# OTA 連合学習におけるコホートベースの送信電力制御と勾配回復

Cohort-based Transmit Power Control and Gradient Recovery for Over-The-Air Federated Learning

**江　易翰**（Yi-Han CHIANG、Ph. D.）
**大阪公立大学　大学院工学研究科**
**　電気電子システム工学専攻　助教**
(Assistant Professor, Osaka Metropolitan University, Department of Electrical and Electronic Systems Engineering, Graduate School of Engineering)
IEEE　電子情報通信学会
研究専門分野：通信工学　情報ネットワーク

**あらまし**

無線通信環境における空中計算（OTA：Over-The-Air Computation）を利用した分散型機械学習であるOTA連合学習（Federated Learning）は、エッジAIや医療IoTなどへの応用が大いに期待されている．しかし、OTA連合学習の勾配集約と回復が無線チャネル品質に大きな影響を受けてしまう問題がある．本研究では、無線チャネル品質が近いエッジデバイスとなったコホートに基づいて送信電力制御と勾配回復の手法を提案する．まず、コホートを構築し、構築されたコホートは次の通信ラウンドでスケジューリングされる．第2に、構築された各コホートにおいて、最も弱い無線チャネル品質を持つエッジデバイスが送信電力バジェットを使い切るように、デバイスは個別に送信電力をスケーリングする．第3に、勾配集約の後、コホートベースで適用された送信電力と無線チャネル品質の情報を使用してローカル勾配を回復する．提案法の性能を評価するために、テスト精度の観点からいくつかの比較手法と比べる．さらに、提案法の性能が様々なシステムパラメータの下でどのように変化するかを検証する．

## １．研究の目的

近年、無線通信環境における無線信号の重畳特性を活用した空中計算（OTA：Over-The-Air Computation）に基づいたOTA連合学習が注目されている．しかし、先行研究では、無線チャネル品質が近いエッジデバイスのサブセット、すなわちコホート（Cohort）を構築することでOTA連合学習の学習性能向上に寄与することを軽視している．本研究では、OTA連合学習におけるコホートベースの送信電力制御と勾配回復の手法を提案し、シミュレーションで比較手法に対する優位性を示す．

## ２．研究の背景

連合学習*1において、最もよく知られたアルゴリズムであるFederated Averaging（FedAvg）[R1]の仕組みは、モデル配布、ローカル学習、モデル集約、勾配平均という4つのステップから構成されている．まずモデル配布では、パラメータサーバ（以下は「サーバ」という）はグローバルモデルを全てのエッジデバイス（以下は「デバイス」という）に配布する．次に、ローカル学習では、各デバイスは自身が所有するローカルデータセットを使用してローカルモデルを更新する．モデル集約では、サーバは選択されたデバイスからローカルモデルを集約し、集約されたローカルモデルの平均に基づいてグローバルモデルを更新する．これらの手順を1回の通信ラウンドとし、予め定義された収束基準が満たされるまで、複数回の通信ラウンドにわたって進行する．

従来の通信手法では、基地局（Base Station）が多数のデバイスから無線信号を受信して処理する場合、まず各デバイスからの無線信号を時空間的に分けて受信し、最後に統計処理がなされる．つまり、通信タスクと計算タスクが完全に分離されている．この方法では、デバイスの数が増加した場合に、膨大な通信リソース消費の恐れがある．一方、無線チャネルのマルチアクセスの特性を活用した空中計算*2は、図1に示すように、無線信号の重ね合わせによって合計または平均を計算することが意図されている．この方法では、どのデバイスからどのデータが送られてきたかを区別することは不可能である．しかし、そのような情報を

# OTA 連合学習におけるコホートベースの送信電力制御と勾配回復

Cohort-based Transmit Power Control and Gradient Recovery for Over-The-Air Federated Learning

