プライバシ保護に優れた分散機械学習における リッチデバイス-エッジサーバ間の連携の検討

高野　紗輝† 　中尾　彰宏†† 　山口　実靖††† 　小口　正人†

† お茶の水女子大学　〒 112–8610 東京都文京区大塚 2–1–1
†† 東京大学　〒 113–8654 東京都文京区本郷 7–3–1
††† 工学院大学　〒 163–8677 東京都新宿区西新宿 1–24–2
E-mail: †saki-t@ogl.is.ocha.ac.jp, ††nakao@nakao-lab.org, †††sane@cc.kogakuin.ac.jp, ††††oguchi@is.ocha.ac.jp

あらまし　従来の分散機械学習分野の研究では，高性能なサーバが全てのデータを管理するものが多い．そのため，個人情報を外部のサーバへ受け渡すことによる情報漏洩の危険性がある．この課題に対し，我々はエッジデバイス上のデータを外部へと一切受け渡さない選択が可能なプライバシ保護に優れた分散機械学習モデルを検討する．本論文では，エッジサーバでの学習をエッジデバイスで引き継ぎ，ユーザの許可を得た結果のみをエッジサーバで統合する提案モデルを，エッジデバイスとして Jetson Nano を用いて実装した．その結果，エッジデバイス上では短時間で個人情報にも対応した結果を，エッジサーバ上ではより精度の高い結果をプライバシを保護しつつ得ることが可能であることが示された．

キーワード　分散機械学習, エッジコンピューティング, Federated Learning, IoT デバイス

1 はじめに

近年，スマートフォンや IoT デバイスの普及および性能向上により，エッジデバイス上で膨大なデータが収集されるようになった．さらに，おすすめ表示や画像認識などと機械学習を用いたデータ活用機会が増加している．エッジデバイスで収集した大量のデータには個人情報等の機密性の高いデータも含まれており，プライバシ保護を強固に行った上で機械学習に用いることが期待されている．

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 [1] [2] [3]，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．遠隔にあるクラウドのサーバと比較して物理的に近い位置で処理を行うことにより，利点として低遅延である点やエッジデバイスで処理を行うことでクラウドサーバにかかる負荷を分散できる点，エッジデバイスからクラウドサーバへ送信されるデータ量を削減し，トラフィックの混雑を解消できる点が挙げられる [4]．このような利点を活かし，エッジコンピューティングはスマートシティ [5] [6] や高度道路交通システム [7] などで IoT アプリケーションに応用され，クラウドコンピューティングでは実装することができなかったリアルタイムに応答するシステムが構築されている．

一方で，上記のコンピューティングモデルでは，全ての機械学習を高性能なサーバ上で行う．そして，性能の低いエッジデバイス側はあくまでデータを収集し，そのデータをサーバに転送するという役割を果たしている．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ている可能性がある．メンバーシップ推論攻撃 [8] やデータポイズニング [9] 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．特に近年，欧州で EU 一般データ保護規則 (GDPR, General Data Protection Regulation) が定められるなど，プライバシ保護への関心が高まっており，エッジデバイスで収集される個人データをサーバへ受け渡すことへの抵抗がさらに大きくなると予想される．その結果，サーバで一般的なデータを用いて学習した結果しか利用することができず，個人情報やエッジデバイスで収集した最新のデータを反映したそれぞれのエッジデバイスに最適な学習結果を得ることが難しくなると考えられる．

近年エッジデバイスの性能向上は著しく，エッジデバイスでのデータ処理能力がさらに上がる事が期待されているため，エッジデバイス上でも重い機械学習処理を行い，上記の課題解決に挑戦する．一方，エッジデバイスの性能はサーバと比較してかなり低いため，性能の高いサーバとの連携が必要になると考えられる．本稿では，一般的なデータを用いてエッジサーバ上で行った学習を，エッジデバイスで収集したデータを用いてエッジデバイス上で引き続き行うことでプライバシ保護に優れた分散機械学習モデルの検討を行う．また，エッジデバイス上での学習結果のうち，個人情報が含まれず，かつユーザの許可を得た結果のみをエッジサーバに集約し，統合することを検討する．本提案モデルは，エッジデバイスで収集したデータに加え，学習結果等の個人情報を一部含む情報も一切エッジサーバ

