrics.

(3) Composition and Post-processing: Several privacy metrics satisfy post-processing, including differential privacy and its variants [12], [14], [23], [28], [29], as well as maximal leakage and its variants [9], [16], [25]–[27]. An adaptive composition property with an additive form is known to hold for differential privacy and its variants [12], [14], [23], [28], [29], as well as pointwise maximal leakage [16]. Additive composition holds for maximal leakage [9] and (α, β)-leakage [25] only when successive outputs from the mechanism(s) are conditionally independent, conditioned on the input data.

III. Notation and Problem Formulation

A data holder has data drawn from a distribution parameterized by a parameter vector θ. The parameter θ is itself a realization of a random variable $\Theta \in \mathbf{\Theta}$ belonging to a finite set $\mathbf{\Theta}$.[2] $\mathbb{P}_\Theta$ represents the prior distribution of the parameter random variable Θ, and can equivalently be viewed as the prior over the input data. We use ν to denote distribution measures.

The data holder aims to protect a secret $g = \mathfrak{g}(\theta)$ where $\mathfrak{g}$ is a function that is fixed and known. g is a realization of random variable $G \in \mathcal{G} \triangleq \{g_1, g_2, \cdots, g_s\}$ (i.e., the secret can take s values). We use $\mathbf{\Theta}_g$ to represent the set of original parameters whose secret values are g, i.e., $\mathbf{\Theta}_g = \{\theta \in \mathbf{\Theta} | \mathfrak{g}(\theta) = g\}$. In §V, we use $\mathbf{\Theta}_i$ to represent $\mathbf{\Theta}_{g_i}, \forall i \in [s]$, for convenience.

The data holder releases data via a data release mechanism $\mathcal{M} = \mathbb{P}_{\Theta'|\Theta}$, which maps input parameter θ to a (possibly random) output parameter $\Theta' \in \mathbf{\Theta}'$ (in general, $\mathbf{\Theta} \neq \mathbf{\Theta}'$). We use $\mathcal{M}(\theta)$ to denote the random distribution parameter Θ' output by mechanism $\mathcal{M}$ with input θ. Given θ', the realization of Θ', the attacker outputs a (possibly random) estimate of the secret, $\hat{G}$. We assume the attacker knows the prior distribution of the data and the data release mechanism $\mathcal{M}$, and has infinite computational power. The overall data sharing and attacker guessing process can be formulated as a Markov chain $G - \Theta - \Theta' - \hat{G}$.

Utility Metric: To analyze the privacy-utility tradeoff for a mechanism $\mathcal{M}$, we define the distortion of $\mathcal{M}$ as the expected total variation (TV) distance between the original and released data, represented by X_Θ and $Y_{\Theta'}$ respectively, under the worst-case prior:

$$\Delta_{\mathcal{M}} = \sup_{\mathbb{P}_\Theta} \mathbb{E}_{\Theta, \Theta' = \mathcal{M}(\Theta)}[D_{\text{TV}}\left(\nu_{X_\Theta} \| \nu_{Y_{\Theta'}}\right)],$$

where D_{TV} is the total variation distance. Since our utility metric considers a worst-case prior distribution, the distortion of mechanisms proposed for attribute privacy [12], distribution privacy [13], or distribution inference [14] can reach the upper bound on distortion. Our goal is to achieve non-vacuous distortion bounds while satisfying the desired properties from §I.

IV. Statistic Maximal Leakage

Statistic maximal leakage (SML) measures the largest increase an adversary can gain in their guess of g; it is a property of a data release mechanism $\mathcal{M}$ and a secret mapping $\mathfrak{g}$. We define it as follows:

$$\Pi_{\mathcal{M}, \mathfrak{g}} = \sup_{\mathbb{P}_\Theta, \mathbb{P}_{\hat{G}|\Theta'}} \log \frac{\mathbb{P}\left(\hat{G} = G\right)}{\sup_{g \in \mathcal{G}} \mathbb{P}_G(g)}, \tag{1}$$

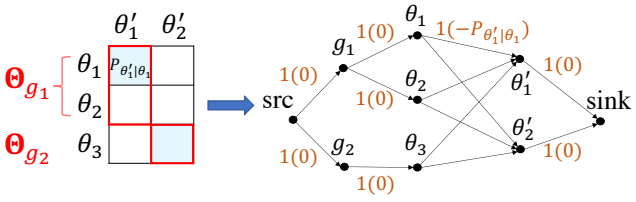
where the supremum is over all prior distributions $\mathbb{P}_\Theta$ over the distribution parameter θ and attack strategies $\mathbb{P}_{\hat{G}|\Theta'}$. The probability in the numerator is over the attacker's randomized estimator, the mechanism, and the secret. Note that the secret function $\mathfrak{g}$ and the data release mechanism $\mathcal{M}$ are fixed in this optimization.