必要としない場合、つまりデータの合計値などの統計的な値だけを必要とする場合には、空中計算はデータの送受信に要する通信リソースを大幅に減少する．空中計算を利用した OTA 連合学習では、サーバ側で行われる計算は平均などの統計演算のみであり、どのデバイスの情報を利用するかを知る必要がないため、通信リソース消費を削減することができる．

## ３．システムモデル

各通信ラウンド$t$において、サーバはまずデバイスのサブセット$\mathcal{J}(t) \subseteq \mathcal{J}$をスケジューリングし、スケジューリングされたデバイスは指定された周波数帯域で同期して更新されたローカルモデルをアップロードする．簡潔にするため、コホートサイズ$|\mathcal{J}(t)| = N$を満たす．また、各デバイス$j$が持つローカルデータセットを$\mathcal{D}_j = \{(\mathbf{x}_j^i, y_j^i)\}$とする．ただし、$\mathbf{x}_j^i$はローカルデータセット$\mathcal{D}_j$の$i$番目の特徴量、$y_j^i$はそれに対応するラベルである．ここで、ローカルデータセットに重複がないことを確実にするために、$j' \neq j''$の場合、$\mathcal{D}_{j'} \cap \mathcal{D}_{j''} = \emptyset$であるとする．また、データプライバシーを保護するため、$\mathcal{D}_j$はサーバに送信されない．その代わりに、データセットの大きさ$|\mathcal{D}_j|$がサーバに共有される．さらに、すべてのデバイスがローカル学習に使用するニューラルネットワークの構造は同一であると仮定する．

各通信ラウンド$t$において、サーバは以下の 4 つのステップに従ってデバイスと通信を行う（図 1 を参照）．

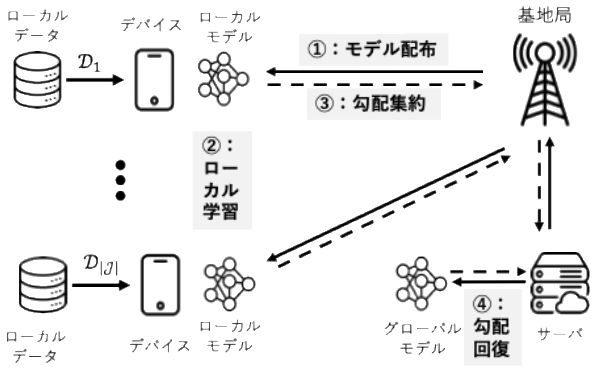


図 1：OTA 連合学習の仕組み

① **モデル配布：**まず、サーバは基地局を通してグローバルモデル$\mathbf{w}(t)$を通信ラウンド$t$でスケジューリングされたコホート$\mathcal{J}(t)$へ送信する．ただし、$\mathbf{w}(0)$はランダムなベクトルで初期化される．

② **ローカル学習：**次に、各デバイス $j \in \mathcal{J}(t)$は、配布されたグローバルモデル$\mathbf{w}_j(t)$と自身の持つデータセット$\mathcal{D}_j$を用いてローカルモデル$\mathbf{w}_j(t)$を以下の式で更新する．

$$\mathbf{w}_j(t) = \mathbf{w}_j(t-1) - \eta \mathbf{g}_j(t), \quad \forall j \in \mathcal{J}(t), 1 \leq t \leq T$$

ここで$\eta$は学習率を表し、ローカル勾配$\mathbf{g}_j(t)$は以下の式で計算される．

$$\mathbf{g}_j(t) = \frac{1}{|\mathcal{D}_j|} \sum_{i \in \mathcal{D}_j} \nabla f\left(\mathbf{w}(t), \mathbf{x}_j^i, y_j^i\right), \quad \forall j \in \mathcal{J}(t), 1 \leq t \leq T$$

