連合学習における分散差分プライバシーに関する一考察

前田 若菜[1,a)]　長谷川 聡[1,b)]　高橋 翼[1,c)]

概要：クライアント群からのデータの収集と統計処理を秘匿した形で実施するセキュアアグリゲーションを前提として，クライアントでの小さいノイズの加算でも厳密かつ十分なプライバシー保護を達成可能な分散差分プライバシー (DDP) が連合学習におけるプライバシーモデルとして注目されている．しかしながら，発展が著しい DDP のメカニズム間の優劣は明確ではなく，連合学習の実用面でどの手法を選択すべきかを判断することは容易ではない．本研究では，連合学習における最先端の DDP メカニズムを比較，実験，および考察し，メカニズム選択の指針を提案した．

キーワード：分散差分プライバシー，セキュアアグリゲーション，連合学習

Federated Learning with Distributed Differential Privacy: Performance Analysis

Wakana Maeda[1,a)]　Satoshi Hasegawa[1,b)]　Tsubasa Takahashi[1,c)]

Abstract: Distributed differential privacy (DDP) is getting attention as a privacy model in federated learning, assuming secure aggregation in which aims to collect and statistically process data from a group of clients in a privacy-preserving manner. This is because DDP can achieve strict and sufficient privacy protection even with the addition of small noise on the client-side. However, the relative merits of the rapidly developing DDP mechanisms are not clear, and it is not easy to determine which method should be chosen for practical applications in federated learning. In this study, we compare, experiment with, and discuss several DDP mechanisms to provide a guideline for selecting DDP mechanisms in federated learning.

Keywords: Distributed Differential Privacy, Secure Aggregation, Federated Learning,

1. はじめに

連合学習 (FL: Federated Learning) は，大規模に分散したクライアント群と中央集権なサーバーが協調して取り組む機械学習のフレームワークである[*1]．FL においても，厳密なプライバシー保護の達成を目的として，差分プライバシー (DP: Differential Privacy)[12] が導入されている．連合型の解析において DP の保証を考えると，サーバーで集約した結果にノイズを加算するセントラル DP(CDP: Central DP) は加算するノイズが少量でよい反面，データを収集するサーバーをクライアント群が信頼する必要がある．一方，サーバーへの信頼を必要としないローカル DP (LDP: Local DP)[18] を前提とする場合には，クライアントで加算するノイズが膨大になり，有用性の維持が困難，といった課題がある．

この課題の解決策の一つとして，分散 DP(DDP: Distributed DP) と呼ばれるプライバシーモデルが提案されている．DDP は，信頼できる計算機構を前提として，クライアントでの小さいノイズの加算でも厳密かつ十分なプライバシー保護を計算機構全体で達成する．このとき必要な計算機構として，クライアント群からのデータの収集と統計処理を秘匿した形で実施するセキュアアグリゲーション (SecAgg: Secure Aggregation) を用いる．SecAgg は，秘

[1] LINE ヤフー株式会社
LY Corporation
[a)] wakana.maeda@lycorp.co.jp
[b)] satoshi.hasegawa@lycorp.co.jp
[c)] tsubasa.takahashi@lycorp.co.jp
[*1] 予測変換の学習 [14] やメッセージングアプリケーションにおけるスタンプの推薦で活用されている．

密分散や TEE[25] を用いて実現し [7][8][15][23][24][27]，クライアントからのデータの収集と統計処理までの過程をサーバーに対して秘密裏に実施する．SecAgg の活用事例として Google と Apple が共同で開発した COVID-19 の Contact Tracing[4] がある．FL でもクライアント群からのデータの収集と統計処理に SecAgg を導入することで，DDP を導入することができる．クライアント上の DDP メカニズムが加算するノイズは，一般的には，SecAgg 後に CDP と等価なプライバシーが達成されるように設計される．このような DDP によれば，プライバシーと有用性を高い水準でバランスした FL の実現が期待できる．

DDP の研究では，分散離散ガウスメカニズム [9] をはじめとして，Skellam Mixuture メカニズム [6]，Poisson Binomial メカニズム [10] といった新しいメカニズムが多数提案されている．しかしながら，発展が著しい DDP のメカニズム間の優劣は明確ではなく，実用面でどの手法を選択すべきかを判断することは容易ではない．

本研究は，FL における DDP メカニズム選択の指針を得るために，いくつかの DDP メカニズムを比較，実験，および考察する．本研究の調査と実験により，以下の事実を確認した．

- 同じパラメータを使用していても，一回のベクトル平均と複数回のベクトル平均を行う FL ではメカニズムが与える有用性への影響が異なる
- メカニズムに依存しない FL の 3 つのパラメータに基づき，最適なメカニズムが選択可能な事例があること

本研究の貢献は，以下の 2 点である．1) 最先端の FL における DDP メカニズムの比較，実験，および考察に取り組み，DDP メカニズム選択の指針を提案した．2) 上述のような事実を実験的に明らかにした．

2. 準備

本研究の調査対象である分散差分プライバシー (DDP) を満たす FL を導入するため，本節では差分プライバシー (DP)，DDP および連合学習 (FL) を説明する．

2.1 差分プライバシー

まず DP を定義し，DP を保証するメカニズムを設計する上で関わりの深い「感度」を導入する．そして，効率よくプライバシー消費を計算可能な Rényi DP(RDP)[22] を導入する．

定義 1. ((ϵ,δ)-DP[12])．$\epsilon \geq 0$ と $0 \leq \delta < 1$ が与えられたとき，あるランダム化メカニズム $\mathcal{M}: \mathcal{D} \to \mathcal{O}$ が (ϵ,δ)-DP を満たすとは，1 レコードしか異ならない任意の 2 つのデータベース $D, D' \in \mathcal{D}$，および任意の出力の部分集合 $S \subseteq \mathcal{O}$ について，次式が成り立つときである．

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in S] + \delta. \qquad (1)$$

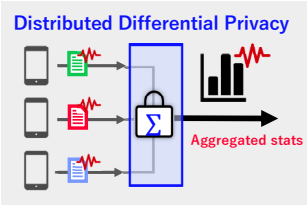

図 1 分散差分プライバシーモデルのイメージ図.
Fig. 1 Distributed Differential Privacy Model Diagram.