に送信しないという特徴を持つため，プライバシ保護に優れている．さらに，エッジデバイスとして Jetson Nano を用いた実験を行い，エッジサーバ上での学習をエッジデバイス上で引き継ぐことで，短時間でエッジデバイス，エッジサーバ共により精度の高い学習結果を得ることが可能となることを示した．

本稿の貢献を以下に示す．

（1） 個人情報をサーバへと送る従来手法で生じているプライバシの問題を解決しつつ，個人情報を活用することが可能となるモデルを提案し，実装した．
（2） シミュレーションではなく Jetson Nano という実機を用いて実験を行い，モデルの有効性を示した．
（3） エッジサーバと連携することにより，エッジデバイス上での学習時間を大幅に短縮し，より良い精度を得ることが可能であることを示した．
（4） エッジデバイス上での学習をエッジサーバへ安全にフィードバックすることで，エッジサーバ上においてより多くのデータが反映された精度の高い結果を得ることが可能であることを示した．

本稿の構成は以下の通りである．第 2 章においてエッジデバイス上で機械学習処理を行う分散機械学習の 1 つである Federated Learning を紹介する．第 3 章で研究課題とその解決手法を提案し，第 4 章で Jetson Nano を用いた実装及び評価を行う．最後に第 5 章でまとめる．

2 関連研究 (Federated Learning)

デバイス上でも機械学習を行うモデルとして，Federated Learning (連合学習) という分散型機械学習モデルが提案された [10] [11] [12]．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．Federated Learning はエッジコンピューティングとは異なり，プライバシに配慮しながらエッジデバイスの情報をクラウドに集約し，クラウドが一括管理するコンピューティングモデルとなっている．

論文 [13] では，Federated Learning を Google キーボードに応用した例が実装されており，デバイスの持つ固有のデータを受け渡すことなく，デバイス - クラウド間にまたがる分散機械学習が可能であることが示されている．その他にも，Federated Learning は機密性の高いデータを扱う医療現場における情報共有 [14] やヘルスケアアプリケーション [15]，自動車運転時における通信 [16]，スマートシティでのセンサから取得したデータの利用 [17] といった様々な分野での応用が期待され，近年盛んに研究が行われている．

一方で，Federated Learning は安全と言われているが，ユーザには学習結果をサーバへ集約するかどうかの選択権が与えられておらず，必ず学習後のパラメタをデバイスの外部へと送信している．パラメタには学習データの情報が多少含まれているため，生データはデバイスの外部へと受け渡していないものの，情報の一部を受け渡していることになる．そのため，Federated Learning における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 [18]．一例として，デバイスで学習した画像をクラウドに送信したパラメタから鮮明に復元可能であることが示されている [19] [20]．つまり，個人情報の漏洩の心配が全くないとは言い切れない．そのため，情報漏洩の許されない機密性の高いデータを従来の Federated Learning の学習に用いることは好ましくない．

3 課題と解決手法の提案

3.1 研究課題

3.1.1 プライバシ保護

最重要な課題としてプライバシ保護がある．従来のエッジコンピューティングの手法ではサーバに全てまたは一部の情報を集約し，サーバが一括管理する．また，Federated Learning ではエッジデバイスで学習した全てのデータをクラウドサーバに転送しており，ユーザ側にサーバへフィードバックを行うかどうかの選択権が与えられていない．そのため，情報漏洩の危険性がある．プライバシ保護の観点からすると，暗号化されたパラメタを含め，エッジデバイスで収集した個人情報を一切サーバへ受け渡さない手法が安心である．