SML bears some similarities with maximal leakage [9] and worst-case min-entropy leakage [8], though they do not simultaneously satisfy all three desirable properties we propose. Maximal leakage $\mathcal{L}_{ML}$ is defined as follows:

$$\mathcal{L}_{ML} = \sup_{\mathbb{P}_{G|\Theta}, \mathbb{P}_{\hat{G}|\Theta'}} \log \frac{\mathbb{P}\left(\hat{G} = G\right)}{\sup_{g \in \mathcal{G}} \mathbb{P}_G(g)}.$$

It requires the prior distribution over input data and assumes the secret function $\mathfrak{g}(\theta)$ is unknown. Worst-case

[2] In practice, the value of the distribution parameter is bounded, and we can only estimate it to within finite precision. Therefore, we model the parameter set as finite, which also allows us to easily apply our techniques to a finite, tabular dataset (§V). We leave the extension to infinite, continuous parameter sets to future work.



Policy Matrix Minimum-Cost Flow Problem

Fig. 1: Given a mechanism $\mathcal{M} = \mathbb{P}_{\Theta'|\Theta}$, the left subfigure shows a policy matrix. For each column j, the red outlined region indicates rows of parameters with secret g maximizing $\mathbb{P}_{\Theta'|\Theta}\left(\theta'_j|\theta_g\right)$. The blue cell lies in the row of θ_g. When the mechanism $\mathcal{M}$ is deterministic, SML calculation can be converted to a min-cost flow problem (right). The constructed directed graph contains three columns of nodes (representing G, Θ, Θ' respectively) between the source and sink nodes. The capacity of all edges are 1, and only the edges between nodes in Θ and Θ' columns have non-zero cost ($-\mathbb{P}_{\Theta'|\Theta}\left(\theta'_k|\theta_j\right)$ between θ_j and θ'_k). Edges are annotated as: Capacity (Cost).

min-entropy leakage $\mathcal{L}_{MEL}$ treats the entire input data distribution as the secret to protect; for the Markov chain $\Theta - \Theta' - \hat{\Theta}$, it is defined as

$$\mathcal{L}_{MEL} = \sup_{\mathbb{P}_\Theta, \mathbb{P}_{\hat{\Theta}|\Theta'}} \log \frac{\mathbb{P}\left(\hat{\Theta} = \Theta\right)}{\sup_{\theta \in \mathbf{\Theta}} \mathbb{P}_\Theta\left(\theta\right)}.$$

Under a fixed prior, maximal leakage can have a smaller value than SML since SML considers the worst-case prior. However, one can construct $\mathcal{M}$ and $\mathfrak{g}$ for which maximal leakage, with a worst-case prior, achieves its largest possible value (i.e., $\min\{\log|\Theta'|, \log|\Theta|\}$) while SML is 0. The following property shows that SML is upper- and lower-bounded by both maximal leakage with a worst-case prior and worst-case min-entropy leakage, by up to an additive factor that depends on the secret and the parametric family.

Property 1 (Relation to Maximal Leakage and Min-Entropy Leakage)**.** *Statistic maximal leakage* $\Pi_{\mathcal{M},\mathfrak{g}}$ *satifies*

$$\mathcal{L}_{MEL} - \sup_{g\in\mathcal{G}} \log|\mathbf{\Theta}_g| \;\leq\; \Pi_{\mathcal{M},\mathfrak{g}} \;\leq\; \mathcal{L}_{MEL},$$
$$\sup_{\mathbb{P}_\Theta} \mathcal{L}_{ML} - \sup_{g\in\mathcal{G}} \log|\mathbf{\Theta}_g| \;\leq\; \Pi_{\mathcal{M},\mathfrak{g}} \;\leq\; \sup_{\mathbb{P}_\Theta} \mathcal{L}_{ML}.$$

A. Computation

Computing SML is more convenient under an alternative form, which shows that we only need to search over a restricted class of priors that assign nonzero probability mass to at most one parameter $\theta \in \mathbf{\Theta}_g$, for each secret $g \in \mathcal{G}$. Under such a prior, for a fixed $g \in \mathcal{G}$, there is at most one $\theta \in \mathbf{\Theta}$ such that $\mathbb{P}_{\Theta|G}\left(\theta|g\right) > 0$; we use θ_g to denote this value.