定義 2. (感度)．1 レコードしか異ならない任意の 2 つのデータベース $D, D' \in \mathcal{D}$ に対する関数 f の $\mathcal{L}_p$ ノルムの感度 S_f は以下のように表される．

$$S_f = \sup_{D,D'} ||f(D) - f(D')||_{\mathrm{p}}. \qquad (2)$$

なお，DP では，メカニズムを通してデータベースにアクセスする度にプライバシーを消費すると考える．複数回のアクセスを想定する場合，全体を通してどのくらいのプライバシー消費があったかを計算する必要がある．そこで，効率よくプライバシー消費を計算するために，次式のような $(\alpha, \epsilon(\alpha))$-RDP[22] を導入する．

定義 3. ($(\alpha,\epsilon(\alpha))$-RDP[22])．あるランダム化メカニズム $\mathcal{M}: \mathcal{D} \to \mathcal{R}$ が $(\alpha, \epsilon(\alpha))$-RDP を満たすとは，1 レコードしか異ならない任意の 2 つのデータベース $D, D' \in \mathcal{D}$ について，次式が成り立つときである．

$$\frac{1}{\alpha - 1} \log \left(\mathbb{E}_{x \sim Q} \left[\left(\frac{M(D)}{M(D')} \right)^{\alpha} \right] \right) \leq \epsilon(\alpha). \qquad (3)$$

$(\alpha, \epsilon(\alpha))$-RDP において，$T$ 回アクセスした場合のプライバシー消費は $(\alpha, T\epsilon(\alpha))$-RDP である [22]．また，$(\alpha, \epsilon(\alpha))$-RDP は (ϵ, δ)-DP に変換可能である [5]．

2.2 分散差分プライバシー

DDP のプライバシーモデルを示し，DDP で使用される暗号プロトコルのセキュアアグリゲーション (SecAgg) について説明する．

2.2.1 プライバシーモデル

DDP(**図 1**) は，各クライアントのノイズが少量でも，十分なプライバシー保護を達成可能な技術である．クライアントのノイズは，一般的には再生性[*2]を持つような確率分布 (例えばガウス分布や二項分布) に従うノイズを前提とし，各クライアントの送信するノイズの総和が，セントラル DP(CDP) と等価なプライバシーを達成できるように設計される．なお，DDP はノイズの総和により DP を保証することから，次節に示すセキュアアグリゲーションのような，サーバが集約した値しか得ることができない信頼できる計算機構との組み合わせが必要となる．

2.2.2 セキュアアグリゲーション

SecAgg とは，暗号化したまま総和を計算するプロトコ

[*2] 同じ分布に従う独立な確率変数の和は、元の分布に従う性質

ルのことをいう．SecAgg の実現方法として，セキュアマルチパーティ計算 (MPC: Multi-Party Computation) や Trusted Execution Environment(TEE)[25] の使用がある．

特に前者の計算 [7][8][15] は，有限体上で構成されていることから, クライアントが送信するデータは有限体上の離散値であることが前提となる．

2.3 差分プライバシーを満たす連合学習

本研究では, FL の著名なアルゴリズムである Federated stochastic gradient descent(FedSGD)[20] を基にした DP を満たす連合学習を対象とする．FedSGD について説明したのちに, CDP および DDP を満たす FL を導入する．

2.3.1 FedSGD

FedSGD は, 一般的な機械学習で用いられる SGD を, クライアント群とサーバーが協調し学習を行う分散型機械学習に拡張した技術である [20]．サーバ側は t 回目の学習において，クライアント群からサンプリングしたグループサイズ n のクライアント集合 S_t を取得し，それらに d 次元のモデルのパラメータ $w_t \in \mathbb{R}^d$ を送信する．クライアント $k \in S_t$ は，受信したパラメータ w_t と手元の学習データを用いて勾配 $g_t^{(k)}$ を計算し，これをサーバへ送信する．サーバは，クライアント集合 S_t から得た勾配から平均値 $\bar{g}_t$ を算出し，これに学習率 η を乗算したものを使ってパラメータを w_{t+1} へ更新する．これを繰り返すことでパラメータ w を学習する．

2.3.2 セントラル差分プライバシーを満たす連合学習

先行研究 [21] をはじめとする CDP を満たす FL は, 一般的な機械学習において CDP を保証するために提案された DPSGD[1] をベースとする．DPSGD は，勾配の平均値の計算にガウスメカニズムを適用することで DP を満たすものである．なお，勾配のノルムは $[0, \infty)$ の値を取りうるため，感度が無限になってしまうことから，勾配の $\mathcal{L}_2$ ノルムをクリップサイズ Δ_2 以下に制限 ($\mathcal{L}_2$ クリッピング) することで感度を Δ_2 に抑える．

