Differential Privacy: Toward a Better Tuning of the Privacy Budget (ε) Based on Risk

Mahboobeh Dorafshanian[a] and Mohamed Mejri[b]
Department of Computer Science and Software Engineering, Laval University, QC, Canada

Keywords: Differential Privacy, Risk, Data Privacy, Security, Big Data, Privacy Budget, Risk of Data Disclosure.

Abstract: Companies have key concerns about privacy issues when dealing with big data. Many studies show that privacy preservation models such as Anonymization, k-Anonymity, l-Diversity, and t-Closeness failed in many cases. Differential Privacy techniques can address these issues by adding a random value (noise) to the query result or databases rather than releasing raw data. Measuring the value of this noise (ε) is a controversial topic that is difficult for managers to understand. To the best of our knowledge, a small number of works calculate the value of ε. To this end, this paper provides an upper bound for the privacy budget ε based on a given risk threshold when the Laplace noise is used. The risk is defined as the probability of leaking private information multiplied by the impact of this disclosure. Estimating the impact is a great challenge as well as measuring the privacy budget. This paper shows how databases like UT CID ITAP could be very useful to estimate these kinds of impacts.

1 INTRODUCTION

With the increasing expansion of databases, the importance of protecting the personal information of individuals has received more attention. Humans have always tended to preserve their information. They like to set measures to limit undesirable access to their data. Tech companies are aimed to collect and analyze information about their customers which can provide high-quality services. This information can be used in a wide variety of domains, such as statistics (Azencott, 2018; Kim et al., 2018; Subramanian, 2022), learning (Berger and Cho, 2019; Jiang et al., 2018), economics (Dankar and Badji, 2017; Hawes, 2020), etc. (Johnson et al., 2018; Winslett et al., 2012). In fact, this is a delicate position for tech companies to collect and use customers' data while maintaining their privacy. With the California Consumer Privacy Act (CCPA) (Goldman, 2020) effective on January 1, 2020, and General Data Protection Regulation (GDPR) (Regulation, 2018) applied in the EU from May 2018, there is a compelling demand to provide rigorous privacy guarantees for users when analyzing and collecting their usage data. Moreover, many governments set strict policies about how tech companies can collect and share user data. Companies that do not follow these policies can face huge fines. For example, a Belgian court (Gibbs, 2018) in 2018 ordered Facebook to stop collecting data on users' browsing habits on external websites, or face fines of €250,000 a day or up to 100 million euros.

Nowadays, many multinational companies who operate in different areas, like Apple (Greenberg, 2016), Google (GoogleDP, 2018) or US census bureau (Abowd, 2018) have begun to use differentially private algorithms to collect behavioral statistics from their users. In 2016, Apple announced that it would use Differential Privacy algorithms in the iPhone. Google also tries to bring Differential Privacy into practice, as implemented a feature in Chrome that collects behavioral statistics from Chrome browsers. We can find other practical examples in Privitar. These products enable companies to perform meaningful analyses on sensitive data while providing privacy guarantees to their users.

Research Question. This paper aims to answer the following questions:

1. For a query q that has an impact I on privacy disclosure, how to fix the value of ε, so that the risk will be lower than a threshold value R_T?
2. For n queries $q_1, \ldots, q_n$ that have impacts, respectively, $I_1, \ldots, I_n$ on privacy disclosure, how to fix

[a] https://orcid.org/0000-0003-1064-5024
[b] https://orcid.org/0000-0003-4820-3176

Dorafshanian, M. and Mejri, M.
Differential Privacy: Toward a Better Tuning of the Privacy Budget () Based on Risk.
DOI: 10.5220/0011896600003405
In *Proceedings of the 9th International Conference on Information Systems Security and Privacy (ICISSP 2023)*, pages 783-792
ISBN: 978-989-758-624-8; ISSN: 2184-4356

the value of ε so that the global risk will be lower than a threshold value R_T?

3. How can we estimate the impact of the privacy disclosure related to a query q in the real world?

Outline. The remaining part of this paper is structured as follows: Section 2 gives some preliminaries useful notations, definitions, and results related to Differential Privacy. Section 3 gives an upper bound of privacy budget based on risk and answers questions 1 and 2. Section 4 answers to question 3 by showing how the impact of data leaking could be estimated. Section 5 gives a literature review and Section 6 concludes the paper and gives some perspectives.

