安全性と有用性の実験評価に基づく局所差分プライバシー方式 GRR,OUE,Hadamard Response の最適化

小練大智　藪悠馬　菊池浩明 †

明治大学総合数理学部 †

1　はじめに

Web サービス等の個人活動データは有用である一方, 個人の再識別や属性推定のリスクが伴う. この課題に対し, 統計量推定を可能にする局所差分プライバシー Local Differential Privacy(LDP）が注目されている.

LDP に基づく頻度推定では，Generalized Randomized Response（GRR）[1, 2], Optimized Unary Encoding（OUE）[3]，Hadamard Response（HR）[4] などが広く知られている. しかし，その推定精度は与えられたデータの種類やかけるプライバシー費用に大きく依存する為, それらの使い分けは自明ではない. 本研究では,ϵ とカテゴリ数 k を変化させ, GRR, OUE, HR の推定精度と通信量を比較し, 最適な手法選択の指針を得ることを目的とする.

2　先行研究

2.1　局所差分プライバシー

LDP では，ユーザは元の値 $x \in X$ を直接送信せず，ランダム化機構 Q により生成された出力 z を送信する. 任意の $x, x' \in X$ と任意の出力 z に対して

$$Q(z \mid x) \le e^{\epsilon}\, Q(z \mid x')$$

が成り立つとき，Q は ϵ-LDP を満たす.

2.2　摂動化手法

2.2.1　GRR

GRR [1, 2] は最も基本的な LDP メカニズムである.k 値のカテゴリ値 $x \in [k]$ を

$$Q(z \mid x) = \begin{cases} \dfrac{e^{\epsilon}}{e^{\epsilon} + k - 1}, & \text{if } z = x, \\ \dfrac{1}{e^{\epsilon} + k - 1}, & \text{if } z \neq x \end{cases}$$

に従って摂動化する.

2.2.2　OUE

OUE [3] はカテゴリを長さ k のワンホットベクトル $(x_1, \ldots, x_k)$ に符号化し，各ビットを独立にランダム化する方式である．$x_i \in \{0, 1\}$ とすると,

$$Q(z_i = 1 \mid x_i) = \begin{cases} \dfrac{1}{2}, & \text{if } x_i = 1, \\ \dfrac{1}{e^{\epsilon} + 1}, & \text{if } x_i = 0 \end{cases}$$

と摂動化する．0 側にもノイズを分散させることで分散が抑えられ，高プライバシー条件でも安定した推定が得られる.

2.2.3　HR

HR [4] は，k 以上の最小の 2 の冪を $K = 2^{\lceil \log_2(k+1) \rceil}$ とし，$K \times K$ Hadamard 行列 $H_K = (h_{ij})$ を用いてカテゴリを符号化し，符号パターンに基づき 1 つのインデックス z のみを報告する.

カテゴリ x に対応する列集合 C_x を

$$C_x = \{\, i \in \{1, \ldots, K\} \mid (H_K)_{x+1,i} = +1 \,\}$$

と定義する．x の摂動化 $z \in [K]$ は,C_x についての確率

$$Q(z \mid x) = \begin{cases} \dfrac{e^{\epsilon}}{s e^{\epsilon} + K - s} & \text{if } z \in C_x, \\ \dfrac{1}{s e^{\epsilon} + K - s} & \text{if } z \notin C_x, \end{cases}$$

に従って定める．ここで $s = K/2$ とする.

3　評価実験

3.1　実験方法

GRR，OUE，HR を対象に，$\epsilon = 0.1, 1, 2, ..., 10$ および $k = 594, ..., 2580$ と変化させて頻度を推定する．データは研究室で取得した独自アンケートデータ（`lab_data` とする）（性別 × SNS 利用時間：$k = 6$）と，Kaggle の公開データ [5]（`student_data`）（性別 × 専攻：$k = 18$）の 2 種類である．`student_data` では，ϵ 依存を見るための 1000 件固定サンプルを用いた.

---

†

Optimization of Local Differential Privacy Mechanisms Based on Experimental Evaluation of Safety and Utility: GRR, OUE, and Hadamard Response,Daichi Koneri,Yuma Yabu and Hiroaki Kikuchi, School of Interdisciplinary Mathematical Science, Meiji University.

各データに LDP を適用後，手法ごとの推定アルゴリズムによりカテゴリ値 x_i の確率分布 $\hat{p}_i$ を推定する．真の分布 p に対する平均二乗誤差（MSE）

$$\mathrm{MSE} = \frac{1}{k}\sum_{i=1}^{k}(\hat{p}_i - p_i)^2$$

とする. 各条件で 100 回実施して平均値と標準偏差を算出する.