DPSGD に基づく CDP を満たす FL では，クライアントが勾配の $\mathcal{L}_2$ クリッピングを行い，サーバが勾配の平均を求める前に，勾配の合計にノイズを付加する．

2.3.3 分散差分プライバシーを満たす連合学習

具体的な処理は節 3.2 に譲り, 本節では一般化した DDP を満たす FL のアルゴリズムを示す (Algoritm 1)．FedSGD と比較し，クライアントが勾配の計算後に行う処理として，感度を抑えるための勾配クリッピングと，ランダム化がある．なお，既存手法 [2][3][6][10][16] は有限体上で実現される SecAgg を想定しているため，勾配を離散化し有限体上にマッピングする処理が必要である．サーバは SecAgg から安全に集約された勾配の合計値を取得し，メカニズムに応じた方法で勾配の平均値を推定する[*3]．

[*3] メカニズムによって，クライアントで行われた変換に対する逆変換が異なり，また有限体上の合計値の処理も異なる．詳細は各論文 [2][3][6][10][16] を参照されたい．

Algorithm 1 FL with DDP

Require: Leaning rate η, group size n.

Initialize model w_0.
for each round $t = 0, 1, ...$ **do**
  $S_t \leftarrow$ group of n clients sampled
  for each client $k \in S_t$ **do**
    $g_t^{(k)} \leftarrow \text{ComputeGradient}(w_t)$
    $\tilde{z}_t^{(k)} \leftarrow \text{ClientProcedure}(g_t^{(k)})$
  end for
  $\hat{z}_t \leftarrow \text{SecureAggregation}(\tilde{z}_t)$
  $\hat{g}_t \leftarrow \text{EstimateGradientMean}(\hat{z}_t)$
  $w_{t+1} \leftarrow w_t - \eta \hat{g}_t$
end for

3. 分散差分プライバシーメカニズムの比較

まず，既存の分散差分プライバシー (DDP) のメカニズムを整理し評価やその特徴から最先端手法を抽出する．抽出した最先端手法について，メカニズムにおけるクライアントにおける処理と保証する Rényi DP(RDP)[22] を示す．最後に，最先端手法のプライバシー分析の比較を行う．

3.1 各メカニズムの特徴

DDP を達成するために提案されたメカニズムとして, Binomial（cpSGD）[3], Distributed Discrete Gaussian（DDG）[16], Skellam（Sk）[2], Skellam Mixture（SkM）[6], Poisson Binomial（PB）[10] がある．これらの手法の特徴を**表 1** にまとめる．

最初に提案された cpSGD[3] は，プライバシーの分析が差分プライバシー (DP) のみであり，繰り返し学習が行われる連合学習 (FL) において，1 回あたりに加算するノイズが大きく，学習への悪影響が大きい [6]．DDG[16] は RDP を達成する離散ガウス分布 [9] を用いた手法である．cpSGD[3] と比較し，よりタイトなプライバシー消費の計算が可能な RDP を用いており，1 回あたりに加算するノイズを小さくできる．一方，加算に対して閉じていない[*4]ため，特定の条件下ではプライバシーの保証が弱まることが報告されている [2]．これに対し，Sk[2] では，加算に閉じているスケラム分布 (節 3.2.1 参照) を用いて解決している．

入力が離散値に限定されていた cpSGD[3]，DDG[16]，Sk[2] に対し，SkM[6] は実数の入力を許容する．そのため，cpSGD，DDG，Sk で使用された実数の離散化のための確率的量子化 (Stochastic Quantization)[3] や条件付け丸め (Conditonal Rounding)[2][16] が不要になり，離散化によるバイアスを排除することができる．

これら DDG[16]，Sk[2]，SkM[6] では，ノイズは無限の定義域を持つ分布から得られる．そのため，有限体上で実

[*4] 離散ガウス分布から得たノイズの和は，必ずしも離散ガウス分布に従わない．

表 1　分散差分プライバシーの各手法のまとめ
Table 1　Summary of Methods for Distributed Differential Privacy.

Mechanism	Input	Randomize	Input Discretization	Modular Clipping	Rényi Differential Privacy
Binomial[3]	Discrete	Additive Noise	Stochastic Quantization	-	-
Distributed Discrete Gaussian[16]	Discrete	Additive Noise	Conditional Rounding	Required	✓
Skellam[2]	Discrete	Additive Noise	Conditional Rounding	Required	✓
Skellam Mixture[6]	Real	Additive Noise	-	Required	✓
Poisson Binomial[10]	Real	Mapping	-	-	✓

現されるセキュアアグリゲーションを行うには，モジュラークリッピングが必要である．モジュラークリッピングとは，データの値をモジュロ演算を用いて特定の範囲内に制限する操作を指す．データが特定の範囲を超える場合，その値は折り返される．折り返しにより，元のデータの分布が歪む可能性があり，推定値にはバイアスが生じる．このバイアスを排除するために，PB[10] では，二項分布を用いて実数を直接有限な離散値にマッピングする．離散化とモジュラークリッピングが不要なため，これらのバイアスを排除することができる．

SkM については，入力が離散値に限定されている cpSGD[3]，DDG[16]，Sk[2] と比較して，高精度のモデルを学習できることが実験的に示されている [6]．一方，PB については既存手法との実験的な比較はない．そのため，FL に適用した際に SkM と比べて，PB はどういう特性があるのか，どのような点が優れているのか明らかではない．

