明治大学総合数理学部

2024 年度

卒　業　研　究

局所差分プライバシー方式 **GRR** と **OUE** の精度比較

学位請求者　先端メディアサイエンス学科

田 代 大 晴

目次

第 1 章

はじめに

1.1　はじめに

近年, コンピュータの普及に伴い, ユーザの検索履歴や位置情報などの大量のデータが, サービス事業者により効率的に利活用されることが可能となった．例えば [1] では，人々に携帯電話が普及したことによる人流データの利活用や家計簿アプリの情報を利用した消費動向分析まで行われていることが述べられている．しかし, 行動パターンや嗜好を反映しているデータは, 匿名化 [2] などを施して第三者へ送信する必要がある．

しかし，従来の匿名化やマスキングといった技術はデータ分析の技術発展により, 再識別やリンク攻撃に対して脆弱であることが知られている [2]．例えば．データを収集する事業者の内部不正 [3] やインシデントによる個人情報の流出 [4] なども多数起こっている．従って, 企業は個人データの取り扱いにおいて厳格なコンプライアンスを求められている．例えば, アメリカのセンサス局 [8] や Apple[5], Google[6], Microsoft[7] などは局所差分プライバシー (Local Differential Privacy,LDP) を導入し, 個人のプライバシーを保ちながらデータを活用する方法を模索している．

本研究では, 局所差分プライバシー (Local Differential Privacy, LDP) 方式 [10] の代表的なプロトコルのうち Generalized Randomized Response(GRR)[11] と GRR を大きなカテゴリーサイズに向けて最適化した Optimized Unary Encoding(OUE)[11] の 2 つに注目し，カテゴリーサイズによる精度の影響を評価する．実施したアンケートとオープンデータを用いた実験で有用性を評価する．

第 2 章

関連研究

2.1　局所差分プライバシー

局所差分プライバシー [10] は, 個々のデータに対してプライバシー保護を行う手法である．ユーザはデータを集計者の持つサーバに送信する前に, デバイス上でデータを摂動する．この加工により, 個人が特定されるリスクを大幅に低減しつつ, 集計データの有用性を保持する．

X を入力の集合, Y を出力の集合とする．M を入力 $x \in X$ に対して $y \in Y$ を出力するランダムアルゴリズムとする．また，ランダマイズアルゴリズムが保証するプライバシーを表す値としてプライバシー予算 ϵ を定義する．任意の 2 つの入力 $x, x' \in X$ と任意の出力 $y \in Y$ に対して,

$$\Pr[M(x, \epsilon) = y] \leq e^{\epsilon} \Pr[M(x', \epsilon) = y]$$

が成立するとき, ランダムアルゴリズム M は ϵ-局所差分プライバシー (ϵ-LDP) を満たすという [10].

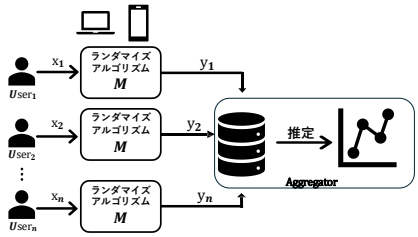

図 2.1　ユーザと集計者の関係

2.2　Generalized Randomized Response(GRR)

GRR[11] は，二択の選択に関するプライバシーを保護する Randomized Reponse[9] を 2 以上の多値属性に拡張したプロトコルである．GRR は，離散値の入力に対して適用する．入力を確率的に他の値に遷移し出力することで，回答者のプライバシーを保護し，集計値の有用性を確率的に保証する．GRR のユーザ側と集計者側のアルゴリズムをそれぞれ Algorithm 1, 2 に示す．

