Adaptive Gaussian Mechanism Based on Expected Data Utility under Conditional Filtering Noise

Hai Liu[1], Zhenqiang Wu[1], Changgen Peng[2], Feng Tian[1] and Laifeng Lu[3]
[1] School of Computer Science, Shaanxi Normal University
Xi'an, 710119 - China
[e-mail: liuhai@snnu.edu.cn, zqiangwu@snnu.edu.cn, tianfeng@snnu.edu.cn]
[2] Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University
Guiyang, 550025 - China
[e-mail: sci.cgpeng@gzu.edu.cn]
[3] School of Mathematics and Information Science, Shaanxi Normal University
Xi'an, 710119 - China
[e-mail: lulaifeng@snnu.edu.cn]
*Corresponding author: Zhenqiang Wu

Received October 21, 2017; revised January 27, 2018; accepted February 14, 2018;
published July 31, 2018

Abstract

Differential privacy has broadly applied to statistical analysis, and its mainly objective is to ensure the tradeoff between the utility of noise data and the privacy preserving of individual's sensitive information. However, an individual could not achieve expected data utility under differential privacy mechanisms, since the adding noise is random. To this end, we proposed an adaptive Gaussian mechanism based on expected data utility under conditional filtering noise. Firstly, this paper made conditional filtering for Gaussian mechanism noise. Secondly, we defined the expected data utility according to the absolute value of relative error. Finally, we presented an adaptive Gaussian mechanism by combining expected data utility with conditional filtering noise. Through comparative analysis, the adaptive Gaussian mechanism satisfies differential privacy and achieves expected data utility for giving any privacy budget. Furthermore, our scheme is easy extend to engineering implementation.

Keywords: Sensitive information, differential privacy, conditional filtering noise, expected data utility, adaptive Gaussian mechanism

This research was supported by the National Natural Science Foundation of China under Grant No. 61602290 and No. 61662009, the Fundamental Research Founds for the Central Universties under Grant No. 2016CBY004, No. GK201704016, No. GK201603093, and No. GK201501008, the Program of Key Science and Technology Innovation Team in Shaanxi Province under Grant No. 2014KTC-18, and the Open Project Fund of Guizhou Provincial Key Laboratory of Public Big Data under Grant No. 2017BDKFJJ026.

http://doi.org/10.3837/tiis.2018.07.027 **ISSN : 1976-7277**

1. Introduction

The end of privacy [1] is the special issue of Science. Privacy as we have known it is ending, and we are only beginning to fathom the consequences. Hence, differential privacy has been proposed by Dwork [2], which is mainly privacy preserving method to individual sensitive data. That is to say, in statistical analysis, for all datasets D_1 and D_2 which differ in only one record, for all adversaries A and all outputs t denoted by $T_A(D)$, according to privacy budget ε which is a random variable, such that

$$\left|\ln\left(\frac{\Pr[T_A(D_1)]}{\Pr[T_A(D_2)]}\right)\right| \le \varepsilon \tag{1}$$

Not only this approach gains an extent of data utility, but also satisfies differential privacy. Thus, differential privacy has widely applications. Differential privacy was used to protect privacy for big data in body sensor networks [3]. Boyd et al [4] proposed a differential privacy computation of classifier evaluation metrics, irrespective of the training setting. With respect to datasets D_1 and D_2 which differ in only one record, Dwork et al [5] extended the Laplace distribution of [2] to the Gaussian distribution $N(0,\sigma^2)$ for which the noise sums primitive yields δ -approximate ε -indistinguishable, where $0<\varepsilon<1$ and $\sigma^2 \ge 2\log(2/\delta)/\varepsilon^2$. Built on [5], Dwork and Roth [6] obtained an alternative for adding Laplace noise was identical to adding Gaussian noise. In [6], let $\varepsilon \in (0,1)$, for $c^2 > 2\ln(1.25/\delta)$, the Gaussian mechanism of parameter $\sigma \ge c\Delta f/\varepsilon$ is (ε,δ) -differential privacy, where Δf is ℓ_1 -sensitivity, please see its definition in section 3. Another, differential privacy is meaningful for $\varepsilon > 1$. However, differential privacy mechanisms generated random independent identical distribution noise. Noise may be very larger, such that the data added considerably large noise is not available. Such as we want a function for counting the number of people aged in (30,40) by differential privacy mechanisms. Under a differential privacy mechanism, random and considerable noise may be added to some ages, so it has no significance to differential privacy statistical analysis in the age interval (30,40). In this situation, individual sensitive data achieve best differential privacy preserving, but data is lower utility. Hence, differential privacy preserving method is limited because of existing data excessive or mild distortion by adding random and large noise.

So, in this paper, we achieved adaptive differential privacy based on expected data utility under conditional filtering noise, where privacy budget $\varepsilon > 0$ is arbitrary in our scheme. Firstly, we made conditional filtering for noise that absolute value satisfied interval (0.5,1.5). Secondly, we gave the definition of expected data utility. Thirdly, we proposed the adaptive Gaussian mechanism by combining expected data utility with conditional filtering noise. Through comparative analysis, given the privacy budget ε , our scheme guarantees the differential privacy and achieves the expected utility of noise data. Another, our scheme is also effective. In practice, we can achieve adaptive differential privacy by specifying the expected data utility based on our scheme.

Based on Gaussian mechanism, we generalize privacy budget $\varepsilon \in (0,1)$ to $\varepsilon > 0$, such that the privacy budget has a widely range. We present adaptive differential privacy by combining expected data utility with conditional filtering noise. Then through analysis, we demonstrate this approach is reasonable, and it can achieve expected data utility, while satisfying differential privacy. Our main contributions are as following:

(1) We required the absolute value of conditional filtering noise belonging to the interval (0.5,1.5), and we gave the definition of expected data utility according to the absolute value of relative error. Then, we proposed an adaptive Gaussian mechanism by combining expected data utility with conditional filtering noise.

(2) We proved that adaptive Gaussian mechanism achieves differential privacy and ensures expected data utility under any privacy budget.

(3) In engineering applications, we constructed applications framework using adaptive Gaussian mechanism in interactive and non-interactive environment. We conducted an instance analysis, which shows that our scheme ensures differential privacy and achieves expected data utility.

The rest of this paper is organized as follows. We introduce related work of data utility to noise response under differential privacy in the section 2. In the section 3, we introduce the preliminaries of differential privacy. In the section 4, we give the definitions of conditional filtering noise and expected data utility based on the absolute value of relative error. Then we present adaptive Gaussian mechanism, it can achieve the expected data utility of noise response results, while satisfying differential privacy. We give non-interactive and interactive applications framework using adaptive Gaussian mechanism, and present the corresponding protocols in the section 5. Section 6 conducts a comparative experimental analysis of privacy and expected data utility. In section 7, we would safely make a conclusion about this paper.

2. Related Work

How to build valuable statistical analysis and protect confidentiality is mutually contrary. Therefore, it is necessary that designing protocols for balancing privacy versus utility in using personal sensitive data. Goroff [7] started traditional methods whose strengths and shortcomings motivate more recent approaches. How to balance the utility and the privacy will allow everyone to benefit from big data science and protect sensitive information that enormous and growing stores of everyone.

Today, in an era of data explosion, the traditional methods for privacy preserving no longer work, hence, privacy preserving by controlling the use of data is gaining more attention [8]. Thus, Gaussian mechanism has studied to make sure that the tradeoff between privacy and utility. Another, Gaussian mechanism has had widely applications.