Proposition 1. *Statistic maximal leakage satisfies*

$$\Pi_{\mathcal{M},\mathfrak{g}} = \sup_{\mathbb{P}_{\Theta|G}\in\{0,1\}} \log \sum_{\theta'\in\mathbf{\Theta}'} \sup_{g\in\mathcal{G}} \mathbb{P}_{\Theta'|\Theta}\left(\theta'|\theta_g\right).$$

Based on Prop. 1, we explain how to compute SML. For concreteness, Fig. 1 (left) illustrates an example with $\mathbf{\Theta} = \{\theta_1, \theta_2, \theta_3\}, \mathbf{\Theta}' = \{\theta'_1, \theta'_2\}$, and $\mathbf{\Theta}_{g_1} = \{\theta_1, \theta_2\}, \mathbf{\Theta}_{g_2} = \{\theta_3\}$. Given a mechanism $\mathcal{M} = \mathbb{P}_{\Theta'|\Theta}$, we can construct a policy matrix where the value in the i-th row and j-th column is $\mathbb{P}_{\Theta'|\Theta}\left(\theta'_j|\theta_i\right)$. First, fix a prior $\mathbb{P}_\Theta$ such that $\mathbb{P}_{\Theta|G} \in \{0,1\}$. Fig. 1 illustrates a case where the prior satisfies $\theta_{g_1} = \theta_1, \theta_{g_2} = \theta_3$ (θ_g is defined above Prop. 1). Next, fix a column θ'_j in the policy matrix. We can now find a secret value $\tilde{g} \in \mathcal{G}$ that maximizes $\mathbb{P}_{\Theta'|\Theta}\left(\theta'_j|\theta_{\tilde{g}}\right)$—i.e., $\tilde{g}$ is the maximum likelihood secret for an observed output θ'_j. For each column, the red outline denotes the input parameters in $\mathbf{\Theta}_{\tilde{g}}$. Our example mechanism satisfies $\arg\sup_{\tilde{g}} \mathbb{P}_{\Theta'|\Theta}\left(\theta'_1|\theta_{\tilde{g}}\right) = g_1, \arg\sup_{\tilde{g}} \mathbb{P}_{\Theta'|\Theta}\left(\theta'_2|\theta_{\tilde{g}}\right) = g_2$. For each column, the blue square is the intersection of the red region with the row of $\theta_{\tilde{g}}$. We finally sum the likelihoods of all the blue squares. Our goal is to find the worst-case prior and calculate the maximum value of $\log\sum_{\theta'\in\mathbf{\Theta}'} \sup_{\tilde{g}\in\mathcal{G}} \mathbb{P}_{\Theta'|\Theta}\left(\theta'|\theta_{\tilde{g}}\right)$. Worst-case, this can be done in time exponential in the number of input parameters $|\mathbf{\Theta}|$ by enumerating all feasible $P_{\Theta|G}$.

Computation via Min-Cost Flow: When the mechanism $\mathcal{M}$ is *deterministic*, i.e., $\mathbb{P}_{\Theta'|\Theta} \in \{0,1\}$, SML calculation process can be converted to a min-cost flow problem [30]. Given a directed graph where each edge is assigned a capacity and a cost, the min-cost flow problem aims to design a network flow satisfying the capacity constraint of each edge, while achieving the minimum cost. The final cost of the min-cost flow has a one-to-one correspondence to the SML of the underlying problem.

To construct the network, we start with a source and a sink node, and create three columns of nodes between them. The first G-column contains all potential secret values (g_1, g_2 in Fig. 1). The capacity of the edge between the source and each node in the G-column is 1 and the cost is 0. The second Θ-column contains all possible input parameter values (θ_1, θ_2, θ_3 in Fig. 1). There is an edge between node g_i and θ_j iff $\theta_j \in \mathbf{\Theta}_{g_i}$. The capacity of each edge is 1 and the cost is 0. The third Θ'-column contains all possible released parameter values (θ'_1, θ'_2 in Fig. 1). This column is fully connected with the second column. The capacity of the edge between θ_j and θ'_k is 1 and the cost is $-\mathbb{P}_{\Theta'|\Theta}\left(\theta'_k|\theta_j\right)$.