ここまでの議論を通じて, 他手法と比較して精度の観点で実験的に優れていることが示されている SkM, 他の手法と比較し離散化等による誤差が生じえない PB の 2 つを最先端手法とみなす．これらの実験，および考察を行い，FL に DDP を適用する際のメカニズムの選定指針を探る．

3.2　最先端手法の比較

DDP の最先端手法である SkM[6] と PB[10] の特性を明らかにするために比較を行う．まず，クライアントにおける処理とメカニズムが保証する RDP を示す．そして，プライバシー分析の比較を行う．

3.2.1　Skellam Mixture メカニズム

クライアントにおける処理は次のとおりである．

(1) 勾配をランダム回転させ[*5]，
(2) スケールパラメータ s でスケーリングした勾配を，クリッピングを行うために変換したヘルパーベクトル [6] を介して Δ_1 と Δ_∞ に $\mathcal{L}_1$，$\mathcal{L}_\infty$ クリッピングし[*6]，
(3) 勾配の小数部分を確率 p としたベルヌーイ分布 $Ber(p)$ から得たサンプルとスケラム分布 $Sk(\lambda,\lambda)$ から得たノイズを勾配に加算し，
(4) 勾配の整数部分をモジュラス m にてモジュラークリッピングする．

スケラム分布 [2][6] とは，λ をパラメータとする離散確率分布であり, 再生性[*2]を有する. 詳細は [2][6] に委ねる．

SkM が保証する $(\alpha,\epsilon(\alpha))$-RDP は次のとおりである．

$$\epsilon(\alpha)=\frac{1.2\alpha+1}{2}\cdot\frac{\Delta_1}{2n\lambda} \tag{4}$$

ただし，n はグループサイズである．

3.2.2　Poisson Binomial メカニズム

クライアントにおける処理は次のとおりである．

(1) 勾配を Δ_2 に $\mathcal{L}_2$ クリッピングし，
(2) それを Kashin の表現[*7] [17][19] に変換し，
(3) $\gamma\cdot d(\gamma>1)$ 次元の Kashin の係数を二項分布 $Bin(m-1,p)$ の成功確率 p へパラメータ $\theta\in[0,1/4]$ を利用して変換し，この分布からサンプルを得ることで範囲 $[0,m-1]$ の離散値にランダム化する．

PB が保証する $(\alpha,\epsilon(\alpha))$-RDP は次のとおりである．

$$\epsilon(\alpha)=\gamma C_0\frac{d\alpha m\theta^2}{(1-2\theta)^4 n} \tag{5}$$

ただし，d は入力された勾配の次元数，m は有限体の位数，n はグループサイズ，$C_0>0$ は普遍定数である．

3.2.3　プライバシー分析の比較

SkM[6]，PB[10] はそれぞれ異なるクリッピングおよびランダム化処理に基づいており, 保証する RDP の値も異なる．一方，共通する点として, グループサイズ n に対して線形に $\epsilon(\alpha)$ が減少することが挙げられる. 両メカニズムとも, グループサイズを増やすことでプライバシー強度を高めることができる．

特筆すべき点として, PB の $\epsilon(\alpha)$ がモデルパラメータの次元数 d や有限体の位数に関わる m に対して線形に増えることに対し，SkM は, それらの値が明示的に $\epsilon(\alpha)$ の計算に含まれていない．これらの違いが, ベクトル平均を計算するタスクにおいてどのように影響するかは, 次節以降の実験で確かめることとする．

4.　実験

分散差分プライバシー (DDP) を満たす連合学習 (FL)

[*5] Walsh-Hadamard 変換を行い，パブリックなランダム符号ベクトルで符号反転を行う．
[*6] ヘルパーベクトルへの変換方法，クリッピング方法は [6] の Algoritm 5 を参照
[*7] Kashin の表現はベクトルを動的範囲を最小化するようにフレーム展開 (d 次元ベクトルを，$\gamma\cdot d(\gamma>1)$ 次元の係数ベクトルを用いて表現) するもので，情報が係数の各成分に均等に分散されつつ各成分の大きさも小さいため，ノイズに対するロバスト性を高めることができる．

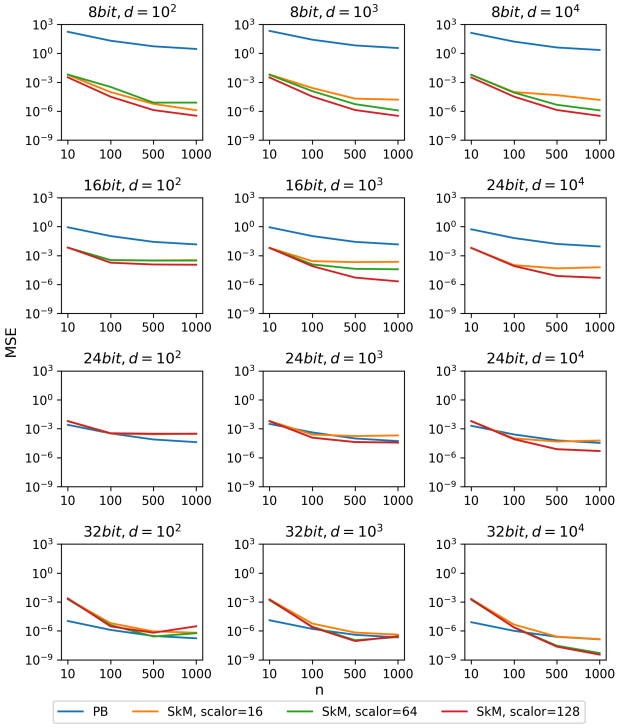

図 2　ベクトル平均推定の評価
Fig. 2　Evaluation on distributed mean estimation

において，メカニズム選択の指針を得るために，Skellam Mixture メカニズム (SkM)[6] と，Poisson Binomial メカニズム (PB)[10] の比較実験を行う．

