リッチクライアント-エッジサーバ間における プライバシ保護に優れた分散機械学習

理学専攻・情報科学コース　高野紗輝

1　はじめに

スマートフォンや IoT デバイス等のエッジデバイス上で収集される大量のデータを，個人情報を保護しながらクラウドやエッジサーバ上での機械学習に用いることが盛んに研究されている．エッジデバイスで収集したデータを活用する既存手法としては，クラウドコンピューティングやエッジコンピューティング [1] が挙げられる．特にエッジコンピューティングは，ネットワークエッジにエッジサーバを配置し，データ処理を最大限エッジで行うコンピューティングモデルであり，エッジデバイスで収集される大量のデータを高速に処理する手法として近年注目されている．しかし，エッジデバイスで収集するデータには個人情報等の機密性の高い情報が含まれる可能性がある．メンバーシップ推論攻撃やデータポイズニングなどの攻撃手法によりこれらの個人情報が漏洩や改竄される危険性があり，データをエッジデバイスの外部へと持ち出すことに対してプライバシの問題が生じる．

そこで，エッジデバイスで収集されるデータのプライバシ保護を強固にしつつ，これらのエッジサーバに渡すことのできない機密性の高いデータも含めた機械学習を行うことを目指す．本研究では，エッジサーバと連携しつつエッジデバイス上でも機械学習を動かすことで，プライバシ保護に優れた分散機械学習の実現を目指す．

2　関連研究 (Federated Learning)

Federated Learning [2] は，近年登場した高性能なエッジデバイス上でも機械学習処理を行い，プライバシに配慮した分散機械学習モデルとして注目されている．Federated Learning では，まずクラウド上のデータで学習を行って得られた学習モデルを各デバイスに配布し，各デバイスはそれぞれが収集した固有のデータを利用してさらに学習を進めた上で変更点の情報のみをクラウドに送信する．そして，クラウドは各デバイスから収集した変更点を平均化し，元の学習モデルを改善してより良いモデルを作成する．このように各デバイスで収集した生データをデバイスの外部に受け渡さないため，プライバシを担保しつつデバイス上にあるデータを機械学習に活用することが可能となる．しかし，Federated Learning では，学習結果をクラウドが集約・一括管理するため，生データはデバイスの外部へと受け渡さないものの，情報の一部を受け渡す必要がある．そのため，プライバシの保護は十分であるとは言えず，クラウドに送信されるパラメタからデバイスで収集したデータを解読することが可能である [3].

3　提案モデル

修士論文に示す 3 つの提案モデルのうち，エッジデバイスでの学習をエッジサーバへと安全にフィードバックする提案モデルを以下に示す．図 1 に示す通り，エッジサーバにおいて，あらかじめ一般的なデータを用いて学習を行い，学習の重みを保存したチェックポイントファイルを作成しておく．スマートフォンなどのエッジデバイスが移動し，エッジサーバに接続すると，エッジサーバ上で作成されたチェックポイントファイルを受け取る．このチェックポイントファイルを読み込み，エッジデバイスで収集したデータを用いて学習を再開する．この際，個人情報を含むデータを用いた学習と個人データを含まないデータを用いた学習の 2 通りの学習を行う．そして，ユーザの許可を得た個人情報の含まない学習結果のみをエッジサーバへフィードバックする．エッジサーバでは集約された複数の学習結果を統合し，エッジサーバ上のモデルを更新する．このモデルでは，エッジデバイスで収集した個人情報はエッジデバイス内のみで処理を行い，ユーザの許可を得た結果以外の情報をエッジサーバへ一切渡さないという特徴を持つため，安全にエッジデバイス上のデータを活用することができる．

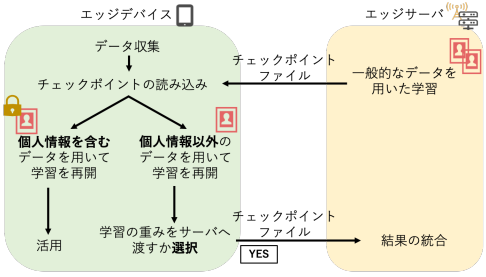

図 1: 提案モデル

4　実験

4.1　実験環境

使用したエッジサーバの性能は，OS Ubuntu 18.04 LTS，CPU Intel Core i7-8700，GPU GeForce RTX 2080Ti，Memory 32 Gbyte であり，エッジデバイスとして使用した Jetson Nano の性能は，OS Ubuntu 18.04 LTS，CPU Quad-core ARM A57 @ 1.43 GHz，GPU 128-core Maxwell，Memory 4 GB 64-bit LPDDR4 25.6 GB/s である．Jetson Nano は GPU を搭載した小型 AI コンピュータボードであり，近い将来スマートフォンや様々な IoT デバイスがこのような性能を持つことが期待される．

以下では一例として，エッジデバイスとして Jetson Nano を 2 台用意し，提案モデルを実装する．実験データには，実際のアプリケーションなどで使用されることが想定される機密性が高い顔画像を使用する．インターネット上より収集した 30 人分の jpg 画像を人物ごとにフォルダ分けし，そのうち 2 割を test データ，残りを train データとして表 1 のように重複のないように分配する．エッジデバイスには個人情報が含まれることが想定されるため，上記の 30 人とは異なる人物をそれぞれのエッジデバイスに train データ 48 枚，test データ 12 枚となるように加える．さらに，train デー

