2023年2月12日

修士論文発表会

EMアルゴリズムを用いた Key-Valueデータについての 局所差分プライバシプロトコルの提案

堀込 光

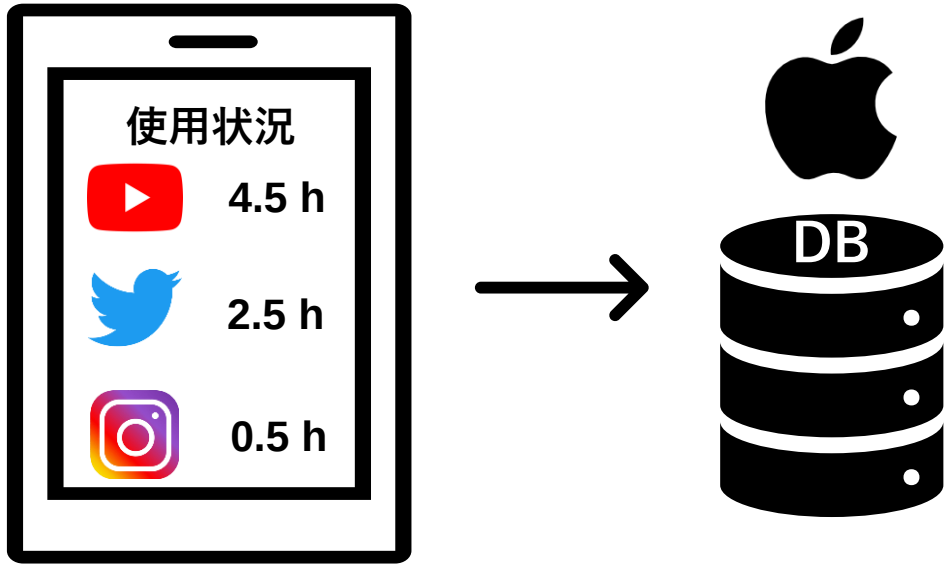
斉藤（菊池）研究室

研究背景

Appleがお客様から収集する個人データ

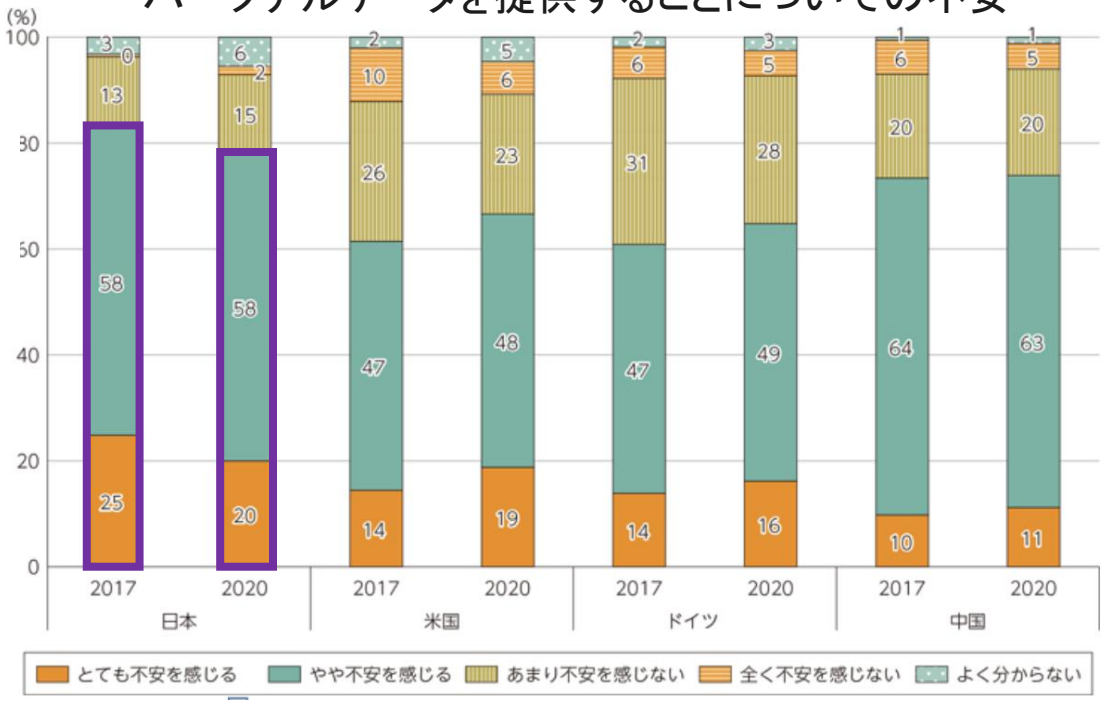
- **使用状況データ。**Appleサービス内でのアプリケーションの起動など、Appleの製品やサービスにおけるお客様のアクティビティと使用に関するデータ（閲覧履歴、検索履歴、製品の操作、クラッシュデータ、パフォーマンスなどの診断データ、その他の使用状況データを含む）