評価として，FL のような複数回のベクトル平均に基づく計算の評価だけでなく，1 回のベクトル平均の評価も行うことで，SkM および PB の特性を詳細に明らかにする．

いずれの実験で共通のパラメータは次のとおりである．プライバシーパラメータ $\delta = 10^{-6}$, $\mathcal{L}_2$ のクリップサイズは $\Delta_2 = 1$ とする．SkM のパラメータとして, $\mathcal{L}_1$ のクリップサイズ $\Delta_1 = (s \cdot \Delta_2)^2$[6] とし，$\mathcal{L}_\infty$ のクリップサイズは [6] の (3) 式を使って求める．SkM のパラメータ λ と PB のパラメータ θ はプライバシーパラメータ ϵ の値に応じて算出したものを用いる．

4.1　ベクトル平均

プライバシーパラメータ $\epsilon = 1$，グループサイズ $n = \{10, 100, 500, 1000\}$ にて次元 $d = \{10^2, 10^3, 10^4\}$ の 0 ベクトルの平均値推定を行い，平均二乗誤差 (MSE: Mean Squared Error）の 10 回試行の平均を比較する．有限体の位数 $m = \{2^8, 2^{16}, 2^{24}, 2^{32}\}$ とし，以降それぞれ 8，16，24，32bit と呼ぶこととする．

結果を**図 2** に示す．PB は次元 d の影響をうけず，グループサイズ n が大きくなるにつれて MSE が減少した．また，有限体の位数 m が大きくなるにつれて，MSE が減少した．SkM は次元 d の影響をうけ，グループサイズ n が大きくなっても MSE が減少しないことがあった．異なるスケールパラメータ s であっても同様の MSE を示すことがあった．

4.2　連合学習

4.2.1　実験設定

データセット　実験データとして MNIST[11] と Fashion-MNIST(FMNIST)[26] を用いる．いずれも 10 クラス分類タスクであり，MNIST は手書き数字を 0 から 9 に，FMNIST は服の種類を分類する．クライアントが持つ訓練データの数をクライアントによらず一律でそれぞれ 5 つとし，テストデータは 1 つとする．全クライアント数を 10^6 とし，クライアントのもつデータを復元抽出によって得る．

モデルとパラメータ　CNN モデルを用い，そのアーキテクチャは文献 [13] の TABLE XIII を参考にした．モデルのパラメータ数について，[13] が最大になるようにし (Large)，次に第一畳み込み層のフィルターを 8，第二畳み込み層のフィルターを 16，全結合層のユニット数を 16 にしたモデル (Middle)，それぞれのフィルターを 4,8 にしユニット数を 8 にしたモデル (Small)，それぞれのフィルターを 2,4 にしユニット数を 4 にしたモデル (Tiny) を用いる．それぞれのモデルのパラメータ数 d は，9,594，2,754，870，312 である．グループサイズ $n = \{10^2, 10^3\}$，モデルの更新回数 (Round) を 250 とし，学習率は 0.1 とする．モデルの精度評価として，3 回試行による分類精度 (Accuracy) の平均を用いる．

メカニズム　ベースラインとして DPSGD[1] を用い, SkM と PB を比較する．プライバシーパラメータ $\epsilon = \{0.5, 1, 3\}$ とし，有限体の位数 m を 8bit，16bit とする．なお，SkM のスケールパラメータ $s = 64$ とする．

4.2.2　実験結果

モデル Large にて，グループサイズ n, プラバシーパラメータ ϵ を変化させたときの結果を**図 3** に示す．

SkM は，16bit のときいずれの条件でも精度が高かった．8bit のときは，グループサイズ $n = 10^3$ のときに精度が 10%前後であり, うまく学習できていないことがわかった．PB は，グループサイズ $n = 10^2$ のとき，SkM に比べて 30〜40%ほど精度が低かった．グループサイズ $n = 10^3$ のときは，PB と 16bit の SkM は同程度の精度を示した．

グループサイズ $n = 10^2$ のときのプライバシーパラメータ ϵ の影響は PB の方が SkM より大きかった．プライバシーパラメータ $\epsilon = 1$ の SkM の精度を基準すると，SkM では，16bit のときプライバシーパラメータ $\epsilon = \{0.5, 3\}$ いずれでも精度は同等であった．8bit のとき，プライバシーパラメータ $\epsilon = 0.5$ で精度が 0.9 倍，プライバシーパラメータ $\epsilon = 3$ で精度は同等だった．一方，PB では，プライバシーパラメータ $\epsilon = 1$ の PB の精度を基準とすると，16bit のとき，プライバシーパラメータ $\epsilon = 0.5$ で精度が 0.7〜0.8 倍，プライバシーパラメータ $\epsilon = 3$ で精度が 1.4〜1.5