ユーザは，d 種類の選択肢から一つの値を選び，d とプライバシー予算 ε に基づいた遷移確率で値を変化させ送信する．

集計者は，n 人のユーザから出力を収集し，各 $i \in [d]$ の度数を推定する．収集した出力の中で i の度数を f_i' とし，i の真の度数を f_i とする．最尤推定法では，ユーザ全体のうち i を保有する平均 $f_i p$ 人が i を出力し，i 以外を保有する平均 $(n - f_i)q$ 人が $\frac{1}{d-1}$ の確率で i を出力するため f_i' の期待値は，

$$f_i' = f_i p + \frac{(n - f_i)q}{d - 1}$$

となる．上式から真の度数 f_i の最尤値は，

$$L[f_i] = \frac{f_i'(d-1) - nq}{p(d-1) - q} = \frac{f_i'(d-1) - n(1-p)}{dp - 1}$$

Algorithm 1 GRR-User

Input: User's selected value v from d possible values, domain size d, privacy budget ϵ

Output: Privacy-protected value y using GRR

Encode v: Set $x = v$

Perturb x:

$$y = \begin{cases} x & \text{w/p } p = \frac{e^\epsilon}{e^\epsilon + d - 1}, \\ x' \in [d] \setminus \{x\} & \text{w/p } q = \frac{1}{e^\epsilon + d - 1}. \end{cases}$$

Algorithm 2 GRR-Aggregator

Input: Reported values from users $\{y_1, y_2, \ldots, y_n\}$, domain size d, privacy budget ϵ

Output: Maximum likelihood estimate $L[f_i]$ for each value $i \in [d]$

$p \leftarrow \frac{e^\epsilon}{e^\epsilon + d - 1}$, $q \leftarrow \frac{1}{e^\epsilon + d - 1}$

Initialize: $f_i' \leftarrow 0 \ \forall i \in [d]$

for each user $j = 1$ to n **do**

    $f_{y_j}' \leftarrow f_{y_j}' + 1$

end for

for $i = 1$ to d **do**

    $L[f_i] \leftarrow \frac{f_i'(d-1) - n(1-p)}{dp - 1}$

end for

Return $\{L[f_i] \mid i \in [d]\}$

2.2.1 数値例

離散値の集合を $d = 3$ の場合で考える．$|d|$ はユーザが選択可能な値の種類の数を示し，例えば選択肢は $\{A, B, C\}$ とする．プライバシー予算を $\epsilon = 2$ とし，ユーザ数は $n = 10$ とする．

■ステップ 1: 真の入力分布　ユーザが選択した真の値を以下のように仮定する：

$$\text{入力: } \{A, A, A, B, B, C, C, C, C, C\}.$$

したがって，真の度数 f_i は次のようになる：

$$f_A = 3, \quad f_B = 2, \quad f_C = 5.$$

■ステップ 2: 確率の計算　GRR プロトコルでは，以下の確率が計算される：

$$p = \frac{e^{\epsilon}}{e^{\epsilon} + d - 1} = \frac{e^2}{e^2 + 2} \approx 0.737,$$

$$q = \frac{1}{e^{\epsilon} + d - 1} = \frac{1}{e^2 + 2} \approx 0.107.$$

■ステップ 3: 出力データの収集　各ユーザが GRR プロトコルを適用し，摂動した値を送信する．仮に以下のような摂動結果が得られたとする：

$$\text{出力: } \{A, A, C, B, B, C, C, A, C, C\}.$$

収集された出力の頻度（f_i'）を計算すると：

$$f_A' = 3, \quad f_B' = 2, \quad f_C' = 5.$$

■ステップ 4: 真の度数の推定　GRR プロトコルでは，真の度数 $L[f_i]$ を以下の式で推定する：

$$L[f_i] = \frac{f_i'(d-1) - n(1-p)}{dp - 1}.$$

それぞれ計算すると：

$$L[f_A] = \frac{3 \cdot 2 - 10 \cdot (1 - 0.737)}{3 \cdot 0.737 - 1} \approx 2.843,$$

$$L[f_B] = \frac{2 \cdot 2 - 10 \cdot (1 - 0.737)}{3 \cdot 0.737 - 1} \approx 1.374,$$

$$L[f_C] = \frac{5 \cdot 2 - 10 \cdot (1 - 0.737)}{3 \cdot 0.737 - 1} \approx 5.78.$$