ただし、$\nabla f\left(\mathbf{w}(t), \mathbf{x}_j^i, y_j^i\right)$はロス関数を示す．

③ **モデル集約：**ローカル学習が完了すると、$\tilde{\mathbf{g}}_j(t) = \mathbf{g}_j(t) / \|\mathbf{g}_j(t)\|$のようにローカル勾配の正規化を行う．そして、デバイス$j$は以下の式で送信信号を構成する．

$$\mathbf{s}_j(t) = \sqrt{p_j(t)} \tilde{\mathbf{g}}_j(t), \forall j \in \mathcal{J}(t), 1 \leq t \leq T$$

ただし、$p_j(t) \leq P_j^{\max}$は送信電力を表す．その後、コホート内のデバイスは基地局へ同時に信号を送信し、受信信号は以下のように表される．

$$\mathbf{r}(t) = \sum_{j \in \mathcal{J}(t)} h_j(t) \mathbf{s}_j(t) + \mathbf{z}(t), \forall 1 \leq t \leq T$$

ここで、$h_j(t)$はデバイス$j$と基地局間の無線チャネルを、$\mathbf{z}(t) \sim \mathcal{CN}(0, \sigma^2)$は雑音電力$\sigma^2$の加算性白色ガウス雑音を表す．

④ **勾配回復：**勾配集約が完了すると、受信信号はサーバへ転送される．ローカル勾配を受信信号から適切に復元するために、グローバル勾配を以下の式から得る．

$$\mathbf{g}(t) = \frac{\mathbf{r}(t)}{\xi(t)}, \forall 1 \leq t \leq T$$

ここで、$\xi(t)$は勾配回復係数を表す．最後に、グローバルモデルが以下のように更新される．

$$\mathbf{w}(t) = \mathbf{w}(t-1) + \mathbf{g}(t-1), \forall 1 \leq t \leq T$$

# OTA 連合学習におけるコホートベースの送信電力制御と勾配回復

Cohort-based Transmit Power Control and Gradient Recovery for Over-The-Air Federated Learning

## ４．提案法

通常の連合学習とは異なり、最大送信電力制約や無線チャネルの品質は、OTA 連合学習の学習性能に大きな影響を与える．OTA 連合学習の通信ラウンド間で、デバイスと基地局間の無線チャネル品質は時変であり、またデバイス間での無線チャネル品質のばらつきとデバイスの最大送信電力制約により、より良い無線チャネル品質を持ったデバイスからの信号が減衰してしまう．これらの問題に対処するため、本研究では動的にコホートを構築し、コホートベースの送信電力制御と勾配回復（CRAIC：Cohort-based Transmit Power Control and Gradient Recovery）アルゴリズムを行う手法を提案する．提案法は以下のコホート構築、送信電力制御、勾配回復の 3 つのステップから構成される．

① **コホート構築：**まず、デバイスの集合$\mathcal{J}$からランダムに一つの$j^*$を選択する．次に、

$$\left|\left|h_j(t)\right| - \left|h_{j^*}(t)\right|\right|, \forall j \in \mathcal{J}$$

を計算し、昇順ソートしたものを$Q(t)$とする．$Q(t)$の先頭から$N$個の値に対応する$j$をコホート$\mathcal{J}(t)$とする．

② **送信電力制御：**各通信ラウンド$t$においてコホート$\mathcal{J}(t)$内のデバイスは異なる最大送信電力やチャネル係数を持っている．ローカル勾配はグローバルモデルに対する変化量を表すので、デバイス間で、送信されるローカル勾配の相対的な大きさのオーダーを保持しながら適切に集約されるように、スケジューリングされたデバイスの送信電力を調整する必要がある．しかし一方で、コホート全体の送信電力が不足している場合、信号がノイズのレベルまで減衰し、発散してしまう．そこで、以下のように送信電力制御の手法を設計する．以下で$\sqrt{P_j(t)}\left|h_j(t)\right|$を合成チャネル係数と呼ぶことにする．