3.1.2 エッジデバイスで得た学習結果のフィードバック

先行研究 [21] では，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エッジデバイスへと送信し，エッジデバイス上で個人情報に関する学習を引き続き行う分散機械学習モデルの提案を行った．このモデルでは，エッジデバイスからエッジサーバ方向へのデータのやり取りが一切行われないため、情報漏洩の心配が一切ない．しかし，エッジデバイスで学習した結果をエッジデバイス上以外で利用しないため，その有用性は限定的である．プライバシを強固に保護しつつ，エッジサーバに学習結果を集約し，統合していくことが期待される．また、集約・統合を行う従来の Federated Learning ではデバイスから学習結果を集約した場合とそうでない場合での学習精度の詳細な考察がなされていない．本稿では，フィードバックによる精度の向上について実機を用いた実験によって示す．

3.2 提案モデル

従来の IoT デバイスなどのエッジデバイスと比較し，データ処理能力や通信能力が格段に高い小型デバイスが登場し始めている．その結果，データ収集と結果表示のみを行うだけのクライアントではなく，複雑なデータ処理を行うことも想定するリッチクライアントの使用が期待されている．このような機械学習等のある程度複雑な処理も行うことが可能な高性能なデバ

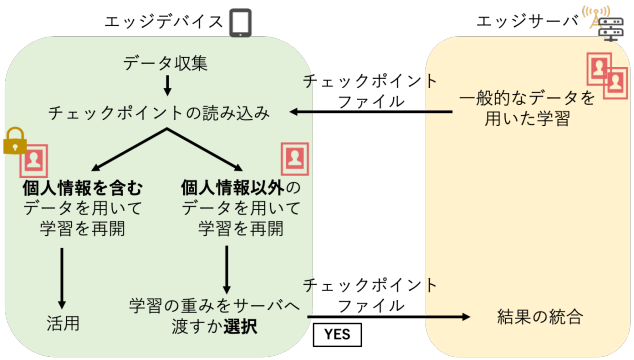

図 1　提案モデル

イスが登場したため，上記の課題の解決を目指したリッチクライアントに適した分散機械学習モデルを提案する．ここでは，従来のエッジコンピューティングモデルにおいてエッジサーバ上で行っていたタスクの一部をエッジデバイスへとオフロードすることで，エッジデバイス上でも機械学習処理を行い，個人情報の含まれない学習結果をエッジサーバへと送信するか否かを選択可能とすることで，プライバシ保護を強固に行う手法を提案する．

本論文においては，エッジサーバ上では一般的なデータに関して最新のデータも含めて精度の高い結果を得ることが目標であり，エッジデバイス上では一般的なデータに加えて個人データに関しても精度の良い結果を得ることが目標である．そして，エッジデバイスでは端末の所有者の個人データに加え，ローカルで収集した最新の一般的なデータが得られることを想定する．

提案モデルの概要図を図 1 に示す．あらかじめエッジサーバ上において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，エッジデバイスは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情報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，個人情報を含まないデータを用いた学習のうちユーザの許可を得た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．

3.3　想定されるアプリケーション

スマートフォンや IoT デバイスなどの端末では，個人を特定可能な情報を含む記念撮影などの画像と風景のみを撮影した画像の両方が収集されると考えられる．例えばこれらの画像は観光スポットの混雑予想などに用いることが想定され，個人を特定可能な画像を含めて学習を行うことでリアルタイムな混雑状況を反映した結果を得ることができる．さらに個人を特定可能な画像を除いて学習を行なった結果をエッジサーバへと送信し，エッジサーバのモデルを更新することで，より正確な混雑予想が可能なアプリケーションが構築できる．その他にも，玄関先に取り付けられた防犯カメラの動画像を用いた道路状況予測アプリケーションや，ペットの健康状態を画像から判断するアプリケーションなど様々な場面への応用が期待される．

表 1　実験で用いる 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 (Colin Powell)	48 枚 (George W Bush)

4　提案手法の実装と評価

4.1　データセット

本評価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用いて行う．インターネット上より有名人の jpg 画像を収集し，人物毎にフォルダ分けを行う．それぞれの画像の顔抽出を行い，適切に抽出を行うことのできていない画像を取り除いた後，各フォルダの 8 割を train データとする．残りの画像を test データとする．ここでは，30 人分の画像を収集し，それぞれ train データ 48 枚，test データ 12 枚となるようデータセットを作成した．

