計算資源制約下における差分プライバシー強化手法の提案と有効性の検証

田口 魁人[1,a)]　櫻井 幸一[2,b)]　飯田 全広[3,c)]

概要：局所差分プライバシーは取得データを管理先に送信する際にノイズを加えることでプライバシーを保護する概念である．本研究では MEC(Multi-access Edge Computi ng) デバイスを用いて, ノイズ処理を行うことを想定し, 入力されるデータをストリーム処理するうえで安定した速度を保つためハードウェアによる実装を目指す．ハードウェア実装の先行研究として, 低解像度, 固定小数点という特徴を持った局所差分プライバシー回路が提案された．低解像度の制約はノイズ分布の幅や各ノイズの間隔に影響を与え, プライバシー損失が起こる．これに対し先行研究では分布の幅を制限する閾値処理法によりプライバシー損失を抑えている．我々は, 整数値出力の差分プライバシー回路において, 低解像度を原因としたプライバシー損失が起こること, 固定小数点の制約により使用可能なプライバシー変数に制限があることを示した．また, 低解像度を原因としたプライバシー損失を防ぐ閾値決定法を示した．本研究では, ソフトウェアシミュレーションにより評価をおこない, プライバシー保護が強い状況下で先行研究と同等の有用性がもつことを示す．

キーワード：差分プライバシー, 低解像度, 固定小数点, Multi-access Edge Computing, ハードウェア

Proposal and Validation of a Differential Privacy Enhancement Method under Computational Resource Constraints

Kaito Taguchi[1,a)]　Kouichi Sakurai[2,b)]　Masahiro Iida[3,c)]

Abstract: Local Differential Privacy (LDP) is the concept of adding noise to acquired data before submitting it to a data curator. In this paper, we assume noise processing on a MEC(Multi-access Edge Computing) device and aim at a hardware implementation to maintain stable performance in the stream processing of input data. Previous work proposes an LDP circuit with low resolution and fix-point. The constraint of low resolution affects the width of the noise distribution and the spacing between each noise spacing, and privacy loss occurs. To overcome this limitation, they propose thresholding that limits the width of the distribution. In the integer-valued LDP circuit, we mentioned privacy loss caused by low resolution remains and the availability of privacy variable limits. Also, we proposed a method of searching for a threshold value to avoid privacy loss due to low resolution. In this paper, we evaluate the difference between previous work and our proposal and show our proposal has a utility equivalent to the previous one under strong privacy.

Keywords: differential privacy, low resolution, fixed-point, Multi-access Edge Computing, Hardware

1. はじめに

プライバシー保護は近年のデータ利用における一つの課題である．局所差分プライバシーは端末などのエッジデバ

[1] 九州大学大学院システム情報科学府
Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University
[2] 九州大学大学院システム情報科学研究院
Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University
[3] 熊本大学大学院先端科学研究部
Faculty of Advanced Science and Technology, Kumamoto University

[a)] taguchi.kaito.834@s.kyushu-u.ac.jp
[b)] sakurai@inf.kyushu-u.ac.jp
[c)] iida@cs.kumamoto-u.ac.jp

イス上で得られたデータにノイズを付与し, これをクラウドへ送信することで個々のデータのプライバシーを保護する.

計算リソースの限られたエッジデバイスでは, デバイス側だけで処理を完結させることが難しく, そのためクラウド側でも処理を行うエッジコンピューティングが主流である. しかし, デバイスの量が増えることでクラウドに負荷がかかることやクラウド側の端末の位置によっては物理的な距離による遅延が発生しリアルタイム処理に向かない. これに対し MEC(Multi-access Edge Computing) では, クラウドとエッジデバイスの中間に位置するデバイスをエッジ端末の近くに配置することで以上の課題を解決する.

我々は MEC デバイス上で局所差分プライバシーの処理を行うことを想定し, ストリーム処理において利点のある FPGA による実装を目指す [3][4][5].

ハードウェア実装の局所差分プライバシー回路の先行研究として Choi らの研究 [7] がある. 彼らは, 固定小数点と低解像度の制約によりプライバシー損失が起こることを示した. これに対し, 閾値を設定して分布を制限する閾値処理法 (thresholding) と再抽出法 (resampling) が提案された. また, プライバシー流出に上限を設けるため予算管理アルゴリズムを導入した. これらを構成要素とし, 高い有用性と低いオーバーヘッドをもつ DP-Box が提案された.