まず、コホート$\mathcal{J}(t)$を最大送信電力で送信した場合の合成チャネル係数の 2 乗の値で昇順にソートして$\mathcal{J}^\circ(t) = \{j_1^\circ, \dots, j_N^\circ\}$を作成する．次に、$\mathcal{J}^\circ(t)$の先頭から$c$番目のデバイスまでの合成チャネル係数の 2 乗の合計値を次のように定義する．

$$\Omega_c(t) = \sum_{n=1}^{c} (N-n+1)\left(\sqrt{P_{j_N^\circ}^{\max}}\left|h_{j_N^\circ}(t)\right|^2 - \sqrt{P_{j_{n-1}^\circ}^{\max}}\left|h_{j_{n-1}^\circ}(t)\right|^2\right), \forall 1 \le c \le N, 1 \le t \le T$$

次に、$c = \{N, \dots, 1\}$に対して、あらかじめ決めたノイズに対する値$\Gamma\sigma^2$を上回る場合は、$c^*$を定義し、下回るまでこの処理を繰り返す．そして、$c = \{1, \dots, c^* - 1\}$のデバイスは最大送信電力を用い、残りのデバイスは次のように送信電力を決定する．

$$p_{j_c^\circ}(t) = \frac{\Gamma\sigma^2 - \Omega_{c^*-1}(t)}{(N - c^* + 1)\left|h_{j_c^\circ}(t)\right|^2} + \frac{P_{j_{c^*-1}^\circ}^{\max}\left|h_{j_{c^*-1}^\circ}(t)\right|^2}{\left|h_{j_c^\circ}(t)\right|^2}, \forall c^* \le c \le N, 1 \le t \le T$$

ここで、$c = N$の場合にも$\Gamma\sigma^2$を下回る場合は、コホート$\mathcal{J}(t)$内の全てのデバイスが最大送信電力を用いる．このように送信電力制御を設計することで、コホート全体の送信電力が不足している場合は最大送信電力を用い、そうでない場合は、最大限デバイス間の合成チャネル係数の値を一定にすることで、各デバイス勾配を適切なオーダーを保ったまま集約することを目指す．

③ **勾配回復：**勾配の集約は、無線チャネル品質やノイズによって大きく変動する可能性がある．集約されたローカル勾配が元の勾配より小さすぎると、OTA 連合学習プロセスの収束が遅くなる可能性がある．逆に、集約されたローカル勾配が大きすぎると、学習データの分布をうまく表現できなくなり、OTA 連合学習プロセスが発散してしまう可能性がある．したがって、ローカル勾配を元の勾配と同程度の大きさに回復することが重要であり、それによって集約されたローカル勾配をローカル学習に適切に利用できるようにする．このため、通信ラウンド$t$における勾配回復係数を以下のように定義する．

# OTA 連合学習におけるコホートベースの送信電力制御と勾配回復

Cohort-based Transmit Power Control and Gradient Recovery for Over-The-Air Federated Learning

$$\xi(t) = \frac{1}{|\mathcal{J}(t)|}\sum_{j\in\mathcal{J}(t)}\sqrt{p_j(t)}\left|h_j(t)\right|, \forall 1 \le t \le T$$

## ５．シミュレーション

以下では、デフォルトのパラメータ設定を示す．128 個のデバイスは、基地局から測って$d^{\min} = 35\text{m}$と$d^{\max} = 1\text{km}$の範囲に一様に分布している．コホートサイズは 16、通信ラウンド数は 100 とする．各デバイスは 3 つの全結合層からなるニューラルネットワークを用いてローカルモデルを更新する．なお、最適化手法には Adam、ロス関数に Cross-Entropy、バッチサイズは 64、学習率には 0.0001 を用いる．また、デバイス間の学習データのラベルの分布がどの程度異なっているかを示す．MNIST データセット[R2]をデバイスへ振り分ける際、ディリクレ分布（Dirichlet Distribution）[R3]という確率分布に従って配分する．ディリクレ分布のパラメータ$\alpha$はデータの不均一度を示す．$\alpha \to \infty$の場合、デバイス間の配分された学習データのラベル分布は独立同分布（IID：Independent and Identically Distributed）に近づく．一方、$\alpha \to 0$の場合は、ラベル分布は non-IID 度合いが強まる．