From Prop. 1, we know that among all the distributions we are optimizing over, there is only one θ_g satisfying $\mathbb{P}_{\Theta|G}\left(\theta_g|g\right) = 1 > 0, \forall g \in \mathcal{G}$. For any *deterministic* mechanism, there is only one $\theta' \in \mathbf{\Theta}'$ satisfying $\mathbb{P}_{\Theta'|\Theta}\left(\theta'|\theta\right) = 1 > 0, \forall\theta \in \mathbf{\Theta}$. Finally, for each $\theta' \in \Theta'$, we can only select one $\tilde{g} \in \mathcal{G}$ to calculate $\mathbb{P}_{\Theta'|\Theta}\left(\theta'|\theta_{\tilde{g}}\right)$ for the SML calculation, based on Prop. 1. Hence, we set the capacity of all edges as 1. It is known that there exists a min-cost network flow such that the flow of each edge is either 1 or 0 [30]. In that case, for all nodes $\theta \in \mathbf{\Theta}_g$ in Θ column, only one can accept one unit of flow from g, and this node is θ_g. For each node in the Θ' column, it can also only accept one unit of flow from $\theta_g, \forall g \in \mathcal{G}$. Therefore, the min-cost flow problem under our constructed network shares the same objective as Prop. 1. Importantly, the min-cost flow problem can be solved efficiently in polynomial time in $|\mathbf{\Theta}|\cdot|\mathbf{\Theta}'|$ [30].

In our example, we allocate 1 unit of flow from the source to each of g_1 and g_2. For g_1, the full flow goes either to θ_1 or θ_2; the selected node is dubbed θ_{g_1} under $\mathbb{P}_{\Theta'|\Theta}$. The

flows from θ_{g_1} and θ_3 then go to θ'_1 and θ'_2, respectively. This is because θ'_1 has g_1 as its ML secret, and θ'_2 has g_2 as its ML secret under $\mathbb{P}_{\Theta'|\Theta}$. Finally, the flows merge to the sink. The log of the negative total cost of the flow is the SML.

B. Properties of Statistic Maximal Leakage

We show that SML satisfies two natural desired properties: adaptive composition and post-processing. Adaptive composition bounds the total leakage of releasing multiple results from one or more possibly adaptive mechanisms applied sequentially over the same data.

Theorem 1 (Adaptive Composition)**.** *Suppose a data holder sequentially applies m mechanisms $\mathcal{M}_1, \ldots, \mathcal{M}_m$, where $\forall i \in [m]$, the ith mechanism is a function of the input data θ and all of the previous outputs, which we denote as $\theta'^{(1)}, \ldots \theta'^{(i)}$. That is, $\mathcal{M}_i(\theta, \theta'^{(1)}, \ldots, \theta'^{(i-1)}) = \theta'^{(i)}$. Suppose $\forall i \in [m]$, mechanism $\mathcal{M}_i$ satisfies a statistic maximal leakage guarantee with respect to $\mathfrak{g}$ of $\Pi_{\mathcal{M}_i,\mathfrak{g}}$. Let $\boldsymbol{\mathcal{M}} = \mathcal{M}_1 \circ \mathcal{M}_2 \circ \ldots \circ \mathcal{M}_m$ denote the composition of these adaptively chosen mechanisms. The SML with respect to an adversary that can see all intermediate outputs $\theta'^{(1)}, \theta'^{(2)}, \ldots, \theta'^{(m)}$ can be bounded as $\Pi_{\boldsymbol{\mathcal{M}},\mathfrak{g}} \leq \sum_{i\in[m]} \Pi_{\mathcal{M}_i,\mathfrak{g}}$.*

Thm. 1 shows that statistic maximal leakage degrades additively when one or more possibly adaptive mechanisms are applied multiple times sequentially to the original dataset. This additive result is similar in form to analogous composition results for other privacy metrics, including pointwise maximal leakage [16, Thm. 13] and differential privacy [31, Thm. III.1]. In particular, we note that the result for pointwise maximal leakage does not imply ours, nor vice versa. Pointwise maximal leakage requires knowledge of the prior distribution of the input data and assumes the secret function is unknown; the composition property for the worst-case secret function or prior does not imply the composition property for an arbitrary secret or prior.