To enhance accuracy of noise data, Ny and Pappas [8] conducted the filtering approximation set-ups of differential privacy for input perturbation and output perturbation. They defined $k(\varepsilon,\delta)=(K+\sqrt{K^2+2\varepsilon})\big/2\varepsilon$ and $K=p^{-1}(\delta)$, where $p(x)=\int_x^{\infty}e^{-u^2/2}du\big/\sqrt{2\pi}$. Then in input perturbation, the white Gaussian noise $X\sim N(0,\sigma^2)$ is added to every input signal, where $\sigma=k(\varepsilon,\delta)d(D_1,D_2)$. And in output perturbation, the Gaussian mechanism is defined $M(D)=f(D)+X$, where $X\sim N(0,\sigma^2)$, $\sigma\geq k(\varepsilon,\delta)\max_{1\leq i\leq n}\{\|f\|_\infty d(D_1,D_2)\}$, and $d(D_1,D_2)$ is Haming distance between datasets D_1 and D_2 . Both of them are (ε,δ) -differential privacy, and they should be evaluated relying on the query function vector f of every query function f_i for all $1\leq i\leq n$ and the number n of data records in the datasets, as none of the error bound is better than others in all circumstances. Their differential privacy filtering mechanism is similar to the Gaussian mechanism [6], but the formulation

$k(\varepsilon,\delta)$ is not rigorous proved from mathematical theory. It is common that the data curators have quite different expectations regarding the acceptable level for privacy to their data. Differential privacy may lead to insufficient privacy preserving for users or over-preserving for others. According to specify a personal privacy requirement for their data, through defining privacy specification is a mapping $\Phi:U\to R$ from users to personal privacy preferences. Joragensen et al [9] proposed the personalized differential privacy. Thus, our scheme is different from this, because of achieving adaptive differential privacy based on expected data utility for giving any privacy budget. Soria-Comas et al [11] proposed individual differential privacy, an alternative differential privacy notion that offers the same privacy guarantees as standard differential privacy to individuals (even though not to groups of individuals). Individual differential privacy allows more analytical accuracy by using local sensitivity. Wang and Anandkumar [12] presented a noise calibrated tensor power method with efficient privacy guarantees by using Gaussian mechanism.

In many applications, the matrix containing sensitive information about individuals, Hardt and Roth [13] gave significant improvements in accuracy over randomize response with Gaussian mechanism under the natural and necessary assumption that the matrix has low coherence. Dwork et al [14] achieved optimal bounds for privacy preserving principal component analysis using Gaussian mechanism, resulting in an improved regret bound. They ensured Gaussian mechanism was (ε,δ)-differential privacy by adding independently drawn random noise from $N(0,2\ln(1.25/\delta)\Delta f^2/\varepsilon^2)$.

In privacy preserving queries, Nikolov et al [15] study tradeoff between utility and privacy in the context of linear queries over histograms. To achieve (ε,δ)-differential privacy of linear queries is adding appropriately scaled independent Gaussian noise to each query, where Gaussian noise subjects to $N(0,\sigma(1+\sqrt{2\ln(1/\delta)})/\varepsilon)^d$ for a $d\times N$ queries matrix $A=(a_i)_{i=1}^N$ consisting of query a_i for $\forall i:\|a_i\|_2\le\sigma$. Li et al [16] described the matrix mechanism, an algorithm for answering a workload of linear counting queries that adapt the noise distribution to properties of the provided queries. Given a workload, the mechanism uses a different set of queries, called a query strategy, which is answered using a standard Laplace or Gaussian mechanism.

In deep learning based on neural networks, the training of models requires large, representative datasets, which may be crowdsourced and contain sensitive information. Training models of deep learning should not expose private information in these datasets, so Abadi et al [17] developed new algorithmic techniques for learning and a refined analysis of privacy costs within Gaussian mechanism. For strongly convex and smooth objectives, Zhang et al [18] proved that gradient descent with output perturbation using Gaussian mechanism not only achieves nearly optimal utility, but also significantly improves the running time of previous state-of-the-art private optimization algorithms. For non-convex but smooth objectives, they proposed a random round private stochastic gradient descent algorithm, which provably converges to a stationary point with privacy guarantee.

Gaussian mechanism has increasingly widely applications, such as it was used in differential privacy recommended system: building privacy into the Netflix Prize contenders [19]. With the demand of ridesharing services increasingly sharply, serious privacy concerns have become a major barrier against its further development, so Tong et al [20] proposed a scheduling protocol based on joint differential privacy by combing Gaussian mechanism and Laplace mechanism for the purpose of protecting users' location privacy and minimizing vehicle miles in the system.

Differential privacy is a formal method to ensure privacy. It has been successfully applied in a range of data analysis tasks. Despite much recent work, the randomness of noise not ensures the expected data utility. Hence, based on contemporary work, we present the adaptive Gaussian mechanism, which achieves the expected data utility and differential privacy by combining expected data utility with conditional filtering noise. This has been an important significance to privacy data analysis task in engineering applications.

3. Differential Privacy

In this section, we introduce the mainly preliminaries of differential privacy [6].

Let $D=\{r_1,r_2,\ldots,r_n\}$ denote a dataset consisted of n records, where $r_i(1\le i\le n)$ denotes the i th record. To define the notion of the adaptive Gaussian mechanism based on expected data utility under conditional filtering noise, we need to several different primitive definitions.

Definition 1: (Adjacent Datasets) It is to say two datasets D_1 and D_2 are adjacent datasets if they have the same size and identical except for a single record. So the Hamming distance $d(D_1,D_2)$ between two datasets D_1 and D_2 is 1.

Throughout this paper, we will use the notion of differential privacy under Hamming distance $d(D_1,D_2)=1$ for adjacent datasets D_1 and D_2.

Definition 2: (ε-Indistinguishable) A mechanism is ε-indistinguishable if for all adjacent datasets D_1 and D_2, for all adversaries A and all transcripts t denoted by $T_A(D_i)(i\in\{1,2\})$, there is

$$\left|\ln\left(\frac{\Pr[T_A(D_1)=t]}{\Pr[T_A(D_2)=t]}\right)\right|\le\varepsilon \tag{2}$$

Definition 3: (δ-Approximate ε-Indistinguishable) A mechanism is δ-approximate ε-indistinguishable if for all adjacent datasets D_1 and D_2, all adversaries A, and all transcripts t denoted by $T_A(D_i)(i\in\{1,2\})$, there is

$$\left|\ln\left(\frac{\Pr[T_A(D_1)=t]-\delta}{\Pr[T_A(D_2)=t]}\right)\right|\le\varepsilon \tag{3}$$

According to **Definition 3**, the definition of (ε,δ)-differential privacy is as following.

Definition 4: ((ε,δ)-Differential Privacy) Given $\varepsilon\ge 0$, a randomized algorithm M is (ε,δ)-differential privacy, if for any two adjacent datasets D_1 and D_2, and any outputs $S\subseteq Range(M)$ of M, so $M(D_1)\in S$ and $M(D_2)\in S$, such that

$$\Pr[M(D_1)]\le e^{\varepsilon}\Pr[M(D_2)]+\delta \tag{4}$$

where $\delta\in[0,1]$ is any probability value without satisfying differential privacy, and if $\delta=0$, M is $(\varepsilon,0)$-differential privacy algorithm.

Next, we firstly give the definition of sensitivity before giving Gaussian mechanism.

For any query function $f:D\to R^k$ about a dataset D, the ℓ_1-sensitivity of f is

$$\Delta f=\max_{d(D_1,D_2)=1}\|f(D_1)-f(D_2)\|_1 \tag{5}$$

for all adjacent datasets D_1 and D_2.

Theorem 1. (Gaussian Mechanism) Let $\varepsilon>0$, Gaussian mechanism with parameter $\sigma\ge\Delta f\sqrt{2\ln(1.25/\delta)}\big/\varepsilon$ is (ε,δ)-differential privacy.

4. Adaptive Differential Privacy Mechanism

This section explicitly details adaptive Gaussian mechanism framework, analyzes its privacy and expected data utility from mathematics, and shows how to achieve the adaptive Gaussian mechanism in practice.

4.1 Adaptive Differential Privacy Framework

To present the adaptive differential privacy framework, we firstly give definitions of conditional filtering noise and expected data utility.

Definition 5: (Conditional Filtering Noise) In Gaussian mechanism, the Gaussian noise $X \sim N(0,\sigma^2)$ and noise $Y = X - k\sigma$ satisfied $0.5 < |Y| < 1.5$ by filtering condition on $-1.5 < Y < -0.5$ and $0.5 < Y < 1.5$.

Especially, $k \in [0,2]$ is a real number, which ensures the monotonicity of adaptive Gaussian mechanism. Here, monotonicity means a privacy metric decreases or increases as the privacy budget increases. Next, we define the expected data utility according to the absolute value of relative error $(x - x')/x$, where x' is the approximate value of x.

Definition 6: (Expected Data Utility) The $f(D)+u$ is the approximate value of $f(D)$, where u is known as utility factor. The expected data utility to be $U = 1 - E(E \in [0,1])$ if the absolute value of relative error is $E = |(f(D)+u) - f(D)|/|f(D)| = u/|f(D)|$.