提案法の性能を評価するため、以下の比較手法と比べる．

- Error-Free（EF）：EF は FedAvg を基にしており、各通信ラウンドにおいて$\mathcal{J}$からランダムに$N$個のデバイスが選択され、ローカル勾配は完全な状態で集約される．空中計算がないため、パスロスやノイズも存在せず、送信電力制御や勾配回復も必要ない．
- Cohort-based Power Scaling with Adaptive Transmit Powers（CPS-A）：CPS-A は空中計算において、各通信ラウンドにおいて$\mathcal{J}$からランダムに$N$個のデバイスが選択され、最大送信電力制約に基づいて送信電力制御される．チャネル反転を基にした送信制御アルゴリズム[R4]と同じである．勾配集約の後、提案した勾配回復が適用される．
- Cohort-based Power Scaling with Maximum Transmit Powers（CPS-M）：CPS-M は CPS-A と似ているが、各通信ラウンドにおいて選択された全てのデバイスが自身の持つ最大送信電力で送信する．つまり、送信電力制御は行われない．

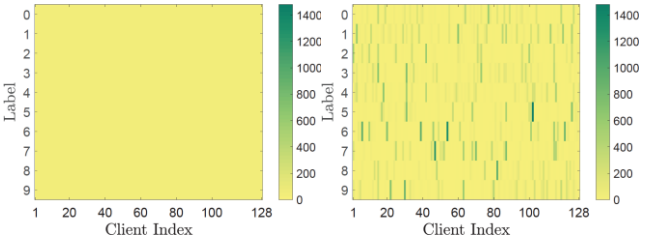


図 2：ディリクレ分布によるトレーニングデータの振り分け（左：$\alpha = 10^{20}$、右：$\alpha = 0.1$）

図 2 では、ディリクレ分布によって MNIST データセットの学習サンプルが 128 個のデバイスにどのように分配されるかを示している．図 2 の縦軸は MNIST データセットの 0 から 9 までのラベルを、横軸は 1 から 128 番目までのデバイスを示し、緑色が濃い部分に多くの学習サンプルが振り分けられている．$\alpha = 10^{20}$の場合、各ラベルの学習サンプルがほぼ均等にデバイスに分布するため、データ分布はほぼ IID となる．逆に$\alpha = 0.1$では、各ラベルの学習サンプルがデバイス間でより不均一になり、データ分布がより non-IID になる傾向があることがわかる．

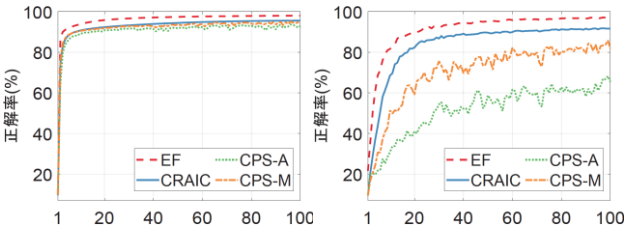


図 3：CRAIC と他の送信電力制御手法との正解率の比較（左：$\alpha = 10^{20}$、右：$\alpha = 0.1$）

図 3 に、CRAIC と他の送信電力制御手法との正解率の比較結果を示す．IID、non-IID ともに CRAIC は CPS-A と CPS-M に対して高い正解率を達成している．これは、CRAIC によって、類似したチャネル係数を持つデバイスがコホートに選択され、かつノイズに対する適切な送信電力を適用できているからであると考え

られる．さらに、CRAIC は EF に近い正解率を達成することができている．なぜなら、それぞれの大きさのオーダーをあまり変えないようにローカル勾配を集約することができ、その一方で、集約されたローカル勾配は元のものと同じようなレベルまで回復することができるからである．