・サービス企業は，プライバシー情報を収集している



Appleプライバシーポリシー
https://www.apple.com/legal/privacy/jp/

サービス・アプリケーションの利用に当たって
パーソナルデータを提供することについての不安

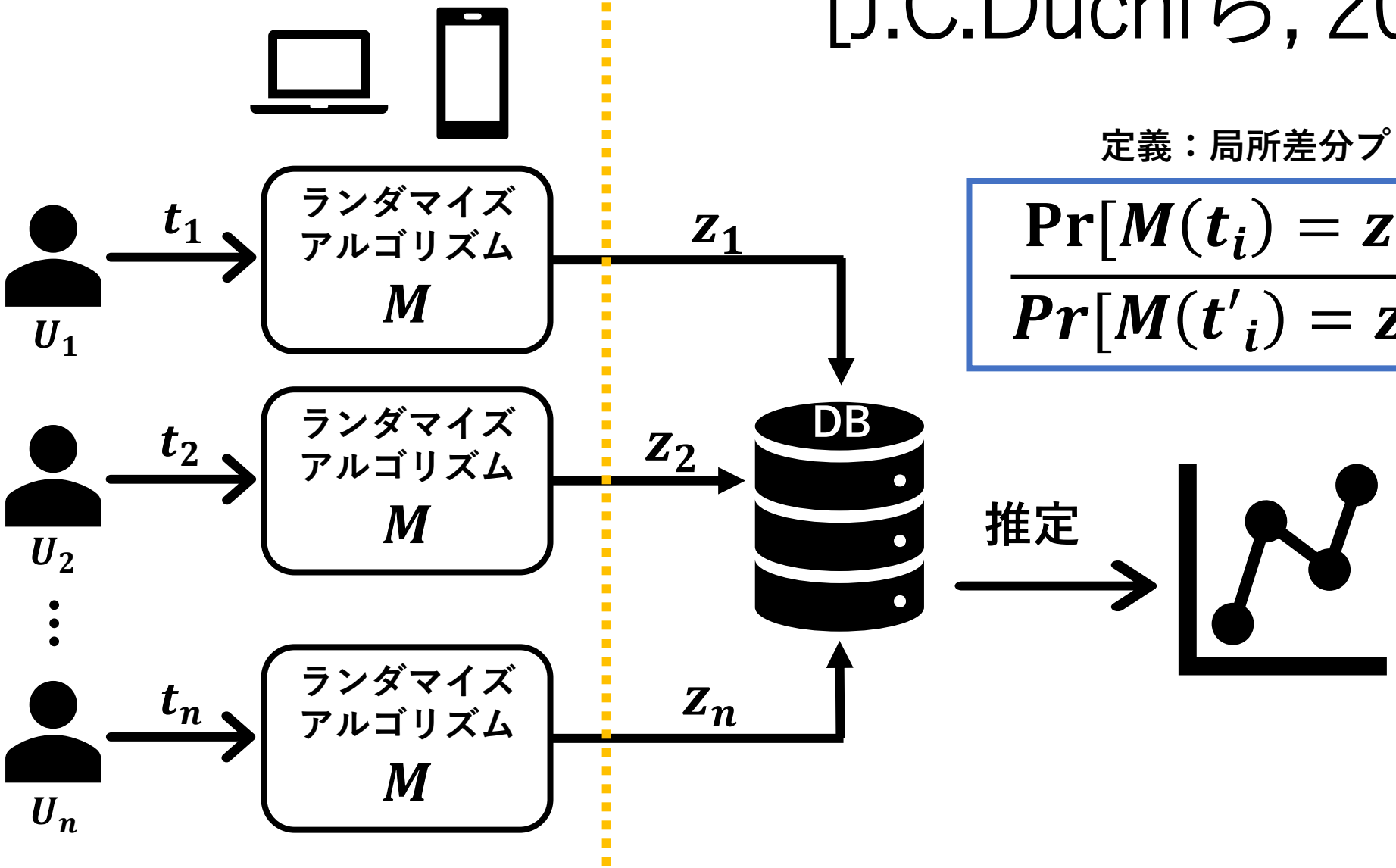


(出典)総務省(2022)
「データの流通環境等に関する消費者の意識に関する調査研究」

・日本人の約8割の利用者は自身の情報を提供することに不安を抱いている

局所差分プライバシ（LDP）

[J.C.Duchiら, 2013]



定義：局所差分プライバシ

$$\frac{\Pr[M(t_i) = z_i]}{Pr[M(t'_i) = z_i]} \leq e^{\varepsilon}$$

PrivKV [Q.Yeら, 2019]

$$\begin{cases} p_1 = \dfrac{e^{\varepsilon_1}}{1+e^{\varepsilon_1}} \\ q_1 = \dfrac{1}{1+e^{\varepsilon_1}} \end{cases} \quad \begin{cases} p_2 = \dfrac{e^{\varepsilon_2}}{1+e^{\varepsilon_2}} \\ q_2 = \dfrac{1}{1+e^{\varepsilon_2}} \end{cases}$$

$$\varepsilon = \varepsilon_1 + \varepsilon_2$$

入力：$< k'_a, v'_a >$

$< YouTube, 0.5 >$

$< k'_{YouTube}, v'_{YouTube} > \rightarrow < 1, 0.5 >$

$< YouTube, * >$

$< k'_{YouTube}, v'_{YouTube} > \rightarrow < 0, random[-1, 1] >$

$< k'_a, v^*_a >$

valueの2値化

$$v^*_a = \begin{cases} 1 & w/p \quad \dfrac{1+v'_a}{2} \\ -1 & w/p \quad \dfrac{1-v'_a}{2} \end{cases}$$

$$E[v^*_a] = \frac{1+v'_a}{2} - \frac{1-v'_a}{2} = v'_a$$

$< k'_a, v^+_a >$

Randomized Response

$$v^+_a = \begin{cases} v^*_a & w/p \quad p_2 \\ -v^*_a & w/p \quad q_2 \end{cases}$$