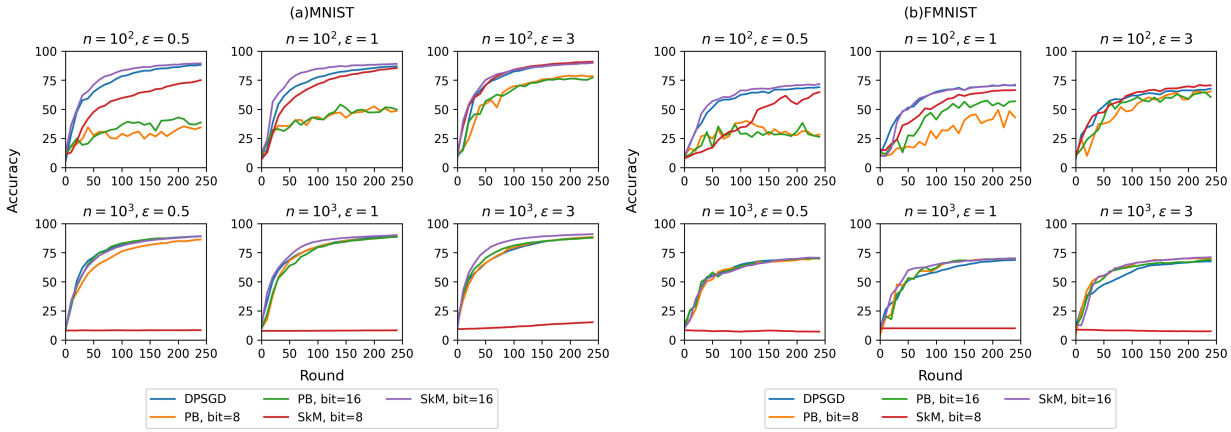

図 3 パラメータ (グループサイズ n, プラバシーパラメータ ϵ) をかえた評価.
データセットは (a)MNIST (b)Fashion-MNIST.

Fig. 3 Evaluation on (a)MNIST and (b)Fashion-MNIST with varying group size n and privacy parameter ϵ

倍であった．8bit のとき，プライバシーパラメータ $\epsilon = 0.5$ で精度が 0.6 倍，プライバシーパラメータ $\epsilon = 3$ で精度が 1.3〜1.4 倍であった．

グループサイズ $n = 10^3$ のとき，プライバシーパラメータ ϵ の影響は SkM，PB いずれもかなり小さかった．

プライバシーパラメータ $\epsilon = 1$ にて，モデルのパラメータ数 d とグループサイズ n を変化させたときの結果を**図 4** に示す．SkM は，モデル Tiny にていずれの bit でも精度が PB よりも低かった．モデル Small については，グループサイズ $n = 10^2$ かつ 8bit のときは PB よりも低かった．PB は，グループサイズ $n = 10^3$ のときパラメータ数 d にかかわらず DPSGD と同等の精度を示した．

5. 考察

Skellam Mixture メカニズム (SkM)[6] と Poisson Binomial メカニズム (PB)[10] の連合学習 (FL) における特性を実験結果から考察する．その後，パラメータ調整という観点でメカニズムを考察し，最後に FL に分散差分プライバシー (DDP) を適用する際のメカニズムを選ぶ指針を提案する．

5.1 Skellam Mixture メカニズムの特性

SkM は，ベクトル平均ではグループサイズにかかわらず低 bit の有限体の位数においても誤差に対して大きな影響はなかった．しかし，FL においては，グループサイズ 1,000 のような大きい場合には特に学習に悪影響を与えた．この事象を考察するために，8bit におけるスケールパラメータの影響を実験したものを**図 5** に示す．SkM はスケールパラメータ $s = \{64, 128\}$ のように大きいときははほとんど学習できていないが，スケールパラメータ $s = \{8, 16, 32\}$ のような極端に小さすぎないときは学習できている．SkM におけるスケールパラメータは学習の効率とモジュラークリッピングによる誤差とのバランスをとるものである．そのため，スケールパラメータが小さいほど勾配の $\mathcal{L}_\infty$ ノルムのクリップサイズも小さくなり ([6] の (3) 式を参照)，勾配の情報が大きく抑制されて学習が進みづらくなる一方でモジュラークリッピングによる誤差を減らすことができる．逆にスケールパラメータが大きいほど勾配の $\mathcal{L}_\infty$ ノルムのクリップサイズは大きくなり，勾配の情報が大きくは抑制されずに学習が進みやすくなる一方でモジュラークリッピングによる誤差は増える．特に，有限体の位数が低 bit の状況において，モジュラークリッピングによる誤差が多くなり，グループサイズ分の誤差が蓄積していたため学習がうまくいかなかったと考えられる．

次元数については，ベクトル平均と FL いずれでも PB より影響を受けやすかった．これは節 3.2 でも言及したとおり，PB はノイズの設定に次元数を考慮している ((5) 式) 一方で SkM は考慮していない ((4) 式) ため，次元に応じたノイズ調整がなされていないからだと考えられる．

5.2 Poisson Binomial メカニズムの特性

PB は，ベクトル平均では bit の大きさが誤差に影響していたが，FL では必ずしも精度に影響があるとは限らなかった．それよりはグループサイズの影響が強く，グループサイズ 100 のような小さい値ではプライバシー強度が高いほど学習が進みづらく，一方でグループサイズ 1,000 のような大きい値ではプライバシー強度が高くても DPSGD[1] と同程度の精度を示した．PB はモジュラークリッピングが不要である分ランダマイズによるノイズが大きくなっていると考えられ，ノイズの影響を小さくするためには大きなグループサイズが必要だと考えられる．