後述の実験 1，2 では，上記の train データをエッジサーバ 1 台とエッジデバイス 2 台 (エッジデバイス 1，エッジデバイス 2) に表 1 のように分配して用いる．ここでは，各人物の枚数がエッジサーバ：エッジデバイス 1：エッジデバイス 2 = 2：1：1 となるようにし，重複のないように分配する．結果，train データをエッジサーバは 1 人につき 24 枚ずつ，エッジデバイスはそれぞれ 1 人につき 12 枚ずつ保持することとなる．このデータを公開されている一般的なデータとする．一方で，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，上記の人物とは異なる人物をそれぞれのエッジデバイスに加える．エッジデバイス 1 には Colin Powell，エッジデバイス 2 には George W Bush の顔画像を加える．エッジデバイスではその持ち主の写真が多く収集されると考えられるため，それぞれ train データは 48 枚，test データは 12 枚となるように追加する．

さらに train データは，1 つの画像データに偏って学習してしまうことや過学習を防ぐため，ぼかし等により 9 倍にして使用する．

4.2　実 験 環 境

実験で使用したエッジサーバの性能を表 2 に，エッジデバイスとして使用した Jetson Nano の性能を表 3 に示す．

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．しかし，性能はエッジサーバと比較すると劣り，GPU のコア数がエッジサーバは 4352 コアであるのに対し，Jetson Nano は 128 コアと大きな差がある．

本実験では分散処理に適している TensorFlow を機械学習に

表 2　エッジサーバの性能

OS	Ubuntu 18.04 LTS
CPU	Intel Core i7-8700
GPU	GeForce RTX 2080Ti
Memory	32Gbyte

表 3　エッジデバイス (Jetson Nano) の性能

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

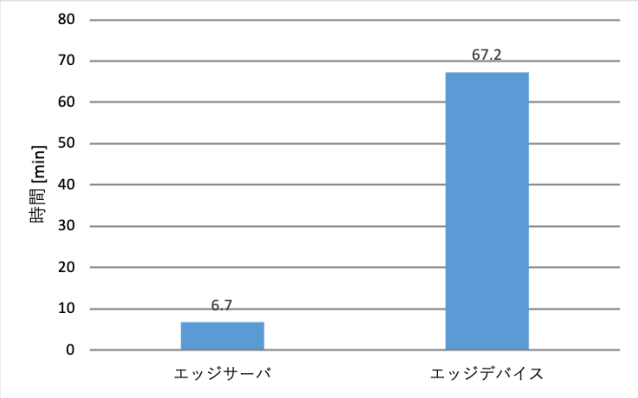

図 2　エッジサーバとエッジデバイスによる実行時間の比較

使用し，Jetson Nano - エッジサーバ間はイーサネットで接続する．

4.3　予 備 実 験

エッジサーバとエッジデバイスにおいて機械学習処理を行った際の実行時間を比較する．作成したデータセットの train データ (1 人あたり 48 枚) 全てをエッジサーバとエッジデバイスそれぞれに与え，ぼかし等により 9 倍にして使用する．エッジサーバ，エッジデバイス共に 30 クラス分類の精度が 75 % となるよう学習した結果を図 2 に示す．

エッジデバイス上でもエッジサーバと同等精度の学習を行うことができるものの，およそ 10 倍の時間を要し，75 % の精度を得るために 1 時間以上の学習が必要となる． このことから，エッジデバイスは低速ではあるが，エッジデバイス内のみでも十分学習可能であることが分かり，プライバシが非常に重要なデータもそのような形で学習に用いることができる．しかし，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と考えられる．

4.4　実験 1 (エッジデバイス：1 台)

4.4.1　実験概要 (実験 1)

