

Transmissão Inteligente de Parâmetros no Aprendizado Federado: Gerenciando o *trade-off* entre o desempenho e o consumo energético em Redes IoT sem Fio

Renan R. de Oliveira^{1,2}, Pedro Augusto S. Belo¹,
Nickolas Carlos C. Silva¹ e Antonio Oliveira-Jr^{1,3}

¹Instituto de Informática – Universidade Federal de Goiás (UFG), GO, Brasil

²Instituto Federal de Goiás (IFG), GO, Brasil

³Fraunhofer Portugal AICOS, Porto, Portugal

renan.rodrigues@ifg.edu.br, pedrobelo@discente.ufg.br,
nickolascarlos@egresso.ufg.br, antoniojr@ufg.br

Abstract. *Federated Learning (FL) allows devices to train a global machine learning model without sharing data, preserving users privacy. FL faces significant challenges related to energy cost, especially in resource-limited environments, such as wireless networks and emerging Internet of Things (IoT) solutions. This work presents an intelligent parameter transmission strategy in the context of FL in wireless IoT networks. The proposed approach is based on the relative percentage difference between the old weights and the new updated weights of the local models, using a defined threshold to decide on the need for transmission. The results demonstrate that conditional transmission reduces the number of transmissions without significantly compromising the accuracy of the global model.*

Resumo. *O Aprendizado Federado (Federated Learning - FL) permite que dispositivos treinem um modelo global de aprendizado de máquina sem o compartilhamento de dados, preservando a privacidade dos usuários. O FL enfrenta desafios significativos relacionados ao custo energético, especialmente em ambientes com recursos limitados, como redes sem fio e soluções emergentes da Internet das Coisas (Internet of Things - IoT). Este trabalho apresenta uma estratégia de transmissão inteligente de parâmetros no contexto do FL em redes IoT sem fio. A abordagem proposta é baseada na diferença percentual relativa entre os pesos antigos e os novos pesos atualizados dos modelos locais, utilizando um limiar definido para decidir sobre a necessidade de transmissão. Os resultados demonstram que a transmissão condicional reduz o número de transmissões sem comprometer significativamente a acurácia do modelo global.*

1. Introdução

A evolução do Aprendizado de Máquina (*Machine Learning* - ML) e o aumento da capacidade computacional impulsionaram inovações na automação de processos, previsão de tendências e tomada de decisões complexas. No entanto, as abordagens tradicionais de ML que dependem da centralização dos dados em servidores para o treinamento de modelos enfrentam limitações significativas. A necessidade de enviar grandes volumes de

dados para servidores não apenas aumenta os custos de transmissão, mas também coloca em risco a privacidade dos usuários. Portanto, esta abordagem pode tornar-se inadequada em setores com rigorosas exigências de privacidade e à necessidade de processamento em tempo real [Nakayama and Jenó 2022].

O FL tem se destacado como uma solução promissora para o treinamento distribuído de modelos de ML, especialmente em cenários onde a privacidade e a eficiência na utilização de recursos são essenciais [McMahan et al. 2016]. Essa abordagem permite que múltiplos dispositivos colaborem no treinamento de um modelo global sem a necessidade de compartilhar dados dos dispositivos e preservando a privacidade dos usuários. O FL em redes sem fio apresenta vantagens com relação ao ML centralizado, pois a transmissão dos parâmetros do modelo de ML em vez dos dados de treinamento entre os dispositivos e a Estação Base (*Base Station* - BS) pode economizar energia, recursos de rede e latência da comunicação [Yang et al. 2022]. Além do mais, a computação de borda integrada à infraestrutura da rede, conforme proposto na abordagem *Multi-Access Edge Computing* (MEC) [ETSI 2022], contribuiu para a implantação de servidores de parâmetros de FL na borda da rede, ou seja, mais próximo dos dispositivos onde os modelos são treinados.