$< k^*_a, v^+_a >$

Randomized Response

$k'_a = 1$のとき，

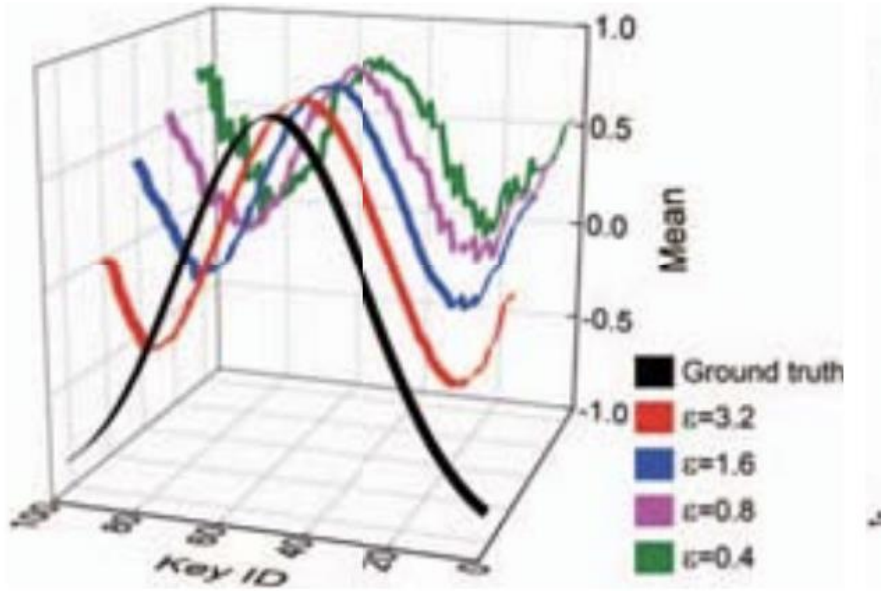
$$< k^*_a, v^+_a > = \begin{cases} < 1, v^+_a > & w/p \quad p_1 \\ < 0, 0 > & w/p \quad q_1 \end{cases}$$

$k'_a = 0$のとき，

$$< k^*_a, v^+_a > = \begin{cases} < 0, 0 > & w/p \quad p_1 \\ < 1, v^+_a > & w/p \quad q_1 \end{cases}$$

PrivKVの問題点：推定精度

$< k'_a, v'_a >=< 0, random[-1, 1] >$



PrivKV：平均値の推定値分布

k_1	v_1
1	-1
0	*
0	*
⋮	
0	*
0	*
m_1	-1

→

k_1^*	v_1^+
1	-1
1	-1
0	0
⋮	
1	1
0	0
$\hat{m}_1$	0

valueの平均値推定

$n'_1 = count(v_a^+ = 1)$
$n'_2 = count(v_a^+ = -1)$
$N = n'_1 + n'_2$

$\mathrm{L}(\hat{n}_1) = \frac{N(p_2-1)+n'_1}{2p_2-1}$
$\mathrm{L}(\hat{n}_2) = N - \hat{n}_1$

key= aの推定平均値$\hat{m}_a$は,

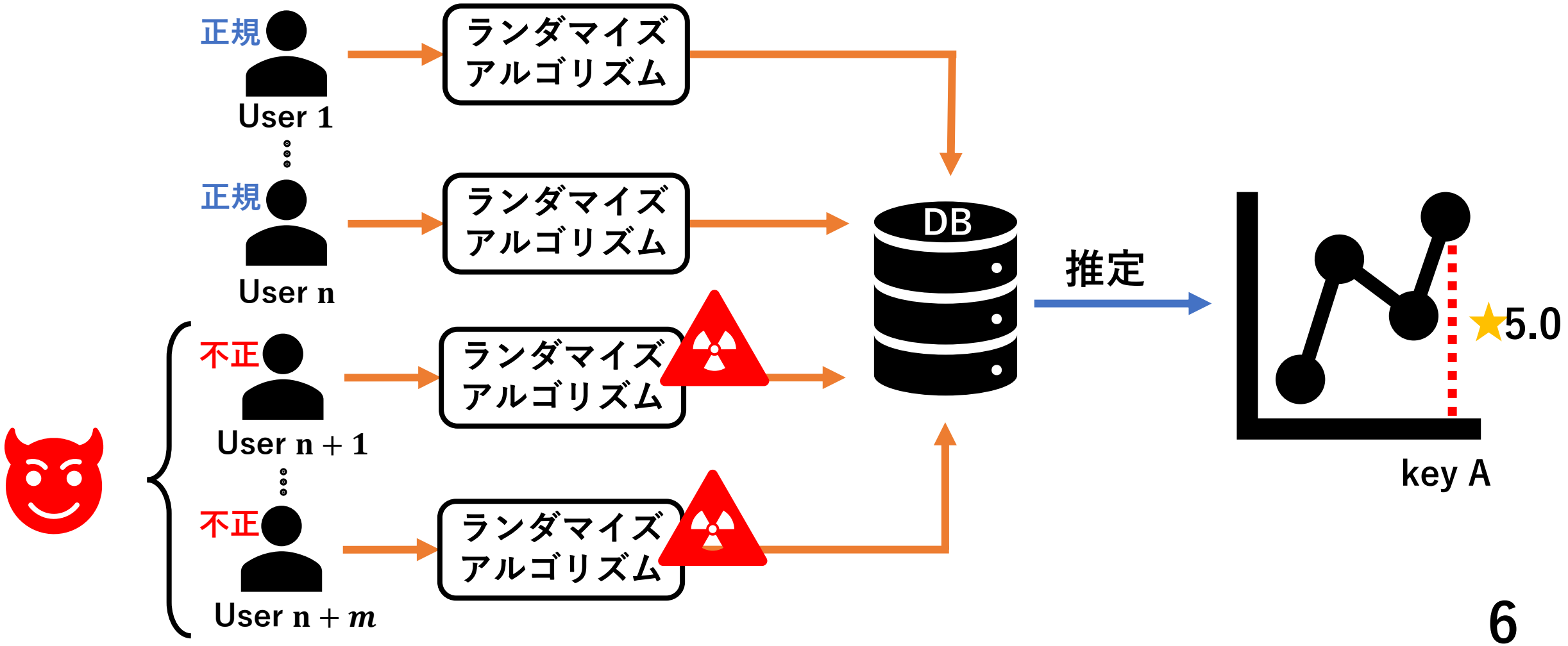
$\hat{m}_a = \frac{\hat{n}_1-\hat{n}_2}{\hat{n}_1+\hat{n}_2}$