エッジサーバとエッジデバイス 1 のみを用いて提案モデルを実行する．

まず初めに，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 epoch 数を 150 として十分に学習を行う．エッジサーバの性能は高く，短時間で多くの学習を行うことが可能であるため，エッジサーバの持つデータにおいて学習の上限となる精度を得ることが可能な epoch 数を設定する．そして，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．エッジデバイスは受け取った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．個人の顔画像を含むデータを用いた学習，個人の顔画像以外のデータを用いた学習の順に epoch 数を 20 としてエッジデバイス上で機械学習処理を行う．本実験では，個人の顔画像以外のデータを用いた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送信する．エッジサーバでは，初めにエッジサーバで学習した結果とエッジデバイス 1 から受け取った学習結果を統合する．ここでは，学習の重みの平均を取る．

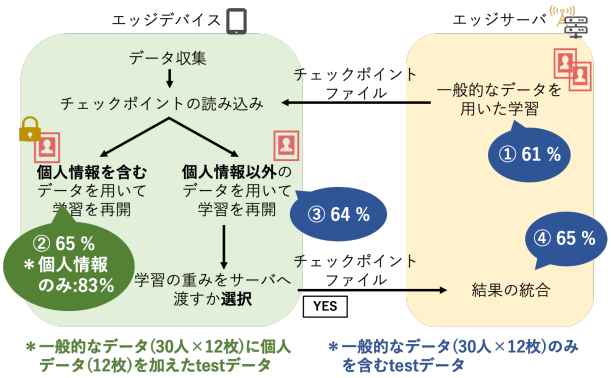

図 3　エッジデバイス 1 台で提案モデルを実行した際の学習精度 (実験 1)

4.4.2　実験結果 (実験 1)

エッジサーバ上での学習後にエッジサーバ上で計測した精度 (①)，学習を引き継ぎ，個人情報を含めたデータで学習した後にエッジデバイス上で計測した精度 (②)，個人情報を含まないデータで学習した後にエッジデバイス上で計測した精度 (③)，エッジサーバで学習結果を統合した後にエッジサーバ上で計測した精度 (④) を図 3 に示す．精度の計測は，個人情報を含めたデータで学習した後の精度 ② は一般的なデータにそのエッジデバイスの個人情報を加えた test データを用い，それ以外は一般的なデータのみの test データを用いる．

エッジサーバ上では ① で示すように一般的なデータに対して 61 % まで学習することが可能であった．得られたチェックポイントファイルをエッジデバイスへと渡した後，エッジデバイス上で個人情報が含まれるデータを用いて学習し，個人情報を含む test データを用いて精度を計測すると ② で示すように 65 % となった．詳細を示すと，個人情報に関しては 83 %，一般的なデータのみに関しては 65 % の精度となっている．個人情報に対しては高い精度で判別が可能であり，一般的なデータに対しても識別応可能という結果となった．一方で，個人情報を含めずに学習を再開し，一般的なデータのみを用いて精度を計測すると ③ で示すように 64 % となった．一方で，この結果は個人情報を全く学習していない結果であるため，個人情報を判別することは一切できなかった．そのため，個人に関する情報は全く含められていない結果となる．この結果をエッジ

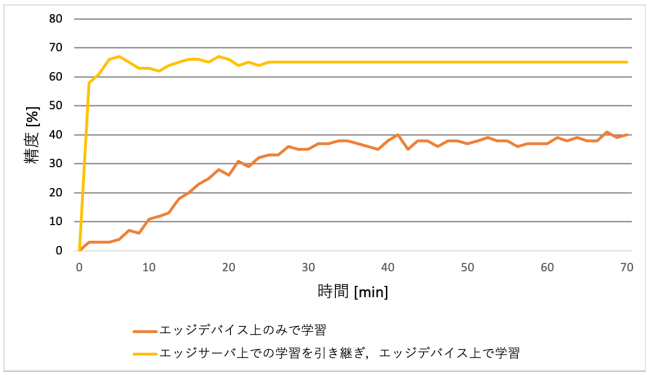

図 4　学習精度と時間の関係

サーバに転送し，初めにエッジサーバで学習した結果と統合すると，④ で示すように 65 % と ① の 61 % から精度の向上が確認できた．

4.4.3　考察 (エッジサーバとの連携の効果)