Theorem 2 (Post-Processing)**.** *Let $\mathcal{M}$ be a data release mechanism whose SML is $\Pi_{\mathcal{M},\mathfrak{g}}$. Let $\widetilde{\mathcal{M}}$ be an arbitrary (possibly randomized) mechanism defined by $\mathbb{P}_{\Theta''|\Theta'}$. Then the SML of $\widetilde{\mathcal{M}} \circ \mathcal{M}$ is $\Pi_{\widetilde{\mathcal{M}}\circ\mathcal{M},\mathfrak{g}} \leq \Pi_{\mathcal{M},\mathfrak{g}}$.*

Thm. 2 shows that applying an arbitrary (possibly randomized) mechanism to the output of a mechanism that satisfies statistic maximal leakage will not degrade statistic maximal leakage.

V. Mechanism Design for Tabular Data

We next study two natural mechanisms for releasing *tabular* data under SML: randomized response and quantization mechanism. Our goal is to understand (a) if each of these satisfies a SML guarantee, and (b) if so, which one has a better privacy-utility tradeoff?

The data holder holds a tabular dataset $\mathcal{D}$ with n rows and c columns, i.e., n samples with c attributes for each sample. Let Γ be the set of combinations of attributes for samples existing in the original dataset $\mathcal{D}$, where $d \triangleq |\Gamma|$. For example, suppose our dataset has binary columns "Above age 18?" and "Registered to vote in the U.S."? and only includes samples with attributes "(Yes, Yes)" and "(Yes, No)"; then $d = 2$.

We assume there is some unknown true feasible set of attribute combinations Γ^*, where $d^* \triangleq |\Gamma^*|$ and $\Gamma \subseteq \Gamma^*$. In our example, $d^* = 3$ because in the U.S., voters must be at least 21 years of age. We use $\hat{\Gamma}^*$ to denote the data holder's estimate of the true support Γ^* (e.g., from public data), where $\hat{\Gamma}^* \subseteq \Gamma^*$[3] and $\hat{d}^* \triangleq \left|\hat{\Gamma}^*\right|$. In our voting example, Γ^* can be accurately estimated based on public information, i.e., $\hat{\Gamma}^* = \Gamma^*$ The data release mechanism is designed such that samples with attribute combinations in $\Gamma \cup \hat{\Gamma}^*$ may exist in the released dataset $\mathcal{D}'$. We illustrate the relation between Γ, Γ^*, and Γ_0 via a venn diagram in Fig. 2. Suppose the released dataset $\mathcal{D}'$ has the same size as the original dataset

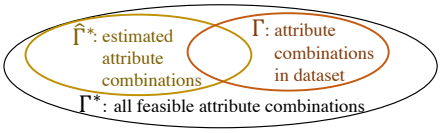


Fig. 2: Relation between Γ, Γ^*, and Γ_0.

Suppose the set of secret values admits a total ordering without loss of generality that $g_1 < g_2 < \cdots < g_s$. For convenience, we use $\boldsymbol{\Theta}_i$ to represent $\boldsymbol{\Theta}_{g_i}$, the parameter set with secret value g_i, $\forall i \in [s]$.

Randomized Response (RR): RR is a widely-used mechanism in the DP literature for discrete distributions [19]. We first consider a RR mechanism $\mathcal{M}_{\text{RR}}$ that outputs the original distribution parameter with some probability, and otherwise releases a different distribution's parameters uniformly at random over the parameter set. Specifically, $\mathcal{M}_{\text{RR}}$ can be written as follows, $\forall \theta \in \boldsymbol{\Theta}$:

$$\mathbb{P}\left(\mathcal{M}_{\text{RR}}\left(\theta\right) = \theta'\right) = \begin{cases} \frac{e^\epsilon}{|\boldsymbol{\Theta}'|+e^\epsilon-1}, & \theta' = \theta, \\ \frac{1}{|\boldsymbol{\Theta}'|+e^\epsilon-1}, & \theta' \in \boldsymbol{\Theta}' \setminus \{\theta\}. \end{cases}$$