我々は, 先行研究 [7] で提案された閾値で分布を制限した場合値抜けによりプライバシー損失が起こることを示した [2]. また値抜けを回避する閾値決定法を示し, ソフトウェアシミュレーションによる有用性の評価を行った.

本研究では, 先行研究 [7] と我々の提案手法 [2] の設定する閾値の違いに着目する. ソフトウェアシミュレーションにより有用性の比較おこない, プライバシー保護の強い状況下で我々の手法が先行研究 [7] と同等の有用性を持つことを示す.

2. 準備

2.1 局所差分プライバシー (LDP)

LDP[9][10] は各データに確率的メカニズムを適用することでプライバシー保護を実現する概念である.

入力 x_1, x_2 とし, 得られる出力を y とする. 確率的メカニズム $\mathcal{M}$ が式 (1) を満たす時 $\mathcal{M}$ は $\varepsilon - LDP$ であるという. O は出力される可能性のある出力の集合を表す.

$$Pr[\mathcal{M}(x_1) = y \in O] \leq e^{\varepsilon} Pr[\mathcal{M}(x_2) = y \in O] \quad (1)$$

式 (1) は異なる入力から得られる出力分布がどれほど類似度を表す. 式 (1) を変形し得られた式 (2) はプライバシー損失が ε で抑えられることを示している. ここで, ある出力 y' が x_1 を入力として出力可能で x_2 を入力として出力不可能であるとする. このとき, 確率 $Pr[\mathcal{M}(x_2) = y']$ が 0 となるため式 (2) が発散する. すなわち, 二入力のみが想定される場合, 入力がどちらであるかを断定可能である.

$$loss_{x_1,x_2} = \log \frac{Pr[\mathcal{M}(x_1) = y \in O]}{Pr[\mathcal{M}(x_2) = y \in O]} \leq \varepsilon \quad (2)$$

2.2 ラプラスメカニズム

最もよく知られた確率的メカニズムとしてラプラスメカニズム [8] が挙げられる. 取得したデータに対し式 (3) に示すゼロ平均ラプラス分布に従ったノイズを加算する仕組みである.

$$f(x) = \frac{1}{2b} e^{-\frac{|x|}{b}} \quad (3)$$

ラプラス分布の幅 b は式 (4) で定義される. 式中の ε がプライバシーの許容度を表し, 値が小さければプライバシー保護が強く, 値が大きければプライバシー保護が弱くなる.

$$b = \frac{GS}{\varepsilon} \quad (4)$$

GS とは敏感度 (global sensitivity) であり $f : D^n \to \mathbb{R}^d$ としたとき, 式 (5) によって表される.

$$GS = \max_{x,y:d(x,y)\leq 1} \|f(x) - f(y)\|_1 \quad (5)$$

また, ラプラスメカニズムにおいてプライバシー保護が強いことと入力から全く異なる出力が得られる確率が高いことは同じ状態を表すので, プライバシー保護の強さとデータの保有する統計的な意味の間にトレードオフが存在するといえる.

3. 先行研究

Gazeau ら [6] は, 浮動小数点を例として有限精度の実数を用いた場合に, 理想的な実数を使用可能な状況を想定した差分プライバシーが保障されないことを言及した. これに対し丸め込み, 値の例外を作るという方法を用いることで実装レベルで差分プライバシーが保障されることを証明した.

Choi ら [7] は, 固定小数点と低解像度の制約により次の二つの問題が発生することが言及された.

(1) 分布に上限が存在

(2) 確率のない出力が存在

以上の二点により, プライバシー損失が起こることが示され, これに対し分布の上限を制限する再抽出法と閾値処理法が提案された. また, LDP の合成定理に対するプライバシー保護策として予算管理アルゴリズムが提案された. 再抽出法によるオーバーヘッドは論文内である程度低いと見積もられているものの回数の保証がないため本研究では閾値処理法のみを用いる.