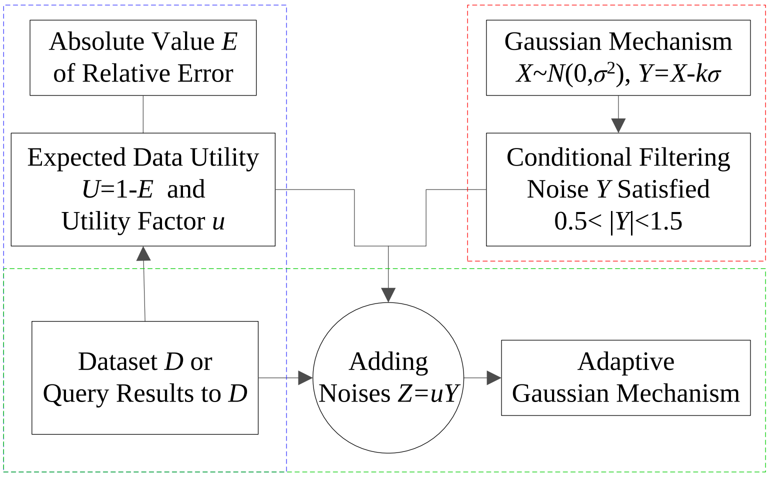
Absolute Value E of Relative Error
Expected Data Utility U=1-E and Utility Factor u
Gaussian Mechanism X~N(0,σ²), Y=X-kσ
Conditional Filtering Noise Y Satisfied 0.5< |Y|<1.5
Dataset D or Query Results to D
Adding Noises Z=uY
Adaptive Gaussian Mechanism

Fig. 1. Adaptive differential privacy framework based on expected data utility under conditional filtering noise

In **Fig. 1,** adaptive differential privacy framework consists of three parts. The first part makes the conditional filtering noise Y satisfied $0.5 < |Y| < 1.5$, where $Y = X - k\sigma$ and $X \sim N(0,\sigma^2)$. The second part computes the utility factor u according to absolute value E of relative error $((f(D)+u) - f(D))/f(D)(u \geq 0)$ of query results. Note that data analyst requests data curator publishing dataset regarded as a simple query function, which treats it as an identity query function $f: D \to D$. Therefore, $U = 1 - E = (|f(D)| - u)/|f(D)|$ is the expected data utility. The third part achieves adaptive Gaussian mechanism by adding noise $Z = uY$ to query results. Why multiplies the conditional filtering noise Y with the utility factor u, see the following section. Since $0.5 < |Y| < 1.5$, then the rounding value of $|Y|$ is 1.

Furthermore, noise value $|Z|$ is approximately equal to the utility factor u for keeping expected data utility to be $U = 1 - E$, that is to say

$$U = 1 - E = \frac{|f(D)| - u}{|f(D)|} \approx \frac{|f(D)| - |uY|}{|f(D)|} \approx \frac{|f(D)| - |Z|}{|f(D)|} \tag{6}$$

Thus, this framework satisfies expected data utility and differential privacy. Please refer to their theory proof of the follow-up section.

4.2 Adaptive Gaussian Mechanism

For any numeric query function $f : D \to R^k$, Gaussian mechanism generates noise directly added to query results. Gaussian noise X_i is independent identical distribution random variables drawn from $N(0,\sigma^2)$ in Gaussian mechanism, however, the Gaussian noise X_i may be very large that leads data is not available. So we construct adaptive Gaussian mechanism, which achieves expected data utility and differential privacy.

According to **Definition 6**, we can achieve expected data utility, but this is not satisfying differential privacy if we regard the utility factor u as noise. So we get noise $Z = uY$ by multiplying the conditional filtering noise Y with the utility factor u. Then we add noise Z to query results. The probability of noise Z corresponding to the probability of noise X satisfies differential privacy. Like this can achieve approximate expected data utility, while satisfying differential privacy. Thus, we have the following definition of adaptive Gaussian mechanism.

***Definition** 7*: (Adaptive Gaussian Mechanism) For conditional filtering noise Y, and utility factor u. For any numeric query function $f : D \to R^k$, the adaptive Gaussian mechanism is defined as follows

$$AGM(D) = f(D) + Z \tag{7}$$

where $Z = uY$, $Y = X - k\sigma$, and X generated by Gaussian mechanism.

In adaptive Gaussian mechanism, provided the privacy budget $\varepsilon > 0$. According to the definition of adaptive Gaussian mechanism, and we have the following theorems.

Theorem 2. Adaptive Gaussian mechanism is (ε,δ)-differential privacy.

Proof of Theorem 2:

We use the proof method of contradiction. Adaptive Gaussian mechanism will return $f(D)+Z$ for a dataset D and a query function f, where $Z = uY$, $Y = X - k\sigma$ and $X \sim N(0,\sigma^2)$. Because of the probability of $Z_i \in Z$ is identical with the probability of $X_i \in X$. Assuming the adaptive Gaussian mechanism is (ε,δ)-differential privacy, then

$$\frac{\Pr(f(D_1)+Z)}{\Pr(f(D_2)+Z)} = \prod\left(\frac{\Pr(f_i(D_1)+Z_i)}{\Pr(f_i(D_2)+Z_i)}\right) = \prod\left(\frac{e^{-X_i^2/2\sigma^2}}{e^{-(X_i+\Delta f)^2/2\sigma^2}}\right) \le e^{\varepsilon} \tag{8}$$

Because of $\sigma \ge \Delta f \sqrt{2\ln(1.25/\delta)}/\varepsilon$ and $\delta \in (0,1)$, there are

$$\frac{-X^2 + (X+\Delta f)^2}{2\sigma^2} \le \varepsilon \tag{9}$$

and

$$X \le \frac{2\Delta f \ln(1.25/\delta)}{\varepsilon} - \frac{\Delta f}{2} \tag{10}$$

Therefore, the Gaussian mechanism is $(\varepsilon,0)$ -differential privacy when noise $X \le 2\Delta f \ln(1.25/\delta)/\varepsilon - \Delta f/2$. Since $Y = X - k\sigma$, computing utility factor u from absolute value E of relative error, we have

$$Z \le \frac{2\Delta fu \ln(1.25/\delta)}{\varepsilon} - \frac{\Delta fu + k\sigma u}{2} \tag{11}$$

So, the adaptive Gaussian mechanism is $(\varepsilon,0)$ -differential privacy by combining the utility factor u and conditional filtering noise $Y(0.5 <| Y |< 1.5)$ when $Z \le 2\Delta fu \ln(1.25/\delta)/\varepsilon - (\Delta fu + k\sigma u)/2$.

Let us partition Z as $Z = R_1 \cup R_2$, where

$$R_1 = \{Z \in R : Z \le \frac{2\Delta fu \ln(1.25/\delta)}{\varepsilon} - \frac{\Delta fu + k\sigma u}{2}\} \tag{12}$$

and

$$R_2 = \{Z \in R : Z > \frac{2\Delta fu \ln(1.25/\delta)}{\varepsilon} - \frac{\Delta fu + k\sigma u}{2}\} \tag{13}$$

For any query results set $S = S_1 \cup S_2$ and defining them as follows

$$S_1 = \{ f(D) + Z : Z \in R_1\} \tag{14}$$

and

$$S_2 = \{ f(D) + Z : Z \in R_2\} \tag{15}$$

Thus, we have

$$\begin{aligned}\Pr(f(D_1) + Z \in S) &= \Pr(f(D_1) + Z \in S_1) + \Pr(f(D_1) + Z \in S_2) \\ &\le \Pr(f(D_1) + Z \in S_1) + \delta \\ &\le e^{\varepsilon} \Pr(f(D_2) + Z \in S) + \delta\end{aligned} \tag{16}$$

Therefore, the adaptive Gaussian mechanism is (ε,δ) -differential privacy.

The hypothesis of $\varepsilon > 0$ is reasonable because that (ε,δ) -differential privacy is symmetric, therefore

$$e^{-\varepsilon}(\Pr(f(D_1) + Z \in S) - \delta) \le \Pr(f(D_2) + Z \in S) \tag{17}$$

Theorem 2 is proved.

Theorem 3. Expected data utility of adaptive Gaussian mechanism is $1 - E$.

Proof of Theorem 3:

Similarly, we use the proof method of contradiction. By Gaussian mechanism generating noise $X \sim N(0,\sigma^2)$, then computing noise $Y = X - k\sigma$. Noise Y required satisfying $0.5 <| Y |< 1.5$ under conditional filtering. So, the rounding value to be 1 of noise absolute value $| Y |$.