■ステップ 5: 真の度数との誤差比較　真の度数 f_i と推定された真の度数 $L[f_i]$ を以下の表にまとめる：

値	真の度数 f_i	推定された度数 $L[f_i]$	誤差
A	3	2.843	-0.157
B	2	1.374	-0.626
C	5	5.78	$+0.78$

■ステップ 6: 結果の解釈　この例では，GRR プロトコルに基づいて摂動されたデータから真の度数を推定した．推定された値は真の値と比較的近いが，プライバシー保護のために若干の誤差が発生している．

2.3 Optimized Unary Encoding(OUE)

Optimized Unary Encoding[11] は，離散値の入力データを Unary Encoding(One-Hot Encoding) をした後に，誤差分散を最小化するよう最適化された遷移確率（0 から 1 に変化する確率）および 維持確率（1 がそのまま維持される確率）を用いてデータを摂動（ランダム化）するプロトコルである．OUE のユーザ側と集計者側のアルゴリズムをそれぞれ Algorithm 3, 4 に示す．

Unary Encoding の利点として，各データポイントを個別のビットで表現するため，情報の損失を最小限に抑えられる点が挙げられる．そのため，大規模なドメインサイズを持つデータに対しても，効率的かつ高精度

な処理が可能になる．Wang らによって，$d > 3e^{\varepsilon} + 2$ で OUE の分散，つまり，推定誤差が GRR より小さくなることが示されている [11].

ユーザーは，各ビットを摂動し送信する．集計者は，n 人のユーザから出力を収集し，各 $i \in [d]$ の度数を推定する．収集した出力の中で i の度数を f_i' とし，i の真の度数を f_i とする．最尤推定法では，ユーザ全体のうち i を保有する平均 $f_i p$ 人が i を出力し，i 以外を保有する平均 $(n - f_i)q$ 人が i を出力するため f_i' の期待値は，

$$f_i' = f_i p + (n - f_i)q$$

となる．上式から真の度数 f_i の最尤値は，

$$L[f_i] = \frac{f_i' - nq}{p - q}$$

Algorithm 3 OUE-User

Input: User's private value v from domain $[d]$, domain size d, privacy budget ϵ
Output: Perturbed binary vector B'
Encode v: Generate a length-d binary vector B such that $B[v] = 1$ and $B[i] = 0$ for $i \neq v$
Perturb B:
for each bit $B[i]$ in B **do**
    If $B[i] = 1$, set $B'[i] = 1$ **w/p** $p = \frac{1}{2}$, otherwise $B'[i] = 0$
    If $B[i] = 0$, set $B'[i] = 1$ **w/p** $q = \frac{1}{e^{\epsilon}+1}$, otherwise $B'[i] = 0$
end for
Return B'

Algorithm 4 OUE-Aggregator

Input: Reported binary vectors from users $\{B_1', B_2', \ldots, B_n'\}$, domain size d, privacy budget ϵ
Output: Maximum likelihood estimate $L[f_i]$ for each value $i \in [d]$
$p \leftarrow \frac{1}{2}$, $q \leftarrow \frac{1}{e^{\epsilon}+1}$
Initialize: $f_i' \leftarrow 0 \ \forall i \in [d]$
for each user $j = 1$ to n **do**
    for $i = 1$ to d **do**
        $f_i' \leftarrow f_i' + B_j'[i]$
    end for
end for
for $i = 1$ to d **do**
    $L[f_i] \leftarrow \frac{f_i' - nq}{p - q}$
end for
Return $\{L[f_i] \mid i \in [d]\}$

2.3.1 数値例

離散値の集合を $d = 3$ の場合で考える．d はユーザが選択可能な値の種類の数を示し，例えば選択肢は $\{A, B, C\}$ とする．プライバシー予算を $\epsilon = 2$ とし，ユーザ数は $n = 10$ とする．