2 PRELIMINARIES

2.1 Formal Differential Privacy

According to Cynthia Dwork's book (Dwork and Roth, 2014):

Definition 1. *(Differential Privacy.) A randomized algorithm $\mathcal{M}$ with domain $\mathbb{N}^{|\mathcal{X}|}$ is (ε,δ)–differential private if for all $\mathcal{S} \subseteq Range(\mathcal{M})$ and for all $x, y \in \mathbb{N}^{|\mathcal{X}|}$ such that $\|x-y\|_1 \leq 1$:*

$$Pr[\mathcal{M}(x) \in \mathcal{S}] \leq exp(\varepsilon) \times Pr[\mathcal{M}(y) \in \mathcal{S}] + \delta$$

This formal definition of Differential Privacy (DP) will guarantee that the manner of the randomized algorithms on similar input databases is likely the same. For well understanding Definition 1, we explain the important notations as follow: $\mathcal{M}$ is a privacy mechanism (probabilistic), $\mathcal{X}$ is a universe set of data types, $\mathbb{N}$ is the set of all non-negative integers, $x \in \mathbb{N}^{|\mathcal{X}|}$ is a dataset in the possible datasets (we also use D, D', y as datasets), $\mathcal{M}(x)$ is the distribution of the outputs of $\mathcal{M}$ given input x, $Range(\mathcal{M})$ is the set of possible outputs of the mechanism, $\mathcal{S} \subseteq Range(\mathcal{M})$ is the subset of possible outputs, ε is the maximum distance between the result of a query on database (x) and database (y), δ is the probability of data leakage and $\|x-y\|_1$ is a L_1 norm which measures how many records are different between x and y.

2.2 Types of "Query Sensitivity" in Differential Privacy

One important parameter that determines how much noise we should add to the ε-differential privacy mechanism is "sensitivity". The sensitivity is the measure to determine how much the outputs would change if we change one entry of data sets. Actually, sensitivity parameterized the amount of noise that is required for the Differential Privacy mechanism. It refers to the impact of changing at most one element on the result of a query.

2.2.1 L_1 and L_2 Norms

The L_1 norm is defined as the sum of the vector's elements. For a vector $V = (V_1, \ldots, V_k)$, we have: $\|V\|_1 = \sum_{i=1}^{k} |V_i|$

For example in two-dimensional space, we call it "Manhattan distance" which is the L_1 norm of the difference between two vectors. $|V_i|$ is an absolute value of the vector's element. If we consider V as a database, $\|V\|_1$ is a measure of the size of the database, that is the number of records it contains. Subsequently, the L1 distance between two datasets V and Z is $\|V-Z\|_1$ and it is a measure to know how many records differ between V and Z. The L_2 norm is defined as the square root of the sum of squares. For a vector $V = (V_1, \ldots, V_k)$, we have:

$$\|V\|_2 = \sqrt{\sum_{i=1}^{k} V_i^2}$$

For example, in two-dimensional space, we called it "Euclidean distance" which is always less than or equal to the L_1 distance. For large databases, according to this formula, the L_2 sensitivity will be significantly much lower than the L_1 sensitivity. So, in real-world applications such as machine learning algorithms, L_2 sensitivity is obviously lower than L_1 sensitivity. The sensitivity of a query can be defined as follows. In fact, we have two types of sensitivity, namely Global sensitivity and Local sensitivity.

2.2.2 Global Sensitivity

Global sensitivity states that if we change one element of any data set, how much is the maximum difference between the outputs of the query? Subsequently, how much noise we should add to the result to satisfy ε-differential privacy requirements? For two data sets x_A, x_B which are different in at most one record and we apply query q on these two data sets then we have:

$$\Delta q_{GS} = \max_{x_A, x_B \subseteq \mathcal{X}} \|q(x_A) - q(x_B)\|_1$$

L_1-norm $\|.\|_1$ is the distance between query results on two databases that are different in just one record, and max define as the maximum result of $q(x_A) - q(x_B)$ for any data sets x_A and x_B. By this definition, for any two neighboring data sets x_A and x_B, the difference between $q(x_A)$ and $q(x_B)$ is at most Δq_{GS}. It is worth mentioning that global sensitivity is independent of the database and just dependent on the query, due to the fact that it is the max difference between the outputs in view of any neighboring data

sets x_A and x_B. This definition has a significant impact on the utility of some queries. For example, consider the sum query on any data set which has arbitrary entries. In this case, the largest difference between the outputs of any query is infinite because there is no upper bound on any input, so the global sensitivity for the sum query is infinite. To solve this issue, we define bounds for the queries. These bounds limit the data sets just to store values less than a predetermined threshold. Consequently, we modify the data set continuously to guarantee that no value exceeds the threshold. Thus, the global sensitivity is dependent on the query and threshold and it is not infinite anymore. Now we have a better definition of global sensitivity. Global sensitivity would be the minimum sensitivity of the query to cover all possible data sets.

2.2.3 Local Sensitivity

We consider any two adjacent databases in global sensitivity, but in local sensitivity, we fix one of the two databases as an actual dataset being queried and considered all its adjacent datasets. For a dataset x which is queried by function q, the local sensitivity is:

$$\Delta q_{LS} = \max_{x_1} \|q(x_1) - q(x_2)\|_1$$

x_1 and x_2 are two adjacent data sets that differ in at most one record. Here, local sensitivity is the maximum difference that changing one record in x_1 can produce and is the minimum sensitivity that is needed for a query to cover the actual data set x_1. We define local sensitivity measures related to the actual data set's size, enabling us to place finite bounds on the sensitivity of some functions which are difficult to set in global sensitivity. The problem with local sensitivity is that it depends on the dataset so the adversary who knows it, may be able to infer some data about the dataset. So, we need to use some auxiliary parameters with local sensitivity. Moreover, even if the adversary does not know the local sensitivity, by comparing just a few query answers, it is possible to determine the scale of the noise. Here we face the question: which one is better, global or local sensitivity? We have many studies and real-world use cases which use both, but it is important to know that local sensitivity is the minimum sensitivity that is needed for the query to cover one fixed (actual) dataset, while global sensitivity is the minimum sensitivity that is needed for the query to cover all possible adjacent datasets.