Quantization Mechanism (QM): QM has also been adopted in privacy-preserving mechanism design [9], [10], [17]. This mechanism partitions the secret values into subsets of size I and uniformly releases a distribution parameter with the secret as the median index of the corresponding bin. Specifically, $\mathcal{M}_{\text{QM}}$ can be written as $\forall k \in \left\{0, 1, \cdots, \left\lceil \frac{s}{I} \right\rceil - 1\right\}, j \in [I], \theta \in \boldsymbol{\Theta}_{kI+j}$:

$$\mathcal{M}_{\text{QM}}\left(\theta\right) \sim \text{Unif}\left(\boldsymbol{\Theta}'_{\text{R}(k)}\right),$$

where $\text{R}\left(k\right) = \left\lfloor \left(k + \frac{1}{2}\right) I \right\rfloor + 1$, and $\boldsymbol{\Theta}'_{\text{R}(k)}$ represents the released parameter set with secret value $g_{\text{R}(k)}$.

When $\hat{\Gamma}^* = \Gamma^*$, i.e., the data holder knows the whole feasible attribute combination set, we analyze and compare the privacy-distortion tradeoffs between RR and QM as follows. Note that in our analysis of QM, we constrain the secret function $\mathfrak{g}$ to be the PMF value for a specific category (e.g., "the fraction of white males of age 32") to simplify the analysis.

[3]We leave the extension to $\hat{\Gamma}^* \not\subset \Gamma^*$ to future work.

Theorem 3. (Privacy and Distortion of Randomized Response) *For any secret function* $\mathfrak{g}$*, the SML and distortion of RR are:*

$$\Pi_{\mathcal{M}_{RR},\mathfrak{g}} = \log \frac{1+sr}{1+r}, \quad \Delta_{\mathcal{M}_{RR}} = \frac{2\left(\hat{d}^* - 1\right)}{\hat{d}^*\left(1+r\right)}.$$

where $r \triangleq \frac{e^\epsilon - 1}{|\Theta'|} = \frac{e^\epsilon - 1}{\binom{n+\hat{d}^*-1}{\hat{d}^*-1}}$.

(Privacy and Distortion of Quantization Mechanism) *The privacy of QM is* $\Pi_{\mathcal{M}_{QM},\mathfrak{g}} = \log \left\lceil \frac{s}{I} \right\rceil$*. When secret is the fraction of a category, the distortion of QM is*

$$\Delta_{\mathcal{M}_{QM}} = 1 + \frac{\hat{d}^* \left\lfloor \frac{I}{2} \right\rfloor - n}{n\left(\hat{d}^* - 1\right)}.$$

(Mechanism Comparison) *When secret is the fraction of a category, for any non-trivial privacy budget* $T < \log s$*, if* $\Pi_{\mathcal{M}_{QM},\mathfrak{g}} = \Pi_{\mathcal{M}_{RR},\mathfrak{g}} \leq T$*, we have* $\lim_{n\to\infty} \frac{\Delta_{\mathcal{M}_{RR}}}{\Delta_{\mathcal{M}_{QM}}} \geq 1$.

From Thm. 3, we know that when the number of samples is large enough, QM *performs at least as well as* RR as long as SML does not achieve its upper bound $\log s$. Intuitively, this is because the output space of RR covers the full support of Θ', while the output space of QM is significantly reduced.

When $\hat{\Gamma}^* \neq \Gamma^*$, i.e., the data holder only partially knows the feasible attribute combination set, we provide the robustness result for the privacy of the mechanisms.

Definition 1 (Robustness to support mismatch)**.** *Consider a tabular dataset* $\mathcal{D}$ *with attribute combination set* Γ*. For any mechanism* $\mathcal{M}$*, let* $\Pi_{\mathcal{M},\mathfrak{g}}$ *be the SML of* $\mathcal{M}$ *if its released dataset only contains samples with attribute combinations in* $\hat{\Gamma}^* \cup \Gamma$*, and* $\Pi^*_{\mathcal{M},\mathfrak{g}}$ *be the SML of* $\mathcal{M}$ *if its released dataset contains samples with attribute combinations in* Γ^**. The mechanism* $\mathcal{M}$ *is* r*-robust if for any* Γ*,* $\Pi_{\mathcal{M}} - \Pi^*_{\mathcal{M}} \leq r\left(d^* - \hat{d}^*\right)$.