■ステップ 1: 真の入力分布　ユーザが選択した真の値を以下のように仮定する：

$$\text{入力: } \{A, A, A, B, B, C, C, C, C, C\}.$$

したがって，真の度数 f_i は次のようになる：

$$f_A = 3, \quad f_B = 2, \quad f_C = 5.$$

■ステップ 2: 確率の計算　OUE プロトコルでは，次の確率が使用される：

$$p = \frac{1}{2}, \quad q = \frac{1}{e^{\epsilon} + 1} = \frac{1}{e^2 + 1} \approx 0.119.$$

■ステップ 3: 出力データの収集　各ユーザが選択した値を One-Hot ベクトルで表現する：

$$\text{例: ユーザが } A \text{ を選択した場合，入力ベクトルは } [1, 0, 0].$$

OUE プロトコルを適用して摂動したベクトルを生成する．摂動結果の例を以下に示す：

$$\text{ユーザ 1 (A): } [1, 0, 0] \to [1, 0, 0],$$
$$\text{ユーザ 2 (A): } [1, 0, 0] \to [1, 0, 1], \ldots$$

10 人のユーザによる出力ベクトルの集計結果は次のようになると仮定する：

$$f'_A = 6, \quad f'_B = 4, \quad f'_C = 7.$$

■ステップ 4: 真の度数の推定　OUE プロトコルでは，真の度数 $L[f_i]$ を以下の式で推定する：

$$L[f_i] = \frac{f'_i - nq}{p - q}.$$

それぞれ計算すると：

$$L[f_A] = \frac{6 - 10 \cdot 0.119}{0.5 - 0.119} \approx 12.62,$$
$$L[f_B] = \frac{4 - 10 \cdot 0.119}{0.5 - 0.119} \approx 7.374,$$
$$L[f_C] = \frac{7 - 10 \cdot 0.119}{0.5 - 0.119} \approx 15.25.$$

■ステップ 5: 真の度数との誤差比較　真の度数 f_i と推定された真の度数 $L[f_i]$ を以下の表にまとめる：

値	真の度数 f_i	推定された度数 $L[f_i]$	誤差
A	3	12.32	+9.32
B	2	7.374	+5.374
C	5	15.25	+10.52

■ステップ 6: 結果の解釈　この例では，OUE プロトコルに基づいて摂動化されたデータから真の度数を推定した．真の度数と推定された値の間に大きな誤差が生じている．これは，OUE プロトコルがカテゴリーサイズの小さなデータに対しては効率的ではないためである．

第 3 章

実験

3.1　実験目的

プライバシー予算とカテゴリーサイズを変化させたときの GRR と OUE の有用性について評価をし，[11] の結果を検討する．データセットに合わせた適切な局所差分プライバシー方式を選択する．

3.2　使用データセット

本実験では，著者が収集したお茶の銘柄に関するレビューデータ Dataset1 とオンラインショッピングサービスの購入履歴のデータセット Dataset2[12] を使用する．

Dataset1 は，14 人の回答者による 5 段階の評価である．図 3.1 にお茶の銘柄の評価値の頻度分布を示す．お茶の銘柄は，0: お〜いお茶，1: 伊右衛門, 2: 綾鷹, 3: 生茶, 4: 明大茶である．

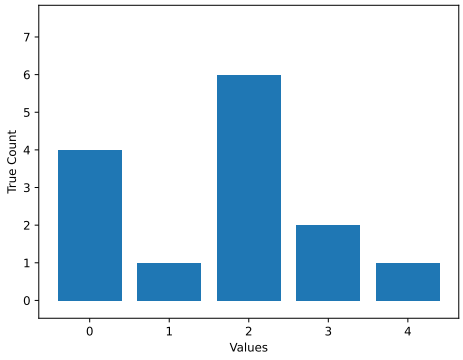

図 3.1　Dataset1 の頻度分布（お茶の銘柄の評価値）

Dataset2 は，49,742 件のズボンの購入履歴である．43 種類のズボンの購入履歴の頻度分布を図 3.2 に示す．

3.3　実験 1:GRR の適用例

真の頻度分布，摂動後の頻度分布，最尤推定後の頻度分布の例を図 3.3 と図 3.4 に示す．実験で使用するプライバシー予算は，$\epsilon = \log(20)$ とする．