Assuming the expected data utility is $1 - E$. E is the absolute value of relative error $(f(D) + u) - f(D)/f(D) = u/f(D)(u > 0)$. So, we get the utility factor u. Despite obtaining the expected data utility $1 - E$ by adding u to query results, this is not satisfying differential privacy. Thus, according to **Definition 7** and **Theorem 2**, we add noise $Z = uY$ to query results for ensuring differential privacy. There is

$$\frac{|(f(D) - Z) - f(D)|}{| f(D)|} = \frac{| Z |}{| f(D)|} = \frac{| uY |}{| f(D)|} = \frac{u | Y |}{| f(D)|} \approx \frac{u}{| f(D)|} = E \tag{18}$$

Therefore, the expected data utility of adaptive Gaussian mechanism is $1 - E$ when the absolute value of relative error for query function f on dataset D is E.

Theorem 3 is proved.

Similarly the properties of differential privacy [6], the adaptive Gaussian mechanism has the properties including group privacy, post processing and composition.

In the following, we give the **Algorithm 1** of the adaptive Gaussian mechanism.

Algorithm 1. Adaptive Gaussian mechanism

```
Input: The privacy budget, $\varepsilon$.
        The probability of without satisfying differential privacy, $\delta$.
        The expected data utility, $U$.
        The $\ell_1$-sensitivity of query function $f$, $\Delta f$.
Output: The noise response results, $f(D)+Z$.
$X \leftarrow N(0,\sigma^2)$ {Generating Gaussian noise}
$Y \leftarrow X - k\sigma$ {The linear function of Gaussian noise}
$Y \leftarrow 0.5 < |Y| < 1.5$ {Conditional filtering noise}
$u \leftarrow E = u/|f(D)|$ {Computing utility factor under expected data utility}
$Z \leftarrow uY$ {Multiplying utility factor with conditional filtering noise}
$f(D)+Z \rightarrow$ Noise response results {Adding noise to query results}
```

5. Applications Framework Using Adaptive Gaussian Mechanism

In non-interactive and interactive framework, we state the applications framework using adaptive Gaussian mechanism.

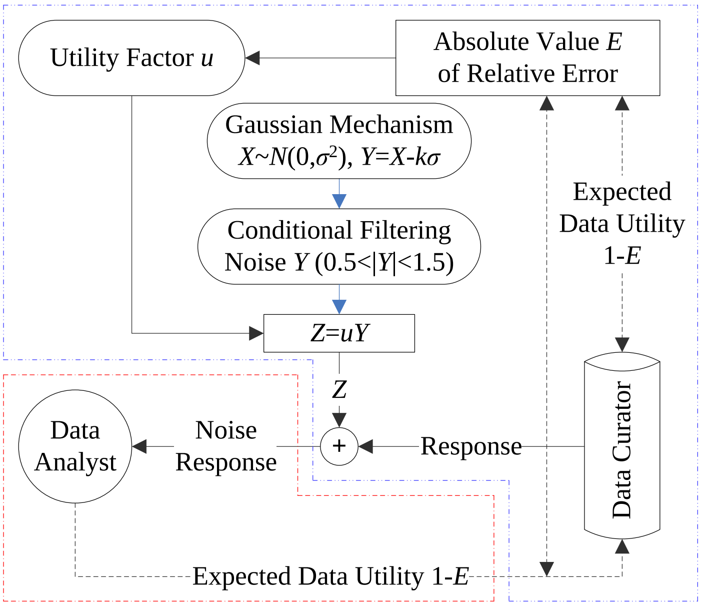
Utility Factor *u*
Absolute Value *E* of Relative Error
Gaussian Mechanism *X*~*N*(0,*σ*²), *Y*=*X*-*kσ*
Conditional Filtering Noise *Y* (0.5<|*Y*|<1.5)
Expected Data Utility 1-*E*
Z=*uY*
Z
+
Data Analyst
Noise Response
Response
Data Curator
Expected Data Utility 1-*E*

Fig. 2. Non-interactive framework using adaptive Gaussian mechanism

Firstly, we give the non-interactive framework using adaptive Gaussian mechanism in **Fig. 2**. This framework consists of two parts. Data analyst publishes expected data utility $U = 1 - E$ to data curator in the first part. In the second part, data curator makes noise response of dataset using adaptive Gaussian mechanism.

In non-interactive framework, we compute the ℓ_1-sensitivity of adjacent dataset D_1 and D_2 is following

$$\Delta f = \max_{d(D_1,D_2)=1} \| D_1 - D_2 \|_1 \tag{19}$$

Fig. 3 gives the protocol of non-interactive framework using adaptive Gaussian mechanism based on computation of the sensitivity. In this protocol, the $E = u/|D|$.

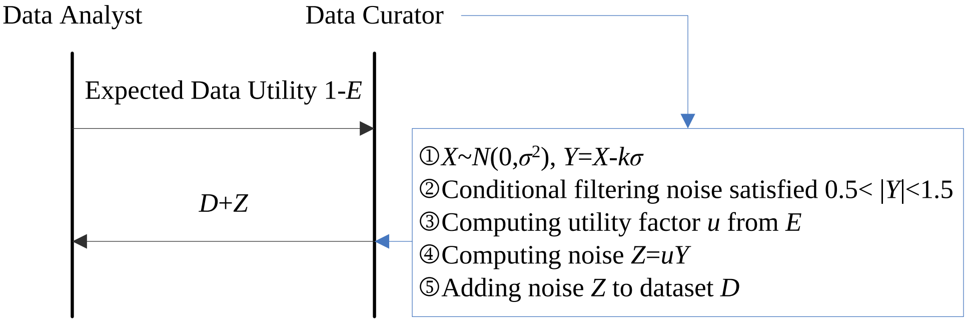
Data Analyst
Data Curator
Expected Data Utility 1-E
D+Z
①X~N(0,σ²), Y=X-kσ
②Conditional filtering noise satisfied 0.5< |Y|<1.5
③Computing utility factor u from E
④Computing noise Z=uY
⑤Adding noise Z to dataset D

Fig. 3. Non-interactive protocol using adaptive Gaussian mechanism

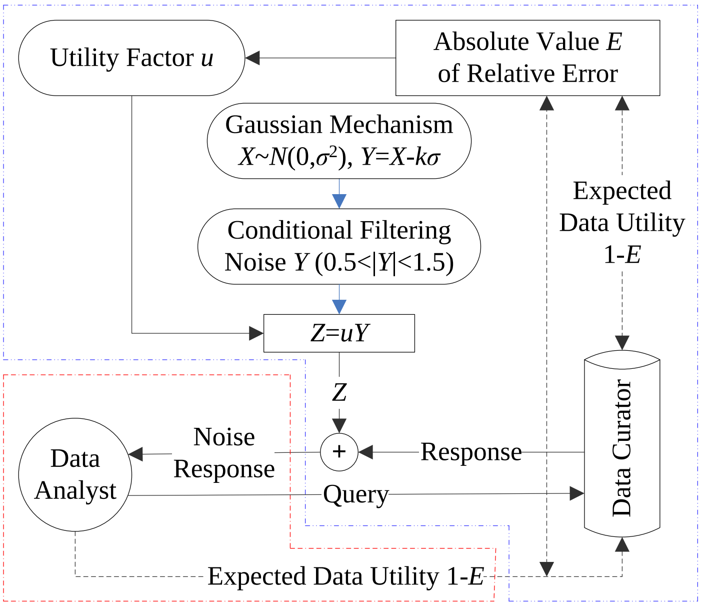
Utility Factor u
Absolute Value E of Relative Error
Gaussian Mechanism X~N(0,σ²), Y=X-kσ
Conditional Filtering Noise Y (0.5<|Y|<1.5)
Expected Data Utility 1-E
Z=uY
Z
+
Data Analyst
Noise Response
Response
Query
Data Curator
Expected Data Utility 1-E

Fig. 4. Interactive framework using adaptive Gaussian mechanism

Next, we give the interactive framework using adaptive Gaussian mechanism in Fig. 4. This framework is slightly different from non-interactive framework. Data analyst publishes expected data utility $U = 1 - E$ of the query results to data curator in the first part. In the second part, data curator sends the noise response of query results according to the data analyst's expected utility.