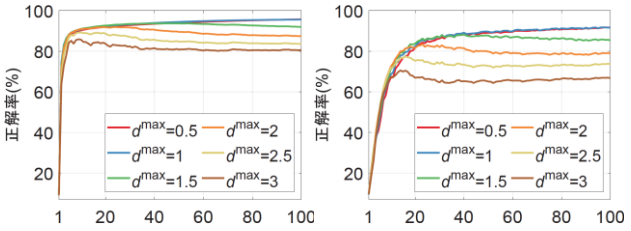


図 4：$d^{\max}$を変化による CRAIC アルゴリズムの性能評価（左：$\alpha = 10^{20}$、右：$\alpha = 0.1$）

図 4 は、デバイスの位置が CRAIC の性能にどのように影響するかを示している．デバイスが基地局に近い位置にある場合（すなわち、$d^{\max}$が小さい場合）、集約されたローカル勾配がノイズよりもはるかに強いため、CRAIC は良好に収束する．デバイスが基地局から遠く離れた場合、性能は低下するものの、CRAIC は IID と non-IID のどちらの設定でも収束性を維持しており、これは広域での OTA 連合学習の実現における CRAIC の可能性を明らかにしている．

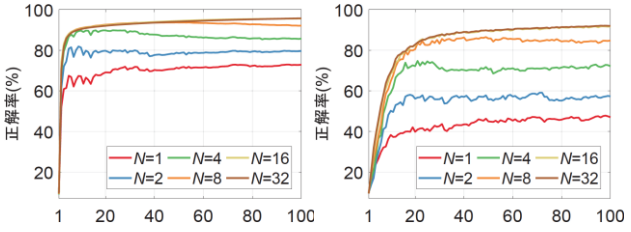


図 5：コホートサイズ$N$の変化による CRAIC アルゴリズムの性能評価（左：$\alpha = 10^{20}$、右：$\alpha = 0.1$）

図 5 はコホートサイズが CRAIC の性能にどのように影響するかを示している．基本的に、コホートサイズが大きいほどテスト精度は高くなるが、これはノイズが集約されたローカル勾配に比べて無視できるほど小さくなるためである．逆にコホートサイズを小さくすると、集約されたローカル勾配の受信信号の強度は必然的に低下する．それにもかかわらず、CRAIC はノイズに対して各コホート内のデバイスの送信電力を可能な限り増加させるコホートを構成することにより収束性を維持している．

## ６．将来展望

ローカルでのデータ保持やデータプライバシーの保護などの特徴を持つため、連合学習はエッジ人工知能等への道を切り開く上で非常に望ましいものである．本研究では空中計算を利用した OTA 連合学習のためのコホートベースの送信電力制御及び勾配回復技術の開発に注目しているものの、実際には調査する価値のある観点が他にもいくつかあり、それらがさらなる研究の課題と機会をもたらしうることは明白である．

## おわりに

本研究では OTA 連合学習におけるコホートベースの送信電力制御と勾配回復に関する問題を検討した．そのために、まずデバイスのチャネル係数に基づいてコホートを構築し、コホート全体の送信電力を考慮しながら、コホート内のチャネル係数に基づいて送信電力制御と勾配回復を実行する CRAIC アルゴリズムを提案した．シミュレーション結果によって、提案法は他の送信電力制御の手法を凌駕し、IID と non-IID の両方の設定においても理想状況の FedAvg に近いテスト精度を示した．さらに、提案法の性能はデバイスが基地局に近い位置にある場合、またはコホートサイズが大きくなった場合にさらに向上することが分かった．

## 用語解説

*1 連合学習（Federated Learning）：従来の機械学習と違って、各デバイスが所有している生データを中央のサーバに集計せず、デバイス同士が協調的に学習を行うことで、よりセキュアな環境で機械学習を実行できるのが特徴である．

*2 空中計算（Over-The-Air Computation）：無線マルチアクセスチャネルの信号重畳の特性を利用し、通信タスクと計算タスクを同時に行うような技術である．

# OTA 連合学習におけるコホートベースの送信電力制御と勾配回復

Cohort-based Transmit Power Control and Gradient Recovery for Over-The-Air Federated Learning


**参考文献**

[R1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. *AISTATS*, Apr. 2017, pp. 1273–1282.