$N = 101$, $n'_1 = 50$, $n'_2 = 51$の時,
$\hat{n}_1 \approx 50$, $\hat{n}_2 \approx 51$となり,
$\hat{m}_1 \approx 0$

未回答を考慮できていないため
度数の小さなkeyでは,
平均値が0に近似する

局所差分プライバシのポイズニング攻撃 [X.cao, et al., 2021]

攻撃者の目的：特定のアイテムに対する推定値を操作すること

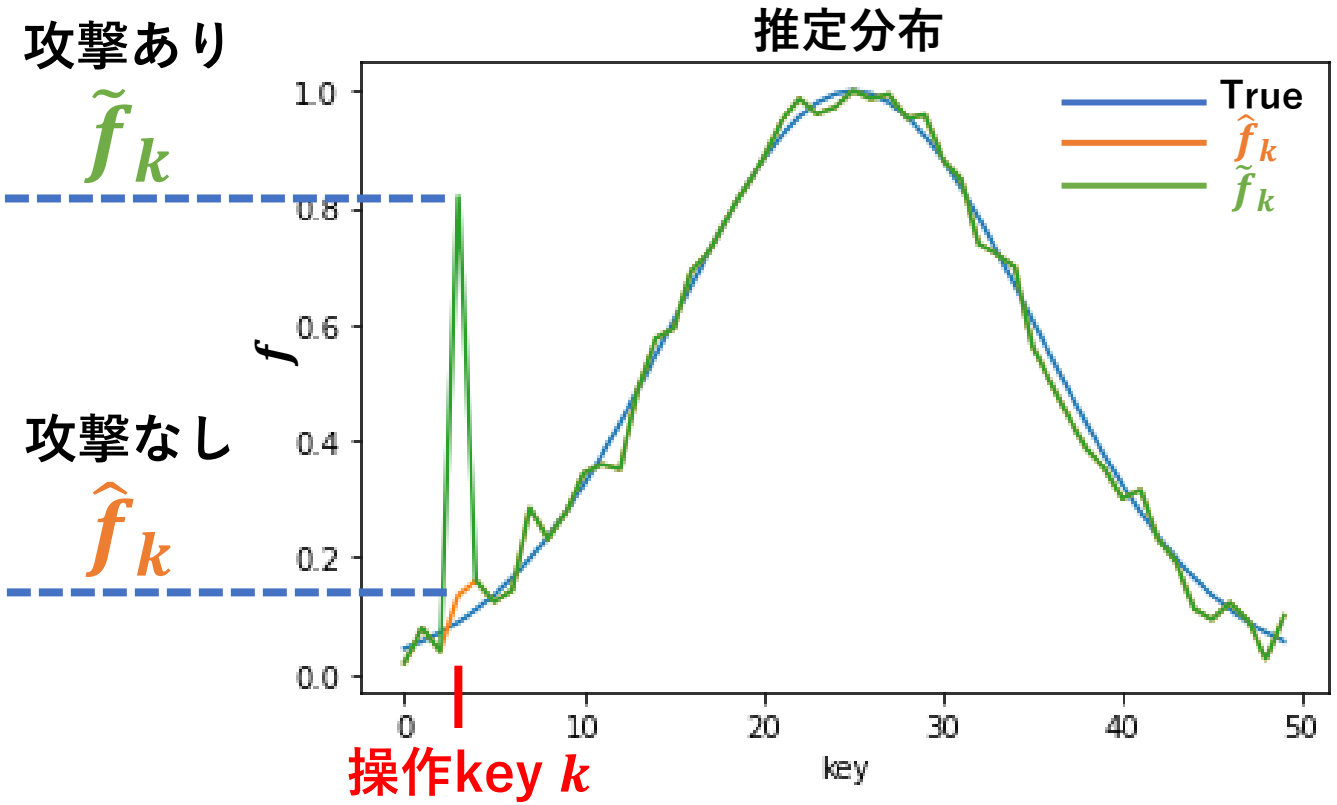


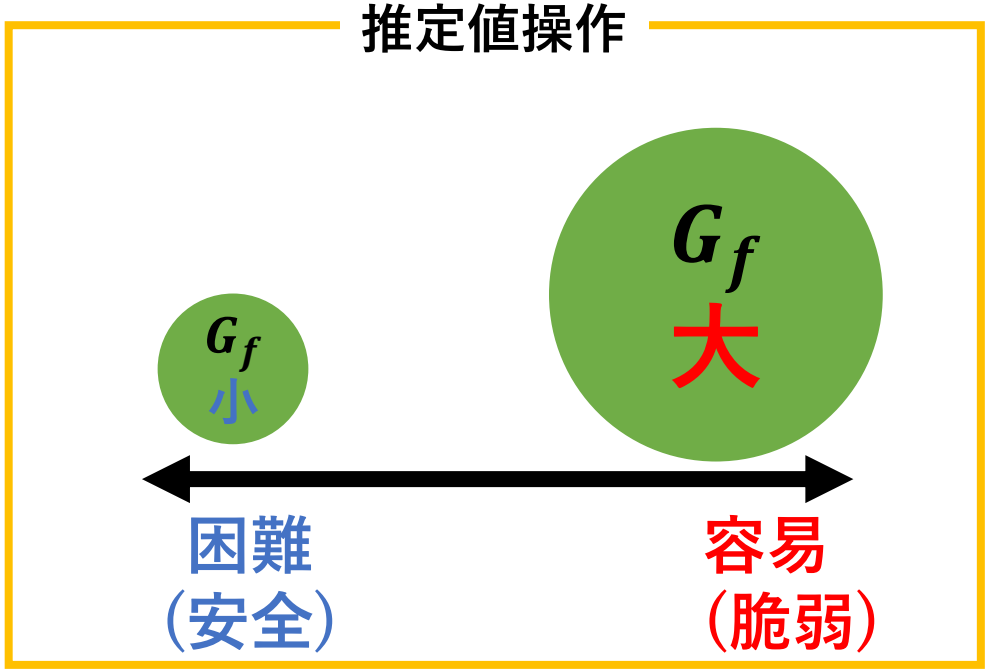
攻撃に対する強度の定量化[X.Cao, et al., 2021]

ポイズニング攻撃による推定度数の変化量 $\Delta\hat{f}_k = \tilde{f}_k - \hat{f}_k$

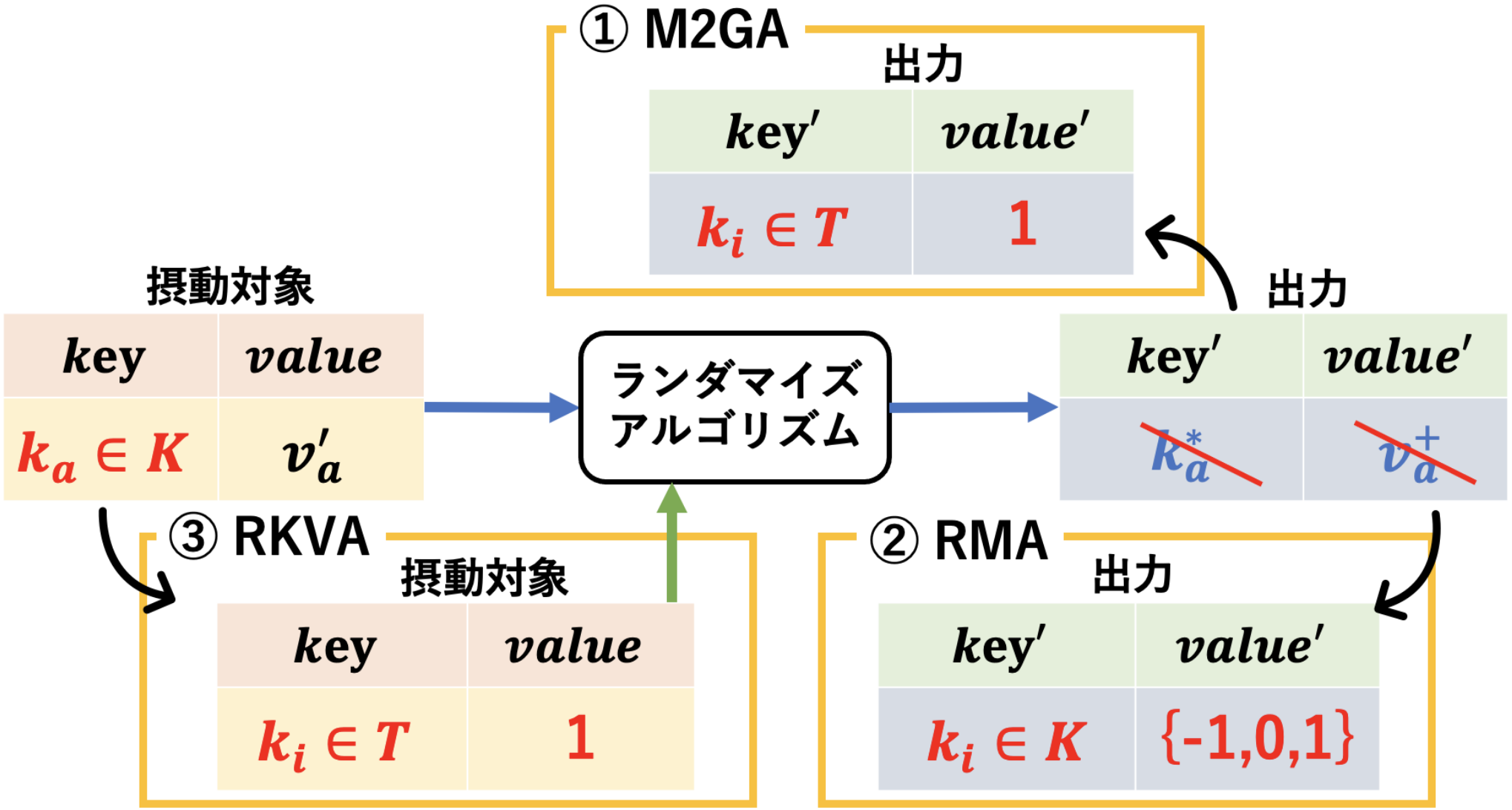
操作対象keyの集合 $T = \{k_1, k_2, \ldots k_r\}$