In this interactive framework, the ℓ_1-sensitivity of query function $f : D \to R^k$ for adjacent datasets D_1 and D_2 is

$$\Delta f = \max_{d(D_1,D_2)=1} \| f(D_1) - f(D_2) \|_1 \tag{20}$$

In **Fig. 5**, we give the protocol of interactive framework using adaptive Gaussian mechanism based on computation of the sensitivity. In this protocol, the $E = u/|f(D)|$.

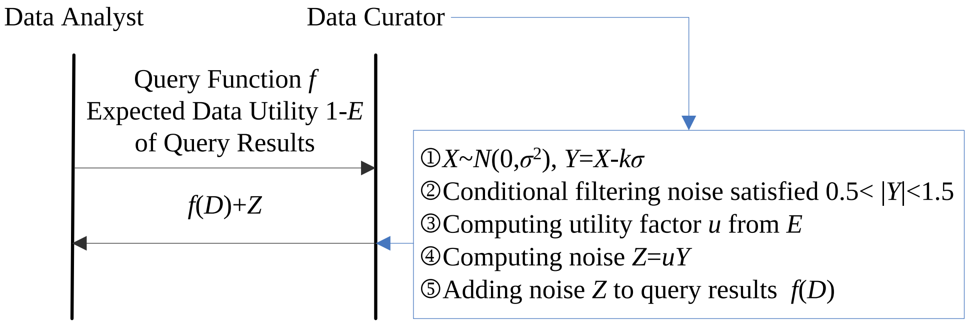
Data Analyst
Data Curator
Query Function f
Expected Data Utility 1-E
of Query Results
$f(D)+Z$
①$X \sim N(0,\sigma^2)$, $Y=X-k\sigma$
②Conditional filtering noise satisfied 0.5< $|Y|$<1.5
③Computing utility factor u from E
④Computing noise $Z=uY$
⑤Adding noise Z to query results $f(D)$

Fig. 5. Interactive protocol using adaptive Gaussian mechanism

6. Experimental Evaluations

In this experimental evaluation, we only make a comparative analysis among Gaussian mechanism, personalized differential privacy [10], individual differential privacy [11], and adaptive Gaussian mechanism in non-interactive framework. We achieve personalized differential privacy and individual differential privacy by using Gaussian mechanism, which are known as personalized Gaussian mechanism (PGM) and Individual Gaussian mechanism (IGM) in this paper, respectively. Similarly, there exists the same experimental analysis in an interactive framework. We will analyze the privacy preserving and expected data utility. In all experiments, let us set $\delta = 0.1$, $\Delta f = 1$, and local sensitivity $LS_f = 0.1$.

6.1 Dataset

We use T-Drive taxi trajectory dataset [21] to evaluate expected data utility and privacy preserving level of adaptive Gaussian mechanism. This is a sample of T-Drive taxi trajectory dataset, which was generated by over 10,000 taxis in a period of one week in Beijing. We chose the taxi ID 7's dataset. We use the average location of its trajectory every day from February 2-8, 2008. So, this trajectory has seven locations.

6.2 Privacy Preserving

We measure the privacy using expected estimation error as following

$$E = \sum p(Y) \| f(D)' - f(D) \|_1 = \sum p(Y) |Y| \tag{21}$$

Here, $f(D)' = f(D) + Y$, Y is noise generated by differential privacy mechanisms. Because of considering the non-interactive framework, the privacy metric is

$$E = \sum p(Y) \| D' - D \|_1 = \sum p(Y) |Y| \tag{22}$$

Here, $D' = D + Y$, Y is noise generated by differential privacy mechanisms.

Another, because trajectory data consists of latitude and longitude, the privacy metric is

$$E = \sum p(Y_1) p(Y_2) |Y_1| |Y_2| \tag{23}$$

Here, noise Y_1 and Y_2 added to latitude and longitude, respectively.

In privacy analysis, we get the average experimental result of repeating 30 times experiments. We directly add noise to the trajectory using Gaussian mechanism, personalized Gaussian mechanism, and individual Gaussian mechanism, see the expected estimation error curve in the **Fig. 6**. In that three Gaussian mechanisms directly add noise to the trajectory, the expected data estimation error decreases as privacy budget increases.

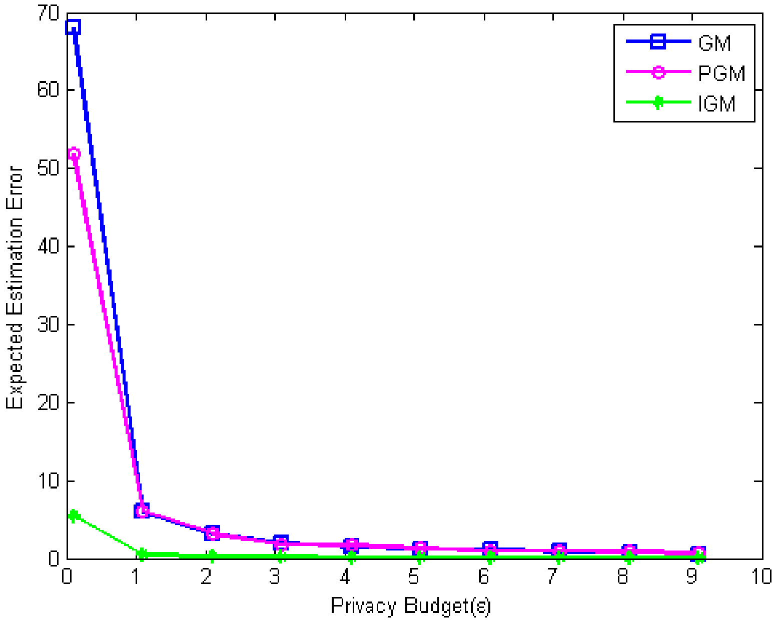

Fig. 6. Expected estimation error of the trajectory using Gaussian mechanism (GM), personalized Gaussian mechanism (PGM), and individual Gaussian mechanism (IGM)

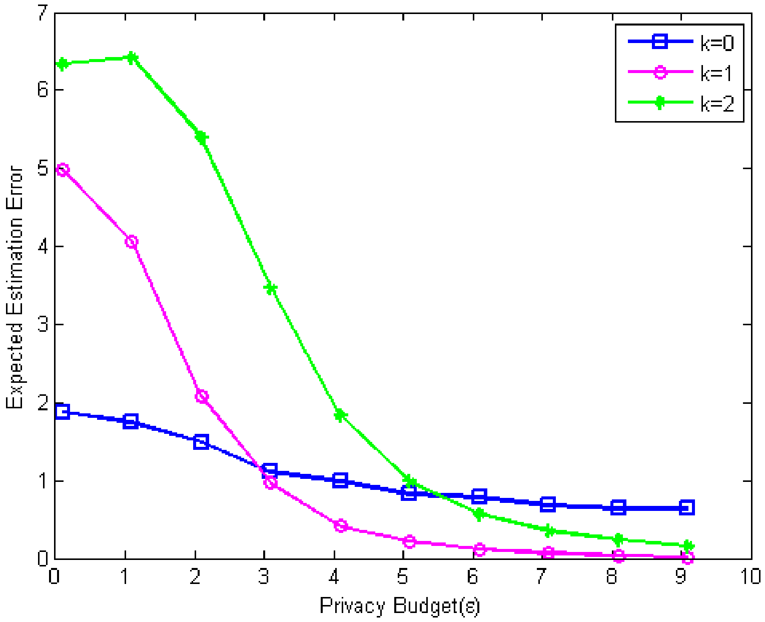

Fig. 7. Expected estimation error of the trajectory using conditional filtering noise of Gaussian mechanism

We get the expected estimation error curve in the **Fig. 7** by only adding conditional filtering noise of Gaussian mechanism to the trajectory. This can achieve differential privacy by adding conditional filtering noise and the expected estimation error decreases as privacy budget increases. When k increases in $Y = X - k\sigma(k \in [0,2])$, the expected estimation error also is rapidly increasing.

Finally, the adaptive Gaussian mechanism satisfies the monotonicity from **Fig. 8**, that is to say, the expected estimation error decreases as privacy budget increases. Another, when the absolute value E of relative error and k increase, the expected estimation error also increases.

In a word, the adaptive Gaussian mechanism can achieve differential privacy and maintains expected data utility.