論文内では分布の上限に対する閾値決定の式として式

Algorithm 1 hl 探索アルゴリズム

```
入力: 降順ソート済み生成可能なノイズ列 noises,
一様乱数のビット数 bit
for i = 1 to bit do
  pre = noises[i − 1]
  post = noises[i]
  if pre − post > 1 then
    hl = noises[i]
  end if
end for
return hl
```

(6) が示された. ここで各変数は, 敏感度を表す d, ノイズの量子化ステップ幅 Δ, 一様乱数のビット数 B_x, ノイズの値 n, プライバシー損失の上限を式 $n\varepsilon$ である.

$$n_{th2} = d + \frac{\Delta}{2} + \frac{d}{\varepsilon}(B_x \log 2 + \log(\exp(-\varepsilon) - \exp(-n\varepsilon))) \tag{6}$$

また, これらを搭載した DP-Box を提案し, 高い有用性と低いオーバーヘッドを持つことを示した.

我々 [2] は, 17bit と低い解像度において先述の 2) 確率のない出力の存在によるプライバシー損失が閾値決定式を用いても発生することを示した. 確率のない出力を出力値の抜けという意味で値抜けと呼び, この値抜けによるプライバシー損失を回避する閾値 hl を求めるアルゴリズム 1 を提案した.

4. 実験

我々の提案手法と Choi ら [7](以下先行研究とする) との有用性の比較を行う. まずは, 先行研究と我々の提案手法における違いを示し, 本実験において着目する特徴を明確にする.

4.1 先行研究 [7] との差異

先行研究では再標本化や予算管理アルゴリズムも導入されており, DP-Box は ASIC(Application Specific Integrated Circuit) である.

一方で, 我々は閾値処理法のみを扱っており, リソースの少ない FPGA への搭載を想定している. 本研究では実装していない点も差異である.

本研究では熊本大学で開発されている FPGA への実装を想定し, 回路規模を検証したうえでソフトウェアシミュレーションを行っている. 熊本大学で開発されている FPGA の特徴として論理セルに SLM[11] を用いるため低面積, 低消費電力である点が挙げられる.

今回の実験ではこれらの違いの中でもそれぞれの設定する閾値で閾値処理法を行った場合の有用性の比較のみを行うこととする.

そのため先行研究と我々の提案手法の違いは設定する閾値のみであり, 我々の閾値はアルゴリズム 1 で求めた hl, 先行研究の閾値は分布の出力可能な上限値とする.

表 1 実験で用いる閾値 [2]

	ε=1	ε=2	ε=3	ε=5	ε=7
提案手法	1427	810	566	366	278
先行研究	2828	1414	942	565	404

表 2 データセット一覧

種類 [単位]
男性の身長 [cm]
女性の身長 [cm]
男性の体重 [kg]
女性の体重 [kg]
男性の最高血圧 [mmHg]
女性の最高血圧 [mmHg]
男性の最低血圧 [mmHg]
女性の最低血圧 [mmHg]
男性の BMI
女性の BMI
ランダム

先行研究の閾値を分布の上限値に設定する理由は次の二つである. 一つ目は先行研究内で示された閾値決定の式が一意でないため, いずれかの閾値を選択する必要があること. 二つ目は選択可能な閾値のうち分布の上限値を設定した場合が最も分布を歪ませないことである. 先行研究での結果から分布に歪みのないほうが有用性が高いと考えられるため, 閾値として分布の上限値を設定する.

それぞれの閾値は表 1 に示す. 我々の提案する閾値 hl は値抜けによるプライバシー損失に対応した閾値であるため分布の歪みは大きいもののプライバシー保護が保証されることに注意いただきたい.

4.2 実験方法

実験には表 2 にしめす 11 種のデータセットを使用する.

また, 有用性は式 (7) に示すように LDP を適用したデータセットの統計値 *ldp-output* と LDP を適用しなかったデータセットの統計値 *original-output* について 100 回分の MAE(平均絶対誤差: Mean Absolute Error) を計算することにより評価する.

$$MAE = \frac{1}{100}\sum_{i=1}^{100} |original\text{-}out_i - ldp\text{-}out_i| \tag{7}$$

また, 本実験では統計値として平均値を用いる.

4.2.1 結果

我々の提案手法で得られた閾値を設定した場合の有用性 $MAE_{proposed}$ を表 3 に, 分布の上限値を設定した場合の有用性 $MAE_{related}$ 表 4 に示す.