2.3 Laplace Mechanism

One of the most popular database queries is numerical queries. In numerical queries, $q : N^{|x|} \rightarrow R^k$, we map the database to k real numbers. Local sensitivity is one of the important parameters which determine how we can accurately answer numerical queries. The local sensitivity determines an upper bound on the noise which we add to the output for preserving privacy. Differential Privacy aims to hide the participation of individuals, so by the local sensitivity, we measure in the worst case how much a single individual's input can influence the output of the dataset.

Definition 2. *(The Laplace Distribution.)(Dwork et al., 2006). The Laplace Distribution with the scale b, is the distribution with probability density function:*

$$Lap(x|b) = \frac{1}{2b} \times exp(-\frac{|x|}{b})$$

The Laplace mechanism uses the noise which is drawn from the Laplace distribution and perturbs each element to compute q. In the Laplace mechanism, noise is scaled to $\frac{1}{\varepsilon}$ which is independent of the size of the database. Actually, the noise is scaled to the [(sensitivity of a (query))/ε], where the sensitivity is equal to the amount that the output of the function will change when its input changes by 1. For instance, the sensitivity of counting queries is always equal to 1.

Definition 3. *(The Laplace Mechanism.)(Dwork et al., 2006). Given any function $q : N^{|x|} \rightarrow R^k$, the Laplace mechanism is define as:*

$$M_L(x, q(.), \varepsilon) = f(x) + (Y_1, \ldots, Y_k)$$

where Y_i are i.i.d random variables drawn from $Lap(\frac{\Delta q}{\varepsilon})$.

By this definition, the Laplace mechanism is (ε,δ)-differential privacy or ε-differentially private, where δ is always equal to 0. The Laplace mechanism is for numeric queries with low sensitivity but it does not have a solution for non-numerical queries.

2.4 UT CID ITAP Dataset

The important question that we faced in the relationship between Differential Privacy and risk is how we can measure the probability and the impact of data disclosure. To address this issue, the Center for Identity at the University of Texas (UT CID) is conducting multi-disciplinary research on critical fraud in the United States. To increase the fundamental understanding of fraud processes, patterns and identity theft, they proposed the risk assessment tool which is called Identity Threat Assessment and Prediction (ITAP)(Zaiss et al., 2019). ITAP collects data on fraud, abuse, and identity theft (from over 6000 identity theft news stories) to investigate many features such as the value of identity attributes, their risk of exposure, and the identified vulnerabilities. The ITAP model finds the most vulnerable identity features to theft, analyzes their importance, and studies

the Personally-Identifying Information (PII) which is more targeted by thieves (more than 50 features about each identity theft incident). It offers identity solutions relevant to financial services, healthcare, consumer services, education, defense, and government.

2.4.1 UT CID Identity Ecosystem

Under the ITAP project, the Identity Ecosystem is developed by the UT Center for Identity (Chang et al., 2021). In fact, the Identity Ecosystem is a Bayesian network representation of a person's identity which analyses how personal identities are built and used in our daily lives. For instance, in the UT CID Identity Ecosystem, we could analyze the security level of an authentication method. By the UT CID Ecosystem, three main real-world queries are answered: 1) the risk of disclosure of a certain PII attribute, 2) the cost/liability of disclosure and 3) the cause of data disclosure. Based on various features in the UT CID Ecosystem, they built the UT CID Identity Ecosystem Graphical User Interface (GUI). With this GUI we can choose the color and size attribute nodes as shown in Figure 1. By this valuable tool, we can analyze the data, model identity theft and abuse, and answer various questions about identity risk and risk management. We will describe more in Section 4.

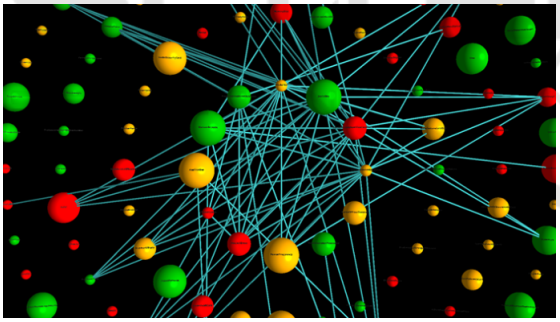

Figure 1: The UT CID Identity Ecosystem. It determines the value and risk of PII attributes. The size of nodes is based on their value and the color of nodes is determined by their risk (Chang et al., 2021).