5.3 メカニズムにおけるパラメータ調整の容易さ

パラメータ調整の観点では，PB の方が容易であるとい

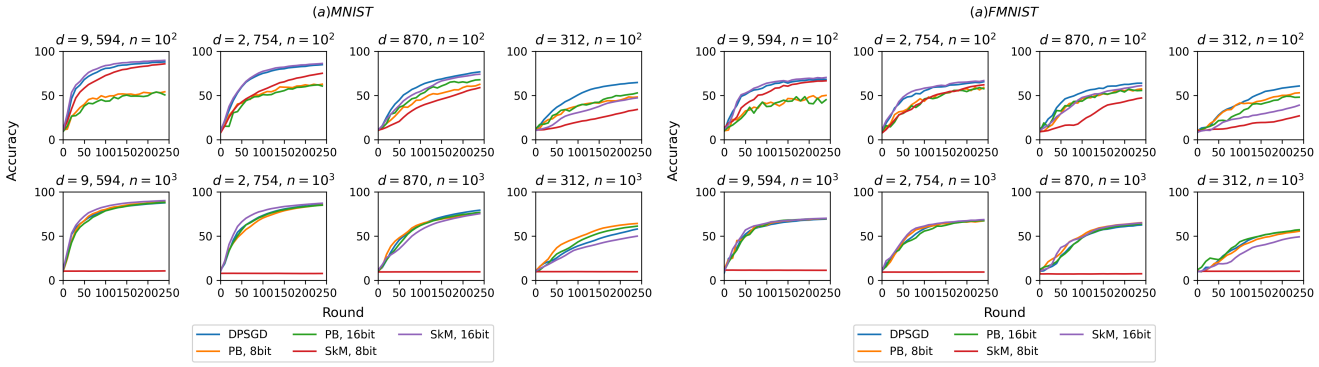

図 4 パラメータ (グループサイズ n，モデルのパラメータ数 d) をかえた評価．
プライバシーパラメータ $\epsilon = 1$，データセットは (a)MNIST (b)Fashion-MNIST．

Fig. 4 Evaluation on (a)MNIST and (b)Fashion-MNIST with privacy parameter $\epsilon = 1$ and varying parameter count of model d and group size n.

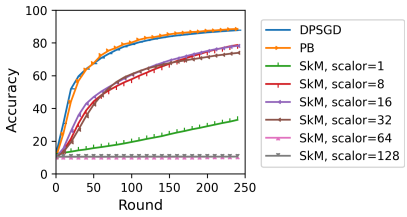

図 5 スケールパラメータ s をかえた SkM の評価．
プライバシーパラメータ $\epsilon = 1$，データセットは MNIST，グループサイズ $n = 10^3$，有限体の位数 8bit，モデル Large．

Fig. 5 Evaluation of SkM on MNIST and with privacy parameter $\epsilon = 1$, group size $n = 10^3$, Large model and order of a 8-bit finite field varying scale parameters s.

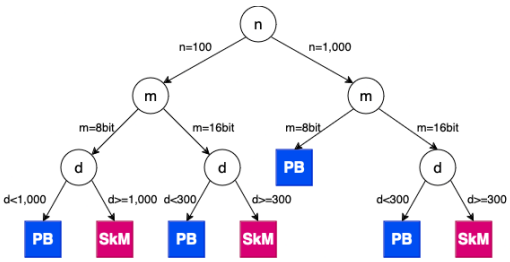

図 6 本実験結果におけるメカニズムの選び方の指針．

Fig. 6 Mechanism selection guidelines from experiments.

える．FL に PB を適用する際にはグループサイズの設定が重要であり，有限体の位数やモデルのパラメータ数による学習への影響は SkM に比べて小さい．SkM は，グループサイズが大きく，有限体の位数を低 bit に設定する場合はスケールパラメータの調整を適切に行わなければ学習をうまくすすめることができない．以上より，SkM は PB に比べて注意深くパラメータ調整をする必要があるといえる．

5.4 メカニズムの選び方の提案

FL おける DDP 適用時のメカニズムの選択指針を提案する (**図 6**)．メカニズムに依存せず，かつメカニズムの効果に影響を与える FL のパラメータ，グループサイズ n，モデルのパラメータ数 d，有限体の位数 m の 3 つを考慮する．

大まかな方針をいえば，モデルのパラメータ数 d が小さい設定であれば PB がよい．モデルのパラメータ数 d が大きい場合，グループサイズ n が小さい設定であれば SkM，グループサイズ n が大きく有限体の位数 m が小さい場合は PB，そうでない場合は SkM がよい．

より詳細な方針を示す．グループサイズ $n = 1,000$ 程度を設定する場合，有限体の位数 m を 8bit のような小さい値に設定するならば，PB がよい．有限体の位数 m を 16bit のような大きい値に設定し，モデルのパラメータ数 d が 300 程度のような極めて小さい際には PB がよい．モデルのパラメータ数 d がより大きいのであれば，SkM がよい．

グループサイズ $n = 100$ 程度を設定する場合，有限体の位数 m を 8bit のような小さい値に設定し，モデルのパラメータ数 d が 1,000 程度のような小さい際には PB がよい．モデルのパラメータ数 d がより大きいのであれば，SkM がよい．有限体の位数 m を 16bit のような大きい値に設定し，モデルのパラメータ数 d が 300 程度のような極めて小さい際には PB がよい．モデルのパラメータ数 d がより大きいのであれば，SkM がよい．

6. おわりに

連合学習 (FL) において，厳密なプライバシー保護を達成するためのプライバシーモデルとして分散差分プライバシー (DDP) を導入することが可能である．様々な DDP メカニズムが提案されているものの，メカニズム間の優劣は明確ではなく，実用面でどの手法を選択すべきかを判断することは容易ではない．我々の調査から，Skellam Mixture メカニズム [6] と Poisson Binomial メカニズム [10] を最先端の手法とみなし，この 2 つのメカニズムの比較，実験評価を行った．メカニズムに依存しない FL の 3 つのパラメータに基づき，最適なメカニズムが選択可能であることを示し，これをふまえて FL に DDP を適用する際のメカニズムの選択の指針を提案した．

参考文献

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 308–318, 2016.