種類はデータセットの種類を, 真値は各データセットに平均値を適用した場合に得られる値, 各表の値は $MAE \pm std$ である. ここで, std は式 (8) で表される MAE の標準

表 3 平均値: 提案手法を用いた $MAE_{proposed}$[2]

種類	真値	ε=1	ε=2	ε=3	ε=5	ε=7
男性の身長	168.22	3.85 ± 3.32	1.99 ± 1.70	1.41 ± 1.18	1.10 ± 0.79	1.30 ± 0.68
女性の身長	154.28	3.84 ± 3.31	1.97 ± 1.69	1.38 ± 1.16	1.00 ± 0.75	1.01 ± 0.63
男性の体重	67.25	3.82 ± 3.25	1.92 ± 1.63	1.27 ± 1.08	0.77 ± 0.64	1.10 ± 0.58
女性の体重	53.64	3.82 ± 3.25	1.91 ± 1.63	1.27 ± 1.08	0.88 ± 0.64	1.78 ± 0.64
男性の最高血圧	128.7	3.83 ± 3.29	1.95 ± 1.67	1.33 ± 1.14	0.86 ± 0.72	0.69 ± 0.53
男性の最低血圧	76.53	3.82 ± 3.26	1.92 ± 1.64	1.28 ± 1.09	0.75 ± 0.64	0.82 ± 0.49
女性の最高血圧	122.97	3.83 ± 3.29	1.95 ± 1.67	1.33 ± 1.13	0.84 ± 0.70	0.64 ± 0.51
女性の最低血圧	72.86	3.82 ± 3.26	1.92 ± 1.64	1.28 ± 1.09	0.76 ± 0.64	0.92 ± 0.52
男性の BMI	23.74	3.82 ± 3.24	1.91 ± 1.62	1.29 ± 1.08	1.40 ± 0.77	4.62 ± 0.58
女性の BMI	22.39	3.82 ± 3.24	1.91 ± 1.62	1.29 ± 1.08	1.45 ± 0.78	4.82 ± 0.58
ランダム	125.59	3.83 ± 3.29	1.94 ± 1.67	1.32 ± 1.13	0.82 ± 0.69	0.56 ± 0.47

表 4 平均値: 閾値に分布の上限を持ちいた $MAE_{related}$

種類	真値	ε=1	ε=2	ε=3	ε=5	ε=7
男性の身長	168.22	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.87 ± 0.73	0.7 ± 0.55
女性の身長	154.28	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.86 ± 0.72	0.66 ± 0.53
男性の体重	67.25	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.87 ± 0.72	0.7 ± 0.55
女性の体重	53.64	3.88 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.86 ± 0.72	0.65 ± 0.53
男性の最高血圧	128.7	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.87 ± 0.72	0.69 ± 0.55
男性の最低血圧	76.53	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.87 ± 0.72	0.69 ± 0.55
女性の最高血圧	122.97	3.88 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.86 ± 0.72	0.67 ± 0.54
女性の最低血圧	72.86	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.86 ± 0.72	0.67 ± 0.54
男性の BMI	23.74	3.88 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.85 ± 0.71	0.61 ± 0.51
女性の BMI	22.39	3.88 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.85 ± 0.71	0.61 ± 0.5
ランダム	125.59	3.89 ± 3.26	1.96 ± 1.67	1.34 ± 1.14	0.87 ± 0.72	0.69 ± 0.54

偏差である.

$$std = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(|original\text{-}out_i - ldp\text{-}out_i| - MAE)^2} \tag{8}$$

また, 両者の差分を表 5 に示す.

$\varepsilon = 1 \sim \varepsilon = 3$ においては両者に大きな差がないことがわかる. しかし, 一部のデータセットにおいてそれより大きな ε をとったとき提案手法の有用性が大きく下がっていることがわかる.

5. 議論

提案手法において有用性が低下することについて, データセットの統計値の真値に近い位置で分布の丸めが行われてる点が原因である. 出力分布はラプラス分布に従うため中央付近の度数が極めて高い. その周辺で丸め込みが起こった場合多くの値が丸め込まれるため平均値に偏りが生まれる. その結果有用性が低下したといえる.