2.5 Risk in Cybersecurity

It is important to correct our view toward the concept of risk in cybersecurity. Cyber risk generally corresponds to any risk of disruption, financial loss, or damage to the reputation of a company due to the failure of its technology system. It could have happened in a variety of ways, such as premeditated and unauthorized breaches of security to gain access to the information. Moreover, poor system integrity causes operational risks. Defectively managed cyber risk can lead to a variety of attacks which consequently compromise the safety of information of individuals. One way to manage and reduce the probability of cyber risk is to apply Differential Privacy methods. Academic achievements show that by applying Differential Privacy mechanisms, the risk of data disclosure is significantly reduced. We have a few studies which are focused on the relationship of Differential Privacy and risk. Tsou et al. (Tsou et al., 2019) use the simple definition of risk.

Definition 4. *Definition of the Risk:*

Risk = Probability of data disclosure (RoD).

As we see, here the definition of risk is limited to the probability of data disclosure and does not consider other important factors. We will bring the complete definition of risk in Section 3.2.

3 AN UPPER-BOUND FOR ε BASED ON RISK

3.1 The Relationship Between ε and the Risk of Data Disclosure

In (Tsou et al., 2019), Tsou et al. proposed a new method for measuring the relationship between the risk of data disclosure and ε. However, as this was a new work in this domain, they used just a simple definition for the risk (the risk of data disclosure). First, they proposed the definition of Differential Privacy as follows (Tsou et al., 2019):

Definition 5. *(Dwork et al., 2006). A randomized function $\mathcal{M}$ is ε–differentialy private if for any datasets x_1, x_2 with at most one different record and any possible outputs $\mathcal{S} \subseteq Range(\mathcal{M})$,*

$$\frac{Pr[\mathcal{M}(x_1) \in \mathcal{S}]}{Pr[\mathcal{M}(x_2) \in \mathcal{S}]} \leq e^{\varepsilon}$$

where the probability Pr depends on $\mathcal{M}$ randomness. Differential Privacy can be implemented by adding Laplace noise into the output of the query or original dataset to perturb the sensitive data of each record. In the dataset, the maximal effect of a record on the output of a query function is global sensitivity.

Definition 6. *(Tsou et al., 2019). For any query q, the stochastic function $\mathcal{M}$,*

$$\mathcal{M}(D) = q(D) + (Lap_1(\frac{1}{\lambda}), Lap_2(\frac{1}{\lambda}), \ldots, Lap_n(\frac{1}{\lambda}))$$

satisfies ε–differential privacy, that $Lap_i(\frac{1}{\lambda})$ are i.i.d Laplace variables with $\lambda = \frac{\varepsilon}{\Delta q}$. According to definition 4, adding Laplace noise into individuals' records

can guarantee their privacy. If $A(j)$ is the real numerical value the j-th individual's data in the original dataset D, by adding Laplace noise to it: $A'(j) = A(j) + Lap_j(\frac{1}{\lambda})$. Due to the value of the Laplace noise $Lap_j(\frac{1}{\lambda})$, we have the variation in the value of $A'(j)$. So, the actual value $A(j)$ would be in the interval $[A'(j) - Lap_j(\frac{1}{\lambda}), A'(j) + Lap_j(\frac{1}{\lambda})]$. According to the Laplace mechanism (Dwork et al., 2006; Eltoft et al., 2006) the parameter $\frac{1}{\lambda}$ generates the noise which is from -∞ to +∞.

The value of Laplace noise is selected randomly and unboundedly, but we can estimate the maximal value of noise $\max(Lap(\frac{1}{\lambda}))$ by a bounded scale parameter and a confidential probability. The detail of this estimation is described in (Tsou et al., 2019).

Theorem 1. *(Tsou et al., 2019). $Max(Lap(\frac{1}{\lambda}))$ is the maximal value of the generated noise related to the value of* ε *and is equivalent to* $-\frac{\varepsilon \times ln(2-2\gamma)}{\Delta q}$.

Δq is the global sensitivity and γ is the confidence probability to estimate noise distribution. In this work, they focus on integer response queries (i.e., queries that return integers such as count). The proof of this theorem is described in (Tsou et al., 2019). Now we can define the RoD according to the maximum value of the noise.

Definition 7. *ROD (Tsou et al., 2019)*
By applying the Laplace noise, the actual value of $A(j)$ would be hidden in the interval $R = [A'(j) - Lap_j(\frac{1}{\lambda}), A'(j) + Lap_j(\frac{1}{\lambda})]$. So, If there are ξ_j values which fall into R, the RoD for the actual value $A(j)$ is equal to $\frac{1}{\xi_j}$ (Tsou et al., 2019) ($\frac{1}{\xi_j}$ is the estimated probability for the actual value $A(j)$).

3.2 New Definition for Risk

When the threat has happened, we can measure the risk associated with two parameters which are the vulnerabilities of data and the impact of this data disclosure. Consequently, the risk's definition is equal to the probability of data disclosure multiplied the impact of this data disclosure.

Definition 8. *Definition of the Risk:*

Risk = Probability of data disclosure (RoD) × Impact of the data disclosure.