実験 1 において，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．すなわち，初めにエッジサーバ上で学習を行い，その結果をエッジデバイスへと送信するエッジサーバとの連携部分の有効性について示す．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 4 に示す．

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結果である．精度は一般的なデータにエッジデバイス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．

4.5　実験 2 (エッジデバイス：2 台)

4.5.1　実験概要 (実験 2)

エッジサーバとエッジデバイス 1 およびエッジデバイス 2 を用いて提案モデルを実行する．実行方法は実験 1 と同様とする．統合部分では，エッジサーバで初めに学習した結果，エッジデバイス 1 および エッジデバイス 2 よりフィードバックされた学習結果の 3 つの重みの平均を取る．

4.5.2　実験結果 (実験 2)

各ステップでの学習精度を図 5 に示す．エッジサーバ，エッジデバイス 1 での学習は実験 1 と同様である．エッジデバイス 2 では，個人情報を含む学習で得た結果の精度を個人情報を含む test データを用いて計測すると 65 %となった．個人情報に関しては 92 % と高い精度で判別することができており，個人情報を含む学習が可能であった．また，個人情報を含まない学習で得た結果の精度を一般的なデータのみを用いて計測すると 65 % となった．そして，エッジサーバで初めに学習した結果とエッジデバイス 1 で個人情報を含まずに学習した結果，エッジデバイス 2 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，統合すると精度は 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反映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，エジデバイス上でさらに良い結果を得ることが期待できる．

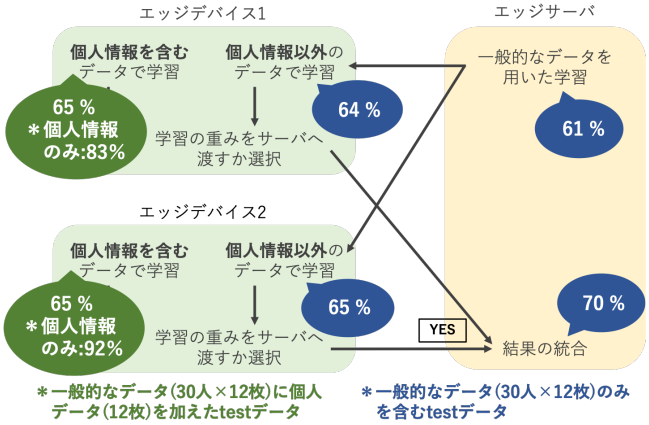

図 5　エッジデバイス 2 台で提案モデルを実行した際の学習精度 (実験 2)

5　まとめと今後の課題

従来のエッジコンピューティングモデルで課題となっている，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学習を可能とすることを目的として，リッチクライアントを用いた分散機械学習モデルの検討を行った．エッジサーバで一般的なデータを用いて学習した結果をエッジデバイスに引き継ぎ，エッジデバイス上で学習した結果のうち個人情報を含まない学習結果をエッジサーバへフィードバックするかをユーザが選択可能なモデルを提案し，エッジデバイス側に Jetson Nano を用いて実験を行った．

その結果，エッジサーバ上において一般的なデータで学習を行い，エッジデバイスが学習を引き継ぐことで，個人情報にも一般的な情報にも対応した学習結果をエッジデバイス上で短時間で得ることが可能であることが示された．本提案モデルでは，個人情報に関わる情報はエッジデバイスの外部へと一切持ち出さないため，研究課題である情報漏洩の恐れのない機械学習が可能となる．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し，統合することで，エッジサーバにおいてより多くのデータで学習した精度の高い結果を得ることが可能であった．

現在はフィードバックの有無を提案したが，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についても検討を行う予定である．

謝　　辞

本研究は一部，JST CREST JPMJCR22M2 の支援を受け

たものである．

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] MG Sarwar Murshed, Christopher Murphy, Daqing Hou, Nazar Khan, Ganesh Ananthanarayanan, and Faraz Hussain. Machine learning at the network edge: A survey. *ACM Computing Surveys (CSUR)*, Vol. 54, No. 8, pp. 1–37, 2021.