今後の展望として, 評価がシミュレーションのみにとどまるため省リソースのハードウェア上で実装を行うことが検討される. また, より一般化した状況でプライバシー損失の起こらない閾値決定法を探索することも重要である.

謝辞 本研究を進めるにあたり御助言御鞭撻を頂いた熊本大学大学院先端科学研究部久我守弘准教授, 相談に乗っていただき回路規模の検証にも協力いただいた熊本大学大学院自然科学教育部中里優弥さんに深く感謝する. 本研究は, JST, CREST, JPMJCR19K1 の支援を受けたものである.

参考文献

[1] 田口魁人, *FPGA* 向け *Local Differential Privacy* 回路の研究, 熊本大学卒業論文, 2022.

[2] 田口 魁人, 櫻井 幸一, 飯田 全広, *IoT* デバイス上の差分プライバシー強化手法の提案と評価, 研究報告コンピュータセキュリティ（CSEC）, 2022-CSEC-98(12), 1-6 (2022-07-12), 2188-8655

[3] S. Biookaghazadeh, Zhao, M. , and Ren, F. , *Are FPGAs Suitable for Edge Computing?.* The USENIX Workshop on Hot Topics in Edge Computing (HotEdge '18). BOSTON, MA, 2018.

[4] Rene Mueller, Jens Teubner, and Gustavo Alonso. 2009. *Data processing on FPGAs.* Proc. VLDB Endow. 2, 1 (August 2009), 910–921.

[5] S. Wu et al., *When FPGA-Accelerator Meets Stream Data Processing in the Edge.* 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 1818-1829.

[6] Ivan Gazeau, Dale Miller, and Catuscia Palamidessi.: *Preserving differential privacy under finite-precision semantics.* Theor. Comput. Sci. 655, PB, 92–108. (2016)

[7] Woo-Seok Choi, Matthew Tomei, Jose Rodrigo Sanchez Vicarte, Pavan Kumar Hanumolu, and Rakesh Kumar.:

表 5 有用性の差分 $MAE_{proposed}$-$MAE_{related}$

種類	真値	ε=1	ε=2	ε=3	ε=5	ε=7
男性の身長	168.22	-0.04	0.03	0.07	0.23	0.6
女性の身長	154.28	-0.05	0.01	0.04	0.13	0.31
男性の体重	67.25	-0.07	-0.04	-0.07	-0.09	0.44
女性の体重	53.64	-0.06	-0.05	-0.07	0.02	1.13
男性の最高血圧	128.7	-0.05	-0.01	-0.01	-0.01	0
男性の最低血圧	76.53	-0.06	-0.04	-0.06	-0.11	0.15
女性の最高血圧	122.97	-0.06	-0.01	-0.01	-0.03	-0.05
女性の最低血圧	72.86	-0.06	-0.04	-0.06	-0.1	0.25
男性の BMI	23.74	-0.06	-0.05	-0.05	0.55	4.01
女性の BMI	22.39	-0.06	-0.05	-0.05	0.6	4.21
ランダム	125.59	-0.06	-0.02	-0.02	-0.05	-0.13

Guaranteeing local differential privacy on ultra-low-power systems In Proceedings of the 45th Annual International Symposium on Computer Architecture (ISCA ' 18). IEEE Press, 561–574. (2018)

[8] Dwork, C., McSherry, F., Nissim, K., Smith, A. (2006). *Calibrating Noise to Sensitivity in Private Data Analysis*. In: Halevi, S., Rabin, T. (eds) Theory of Cryptography. TCC 2006. Lecture Notes in Computer Science, vol 3876. Springer, Berlin, Heidelberg.

[9] J. C. Duchi, M. I. Jordan and M. J. Wainwright, *Local Privacy and Statistical Minimax Rates,* 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, 2013, pp. 429-438.

[10] S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova and A. Smith, *What Can We Learn Privately?,* 2008 49th Annual IEEE Symposium on Foundations of Computer Science, 2008, pp. 531-540.

[11] M.AMAGASAKI, R.ARAKI, M.IIDA, T.SUEYOSHI, *SLM: A Scalable Logic Module Architecture with Less Configuration Memory*, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2016, Volume E99.A, Issue 12, Pages 2500-2506, Released on J-STAGE December 01, 2016.