It is important to know the probability of data disclosure and its relationship with the impact of disclosing this data. We know that the information does not have the same value and companies have policies to measure the impact of leaking this information on their assets. Actually, companies' strategy to investigate the privacy budget is dependent on the value of the information. The company can estimate the impact of data disclosure (I) and the managers can set a threshold for this risk (R_T), which means the maximum risk that the company can support. One useful framework to estimate the probability and impact of data disclosure is the Identity Ecosystem. This practical tool gives an estimation of these two values, as well as other interesting attributes (we investigate more in Section 4).

3.3 Risk and Privacy Budget ε

In Differential Privacy mechanisms, the level of privacy is controlled by the parameter ε. But it is not easy to find the appropriate value for ε. In (Tsou et al., 2019), they intuitively formulated ε by using the confidence probability of the noise estimation.

Theorem 2. *(Tsou et al., 2019). If ξ is the number of values in the estimated distribution and the $\max(Lap(\frac{1}{\lambda})) \geq \frac{\xi-1}{2}$, we can formulate* ε *as follow:*

$$\begin{aligned} \max(Lap(\tfrac{1}{\lambda})) = -\tfrac{\varepsilon \times ln(2-2\gamma)}{\Delta q} &\geq \tfrac{\xi-1}{2} \\ \Rightarrow \quad \tfrac{\varepsilon \times ln(2-2\gamma)}{\Delta q} &\leq \tfrac{1-\xi}{2} \\ \Rightarrow \quad \varepsilon &\leq \tfrac{\Delta q(1-\xi)}{2 \times ln(2-2\gamma)} \end{aligned}$$

Here, ε is estimated according to its relationship with the risk of data disclosure (RoD). One of the challenges of Differential Privacy is how to fix the ε. Decision-makers cannot understand the meaning of this important parameter. They usually make their decisions based on the risk that involves the impacts. They may have risk thresholds according to which they decide. Therefore, it will be useful to connect the risk threshold to the security budget. The following theorem connects the privacy budget ε to a risk threshold R_T. More precisely, given a query q that may reveal private information that could have a negative impact I, the theorem gives an upper bound for the privacy budget ε based on I and R_T.

Theorem 3. *Let q be a query and I be the impact of its privacy disclosure. Let R_T be a risk threshold (the maximum risk that the company can tolerate). The privacy budget* ε *with Laplace noise needs to be equal or less than*

$$u \times (1 - \frac{I}{R_T})$$

where $u = \frac{\Delta q}{2 \times ln(2-2\gamma)}$.

Proof. From Theorem 2, we have: $\varepsilon \leq \frac{\Delta q \times (1-\xi)}{2 \times ln(2-2\gamma)}$.
Let $u = \frac{\Delta q}{2 \times ln(2-2\gamma)}$, then: $\varepsilon \leq u \times (1-\xi)$. From Definition 7, $RoD = \frac{1}{\xi}$, it follows that:

$$\varepsilon \leq u \times (1 - \frac{1}{RoD}) \quad (1)$$

R_T is the maximum tolerated risk, $RoD \times I \leq R_T$.
It follows that: $\frac{1}{RoD} \geq \frac{I}{R_T}$. Then, $1 - \frac{1}{RoD} \leq 1 - \frac{I}{R_T}$. Since u is a positive value, $u \times (1 - \frac{1}{RoD}) \leq u \times (1 - \frac{I}{R_T})$. From Equation (1), we have:

$$\varepsilon \leq u \times (1 - \frac{1}{RoD}) \leq u \times (1 - \frac{I}{R_T})$$

And finally, we conclude that:

$$\varepsilon \leq u \times (1 - \frac{I}{R_T})$$

□

This Theorem is for single-dimensional data. Now we generalize the theorem to n queries.

Theorem 4. *Let $q_1, \ldots, q_n$ be n queries and $I_1, \ldots, I_n$ be the impacts of their privacy disclosures, respectively. Let R_T be a risk threshold (the maximum risk that the company can tolerate). The global privacy budget ε with Laplace noise is equal or less than*

$$U - \frac{\sum_{i=1}^{n} u_i \times I_i}{R_T}$$

where $U = \sum_{i=1}^{n} u_i$ and $u_i = \frac{\Delta q_i}{2 \times ln(2-2\gamma)}$.

Proof. Form Theorem 3, we have: $\varepsilon_i \leq u_i \times (1 - \frac{I_i}{R_T}) = u_i - \frac{u_i \times I_i}{R_T}$, where $u_i = \frac{\Delta q_i}{2 \times ln(2-2\gamma)}$. From the Differential Privacy composition theorem (Dwork and Roth, 2014), it follows that:

$$\varepsilon = \sum_{i=1}^{n} \varepsilon_i \leq \sum_{i=1}^{n} u_i - \frac{\sum_{i=1}^{n} u_i \times I_i}{R_T}$$

□

3.4 An Example for Measuring the ε

By our definition in the previous section, now we have a new formula for ε which is $\varepsilon \leq u \times (1 - \frac{I}{R_T})$. Simply, we can calculate the value of ε (that more precisely, it is the upper bound for ε). In Table 1, we show that the value of ε is not fixed randomly. In fact, it depends on R_T and I. For example, assume that the $I = 5$ and the manager fixes the value $R_T = 7$, then $\varepsilon = 0.29$ (in the next section, we will bring more details about how to estimate the impact and risk of data disclosure in real-world). We suppose that $\Delta q = 1$ (the global sensitivity) and according to (Tsou et al., 2019), we choose an appropriate value for γ to have a positive value for u (in Theorem 2, $u \leq 1$ and for simplicity in our calculation, we suppose $u = 1$). Given different values for R_T and I, we see that by the large value for R_T, the ε is close to 1. On the other hand, when I is large, the value of ε is close to 0. Obviously, when R_T and I are equal, ε= 0. Now, the important question is how we can measure the risk of data disclosure and the impact of data disclosure in the real world.