No entanto, o FL enfrenta desafios significativos relacionados ao custo energético e à eficiência da comunicação, especialmente em ambientes com recursos limitados, como redes sem fio e soluções emergentes da IoT. Neste caso, a transmissão inteligente de parâmetros dos dispositivos emerge como uma estratégia fundamental para otimizar o desempenho do FL. A comunicação constante dos parâmetros de modelos treinados localmente pode gerar sobrecarga nas redes e consumo excessivo de energia, impactando diretamente a viabilidade da solução.

A comunicação contínua dos parâmetros entre os dispositivos e a BS pode resultar em sobrecarga da rede, alto consumo de energia e maior latência. Neste contexto, o presente trabalho propõe uma estratégia inteligente de transmissão de parâmetros para o FL em redes IoT sem fio, baseada na diferença percentual relativa entre os pesos antigos e os novos pesos atualizados. Utilizando um limiar definido por ε , cada dispositivo avalia se a mudança nos parâmetros é substancial para justificar uma nova transmissão. Se a diferença de pesos estiver abaixo de ε , o dispositivo abstém-se de enviar o modelo recém treinado, evitando comunicações desnecessárias e reduzindo o consumo de recursos na rede. Neste caso, o servidor agregador deve utilizar a versão obsoleta enviada pelo dispositivo em alguma rodada de comunicação anterior.

Os resultados obtidos mostram que a abordagem condicional, embora reduza ligeiramente a acurácia, oferece uma diminuição significativa no uso de energia. Ao considerar o valor de $\varepsilon = 25$, o desempenho do modelo global é de aproximadamente 5% menor na estratégia condicional quando comparada com a estratégia de transmissão total. No entanto, o consumo energético da estratégia condicional é aproximadamente 55% menor quando comparada com a estratégia de transmissão total.

As contribuições centrais deste artigo são elencadas a seguir: (i) apresentam-se as formulações matemáticas que definem o modelo do sistema utilizado para a simulação da rede IoT sem fio para tarefas de FL; (ii) explora-se a intersecção entre redes IoT sem fio e o FL, abordando os desafios relacionados ao gerenciamento eficiente de comunicação e

energia; (iii) apresentam-se avaliações de diferentes estratégias de transmissões e discute-se o impacto de cada abordagem no desempenho do modelo global de FL e no consumo energético dos dispositivos; (iv) disponibiliza-se o código-fonte ¹ da simulação para permitir a reprodução e validação dos resultados.

Para além desta seção introdutória, o restante deste artigo está organizado em seções, conforme descrito a seguir. A Seção 2 discute os trabalhos relacionados. A Seção 3 apresenta os modelos matemáticos que foram utilizados na simulação da rede IoT sem fio para tarefas de FL. A Seção 4 apresenta o cenário de simulação deste trabalho. A Seção 5 discute os resultados e a análise da simulação. Por fim, a Seção 6 apresenta as considerações finais.

2. Trabalhos Relacionados

Na formulação inicial do FL proposta por [McMahan et al. 2016], os autores introduziram o conceito de FL e apresentaram o desempenho do algoritmo *Federated Averaging* (FedAvg) com diferentes configurações, conjuntos de dados e modelos de ML. Em [Asad et al. 2020], são avaliadas diferentes estratégias de FL em termos da acurácia e eficiência de comunicação. No entanto, trabalhos de [McMahan et al. 2016] e [Asad et al. 2020] não abordam o desempenho do FL em ambientes de redes sem fio. O trabalho de [Chen et al. 2021] introduz um modelo de FL para redes sem fio, propondo uma seleção conjunta de dispositivos e alocação de recursos. De maneira similar, o estudo de [Chen et al. 2022] explora a otimização de recursos de comunicação para FL, dividindo o problema em escalonamento de dispositivos e alocação de recursos. Contudo, as pesquisas de [Chen et al. 2021] e [Chen et al. 2022] não abordam o problema do *trade-off* entre o desempenho e o consumo energético dos dispositivos.

Motivado pelos desafios do FL no contexto de redes sem fio e dispositivos IoT onde a eficiência energética é um fator a ser considerado, este trabalho propõe uma estratégia inteligente de transmissão de parâmetros para o FL em redes IoT sem fio, mostrando a eficácia da estratégia de transmissão condicional como uma solução que equilibra o desempenho do modelo global e o consumo energético.