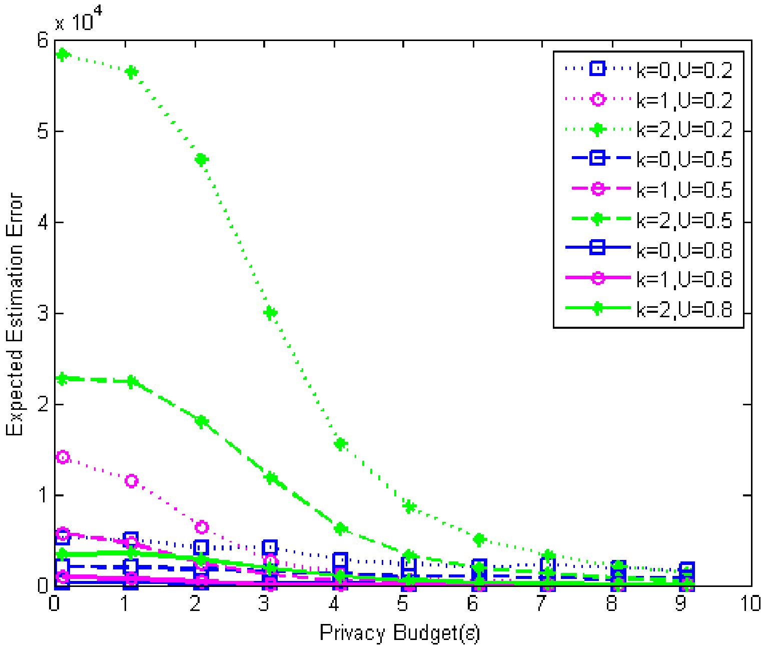

Fig. 8. Expected estimation error of the trajectory using adaptive Gaussian mechanism

6.3 Data Utility

Now, according to **Definition 6** of expected data utility metric, we analyze the data utility of Gaussian mechanism, personalized Gaussian mechanism, and individual Gaussian mechanism, and analyze the expected data utility of adaptive Gaussian mechanism. In all data utility analysis, we also get the average experimental result of repeating 30 times experiments.

In the non-interactive framework using Gaussian mechanism, personalized Gaussian mechanism, and individual Gaussian mechanism, the metric of data utility is $1 - E = (|D| - X)/|D|$ for the absolute value E of relative error. In the non-interactive framework using adaptive Gaussian mechanism, for the absolute value E of relative error, the metric of expected data utility is

$$U = 1 - E \approx \frac{|D| - Z}{|D|} \approx \frac{|D| - uY}{|D|} \tag{24}$$

Thus, we know that the data utility is 1 to every location of the trajectory when the absolute value E of the relative error is 0.

Table 1. Data utility of the trajectory using Gaussian mechanism

Trajectory dataset	$\varepsilon = 0.1$	$\varepsilon = 1.0$	$\varepsilon = 2.0$
Taxi ID 7	(0.8257, 0.5830)	(0.9859, 0.9630)	(0.9922, 0.9807)
	(0.8413, 0.6038)	(0.9832, 0.9596)	(0.9900, 0.9759)
	(0.8646, 0.5730)	(0.9869, 0.9532)	(0.9924, 0.9747)
	(0.8326, 0.5659)	(0.9847, 0.9563)	(0.9915, 0.9763)
	(0.8671, 0.4542)	(0.9884, 0.9643)	(0.9932, 0.9783)
	(0.8851, 0.6133)	(0.9872, 0.9657)	(0.9920, 0.9791)
	(0.8763, 0.5893)	(0.9815, 0.9661)	(0.9950, 0.9791)

Table 2. Data utility of the trajectory using personalized Gaussian mechanism

Trajectory dataset	**Average location**	**Privacy preference**	**Data utility**
Taxi ID 7	(116.7571, 39.8034)	$\varepsilon = 0.1$	(0.8725, 0.5676)
	(116.7140, 39.8456)	$\varepsilon = 2.0$	(0.9924, 0.9805)
	(116.7380, 39.8093)	$\varepsilon = 1.0$	(0.9847, 0.9532)
	(116.4927, 39.8970)	$\varepsilon = 0.5$	(0.9763, 0.9275)
	(116.4595, 39.9290)	$\varepsilon = 1.0$	(0.9857, 0.9675)
	(116.4346, 39.9192)	$\varepsilon = 1.5$	(0.9885, 0.9664)
	(116.5023, 39.9329)	$\varepsilon = 1.0$	(0.9834, 0.9509)

Table 3. Data utility of the trajectory using individual Gaussian mechanism

Trajectory dataset	$\varepsilon = 0.1$	$\varepsilon = 1.0$	$\varepsilon = 2.0$
Taxi ID 7	(0.9843, 0.9626)	(0.9985, 0.9960)	(0.9992, 0.9976)
	(0.9842, 0.9559)	(0.9986, 0.9955)	(0.9993, 0.9973)
	(0.9835, 0.9453)	(0.9982, 0.9960)	(0.9993, 0.9979)
	(0.9856, 0.9534)	(0.9983, 0.9958)	(0.9992, 0.9975)
	(0.9853, 0.9479)	(0.9985, 0.9952)	(0.9990, 0.9975)
	(0.9857, 0.6133)	(0.9983, 0.9941)	(0.9995, 0.9974)
	(0.9836, 0.9543)	(0.9984, 0.9951)	(0.9990, 0.9976)

Table 4. Data utility of the trajectory using conditional filtering noise of Gaussian mechanism

Trajectory dataset	$k = 0, \varepsilon = 0.1$	$k = 0, \varepsilon = 1.0$	$k = 0, \varepsilon = 2.0$	$k = 1, \varepsilon = 0.1$	$k = 1, \varepsilon = 1.0$	$k = 1, \varepsilon = 2.0$	$k = 2, \varepsilon = 0.1$	$k = 2, \varepsilon = 1.0$	$k = 2, \varepsilon = 2.0$
Taxi ID 7	(0.9913, 0.9725)	(0.9913, 0.9772)	(0.9928, 0.9761)	(0.9910, 0.9754)	(0.9916, 0.9749)	(0.9912, 0.9733)	(0.9914, 0.9745)	(0.9911, 0.9728)	(0.9904, 0.9733)
	(0.9917, 0.9746)	(0.9927, 0.9763)	(0.9919, 0.9768)	(0.9917, 0.9741)	(0.9912, 0.9755)	(0.9910, 0.9760)	(0.9913, 0.9755)	(0.9914, 0.9727)	(0.9907, 0.9741)
	(0.9907, 0.9735)	(0.9920, 0.9754)	(0.9925, 0.9799)	(0.9916, 0.9772)	(0.9913, 0.9761)	(0.9916, 0.9764)	(0.9914, 0.9778)	(0.9907, 0.9757)	(0.9908, 0.9746)
	(0.9912, 0.9761)	(0.9915, 0.9731)	(0.9926, 0.9782)	(0.9908, 0.9746)	(0.9912, 0.9732)	(0.9917, 0.9757)	(0.9917, 0.9758)	(0.9913, 0.9752)	(0.9909, 0.9728)
	(0.9920, 0.9748)	(0.9913, 0.9759)	(0.9924, 0.9784)	(0.9917, 0.9747)	(0.9925, 0.9735)	(0.9925, 0.9751)	(0.9917, 0.9750)	(0.9913, 0.9714)	(0.9914, 0.9740)
	(0.9913, 0.9741)	(0.9913, 0.9740)	(0.9910, 0.9768)	(0.9916, 0.9735)	(0.9908, 0.9768)	(0.9907, 0.9772)	(0.9911, 0.9754)	(0.9909, 0.9743)	(0.9912, 0.9739)
	(0.9914, 0.9752)	(0.9919, 0.9765)	(0.9918, 0.9777)	(0.9910, 0.9742)	(0.9911, 0.9746)	(0.9914, 0.9753)	(0.9922, 0.9762)	(0.9915, 0.9740)	(0.9915, 0.9719)

Table 1 shows the data utility variation of the trajectory using Gaussian mechanism, which data utility increases as privacy budget increases. For personalized Gaussian

mechanism, **Table 2** shows the data utility of the average location of the trajectory when privacy specification $\Phi =< 0.1,2.0,1.0,0.5,1.0,1.5,1.0 >$. For example, the privacy preference is $\varepsilon = 0.1$ for the first average location in trajectory dataset, and the privacy preference is $\varepsilon = 2.0$ for the second average location. The privacy preference is $\varepsilon = 1.0$ for the third, fifth, and seventh average location, respectively. **Table 2** demonstrates the data utility of the average location is different under different privacy preferences and the data utility of the average location is approximately equal under identical privacy preference. Similarity, data utility increases as privacy budget increases using personalized Gaussian mechanism. Data utility of the trajectory using individual Gaussian mechanism is shown **Table 3**. Because individual Gaussian mechanism is achieved based on local sensitivity, in contrast to Gaussian mechanism, the data utility is enhanced.