Table 1: An example for measuring the ε.

Impact of data disclosure (I)	Risk threshold (R_T)	ε is equivalent or less than
5	7	0.29
5	5	0
3	6	0.5
0	7	1

4 ESTIMATION OF THE IMPACT OF DATA LEAKING

For measuring the probability of risk of data disclosure and the impact of data disclosure, R. Zaeem et al. have done valuable work (Zaeem et al., 2016). At the University of Texas at Austin, they have designed the Identity Ecosystem. This valuable tool can model identity theft and abuse, analyze the data and consequently answer various questions about identity risk and risk management. The Ecosystem can predict the probability of risk which causes a breach of each Personally Identifiable Information (PII) and calculate a potential monetary value of damage to the PII owner in the situation of identity theft. In the situation that more information is available about the victim or incident, the Ecosystem can update the predicted risk and monetary value according to the risk and value in the real world. They use probabilistic analysis to present the results in the graph-based visualization. As it is shown in Figure 1, in the Ecosystem Graphical User Interface (GUI), nodes are the attributes and edges are the connections between these attributes. The user can use this GUI to interactively play out different scenarios, and graphically see the conclusions about the risk of data disclosure and the potential monetary value of the attributes.

Based on various properties of the attribute such as risk and monetary value, nodes are colored and sized. Figure 1 shows the PII attributes that nodes are colored according to their risk (low-risk attributes colored in green, medium risk in yellow and high risk in red) and are sized according to their monetary value (bigger nodes have the higher monetary value). This GUI can visually show PII attributes, their connections, potential risk, and other interesting values.

In Table 2, we have examples of the sensitivity scores which are assigned to the identity assets. These

scores are assigned according to the prior probability and the monetary loss. For instance, Social Security Number has a higher score among other attributes. Consequently, the risk of disclosing this attribute is more than others.

Table 2: Identity assets and their sensitivity scores.

Identity Asset Name	Prior probability	Loss (USD)	Score
Email Addr.	0.027526	18105024	0.613
Social Security No.	0.096598	27465086	0.938
Passport Info.	0.002565	1252465	0.652
Phone No.	0.017439	4405490	0.605

Related to the risk and managing identity attributes, the Ecosystem can answer three important questions in the situation of disclosing a set of attributes: First, "How does disclosing a set of attributes affect the risk of disclosing other attributes?" The second question is, "What is the source of disclosing data?" And the last question that UT CID Ecosystem can answer is "What is the total cost of disclosing this attribute?" This work is a good example of measuring the risk of data disclosure and the monetary impact of this exposure. More precisely, in the new definition of privacy budget (Theorem 3), we need these two parameters to evaluate the value of noise for the Differential Privacy mechanism.

5 LITERATURE REVIEW

We have many use cases of Differential Privacy techniques in the real world. For example, in the health industry (Azencott, 2018; Kim et al., 2018; Subramanian, 2022), genomics data sharing (Berger and Cho, 2019), location privacy and US census bureau and etc (Abowd, 2018; Hawes, 2020; Jiang et al., 2018; Johnson et al., 2018; Quinton and Reynolds, 2018). For deeply understanding notions about differential privacy, we have an excellent survey and book by Dwork (Dwork et al., 2006; Dwork and Roth, 2014). Dwork and her colleagues proposed several privacy models (Dwork and Lei, 2009; Dwork and Smith, 2010) and discussed many mechanisms.

One of the first researchers who emerges in the field of privacy is done by Adams (Adams, 1999). He conducts three years of research according to users' privacy perceptions of three information multimedia communication environments such as video conferencing, Internet multi-casting, and virtual reality. His research shows that three elements affect the user's perception of privacy: the usage of the information, the level of trust of the user in the information receiver, and the released information sensitivity. In this empirical research, he argues that the risk of data disclosure would relate to the context of the data utilization. Although his valuable research was on real-world cases, he just worked on the risk of data disclosure and did not study the differential privacy concept.

The most detailed discussion on the value of ε and its relationship with RoD is done by Lee and Clifton (Lee and Clifton, 2011). They assume that an attacker has infinite computation power and can obtain arbitrary background knowledge, except for one specific individual. In data set D, there are n rows (n individuals' data) and there is a data set D' which has one less individual, $D' \subseteq D$ and $|D'| = |D| - 1$. The attacker aims to identify a specific individual in D' according to his prior belief on the original data set D. After observing the result, he updates his prior belief depending on whether the outcome was more or less likely if the specific individual had participated. Here, ε controls how much an adversary's belief can change. Subsequently, it is possible to derive a bound on ε in order to keep the adversary's belief below a given threshold. Finally, they obtain posterior belief on D' to calculate the RoD. Although they had new insight into this domain, they just considered the background knowledge of the attackers and did not investigate other attacks such as a linkage attack.