3. Modelo do Sistema

3.1. Modelo de Rede

Considere uma rede IoT sem fio com uma BS conectada diretamente a um servidor agregador de FL conforme proposto na abordagem MEC, com um conjunto $S = \{i_1, i_2, \dots, i_N\}$ de N dispositivos IoT antes do início de uma rodada de treinamento. Cada cliente possui um conjunto de dados \mathcal{P}_i com $n_i = |\mathcal{P}_i|$ amostras armazenadas em seus respectivos dispositivos locais. Estes dispositivos IoT estão conectados a BS por meio de uma conexão sem fio e possuem capacidade para coletar dados e treinar um modelo local para uma determinada tarefa de FL.

3.2. Aprendizado Federado

A ideia do FL é realizar a agregação de um modelo global ao longo de várias rodadas de comunicação com base nos parâmetros dos modelos treinados localmente em cada

¹Disponível em: <https://github.com/PedroASB/federated-learning>

dispositivo, minimizando a função de perda $f(w_{global})$ usando os dados $\mathcal{P} = \cup_{i=1}^N \mathcal{P}_i$. Dessa forma, o processo de treinamento do FL pode ser expresso como

$$\begin{aligned} \min_{w_1, \dots, w_N} f(w_{global}) &\triangleq \min_{w_1, \dots, w_N} \frac{1}{|\mathcal{P}|} \sum_{i=1}^N \sum_{j=1}^{n_i} f(w_i, x_{ij}) \\ &= \min_{w_1, \dots, w_N} \sum_{i=1}^N \frac{n_i}{m} f(w_i) \\ \text{s.a. } w_1 &= w_2 = \dots = w_N = w_{global}, \end{aligned} \quad (1)$$

onde $f(w_i, x_{ij})$ é a função de perda que avalia o desempenho do modelo local w_i por meio da observação da saída produzida pelo treinamento com x_{ij} amostras de dados e $m = \sum_{i \in \mathcal{S}} n_i$. A restrição (1a) garante que os modelos compartilhados entre os dispositivos e a BS devem ser idênticos e explica o objetivo da FL em estabelecer o treinamento de um modelo w_{global} sem transferência de dados.

No início de cada rodada de comunicação t , a BS deve selecionar uma fração aleatória dos N dispositivos, gerando um conjunto \mathcal{S}_t de m dispositivos. Em seguida, a BS utiliza um enlace de comunicação para transmitir o estado atual do modelo w_t para cada dispositivo local. Cada dispositivo treina o modelo local utilizando um algoritmo de otimização baseado na Descida do Gradiente Estocástico (*Stochastic Gradient Descent*) para cada $b \in B$ mini-lotes de \mathcal{P}_i durante E épocas locais, onde $\nabla \ell$ representa o gradiente de ℓ em b . Após a última atualização $w \leftarrow w - \eta \nabla \ell(w; b)$, o dispositivo utiliza um canal de *uplink* do enlace de comunicação para enviar os parâmetros do modelo local w para a BS conectada a um servidor agregador de parâmetros.

Dessa forma, os modelos w_{t+1}^i são recebidos e agregados pelo servidor, gerando o estado atual w_{t+1} do modelo global aplicando a atualização $w_{t+1} \leftarrow \sum_{i \in \mathcal{S}_t} \frac{n_i}{m_t} w_{t+1}^i$, onde $m_t = \sum_{i \in \mathcal{S}_t} n_i$. A agregação trata-se da média ponderada dos parâmetros dos dispositivos, cujos pesos são definidos com base na quantidade de dados utilizada no treinamento local. O processo de treinamento distribuído pode ser repetido por várias rodadas de comunicação até que um determinado valor de convergência seja alcançado ou quando outro critério de parada seja satisfeito [Li et al. 2021].