3.2 実験結果

図 1 に student_data におけるプライバシー費用 ϵ に対する推定誤差, 図 2 にカテゴリ数 k に対する通信量の変化を示す.$\pm\sigma$ で 68% の信頼区間をエラーバーで示す. 図 1 より，$\epsilon \leq 1.85$ の範囲においては OUE,$\epsilon > 1.85$ では,GRR が最小誤差を示した. 一方，GRR と HR を比較すると，$\epsilon = 1.37$ 辺りに境界があることが確認された.

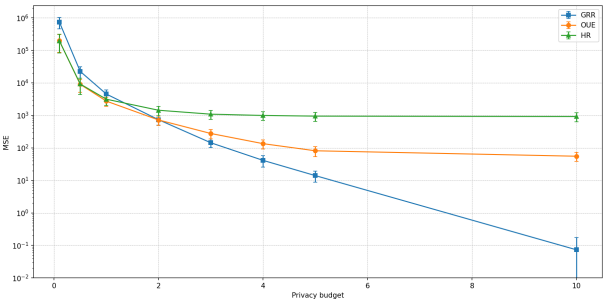

図 1 プライバシー費用 ϵ に対する推定誤差（student_data, k=18 categories,n=1000)

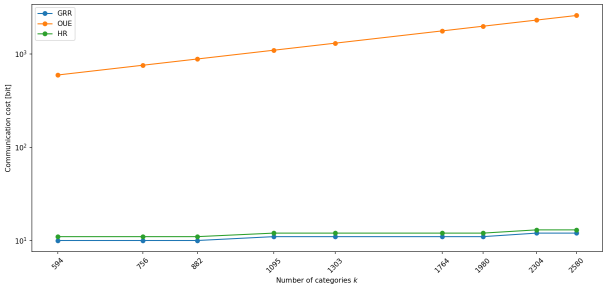

図 2 カテゴリ数 k に対する通信量

表 1 通信量と推定誤差の優位性まとめ（ϵ）

手法	通信量 [bit]	誤差最小の ϵ 条件	最適条件
GRR	$\lceil \log_2 k \rceil$	$1.85 \leq \epsilon$	ϵ が大きい
OUE	k	$\epsilon \leq 1.37$	ϵ が小さい k が小さい
HR	$\lceil \log_2 k \rceil + 1$	$1.37 \leq \epsilon \leq 1.85$	ϵ が小さい k が大きい

3.3 考察

[3] では $k < 3e^{\varepsilon} + 2$ の時に GRR(DE) が最適であることが示されている. 従って, 実験データでは

$$\varepsilon = \ln\left(\frac{16}{3}\right) \approx 1.67$$

の時に GRR の誤差が小さくなるはずだが, 本実験では 1.85 とやや大きい値を示した.OUE は $\epsilon \leq 1.85$ において最も小さい MSE を示したが，通信量がカテゴリ数 k に対して線形に増加するという制約を持つ. 一方,GRR および HR は通信量が $\log k$ に比例するため, カテゴリ数の大きなデータを扱う場合に有利である．ϵ ごとに GRR,OUE,HR の推定精度及び通信量を比較し, 最適な条件を表 1 に整理する.

4 結論

実験データに基づいて, 代表的な LDP 方式の精度と通信量を評価した. 本実験の結果に基づき,$\epsilon \leq 1.85$ かつ小さなカテゴリでは OUE，$\epsilon < 1.37$ かつ大きなカテゴリでは HR，$\epsilon > 1.37$ ではカテゴリ数によらず GRR を選択することが，推定精度と通信量を考慮した結果, 最適であると結論づける.

参考文献

[1] Stanley L. Warner, “Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias,” Journal of the American Statistical Association, 1965.

[2] Peter Kairouz, Sewoong Oh, Pramod Viswanath, “Extremal Mechanisms for Local Differential Privacy,” NeurIPS, 2014.

[3] Tianhao Wang et al., “Locally Differentially Private Protocols for Frequency Estimation,” USENIX Security, 2017.

[4] Jayadev Acharya et al., “Hadamard Response,” AISTATS, 2019.

[5] Ahmjadjad, *Enhanced Student Habits Performance Dataset*, Kaggle, 2025. Available at: https://www.kaggle.com/datasets/ahmjadjad/enhanced-student-habits-performance-dataset/data (accessed 2025-12).