Zhang et al. (Zhang et al., 2022) demonstrate a review and evaluate the open-source differential privacy (DP) tools. They define criteria such as the impact of DP on different functionalities and quantify how different DP tools can be optimally configured to reduce the risk of data disclosure. They propose guidelines to select DP tools according to the user's need and the level of anticipated privacy and utility while working on private data. They openly release their evaluation coding repository, a framework that users can reuse to evaluate privacy tools.

In (Hayes et al., 2022), Hayes et al. propose a framework to compare the adversarial and nominal risk. They use both private and non-private settings in their study. They concentrated risk analysis for robust and private learning to know which parts of differential privacy and adversarial training hurt optimization. Their results show that clipping norm in differential privacy and the size of adversarial perturbation would increase the risk of disclosing data. Nonetheless, they did not apply their new method to real cases.

McClure et al. (McClure and Reiter, 2012) proposed the statistical induction on proportions in synthetic binary data and investigated the relationship between prior beliefs and posterior beliefs for the binary data and synthetic data. They compared the

Table 3: Comparison of Differential Privacy techniques.

Techniques \ Parameters	Privacy budget management	RoD	Impact of revealing information on the risk management
Adams (Adams, 1999)	No	No	No
Lee and Clifton (Lee and Clifton, 2011)	Yes	Yes	No
Dankar and Badgi (Dankar and Badji, 2017)	Yes	No	No
Zhang et al. (Zhang et al., 2022)	Yes	No	No
Hayes et al. (Hayes et al., 2022)	Yes	Yes	No
McClure et al. (McClure and Reiter, 2012)	Yes	Yes	No
Maurizio and Giuseppe (Naldi and D'Acquisto, 2015)	Yes	Yes	No
Yu et al. (Chen et al., 2017)	Yes	No	No
Tsou et al. (Tsou et al., 2019)	Yes	Yes	No
Zaeem et al. (Liau et al., 2019)	Yes	Yes	Yes

prior and posterior probabilities obtained from different levels of ε in an ε-differential privacy private synthesis model. The restriction of this work is the difficulty to extend its analysis for RoD over the one-variable binary/numerical data sets.

Maurizio and Giuseppe (Naldi and D'Acquisto, 2015) defined the RoD in relation to noise pollution. They used a method for choosing ε, which computes how much the actual output of a counting query may be measured from a noise-polluted one. Although their method can be applied to measure the RoD of synthetic data sets, it is limited to counting queries and does not calculate the RoD for a counting query of the joint distribution.

Yu et al. (Chen et al., 2017) proposed an algorithm for choosing an applicable privacy budgetε with a balance between privacy and utility. They used a data-driven algorithm to measure and predict the error of statistical results from the addition of random noise to an original data set. However, they did not investigate in detail the relationship between ε and RoD.

Zaeem et al. (Liau et al., 2019) proposed novel practical research on data privacy, they built a graphical model to represent a complex network for probabilistically dependent data and their correlated random variables and finally performed an inference model. They considered three questions: 1) What is the impact of the exposure risk for the target attributes in correlation to other attributes? 2) What is the most likely source of the exposure of an attribute? and 3) What is the total cost of exposure of an attribute? To answer these questions, they build the Identity Ecosystem based on the Bayesian graph model to answer sophisticated queries such as "how to predict future risk and losses of losing a given set of personal identities".

Table 3, shows a comparison between Differential Privacy techniques according to their relationship with the risk of data disclosure. Our evaluation shows that many studies ignored calculating the privacy budget's value and just used the predetermined value for it. Moreover, just a few works investigate the issue of risk of data disclosure. Nonetheless, they do not consider the full definition of the risk. To the best of our knowledge, just one work evaluated the impact of revealing information on risk management.

6 CONCLUSION

Many companies and institutions are holding huge databases containing private information that could be useful to improve different aspects of human life. However, laws force them to protect their private life. Differential Privacy provides a nice bypass for this restriction. It promises to allow us to take benefits from private information without violating privacy. However, the definition of Differential Privacy is complicated and could not be easily understood by a large part of decision-makers. In particular, the privacy budget is not connected to some metrics with which decision-makers are familiarized, such as risk. This paper gives a theorem providing an upper bound for the privacy budget based on a risk threshold and the impacts of data leaking coming from the involved queries. Another important question addressed by this paper is the evaluation of the impact of data disclosure using the UT CID Identity Ecosystem. We use the Laplace noise in this paper. For our future work, we want to use privacy mechanisms different from the Laplace noises. Moreover, we aim to include the utility (the positive impact) in the new definition of ε. Then we can distribute the privacy budget in a way that we have maximum utility.