操作利得 $G_f = \sum_{k \in T} \mathbb{E}[\Delta\hat{f}_k]$





3つのポイズニング攻撃手法 [Y.Wu, et al., 2022]



本研究について

- 研究目的
  - 小さな誤差で統計値を推定する手法を提案すること
  - 提案手法はポイズニング攻撃に対して安全なのかを調査すること

- 解決手法
  - PrivKVへのEM (Expectation Maximization)アルゴリズムの適用
  - 提案手法へ3種類のポイズニング攻撃を行う

	摂動化	推定
PrivKV	VPP+RR	MLE
本提案		EM

提案手法
EMアルゴリズムの適用：アイデア

摂動の流れ

VPP（valueのランダマイズ）　　keyのランダマイズ

$< k'_a, v'_a > \rightarrow v'_a$の2値化 $\rightarrow < k'_a, v^*_a > \rightarrow RR(v^*_a) \rightarrow < k'_a, v^+_a > \rightarrow RR(k'_a) \rightarrow < k^*_a, v^+_a >$

$v'_a = [-1, +1]$　　$v^*_a = \{-1, 1\}$　　$v^+_a = \{-1, 0, 1\}$

・出力 $< k^*_a, v^+_a >$ を用いて $< k'_a, v^*_a >$ の事前確率を推定する．

頻度と平均値を独立に算出するのではなく，key-valueデータのまま考える

$$< k^*_a, v^+_a > = \{< 1, 1 >, < 1, -1 >, < 0, 0 >\}$$

$$< k'_a, v^*_a > = \{< 1, 1 >, < 1, -1 >, < 0, 1 >, < 0, -1 >\}$$

$$< k_a, v_a > = < 0, \_* >$$

EMアルゴリズムを用いることで，
複数の要素分布から最適なパラメータを算出できる

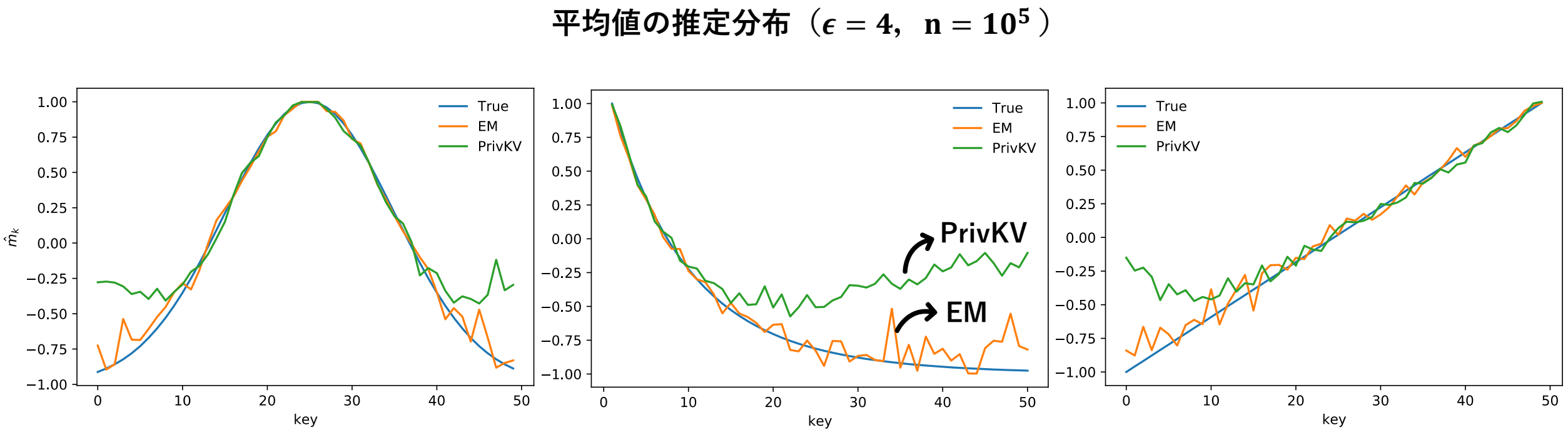
RQ.

- **RQ1. 提案手法はPrivKV, PrivKVMよりも高精度か？**
- **RQ2. 安全性εによって推定誤差に影響はあるのか？**
- **RQ3. 提案手法はポイズニング攻撃(M2GA)に安全なのか？**

評価実験の概要

- データ
  - keyとvalueがガウス分布，べき分布，線形分布に従う合成データ
  -オープンデータセット（Movie Lens，Clothing）
- 評価実験
  - 推定精度
    - 安全性指標εやユーザ数nを変化させ，PrivKV(M)と提案手法でkey-valueデータの度数と平均値を推定し，推定誤差MSEを算出する.
    - この試行を10回行いMSEの平均値をアルゴリズムの評価値とする.
  - ポイズニング攻撃に対する強度
    -偽ユーザの割合b，安全性指標ε，ターゲットkey数rを変化させ，3つのポイズニング攻撃に対する推定値の変化量を算出する.
    -この試行を50回行い平均値を操作利得とする。