3.3. Modelo de Comunicação e Consumo Energético

Em redes sem fio, os recursos podem ser alocados com base em Blocos de Recursos (*Resource Blocks* - RBs), que são unidades formadas por um conjunto de subportadoras contínuas no domínio da frequência e intervalos de tempo no domínio temporal, permitindo a otimização do uso do espectro e a alocação dinâmica e eficiente dos recursos conforme as necessidades de transmissão e as condições do canal [Wong 2012].

Neste contexto, este trabalho considera a técnica de acesso múltiplo por divisão de frequência ortogonal (*Orthogonal Frequency Division Multiple Access* - OFDMA) para o *uplink*, onde cada dispositivo ocupa um RB. De acordo com [Chen et al. 2021], a taxa de *uplink* do dispositivo i transmitindo os parâmetros do modelo w_i para a BS pode ser formulada como

$$c_i^U(r_i, P_i) = \sum_{n=1}^R r_{i,n} B_n^U \mathbb{E} \left(\log_2 \left(1 + \frac{P_i h_i}{I_n + B_n^U N_0} \right) \right), \quad (2)$$

onde $r_i = [r_{i,1}, \dots, r_{i,R}]$ é o vetor de alocação de RBs, R é o número de RBs, $r_{i,n} \in 0, 1$ e $\sum_{n=1}^R r_{i,n} = 1$, com $r_{i,n} = 1$ indicando que o RB n está alocado para o dispositivo i . A largura de banda de cada RB é denotada por B_n^U e P_i é a potência de transmissão do dispositivo i . O ganho do canal entre o dispositivo i e a BS é dado por $h_i = o_i d_i^{-\alpha}$, onde d_i é a distância entre o dispositivo i e a BS, o_i é o parâmetro de desvanecimento de Rayleigh e α é um expoente que afeta como o ganho do canal varia com a distância. $\mathbb{E}(\cdot)$ é a expectativa da taxa de dados em relação a h_i , N_0 é a densidade espectral da potência do ruído e I_n é a interferência em r_n causada por outros dispositivos.

A taxa de dados de *downlink* alcançada pela BS ao transmitir os parâmetros do modelo global para cada dispositivo é dada por

$$c_i^D = B^D \mathbb{E} \left(\log_2 \left(1 + \frac{P_B h_i}{I^D + B^D N_0} \right) \right), \quad (3)$$

onde B^D é a largura de banda utilizada pela BS para transmitir o modelo global para cada dispositivo, P_B é a potência de transmissão da BS e I^D é a interferência causada por outras BSs que não participam da tarefa de FL. Assume-se que os modelos de FL são transmitidos por meio de um único pacote. Dessa forma, o atraso de transmissão entre um dispositivo i e a BS no *uplink* e *downlink* podem ser respectivamente formulados como

$$l_i^U(r_i, P_i) = \frac{S_{pkt}^U}{c_i^U(r_i, P_i)}, \quad (4) \quad l_i^D = \frac{S_{pkt}^D}{c_i^D}, \quad (5)$$

onde S_{pkt}^U e S_{pkt}^D denotam, respectivamente, o tamanho do pacote no *uplink* e no *downlink*. Considera-se que a BS não solicitará aos dispositivos o reenvio de modelos quando estes forem recebidos com erros. Neste caso, conforme apresentado por [Chen et al. 2021], a taxa de erro de transmissão de pacote do *uplink* é dado por

$$q_i^U(r_k, P_i) = \sum_{n=1}^R r_{i,n} \mathbb{E} \left(1 - \exp \left(-\frac{m(I_n + B^U N_0)}{P_i h_i} \right) \right), \quad (6)$$

onde $\mathbb{E}(\cdot)$ é a expectativa da taxa de erro de pacote considerando h_i em r_n , com m sendo um limiar (*waterfall threshold*) que define a qualidade da transmissão. Por fim, o modelo de consumo de energia dos dispositivos para o treinamento e transmissão do modelo w_i pode ser formulados como

$$e_k(r_i, P_i) = \zeta \omega_i \vartheta^2 Z(w_i) + P_i l_i^U(r_i, P_i), \quad (7)$$

onde ϑ , ω_i e ζ referem-se, respectivamente, a frequência do *clock*, o número de ciclos da unidade central de processamento e o coeficiente de consumo de energia de cada dispositivo. Na Equação (7), a primeira parcela representa o consumo de energia para o treinamento do modelo local no dispositivo e a segunda parcela representa o consumo de energia para a transmissão de w_i para a BS.