[3] Weisong Shi, Jie Cao, Quan Zhang, Youhuizi Li, and Lanyu Xu. Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, Vol. 3, No. 5, pp. 637–646, 2016.

[4] Mahadev Satyanarayanan. The emergence of edge computing. *Computer*, Vol. 50, No. 1, pp. 30–39, 2017.

[5] N. Chen, Y. Chen, S. Song, C. Huang, and X. Ye. Poster abstract: Smart urban surveillance using fog computing. In *2016 IEEE/ACM Symposium on Edge Computing (SEC)*, pp. 95–96, 2016.

[6] Bo Tang, Zhen Chen, Gerald Hefferman, Shuyi Pei, Tao Wei, Haibo He, and Qing Yang. Incorporating intelligence in fog computing for big data analysis in smart cities. *IEEE Transactions on Industrial Informatics*, Vol. 13, No. 5, pp. 2140–2150, 2017.

[7] J. Lin, W. Yu, X. Yang, Q. Yang, X. Fu, and W. Zhao. A real-time en-route route guidance decision scheme for transportation-based cyberphysical systems. *IEEE Transactions on Vehicular Technology*, Vol. 66, No. 3, pp. 2551–2566, 2017.

[8] Reza Shokri, Marco Stronati, and Vitaly Shmatikov. Membership inference attacks against machine learning models. *CoRR*, Vol. abs/1610.05820, , 2016.

[9] Sanghyun Hong, Varun Chandrasekaran, Yigitcan Kaya, Tudor Dumitras, and Nicolas Papernot. On the effectiveness of mitigating data poisoning attacks with gradient shaping. *CoRR*, Vol. abs/2002.11497, , 2020.

[10] 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.

[11] Li Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin. A review of applications in federated learning. *Computers & Industrial Engineering*, Vol. 149, p. 106854, 2020.

[12] Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, and Yuan Gao. A survey on federated learning. *Knowledge-Based Systems*, Vol. 216, p. 106775, 2021.

[13] T. Yang, G. Andrew, Hubert Eichner, Haicheng Sun, W. Li, Nicholas Kong, D. Ramage, and F. Beaufays. Applied federated learning: Improving google keyboard query suggestions. *ArXiv*, Vol. abs/1812.02903, , 2018.

[14] Jie Xu, Benjamin S Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research*, Vol. 5, No. 1, pp. 1–19, 2021.

[15] Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N Galtier, Bennett A Landman, Klaus Maier-Hein, et al. The future of digital health with federated learning. *NPJ digital medicine*, Vol. 3, No. 1, pp. 1–7, 2020.

[16] Dongdong Ye, Rong Yu, Miao Pan, and Zhu Han. Federated learning in vehicular edge computing: A selective model aggregation approach. *IEEE Access*, Vol. 8, pp. 23920–23935, 2020.

[17] Ji Chu Jiang, Burak Kantarci, Sema Oktug, and Tolga Soyata. Federated learning in smart city sensing: Challenges and opportunities. *Sensors*, Vol. 20, No. 21, p. 6230, 2020.

[18] Milad Nasr, Reza Shokri, and Amir Houmansadr. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE symposium on security and privacy (SP)*, pp. 739–753. IEEE, 2019.

[19] 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.

[20] Jonas Geiping, Hartmut Bauermeister, Hannah Dröge, and Michael Moeller. Inverting gradients-how easy is it to break privacy in federated learning? *Advances in Neural Information Processing Systems*, Vol. 33, pp. 16937–16947, 2020.

[21] Takano Saki, Nakao Akihiro, Yamaguchi Saneyasu, and Oguchi Masato. Privacy-protective distributed machine learning using rich clients. *2021 International Conference on Emerging Technologies for Communications (ICETC 2021), IEICE Proceedings Series*, Vol. 68, No. C1-4, 2021.