タをぼかし等で 9 倍に加工して使用する．

表 1: train データ (1 人物あたり)

	サーバ	デバイス 1	デバイス 2
一般的なデータ	24 枚ずつ	12 枚ずつ	12 枚ずつ
個人データ	なし	48 枚	48 枚

エッジサーバとエッジデバイスにおいて同等の精度を得るための実行時間を比較すると，エッジデバイスはおよそ 10 倍の時間を要する．低速ではあるものの，エッジデバイス内のみでも十分学習できるが，エッジデバイスのみでの学習には限界があり，エッジサーバとの連携が重要になる．

4.2 実験概要

初めにエッジサーバにおいて個人情報を含まない一般的なデータを用いて十分学習させ，学習の重みを保存したチェックポイントファイルをエッジデバイスへと送信する．そして，エッジデバイス上において個人の顔画像も含むデータを用いた学習と含まないデータを用いた学習の 2 通りの学習を再開する．本実験では，個人情報の含まない学習結果をユーザの許可を得たものとしてエッジサーバへフィードバックする．

4.3 実験結果

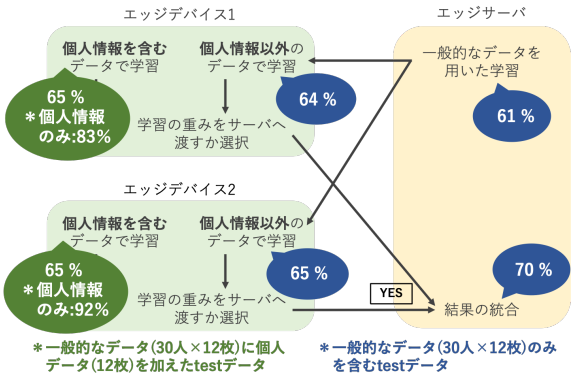

図 2: 実験結果

各ステップでの学習精度を図 2 に示す．エッジデバイスにおいて個人情報を用いて学習を再開することにより，個人情報に対して高い精度で判別が可能となり，一般的なデータに対しても判別可能となった．さらに，初めにエッジサーバ上のみで学習した結果では一般的なデータに対して 61%の精度しか得ることができなかった．一方で，それぞれのエッジデバイスで個人情報を含まずに学習した結果をエッジサーバに集約し，統合すると 70%となった．複数のデバイスで学習した結果が統合されたことにより，より大量のデータを学習した結果を得ることができたため，精度が向上したと考えられる．この新たに得た結果をエッジデバイスに再配布することで，エッジデバイス上でさらに良い結果を得ることが可能になると期待できる．

4.4 エッジサーバとの連携の効果

エッジサーバ上での学習をエッジデバイスで引き継ぐ効果について示す．エッジデバイス 1 でチェックポイントファイルを読み込んだ直後からの時間を横軸として学習精度を図 3 に示す．

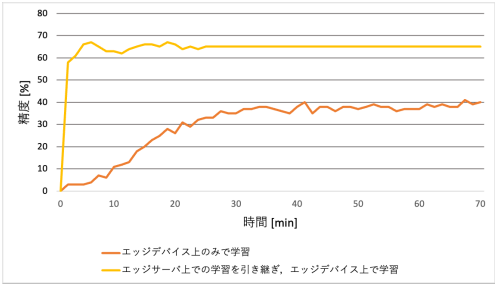

図 3: エッジサーバとの連携の効果

赤のグラフが，エッジサーバの助けを借りずにエッジデバイス上のみで学習を行った際の結果であり，黄色のグラフが，エッジサーバ上で一般的なデータを用いて学習を行った結果を引き継ぎ，エッジデバイス上で学習を再開させた際の結果である．精度は一般的なデータにエッジデバイス 1 の個人情報を加えた test データを用いる．エッジデバイスの性能の低さから，エッジデバイス上で機械学習を動かすにはかなりの時間がかかる．さらに，エッジデバイスで収集された一般的なデータの枚数はエッジサーバに比べ少ないため，精度が低い結果となる．そのため，エッジサーバの助けを借りることが有効であると言える．

5 まとめと今後の課題

プライバシ保護を強固に行なった上でエッジデバイスで収集した個人情報を含めた機械学習を可能とすることを目的として，高性能なエッジデバイスを用いた分散機械学習モデルの検討を行った．Jetson Nano を用いて実装した結果，プライバシ保護が可能であり，エッジデバイス上において短時間で個人情報を反映した結果が得られることが示された．さらに，ユーザの許可を得た個人情報を含まない学習結果をエッジサーバへとフィードバックすることにより，エッジサーバ上においてより多くのデータを反映した精度の高い結果を得ることが可能であった．

今後はフィードバックを行う情報を制限することによる細かい制御の検討を予定している．

参考文献

[1] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella. On multi-access edge computing: A survey of the emerging 5g network edge cloud architecture and orchestration. *IEEE Communications Surveys Tutorials*, Vol. 19, No. 3, pp. 1657–1681, 2017.

[2] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 10, No. 2, pp. 1–19, 2019.

[3] Mengkai Song, Zhibo Wang, Zhifei Zhang, Yang Song, Qian Wang, Ju Ren, and Hairong Qi. Analyzing user-level privacy attack against federated learning. *IEEE Journal on Selected Areas in Communications*, Vol. 38, No. 10, pp. 2430–2444, 2020.