4. Configuração dos Experimentos

4.1. Cenário de Simulação

Considere uma rede IoT sem fio atendendo uma área circular com raio r de 500 metros com uma BS no centro. A BS está diretamente associada a um servidor agregador, com

$N = 100$ dispositivos conectados para uma tarefa de FL, onde 10 dispositivos são selecionados em cada rodada de comunicação. Os dispositivos são distribuídos de maneira uniforme e aleatória com distâncias entre 100 e 500 metros da BS. A largura de banda de *uplink* de cada RB é de 1 MHz com a atribuição de uma interferência distinta e incremental. Cada dispositivo é configurado para transmitir os modelos locais para a BS com a potência de 0.01 W. A largura de banda do *downlink* é de 20 MHz, onde a BS transmite os modelos globais para os dispositivos com uma potência de 1 W. A modelagem de h_i incorpora um efeito de desvanecimento que indica que o ganho do canal diminui conforme aumenta a distância dos dispositivos em relação a BS. Outros parâmetros incluem: $\alpha = 2$, $N_0 = -174$ dBm/Hz, $m = 0.023$ dB, $\vartheta = 10^9$, $\zeta = 10^{-27}$ e $\omega_i = 40$.

4.2. Conjunto de Dados

As tarefas de FL consideram problemas de classificação de imagens utilizando o conjunto de dados de *benchmark* utilizado em pesquisas de FL denominado MNIST, conforme os trabalhos de [McMahan et al. 2016], [Amannejad 2020], [Beutel et al. 2020], [Zhu et al. 2020], [Chen et al. 2021], [Zhao et al. 2022] e [Chen et al. 2022]. Este trabalho utiliza uma variação do MNIST, nomeada como NIID R-MNIST, que tornam os dados heterogêneos de forma non-IID (*non-Independent and Identically Distributed*) como uma característica inerente das redes IoT sem fio. Dessa forma, o conjunto de dados foi dividido em 10 subconjuntos com amostras do mesmo rótulo, utilizando 75% das amostras para o conjunto de treinamento e 25% para o conjunto de teste. Em seguida, cada dispositivo recebeu uma partição de treinamento e teste, onde: 90% das amostras pertencem ao mesmo rótulo e os 10% restantes pertencem igualmente aos demais rótulos; cada imagem é rotacionada em até 45° em sentido horário ou anti-horário; a quantidade final de dados é dada por um fator entre [0.25, 1] da partição inicial.

4.3. Arquitetura de ML

Para a tarefa de classificação do conjunto de dados NIID R-MNIST, considera-se uma arquitetura de rede neural do tipo *Multi-Layer Perceptron* (MLP) com 101.770 parâmetros treináveis, composta por uma camada de entrada, seguida de uma única camada oculta com 128 neurônios, uma função de ativação ReLU e uma camada de saída *softmax*. Além do mais, é utilizado o ADAM (*Adaptive Moment Estimation*) como método de otimização do modelo, *Sparse Categorical Crossentropy* como função de perda, com um *batch size* de tamanho 128. Em cada rodada de comunicação do FL, o modelo local de cada dispositivo selecionado é treinado por 5 épocas locais. Ressalta-se que é possível utilizar diferentes arquiteturas de redes neurais mais profundas com o objetivo de atingir um melhor desempenho do modelo. No entanto, de acordo com [Zhao et al. 2022], mesmo que a precisão do modelo não atinja o estado da arte, isto ainda pode ser suficiente para a avaliação do comportamento de estratégias de FL em diferentes cenários.