ACKNOWLEDGEMENTS

This research is supported by the Beneva Insurance and Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

Abowd, J. M. (2018). Protecting the confidentiality of america's statistics: Adopting modern disclosure avoidance methods at the census bureau. *Census Blogs: Research Matters*.

Adams, A. (1999). The implications of users' multimedia privacy perceptions on communication and information privacy policies. In *Proceedings of Telecommunications Policy Research Conference*, pages 1–23.

Azencott, C.-A. (2018). Machine learning and genomics: precision medicine versus patient privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128):20170350.

Berger, B. and Cho, H. (2019). Emerging technologies towards enhancing privacy in genomic data sharing.

Chang, K. C., Zaeem, R. N., and Barber, K. S. (2021). An identity asset sensitivity model in self-sovereign identities.

Chen, K.-C., Yu, C.-M., Tai, B.-C., Li, S.-C., Tsou, Y.-T., Huang, Y., and Lin, C.-M. (2017). Data-driven approach for evaluating risk of disclosure and utility in differentially private data release. In *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, pages 1130–1137. IEEE.

Dankar, F. K. and Badji, R. (2017). A risk-based framework for biomedical data sharing. *Journal of Biomedical Informatics*, 66:231–240.

Dwork, C. and Lei, J. (2009). Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 371–380.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3876 LNCS:265–284.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407.

Dwork, C. and Smith, A. (2010). Differential privacy for statistics: What we know and what we want to learn. *Journal of Privacy and Confidentiality*, 1(2).

Eltoft, T., Kim, T., and Lee, T.-W. (2006). On the multivariate laplace distribution. *IEEE Signal Processing Letters*, 13(5):300–303.

Gibbs, S. (2018). https://www.theguardian.com/technology.

Goldman, E. (2020). An introduction to the california consumer privacy act (ccpa). *Santa Clara Univ. Legal Studies Research Paper*.

GoogleDP, R. (2018). https://github.com/google/differential-privacy.

Greenberg, A. (2016). Apple's' differential privacy'is about collecting your data—but not your data.(2016). *URL www. wired. com/2016/06/apples-differential-privacy-collecting-data*.

Hawes, M. (2020). Differential privacy and the 2020 decennial census. In *APHA's 2020 VIRTUAL Annual Meeting and Expo (Oct. 24-28)*. APHA.

Hayes, J., Balle, B., and Kumar, M. P. (2022). Learning to be adversarially robust and differentially private. *arXiv preprint arXiv:2201.02265*.

Jiang, Y., Wang, C., Wu, Z., Du, X., and Wang, S. (2018). Privacy-preserving biomedical data dissemination via a hybrid approach. In *AMIA Annual Symposium Proceedings*, volume 2018, page 1176. American Medical Informatics Association.

Johnson, N., Near, J. P., and Song, D. (2018). Towards practical differential privacy for sql queries. *Proceedings of the VLDB Endowment*, 11(5):526–539.

Kim, J. W., Jang, B., and Yoo, H. (2018). Privacy-preserving aggregation of personal health data streams. *PloS one*, 13(11):e0207639.

Lee, J. and Clifton, C. (2011). How much is enough? choosing epsilon for differential privacy. pages 325–340.

Liau, D., Zaeem, R. N., and Barber, K. S. (2019). Evaluation framework for future privacy protection systems: A dynamic identity ecosystem approach. In *2019 17th International Conference on Privacy, Security and Trust (PST)*, pages 1–3.

McClure, D. and Reiter, J. P. (2012). Differential privacy and statistical disclosure risk measures: An investigation with binary synthetic data. *Trans. Data Priv.*, 5(3):535–552.

Naldi, M. and D'Acquisto, G. (2015). Differential privacy: An estimation theory-based method for choosing epsilon. *arXiv preprint arXiv:1510.00917*.

Quinton, S. and Reynolds, N. (2018). Characteristics of digital data.

Regulation, G. D. P. (2018). General data protection regulation (gdpr). *Intersoft Consulting, Accessed in October*, 24(1).

Subramanian, R. (2022). Applications of differential privacy to healthcare. *Available at SSRN 4005908*.

Tsou, Y.-T., Chen, H.-L., and Chang, Y.-H. (2019). Rod: Evaluating the risk of data disclosure using noise estimation for differential privacy. *IEEE Transactions on Big Data*.

Winslett, M., Yang, Y., and Zhang, Z. (2012). Demonstration of damson: Differential privacy for analysis of large data. In *2012 IEEE 18th International Conference on Parallel and Distributed Systems*, pages 840–844. IEEE.

Zaeem, R. N., Budalakoti, S., Barber, K. S., Rasheed, M., and Bajaj, C. (2016). Predicting and explaining identity risk, exposure and cost using the ecosystem of identity attributes. pages 1–8.

Zaiss, J., Nokhbeh Zaeem, R., and Barber, K. S. (2019). Identity threat assessment and prediction. *Journal of Consumer Affairs*, 53(1):58–70.

Zhang, S., Hagermalm, A., and Slavnic, S. (2022). An evaluation of open-source tools for the provision of differential privacy. *arXiv preprint arXiv:2202.09587*.