Table 5. Data utility of the trajectory using adaptive Gaussian mechanism when expected data utility $U = 0.2$

Trajectory dataset	$k = 0, \varepsilon = 0.1$	$k = 0, \varepsilon = 1.0$	$k = 0, \varepsilon = 2.0$	$k = 1, \varepsilon = 0.1$
Taxi ID 7	(0.2239, 0.1579)	(0.2641, 0.2550)	(0.2152, 0.2268)	(0.1687, 0.1723)
	(0.2219, 0.2183)	(0.2448, 0.2139)	(0.1689, 0.3243)	(0.1862, 0.1074)
	(0.1051, 0.1932)	(0.2305, 0.2313)	(0.2779, 0.3499)	(0.2162, 0.1415)
	(0.1684, 0.1709)	(0.2343, 0.1733)	(0.2862, 0.2544)	(0.1754, 0.1466)
	(0.1741, 0.2278)	(0.2089, 0.2065)	(0.2932, 0.3186)	(0.1493, 0.1932)
	(0.2893, 0.2038)	(0.1885, 0.1757)	(0.2050, 0.2795)	(0.1894, 0.1771)
	(0.1802, 0.1241)	(0.1882, 0.2207)	(0.2895, 0.3117)	(0.2176, 0.2608)
$k = 1, \varepsilon = 1.0$	$k = 1, \varepsilon = 2.0$	$k = 2, \varepsilon = 0.1$	$k = 2, \varepsilon = 1.0$	$k = 2, \varepsilon = 2.0$
(0.2564, 0.1274)	(0.1626, 0.1533)	(0.2478, 0.1627)	(0.1778, 0.1910)	(0.1329, 0.1202)
(0.1585, 0.1641)	(0.1737, 0.2385)	(0.1565, 0.2164)	(0.1897, 0.2584)	(0.1278, 0.1711)
(0.2753, 0.2446)	(0.1943, 0.1822)	(0.1724, 0.1684)	(0.1611, 0.1880)	(0.1505, 0.0756)
(0.2057, 0.1689)	(0.1517, 0.2735)	(0.2127, 0.2294)	(0.1635, 0.1559)	(0.1470, 0.1912)
(0.2570, 0.1715)	(0.1902, 0.2624)	(0.1977, 0.1694)	(0.2198, 0.2051)	(0.1243, 0.1847)
(0.1929, 0.2237)	(0.1907, 0.1214)	(0.1551, 0.1533)	(0.2092, 0.1248)	(0.1579, 0.1714)
(0.1841, 0.2562)	(0.3006, 0.1883)	(0.2277, 0.2669)	(0.1862, 0.1631)	(0.2137, 0.1675)

Table 6. Data utility of the trajectory using adaptive Gaussian mechanism when expected data utility $U = 0.5$

Trajectory dataset	$k = 0, \varepsilon = 0.1$	$k = 0, \varepsilon = 1.0$	$k = 0, \varepsilon = 2.0$	$k = 1, \varepsilon = 0.1$
Taxi ID 7	(0.4704, 0.5233)	(0.4555, 0.5032)	(0.5235, 0.5302)	(0.5194, 0.5350)
	(0.5208, 0.5522)	(0.4860, 0.4988)	(0.5512, 0.5061)	(0.5253, 0.4823)
	(0.4766, 0.5025)	(0.4808, 0.4742)	(0.5127, 0.5476)	(0.5070, 0.5052)
	(0.5398,0.4722)	(0.5162, 0.5571)	(0.5260, 0.5267)	(0.5360, 0.4984)
	(0.5107, 0.4923)	(0.4787, 0.5128)	(0.4872, 0.5141)	(0.4988, 0.4749)
	(0.5000, 0.5586)	(0.4876, 0.5341)	(0.5053, 0.5602)	(0.4893, 0.5039)
	(0.4638, 0.4792)	(0.4887, 0.4827)	(0.5236, 0.4780)	(0.4927, 0.5376)
$k = 1, \varepsilon = 1.0$	$k = 1, \varepsilon = 2.0$	$k = 2, \varepsilon = 0.1$	$k = 2, \varepsilon = 1.0$	$k = 2, \varepsilon = 2.0$
(0.4906, 0.4413)	(0.5368, 0.5422)	(0.4649, 0.4907)	(0.4887, 0.4565)	(0.4936, 0.4556)
(0.4881, 0.5103)	(0.5476, 0.4471)	(0.4796, 0.5323)	(0.4859, 0.4847)	(0.4998, 0.4491)
(0.4926, 0.4842)	(0.4590, 0.4877)	(0.4982, 0.5158)	(0.4657, 0.4191)	(0.4583, 0.4561)
(0.5446, 0.4723)	(0.5548, 0.5475)	(0.5289, 0.4659)	(0.4525, 0.4880)	(0.4686, 0.5020)
(0.4695, 0.4890)	(0.5187, 0.5319)	(0.4596, 0.5191)	(0.4810, 0.5157)	(0.4678, 0.4845)
(0.4731, 0.5272)	(0.5283, 0.4813)	(0.5122, 0.4610)	(0.5020, 0.4961)	(0.4508, 0.4669)
(0.4909, 0.5293)	(0.5039, 0.4850)	(0.4832, 0.4346)	(0.4638, 0.4735)	(0.4849, 0.4369)

Next, we observe the data utility variation of the trajectory using conditional filtering noise of Gaussian mechanism in **Table 4**. We find the data utility is approximately equal to 1 for latitude and longitude, since the conditional filtering noise directly added to trajectory data. We also observe the data utility increases as privacy budget increases and k decreases.

Given the absolute value $E = 0.8$, $E = 0.5$, and $E = 0.2$ of relative error, so the expected data utility is $U = 0.2$, $U = 0.5$, and $U = 0.8$, respectively. In the **Table 5**, the expected data utility is approximately equal to 0.2. **Table 6** shows that the expected data utility is approximately equal to 0.5. We observe that the expected data utility is approximately equal to 0.8 from **Table 7**. Therefore, the data utility of adaptive Gaussian mechanism approaches approximately to expected data utility as privacy budget increases and k decreases. We can conclude that adaptive Gaussian mechanism can achieve the same approximate expected data utility for any privacy budget.

We compare the properties of four Gauss mechanisms in **Table 8**. All Gaussian mechanisms satisfy privacy preserving monotonicity. We observe that Gaussian mechanism cannot achieve expected privacy and expected data utility. Personalized Gaussian mechanism can achieve expected privacy according to privacy preferences, but cannot achieve the expected data utility. Individual Gaussian mechanism can enhance data utility, but cannot achieve expected privacy and expected data utility. The adaptive Gaussian mechanism can get expected privacy budget and expected data utility.