4.4. Heurística para Atribuição de RBs no Uplink

O processo de alocação de RBs em redes sem fio é essencial para garantir a eficiência da comunicação entre os dispositivos e a BS. Neste trabalho, utiliza-se uma heurística que visa priorizar dispositivos mais distantes da BS, alocando a esses dispositivos os canais com maior SINR (*Signal-to-Interference-plus-Noise Ratio*). O SINR é um parâmetro fundamental para avaliar a qualidade do canal, considerando a interferência e o ruído.

A heurística para a atribuição de RBs no *uplink* utiliza a estratégia a seguir. Inicialmente, os dispositivos que participarão da rodada atual de FL são escolhidos aleatoriamente. Em seguida, a BS coleta as informações de SINR dos canais de *uplink*. Após a ordenação dos dispositivos considerando a distância da BS, o algoritmo atribui os canais com maior SINR aos dispositivos mais distantes, otimizando a comunicação e compensando a maior atenuação de sinal destes dispositivos. Depois da atribuição de RBs, a BS pode calcular o tempo necessário para cada transmissão e estimar o consumo de energia de cada dispositivo.

5. Análise dos Resultados

Esta seção apresenta os resultados da estratégia de transmissão condicional de parâmetros aplicadas no contexto do FL. A análise se concentra em aspectos relacionados a eficiência energética, sem considerar falhas ou retransmissões devido a perdas de pacotes. Neste contexto, são avaliadas três estratégias:

- **Transmissão Total:** Esta é a versão básica FL, onde todos os dispositivos selecionados sempre transmitem seus parâmetros atualizados após cada rodada de comunicação.
- **Transmissão Condicional:** A transmissão ocorre apenas quando há uma mudança significativa nos parâmetros do modelo local treinado por cada dispositivo considerando ε como um limiar predefinido.
- **Transmissão Aleatória:** A decisão de transmitir é baseada em uma probabilidade predefinida, funcionando como uma linha de base para comparar o desempenho da transmissão condicional.

Na estratégia de transmissão condicional, cada dispositivo decide, com base no cálculo da diferença de pesos, se a transmissão dos novos parâmetros é necessária. A diferença de pesos é obtida como a média percentual das diferenças relativas entre os pesos antigos e os pesos atualizados do modelo local recém-treinado. Esse processo evita comunicações desnecessárias e contribui para a otimização do uso dos recursos de comunicação e do consumo energético dos dispositivos. Neste caso, cada dispositivo utiliza um limiar ε para avaliar se a mudança nos parâmetros é substancial para justificar uma nova transmissão. Em casos de não transmissões, o servidor agregador deve utilizar a versão obsoleta enviada pelo dispositivo em alguma rodada de comunicação anterior.

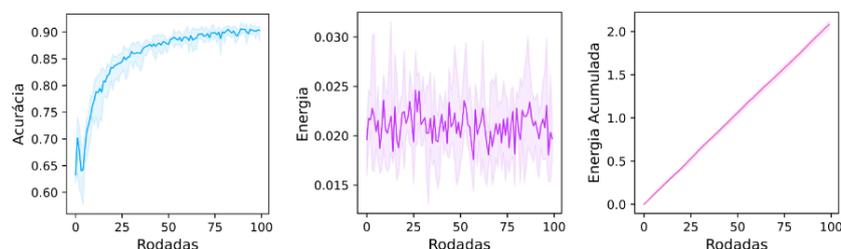


Figura 1. Estratégia de Transmissão Total

A Figura 1 apresenta a acurácia, o consumo energético por rodadas de comunicação e a energia acumulada durante o treinamento do modelo global por 100 rodadas de comunicação. Como não são considerados erros de transmissão, observa-se que a energia acumulada cresce linearmente, refletindo o aumento constante do consumo

energético conforme o avanço das rodadas de comunicação. A estratégia de transmissão total maximiza a contribuição de todos os dispositivos em cada rodada de comunicação, garantindo maior precisão no modelo global. No entanto, isso implica em um aumento significativo no consumo de energia devido ao elevado volume de comunicação.