[R2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[R3] M. Yurochkin, M. Agarwal, S. Ghosh, K. Greenewald, N. Hoang, and Y. Khazaeni, "Bayesian nonparametric federated learning of neural networks," in *Proc. ICML*, Jun. 2019, pp. 7252–7261.

[R4] X. Cao, G. Zhu, J. Xu, and K. Huang, "Optimized power control for over-the-air computation in fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7498–7513, 2020.

**関連文献**

[1] 寺井広大・江易翰・林海・計宇生（2022）．胸部 X 線画像におけるデータ不均一度が連合学習に与える影響に関する分析．令和 4 年電気関係学会関西連合大会講演論文集、220–221.

[2] C. Chen, Y.-H. Chiang, H. Lin, J. C. Lui, and Y. Ji, "Energy harvesting aware client selection for over-the-air federated learning," in *Proc. IEEE GLOBECOM*, 2022, pp. 5069–5074.

[3] 西本賢司・寺井広大・江易翰・林海・計宇生(2023)．モデルポイズニング攻撃に対処する連合学習の開発に関する研究．2023 年電子情報通信学会総合大会—情報・システム講演論文集 2、153.

[4] K. Nishimoto, Y.-H. Chiang, H. Lin, and Y. Ji, "FedATM: Adaptive trimmed mean based federated learning against model poisoning attacks," in *Proc. IEEE VTC-Spring*, 2023, pp. 1–5.

[5] C. Chen, Y.-H. Chiang, H. Lin, J. C. Lui, and Y. Ji, "Joint client selection and receive beamforming for over-the-air federated learning with energy harvesting," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1127–1140, 2023.

[6] 西本賢司・江易翰・林海・計宇生（2023）．モデルポイズニング攻撃に対処する適応的連合学習に関する研究．令和 5 年電気関係学会関西連合大会講演論文集、190–191.

[7] K. Terai, Y.-H. Chiang, H. Lin, and Y. Ji, "Cohort-based power scaling and gradient recovery for over-the-air federated learning," in *Proc. IEEE VTC-Fall*, 2023, pp. 1–5.

[8] Y.-H. Chiang, K. Terai, T.-W. Chiang, H. Lin, Y. Ji, and J. C. S. Lui, "Optimal transport-based one-shot federated learning for Artificial Intelligence of Things," *IEEE Internet Things J.*, vol. 11, no. 2, pp. 2166–2180, 2024.

[9] 寺井広大・江易翰・林海・計宇生（2024）．チャネル時変性に対処した OTA 連合学習におけるコホート作成．2024 年電子情報通信学会総合大会講演論文集、B-15-04.

[10] 西本賢司・江易翰・林海・計宇生（2024）．OTA 連合学習におけるポイズニング攻撃に対する防御手法の提案．2024 年電子情報通信学会ソサイエティ大会講演論文集、B-7-14.

[11] 久森敬太・江易翰・林海・計宇生（2024）．RIS を用いた OTA 連合学習における位相シフト最適化の研究開発．2024 年電子情報通信学会ソサイエティ大会講演論文集、B-7-11.

[12] K. Hisamori, Y.-H. Chiang, H. Lin, and Y. Ji, "Hybrid quantum-classical computing in federated learning with data heterogeneity," in *Proc. IEEE PIMRC*, 2024, pp. 1–5.



**この研究は、令和３年度ＳＣＡＴ研究助成の対象として採用され、令和３～４年度に実施されたものです。**