Table 7. Data utility of the trajectory using adaptive Gaussian mechanism when expected data utility $U = 0.8$

Trajectory dataset	$k = 0, \varepsilon = 0.1$	$k = 0, \varepsilon = 1.0$	$k = 0, \varepsilon = 2.0$	$k = 1, \varepsilon = 0.1$
Taxi ID 7	(0.8170, 0.7904)	(0.8056, 0.8050)	(0.8168, 0.8065)	(0.7968, 0.8011)
	(0.8117, 0.8038)	(0.8100, 0.7940)	(0.8190, 0.8204)	(0.7877, 0.7941)
	(0.7994, 0.7972)	(0.8177, 0.7983)	(0.8151, 0.8069)	(0.7813, 0.7974)
	(0.7985,0.8094)	(0.8020, 0.7961)	(0.8306, 0.7956)	(0.8102, 0.8081)
	(0.8103, 0.7874)	(0.7994, 0.7936)	(0.7952, 0.8120)	(0.7939, 0.8040)
	(0.7852, 0.8034)	(0.8080, 0.8030)	(0.8048, 0.8186)	(0.8029, 0.8075)
	(0.7944, 0.8072)	(0.8056, 0.7805)	(0.8072, 0.8229)	(0.7834, 0.7915)
$k = 1, \varepsilon = 1.0$	$k = 1, \varepsilon = 2.0$	$k = 2, \varepsilon = 0.1$	$k = 2, \varepsilon = 1.0$	$k = 2, \varepsilon = 2.0$
(0.7831, 0.8093)	(0.8022, 0.7967)	(0.8098, 0.8042)	(0.7833, 0.7953)	(0.7928, 0.8031)
(0.7839, 0.8001)	(0.8114, 0.7895)	(0.8195, 0.771)	(0.7764, 0.7858)	(0.7873, 0.7963)
(0.8043, 0.8112)	(0.8105, 0.8129)	(0.8023, 0.7883)	(0.7732, 0.7893)	(0.7867, 0.7829)
(0.8012, 0.7986)	(0.8097, 0.7932)	(0.8027, 0.8073)	(0.7970, 0.7831)	(0.7740, 0.7772)
(0.7924, 0.8107)	(0.7807, 0.8111)	(0.7908, 0.7903)	(0.7942, 0.8030)	(0.8096, 0.7799)
(0.7965, 0.7993)	(0.8206, 0.7968)	(0.8009, 0.8005)	(0.7859, 0.8060)	(0.7822, 0.7681)
(0.7980, 0.7994)	(0.8178, 0.7970)	(0.8030, 0.7894)	(0.8039, 0.7878)	(0.8026, 0.7879)

Table 8. Comparison of various Gaussian mechanisms

Mechanisms	**Monotonicity**	**Expected privacy budget**	**Expected data utility**
Gaussian mechanism	Yes	No	No
Personalized Gaussian mechanism	Yes	Yes	No
Individual Gaussian mechanism	Yes	No	No
Adaptive Gaussian mechanism	Yes	Yes	Yes

7. Conclusion

Existing differential privacy mechanisms cannot achieve the expected data utility. According to definition of differential privacy, differential privacy budget is at most the logarithm of the ratio of the probability distribution of noise added to data query results. In this paper, based on the motivation of achieving expected data utility given any privacy budget, the adaptive Gaussian mechanism of achieving expected data utility is proposed by combining conditional filtering noise with expected data utility. This method can generalize to achieve adaptive Laplace mechanism, and it is good for massive datasets. The adaptive Gaussian mechanism can obtain the expected data utility of statistical analysis under the same privacy budget as Gaussian mechanism, personalized Gaussian mechanism, and individual Gaussian mechanism. In other words, adaptive Gaussian mechanism can achieve expected data utility for any privacy budget. The most important thing that the adaptive Gaussian mechanism is easy applied to engineering practice. In the future work, we will define better expected data utility metrics, and propose a general adaptive differential privacy framework based on these expected data utility metrics. Finally, general adaptive differential privacy mechanisms will be applied to the non-interactive and interactive scenarios.

References

[1] Martin Enserink and Gilbert Chin, "The end of privacy," *Science*, vol. 347, no. 6221, pp. 490-491, January, 2015. Article (CrossRef Link)

[2] Cynthia Dwork, Frank McSherry, Kobbi Nissim and Adam Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. of Theory of Cryptography Conference*, pp. 265-284, March 4-7, 2006. Article (CrossRef Link)

[3] Chi Lin, Zihao Song, Qing Liu, Weifeng Sun and Guowei Wu, "Protecting privacy for big data in body sensor networks: A differential privacy approach," in *Proc. of International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pp. 163-172, November 10-11, 2015. Article (CrossRef Link)

[4] Kendrick Boyd, Eric Lantz and David Page, "Differential privacy for classifier evaluation," in *Proc. of the 8th ACM Workshop on Artificial Intelligence and Security*, pp. 15-23, October 16, 2015. Article (CrossRef Link)

[5] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov and Moni Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Proc. of Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 486-503, May 28-June 1, 2006. Article (CrossRef Link)

[6] Cynthia Dwork and Aaron Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3-4, pp. 211-407, August, 2014. Article (CrossRef Link)

[7] Daniel L. Goroff, "Balancing privacy versus accuracy in research protocols," *Science*, vol. 347, no. 6221, pp. 479-480, January 2015. Article (CrossRef Link)

[8] Suan Landau, "Control use of data to protect privacy," *Science*, vol. 347, no. 6221, pp. 504-506, January, 2015. Article (CrossRef Link)

[9] Jerome Le Ny and George J. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341-354, September, 2013. Article (CrossRef Link)

[10] Zach Jorgensen, Ting Yu and Graham Cormode, "Conservative or liberal? Personalized differential privacy," in *Proc. of 2015 IEEE 31st International Conference on Data Engineering*, pp. 1023-1034, April 13-17, 2015. Article (CrossRef Link)

[11] Jordi Soria-Comas, Josep Domingo-Ferrer, David Sánchez and David Megías, “Individual differential privacy: A utility-preserving formulation of differential privacy guarantees,” *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1418-1429, June, 2017. Article (CrossRef Link)

[12] Yining Wang and Animashree Anandkumar, “Online and differentially-private tensor decomposition,” in *Proc. of Advances in Neural Information Processing Systems*, pp. 3531-3539, December 5-10, 2016. Article (CrossRef Link)

[13] Moritz Hardt and Aaron Roth, “Beating randomized response on incoherent matrices,” in *Proc. of Proceedings of the Forty-Fourth Annual ACM Aymposium on Theory of Computing*, pp.1255-1268, May 19-22, 2012. Article (CrossRef Link)

[14] Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta and Li Zhang, “Analyze gauss: Optimal bounds for privacy preserving principal component analysis,” in *Proc. of Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pp. 11-20, May 31-June 3, 2014. Article (CrossRef Link)

[15] Aleksandar Nikolov, Kunal Talwar and Li Zhang, “The geometry of differential privacy: The sparse and approximate cases,” in *Proc. of the Forty-fifth Annual ACM Aymposium on Theory of Computing*, pp. 351-360, June 1-4, 2013. Article (CrossRef Link)

[16] Chao Li, Gerome Miklau, Michael Hay, Andrew McGregor and Vibhor Rastogi, “The matrix mechanism: Optimizing linear counting queries under differential privacy,” *The VLDB Journal*, vol. 24, no. 6, pp. 757-781, December, 2015. Article (CrossRef Link)

[17] Martín Abadi, Andy Chu, Ian Goodfellow, H. Brendar McMahan, Ilya Mironov, Kunal Talwar and Li Zheng, “Deep learning with differential privacy,” in *Proc. of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 308-318, October 24-28, 2016. Article (CrossRef Link)

[18] Jiaqi Zhang, Kai Zheng, Wenlong Mou and Liwei Wang, “Efficient private ERM for smooth objectives,” in *Proc. of Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pp. 3922-3928, August 19-25, 2017. Article (CrossRef Link)

[19] Frank McSherry and Ilya Mironov, “Differentially private recommender systems: Building privacy into the net,” in *Proc. of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 627-636, June 28-July 1, 2009. Article (CrossRef Link)

[20] Wei Tong, Jingyu Hua and Sheng Zhong, “A jointly differentially private scheduling protocol for ridesharing services,” *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2444-2456, October, 2017. Article (CrossRef Link)

[21] Jing Yuan, Yu Zheng, Xing Xie and Guangzhong Sun, “Driving with knowledge from the physical world,” in *Proc. of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 316-324, August 21-24, 2011. Article (CrossRef Link)

Hai Liu received his M.S. degree from Guizhou University, China, in 2015. Currently, he is a Ph.D. student in School of Computer Science, Shaanxi Normal University, China. His main research interest includes privacy protection.

Zhenqiang Wu received his Ph.D. degree from Xidian University, China, in 2007. He is currently a full professor of Shaanxi Normal University, China. His research interests include computer communications networks, wireless networks, network security, anonymous communication, and privacy protection.

Changgen Peng received his Ph.D. degree from Guizhou University, China, in 2007. He is currently a full professor of Guizhou University, China. His research interests include cryptography, information security, and privacy protection.

Feng Tian received his Ph.D. degree from Xi'an Jiaotong University, China, in 2015. He is currently a lecturer of Shaanxi Normal University, China. His research interest includes location privacy protection.

Laifeng Lu received her Ph.D. degree from Xidian University, China, in 2012. She is currently an associate professor of Shaanxi Normal University, China. Her research interests include privacy protection and network security.