A Figura 2 apresenta os resultados da estratégia de transmissão condicional, variando o valor do limiar ε . Observa-se que a escolha de ε tem um papel crucial na determinação do equilíbrio entre manter a qualidade do modelo e minimizar o dos recursos energéticos. Valores menores de ε podem melhorar a acurácia do modelo, mas com maior custo energético e maior uso dos recursos de comunicação. Por outro lado, valores mais altos de ε reduzem a frequência transmissões de modelo locais, o que pode comprometer o desempenho do modelo global. Para todos os valores de ε , a transmissão condicional reduziu a quantidade de energia conforme o número de rodadas aumenta. Observa-se que a energia total acumulada aumenta continuamente, mas o ritmo de crescimento é menor para valores mais altos de ε e em comparação com a estratégia de transmissão total.

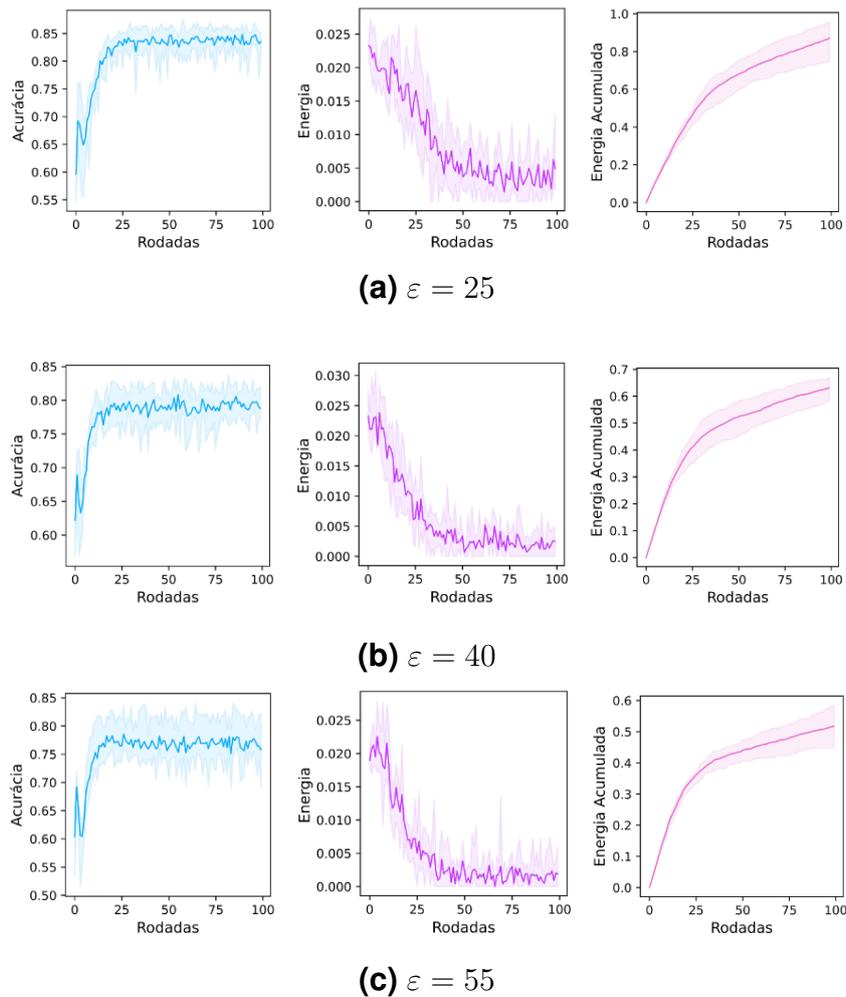


Figura 2. Estratégia de Transmissão Condicional

Os resultados das Figuras 1 e 2 ilustram o *trade-off* clássico em FL quando considerado o desempenho do modelo global e o consumo energético dos dispositivos. Ao

considerar o valor de $\varepsilon = 25$, a Figura 2a mostra que o desempenho do modelo global é de aproximadamente 5% menor na estratégia de transmissão condicional quando comparada com a estratégia de transmissão total. No entanto, o consumo energético da estratégia condicional é aproximadamente 55% menor quando comparada com a estratégia de transmissão total.