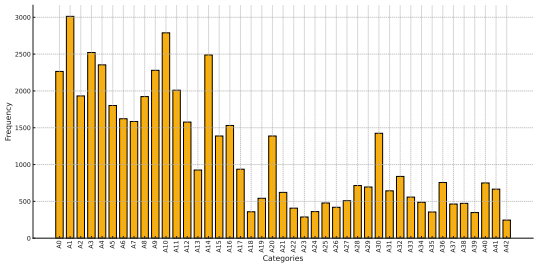

図 3.2　Dataset2 の頻度分布（ズボンの購入履歴）

図 3.3　Dataset1 に GRR を適用した際の頻度分布の例

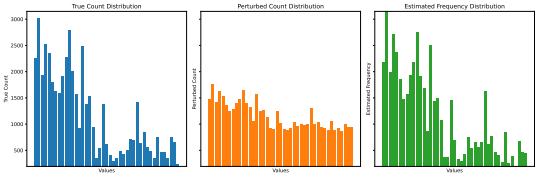

図 3.4　Dataset2 に GRR を適用した際の頻度分布の例

3.4　実験 2:OUE の適用例

真の頻度分布，摂動後の頻度分布，最尤推定後の頻度分布の例を図 3.5 と図 3.6 に示す．実験で使用するプライバシー予算は，$\epsilon = \log(20)$ とする．

図 3.5　Dataset1 に OUE を適用した際の頻度分布の例

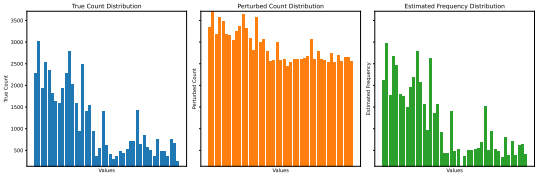

図 3.6　Dataset2 に OUE を適用した際の頻度分布の例

3.5　実験 3:GRR と OUE の精度比較

プライバシー予算を変化させた場合の GRR と OUE の有用性を各プライバシー予算につき，摂動から集計までの流れを 100 回繰り返し，MSE によって比較を行う．図 3.7 と図 3.8 に結果を示す．

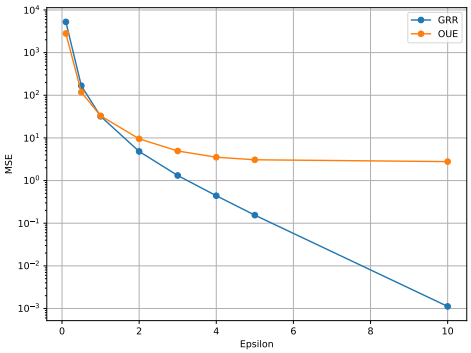

図 3.7　GRR と OUE の誤差比較（Dataset1）

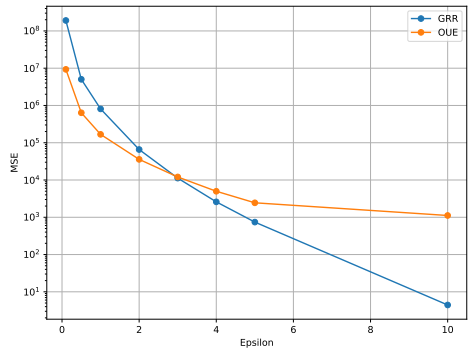

図 3.8　GRR と OUE の誤差比較（Dataset2）

3.6　考察

どちらのデータセットに対しても，プライバシー予算が大きくなると，誤差が小さくなることが確認できる．また，OUE プロトコルが，カテゴリーサイズが大きいデータセットに対しては，頻度分布を精度よく推定できていることが，図 3.6 において確認できる．

GRR と OUE による推定値の誤差分散の比較 [11] より，Dataset1 では常に，Dataset2 では $\epsilon < \ln(41/3) \approx 2.615$ の条件で OUE の MSE が小さくなると予想される．