Proposition 2. *Consider a dataset* $\mathcal{D}$ *with* n *samples. RR is* $\log 3$*-robust if its hyperparameter* ϵ *satisfies* $e^\epsilon - 1 \leq \binom{n+\hat{d}^*-1}{\hat{d}^*-1}/s$*. QM with any interval length* I *is* 1*-robust when the secret is the fraction of a category.*

Prop. 2 indicates that RR is robust to support mismatch when it satisfies certain privacy constraints, and QM is robust under certain secret types.

VI. Empirical Evaluation

We conduct an empirical evaluation on the Census Income dataset [32], which collects information from 48842 individuals about their income, education level, age, gender, and more. The dataset contains 22,381 unique attribute combinations, which we assume to be the whole feasible attribute combination set Γ^*.

We first consider the secret as the fraction of an arbitrary category, e.g., the fraction of white males of age 32, and compare the privacy-utility tradeoffs between RR and QM in Fig. 3, where each point for QM represents a realization of the mechanism with an integer-valued quantization interval, whereas ϵ in RR can be real-valued. The SML for this problem takes values in $[0, 16]$, where the upper bound arises from the limited size of the parameter sets. From Fig. 3, we

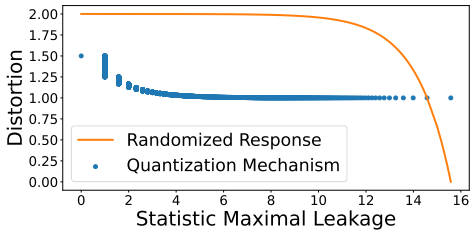


Fig. 3: Privacy-utility trade-offs of RR and QM when the secret is the fraction of an arbitrary category.

observe that for the same privacy guarantee, the distortion of RR is almost twice as large as that of QM under most SML levels, indicating better performance of QM, which is in line with the theoretical insight in Thm. 3.

We next set our secret as the difference between the proportion of white and non-white high-income people (>$50k/yr) within their own race groups, and evaluate the quantization mechanism. We analyze downstream utility by training a random forest classifier on the released data to predict whether an individual has high income. We varied the quantization set size I to achieve different levels of SML from 0 to 2. For each SML level, we conduct the experiment 20 times with independent mechanism outputs, show the averaged ROC curve of the random forests trained on corresponding released datasets. We then compare this with the performance of the random forest trained on the original dataset in Fig. 4. We observe that as SML increases (weaker privacy), AUC (area under the ROC curve) increases, indicating the improvement of the downstream task utility. When SML is as little as 2, the AUC is close to its upper bound (0.89 in raw dataset). When SML is 0 (perfect privacy), the utility drops to 0.75 AUC. Note that a perfect privacy-preserving mechanism can still achieve high utility on this task since we only aim to protect a secret of the dataset, rather than the whole data. These results suggest the promise of QM as a practical data release tool.

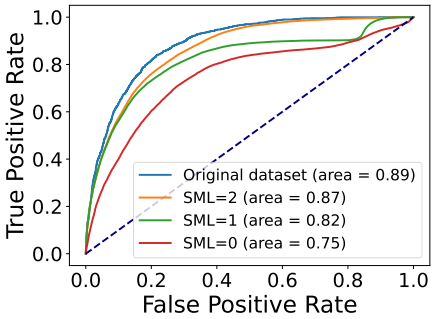


Fig. 4: Comparison of ROC curves of random forests under QM with different SML.

Acknowledgments

The authors gratefully acknowledge the support of NSF grants CIF-1705007, CCF-2338772, and RINGS2148359, as well as support from the Sloan Foundation, Intel, J.P. Morgan Chase, Siemens, Bosch, and Cisco. This material is based upon work supported by the U.S. Army Research Office and the U.S. Army Futures Command under Contract No. W911NF20D0002.

REFERENCES

[1] T. Cho, J.-H. Kim, H.-J. Cho, S.-H. Seo, and S. Kim, "Vulnerabilities of android data sharing and malicious application to leaking private information," in *2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2013, pp. 37–42.

[2] H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 918–923, 1983.

[3] G. Smith, "On the foundations of quantitative information flow," in *FoSSaCS*. Springer, 2009, pp. 288–302.

[4] S. A. Mario, K. Chatzikokolakis, C. Palamidessi, and G. Smith, "Measuring information leakage using generalized gain functions," in *CSF*. IEEE, 2012.