実験結果1
RQ1. 提案手法はPrivKVよりも高精度か？

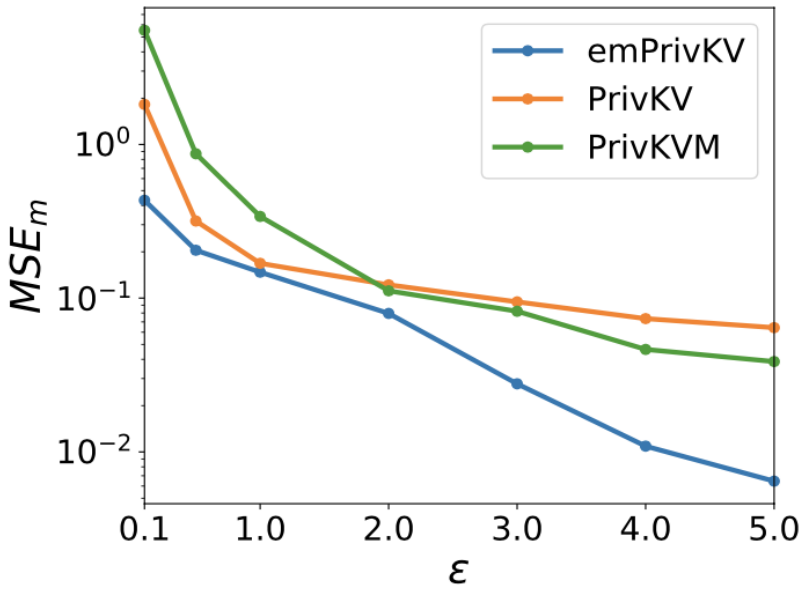


提案手法は低頻度のkeyに対しても高い推定精度

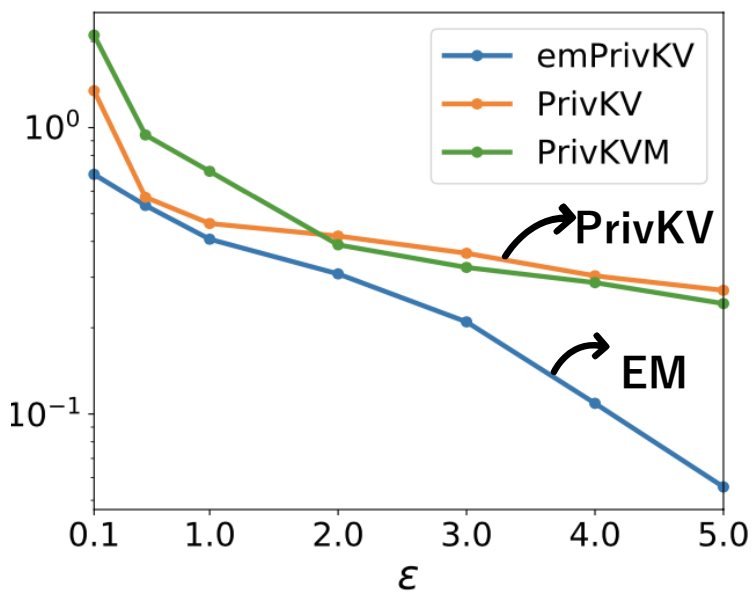
実験結果2
RQ2. 安全性εによって推定誤差に影響はあるのか？

key数=50，ユーザ数 $n = 10^5$

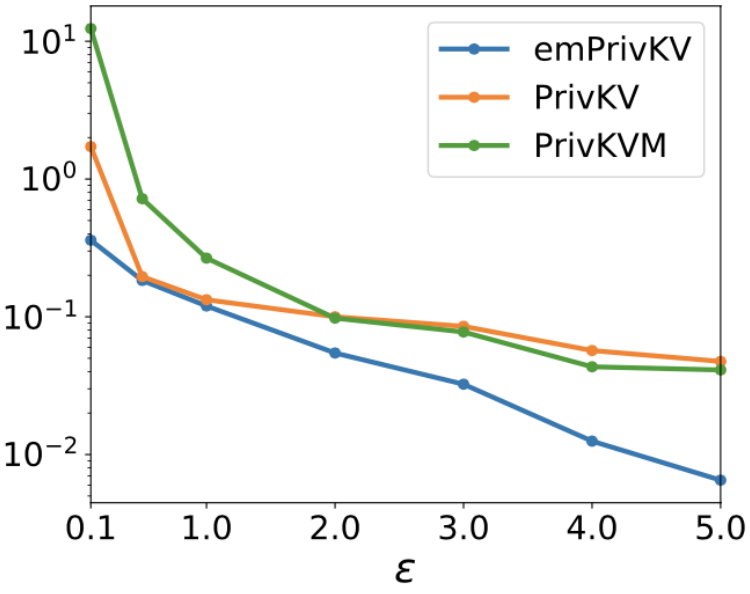
MSE_m（ガウス分布）



MSE_m（べき分布）



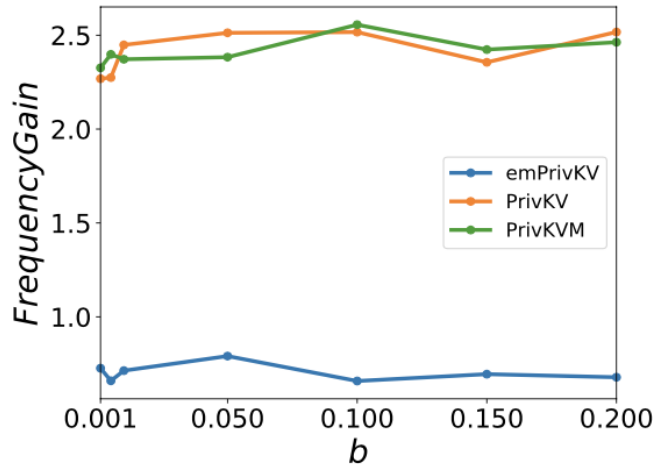
MSE_m（線形分布）

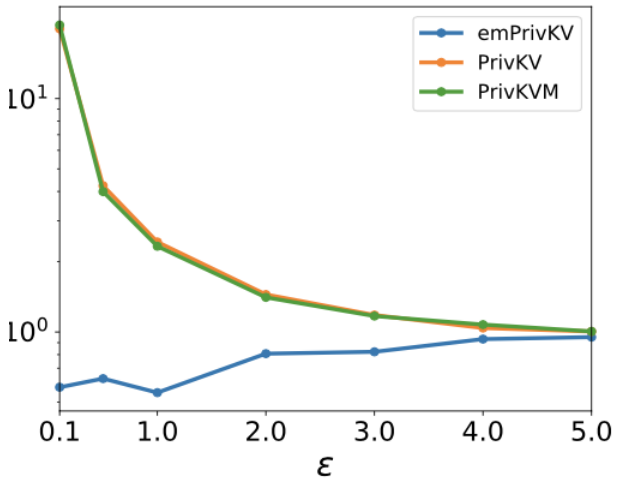


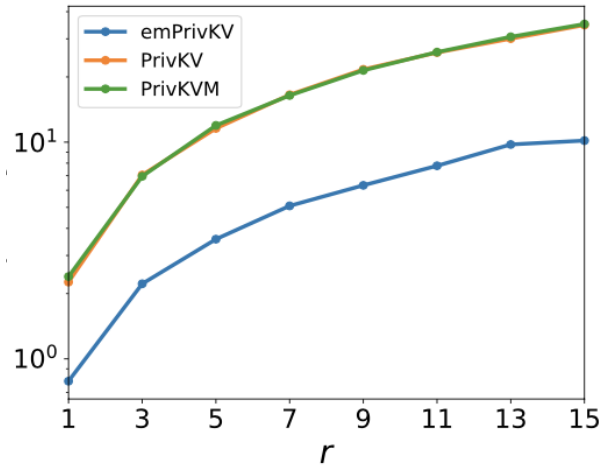
いかなる安全性εでも提案手法の誤差が小さい