Dataset2 においては，プライバシー予算が 2 から 3 の間のときに，OUE と GRR の誤差が逆転していることが図 3.7 から確認できる．

一方で Dataset1 では，理論上は常に GRR の誤差が OUE を下回ると予測されるが，実験結果から ε が 1 を下回っている範囲では，OUE の推定誤差がわずかに小さくなっていることが図 3.8 から分かる．

この結果は，Dataset1 の回答数の少なさと回答の偏りの大きさが，理論通りの結果とならなかった原因であると考えられる．

第 4 章

おわりに

4.1　おわりに

本研究では，局所差分プライバシー方式である GRR と OUE について，自作のデータセットおよびオープンデータを用いて，GRR および OUE の推定誤差を比較し，どちらのプロトコルが優れているかを評価した．結果，カテゴリーサイズが小さく回答数が少ないものには GRR，カテゴリーサイズが大きく回答数も多いものには OUE が比較的有効であることがわかった．

カテゴリーサイズが小さく回答数が多い場合や，カテゴリーサイズが大きく回答数が少ない場合の検証は未実施である．および，データの偏りや回答数に応じた最適なプロトコルの考察も今後の課題とする．

2024.02 Taisei Tashiro.

謝辞

本研究を行うにあたり，多くの方より御指導いただきました．特に明治大学総合数理学部先端メディアサイエンス学科，菊池浩明教授に深く感謝申し上げます．研究についてアドバイスをいただいた菊池研究室の皆様に深く感謝の意を表するとともに，謝辞とさせていただきます．

参考文献

[1] 総務省統計改革実行推進室, “ビッグデータの利活用について” (https://www.soumu.go.jp/main_content/000828345.pdf, 2024 年 12 月参照).

[2] 南和宏, “再識別リスク評価と匿名化の展望”. コンピュータセキュリティシンポジウム 2016 論文集, 2016.2: pp. 166-172, 2016.

[3] 情報処理推進機構,「企業の内部不正防止体制に関する実態調査」報告書 (https://www.ipa.go.jp/security/reports/economics/ts-kanri/20230406.html, 2024 年 12 月参照).

[4] シャープ株式会社, シャープ公式オンラインストア「COCORO STORE」・食材宅配サービス「ヘルシオデリ」における不正アクセスによる個人情報流出に関するお詫びと調査結果のお知らせ (https://corporate.jp.sharp/info/notices/241030-a.html, 2024 年 12 月参照).

[5] Apple Differential Privacy Team, “Differential Privacy Overview” (https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf, 2024 年 12 月参照).

[6] Ú. Erlingsson, V. Pihur, A. Korolova, “RAPPOR :Randomized Aggregatable Privacy-Preserving Ordinal Response”, ACM, pp. 1054-1067, 2014.

[7] Microsoft, “Collecting telemetry data privately”(https://www.microsoft.com/en-us/research/blog/collecting-telemetry-data-privately/, 2024 年 12 月参照).

[8] M. Christ, S. Radway, and S. M. Bellovin. “Differential Privacy and Swapping: Examining De-Identification’s Impact on Minority Representation and Privacy Preservation in the U.S. Census”, 2022 IEEE Symposium on Security and Privacy (SP), pp. 457-472, 2022.

[9] S. L. Warner, “Randomized response: A survey technique for eliminating evasive answer bias”, Journal of the American Statistical Association, pp. 63-69, 1965.

[10] J. C. Duchi, M. I. Jordan, M. J. Wainwright, “Local privacy and statistical minimax rates”, FOCS, pp. 429-438, 2013.

[11] T. Wang, Li, J. Blocki and N. Li. “Locally differentially private protocols for frequency estimation.” 26th USENIX Security Symposium (USENIX Security 17). pp.729-745, 2017.

[12] “Clickstream Data for Online Shopping.” UCI Machine Learning Repository, https://doi.org/10.24432/C5QK7X, 2019.