[5] H. Wang, L. Vo, F. P. Calmon, M. Médard, K. R. Duffy, and M. Varia, "Privacy with estimation guarantees," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8025–8042, 2019.

[6] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard, "From the information bottleneck to the privacy funnel," in *IEEE ITW 2014*. IEEE, 2014, pp. 501–505.

[7] A. Zamani, T. J. Oechtering, and M. Skoglund, "Bounds for privacy-utility trade-off with non-zero leakage," in *IEEE ISIT 2022*. IEEE, 2022, pp. 620–625.

[8] C. Braun, K. Chatzikokolakis, and C. Palamidessi, "Quantitative notions of leakage for one-try attacks," *Electronic Notes in Theoretical Computer Science*, vol. 249, pp. 75–91, 2009.

[9] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1625–1657, 2019.

[10] Z. Lin, S. Wang, V. Sekar, and G. Fanti, "Summary statistic privacy in data sharing," *arXiv preprint arXiv:2303.02014*, 2023.

[11] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *ACM CCS 2016*, 2016, pp. 308–318.

[12] W. Zhang, O. Ohrimenko, and R. Cummings, "Attribute privacy: Framework and mechanisms," in *FACCT*. ACM, 2022.

[13] Y. Kawamoto and T. Murakami, "Local obfuscation mechanisms for hiding probability distributions," in *Computer Security–ESORICS 2019, Luxembourg, September 23–27, 2019, Proceedings, Part I 24*. Springer, 2019, pp. 128–148.

[14] A. Suri and D. Evans, "Formalizing and estimating distribution inference risks," *arXiv preprint arXiv:2109.06024*, 2021.

[15] S. Wang, R. Wei, M. Ghassemi, E. Kreacic, and V. K. Potluru, "Guarding multiple secrets: Enhanced summary statistic privacy for data sharing," in *Privacy Regulation and Protection in Machine Learning*.

[16] S. Saeidian, G. Cervia, T. J. Oechtering, and M. Skoglund, "Pointwise maximal leakage," in *IEEE ISIT 2022*. IEEE, 2022, pp. 626–631.

[17] F. Farokhi, "Development and analysis of deterministic privacy-preserving policies using non-stochastic information theory," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2567–2576, 2019.

[18] A. Chaudhuri and R. Mukerjee, *Randomized response: Theory and techniques*. Routledge, 2020.

[19] P. Kairouz, S. Oh, and P. Viswanath, "Extremal mechanisms for local differential privacy," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 492–542, 2016.

[20] Y. Wang, X. Wu, and D. Hu, "Using randomized response for differential privacy preserving data collection." in *EDBT/ICDT Workshops*, vol. 1558, 2016, pp. 0090–6778.

[21] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Privacy-aware guessing efficiency," in *ISIT*. IEEE, 2017.

[22] B. Rassouli and D. Gündüz, "Optimal utility-privacy trade-off with total variation distance as a privacy measure," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 594–603, 2019.

[23] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer, 2006, pp. 265–284.

[24] A. Suri, Y. Lu, Y. Chen, and D. Evans, "Dissecting distribution inference," in *First IEEE Conference on Secure and Trustworthy Machine Learning*, 2023.

[25] A. Gilani, G. R. Kurri, O. Kosut, and L. Sankar, "(α, β)-leakage: A unified privacy leakage measure," 2023.

[26] J. Liao, O. Kosut, L. Sankar, and F. du Pin Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8043–8066, 2019.

[27] G. R. Kurri, L. Sankar, and O. Kosut, "An operational approach to information leakage via generalized gain functions," 2022.

[28] A. Ghosh and R. Kleinberg, "Inferential privacy guarantees for differentially private mechanisms," in *ITCS*, 2017.

[29] M. Chen and O. Ohrimenko, "Protecting global properties of datasets with distribution privacy mechanisms," in *AISTATS*. PMLR, 2023.

[30] L. R. Ford, "Flows in networks," 2015.

[31] C. Dwork, G. N. Rothblum, and S. Vadhan, "Boosting and differential privacy," in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. IEEE, 2010, pp. 51–60.

[32] B. Becker and R. Kohavi, "Adult," UCI Machine Learning Repository, 1996, DOI: https://doi.org/10.24432/C5XW20.