[2] Naman Agarwal, Peter Kairouz, and Ziyu Liu. The skellam mechanism for differentially private federated learning. *Advances in Neural Information Processing Systems*, 34:5052–5064, 2021.

[3] Naman Agarwal, Ananda Theertha Suresh, Felix Xinnan X Yu, Sanjiv Kumar, and Brendan McMahan. cpsgd: Communication-efficient and differentially-private distributed sgd. *Advances in Neural Information Processing Systems*, 31, 2018.

[4] Apple. Privacy-Preserving Contact Tracing. `https://covid19.apple.com/contacttracing`, 2020. Accessed on August 23, 2024.

[5] Borja Balle, Gilles Barthe, Marco Gaboardi, Justin Hsu, and Tetsuya Sato. Hypothesis testing interpretations and renyi differential privacy. In *International Conference on Artificial Intelligence and Statistics*, pages 2496–2506. PMLR, 2020.

[6] Ergute Bao, Yizheng Zhu, Xiaokui Xiao, Yin Yang, Beng Chin Ooi, Benjamin Hong Meng Tan, and Khin Mi Mi Aung. Skellam mixture mechanism: a novel approach to federated learning with differential privacy. *Proc. VLDB Endow.*, 15(11):2348–2360, jul 2022.

[7] James Bell, K. A. Bonawitz, Adrià Gascón, Tancrède Lepoint, and Mariana Raykova. Secure single-server aggregation with (poly)logarithmic overhead. Cryptology ePrint Archive, Paper 2020/704, 2020. `https://eprint.iacr.org/2020/704`.

[8] Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth. Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 1175–1191, New York, NY, USA, 2017. Association for Computing Machinery.

[9] Clément L Canonne, Gautam Kamath, and Thomas Steinke. The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688, 2020.

[10] Wei-Ning Chen, Ayfer Ozgur, and Peter Kairouz. The poisson binomial mechanism for unbiased federated learning with secure aggregation. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 3490–3506. PMLR, 17–23 Jul 2022.

[11] Li Deng. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine*, 29(6):141–142, 2012.

[12] Cynthia Dwork. Differential privacy. In *Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II*, ICALP'06, page 1–12, Berlin, Heidelberg, 2006. Springer-Verlag.

[13] Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, Ananth Raghunathan, Shuang Song, Kunal Talwar, and Abhradeep Thakurta. Encode, shuffle, analyze privacy revisited: Formalizations and empirical evaluation. *arXiv preprint arXiv:2001.03618*, 2020.

[14] Andrew Hard, Kanishka Rao, Rajiv Mathews, Swaroop Ramaswamy, Françoise Beaufays, Sean Augenstein, Hubert Eichner, Chloé Kiddon, and Daniel Ramage. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*, 2018.

[15] Swanand Kadhe, Nived Rajaraman, O Ozan Koyluoglu, and Kannan Ramchandran. Fastsecagg: Scalable secure aggregation for privacy-preserving federated learning. *arXiv preprint arXiv:2009.11248*, 2020.

[16] Peter Kairouz, Ziyu Liu, and Thomas Steinke. The distributed discrete gaussian mechanism for federated learning with secure aggregation. In *International Conference on Machine Learning*, pages 5201–5212. PMLR, 2021.

[17] B Kashin. Section of some finite-dimensional sets and classes of smooth functions (in russian) izv. *Acad. Nauk. SSSR*, 41:334–351, 1977.

[18] Shiva Prasad Kasiviswanathan, Homin K Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.

[19] Yurii Lyubarskii and Roman Vershynin. Uncertainty principles and vector quantization. *IEEE Transactions on Information Theory*, 56(7):3491–3501, 2010.

[20] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[21] H Brendan McMahan, Daniel Ramage, Kunal Talwar, and Li Zhang. Learning differentially private recurrent language models. *arXiv preprint arXiv:1710.06963*, 2017.

[22] Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pages 263–275. IEEE, 2017.

[23] Fan Mo, Hamed Haddadi, Kleomenis Katevas, Eduard Marin, Diego Perino, and Nicolas Kourtellis. Ppfl: privacy-preserving federated learning with trusted execution environments. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '21, page 94–108, New York, NY, USA, 2021. Association for Computing Machinery.

[24] John Nguyen, Kshitiz Malik, Hongyuan Zhan, Ashkan Yousefpour, Mike Rabbat, Mani Malek, and Dzmitry Huba. Federated learning with buffered asynchronous aggregation. In Gustau Camps-Valls, Francisco J. R. Ruiz, and Isabel Valera, editors, *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pages 3581–3607. PMLR, 28–30 Mar 2022.

[25] Mohamed Sabt, Mohammed Achemlal, and Abdelmadjid Bouabdallah. Trusted execution environment: what it is, and what it is not. In *2015 IEEE Trustcom/BigDataSE/ISPA*, volume 1, pages 57–64. IEEE, 2015.

[26] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.

[27] Lingchen Zhao, Jianlin Jiang, Bo Feng, Qian Wang, Chao Shen, and Qi Li. Sear: Secure and efficient aggregation for byzantine-robust federated learning. *IEEE Transactions on Dependable and Secure Computing*, 19(5):3329–3342, 2022.