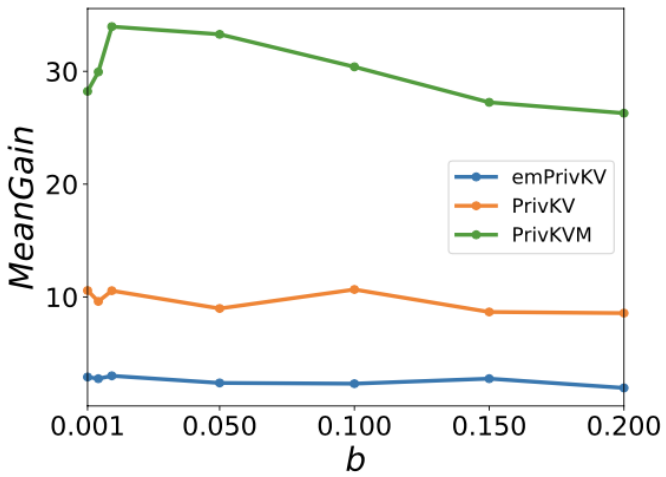
実験結果3　M2GAによる操作利得

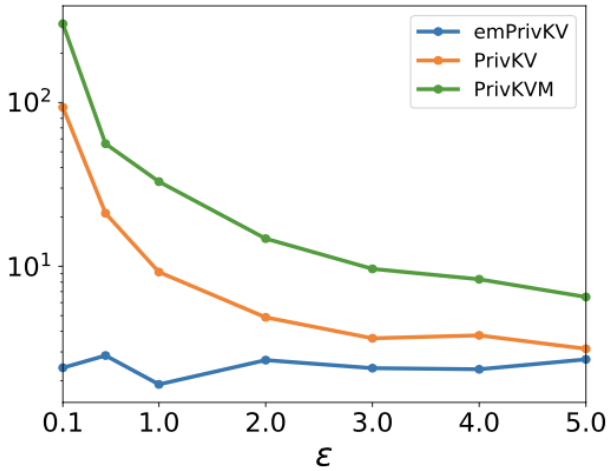
データ：MovieLensデータセット
初期パラメータ：$b = 0.05$, $\varepsilon = 1$, $r = 1$

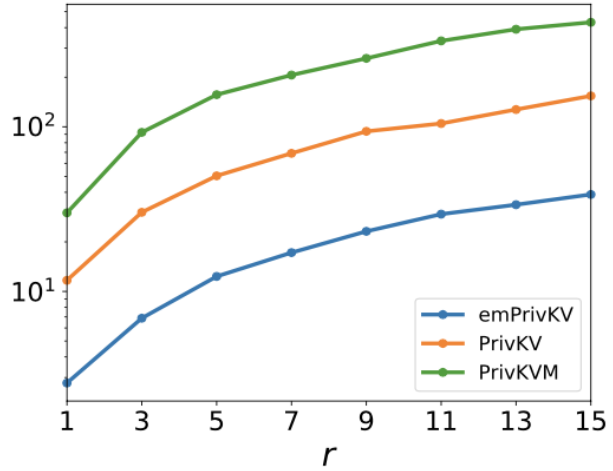












まとめ

- 局所差分プライバシアルゴリズムは最尤推定法を用いて統計値推定を行うため推定誤差が大きい.
- 多次元データの局所差分プライバシプロトコルPrivKVに着目し，統計値を小さな誤差で推定するため，EMアルゴリズムを適用する手法を提案した.
- どのパラメータを操作しても，提案手法の推定精度がよく，ポイズニング攻撃に対しても頑強であった.
- これより，一般的なユースケースにおいて提案手法が有効であると結論付ける.