Por fim, a Figura 3 apresenta os resultados da estratégia de transmissão aleatória, que é utilizada como um linha de base para comparação com do desempenho e estabilidade da estratégia de transmissão condicional. A quantidade de transmissões foram baseadas nos números de transmissões observados nas estratégias condicionais com valores de $\varepsilon = 25$ de 25, 40 e 55. Em todos os casos, quando comparada com a estratégia condicional, a estratégia de transmissão aleatória apresentou uma acurácia inferior e com pouco decaimento da curva referente a energia consumida ao longo das rodadas de comunicação.

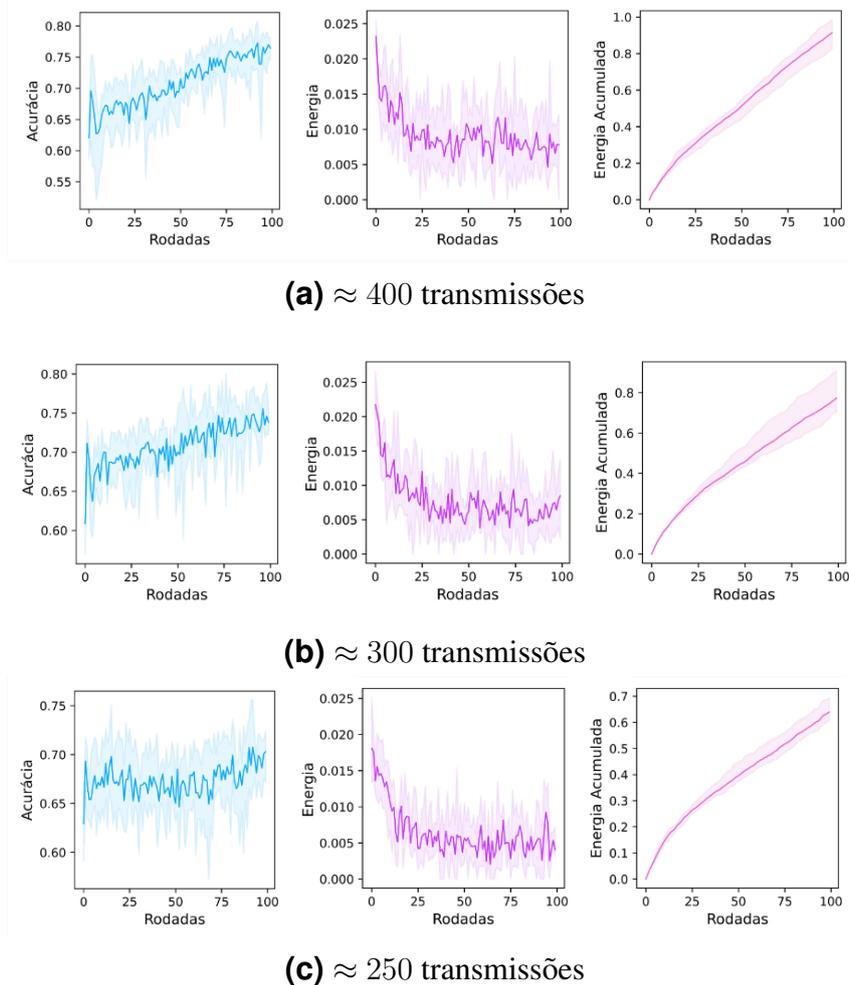


Figura 3. Estratégia de Transmissão Aleatória, onde é simulada a quantidade de transmissões observada na transmissão condicional com o valor de ε igual a 25, 40 e 55.

6. Considerações Finais

Este trabalho propõe uma estratégia inteligente de transmissão de parâmetros para o FL em redes IoT sem fio, baseada na diferença percentual relativa entre os pesos antigos e os

novos pesos atualizados dos modelos locais dos dispositivos. O estudo destacou a eficácia da estratégia de transmissão condicional como uma solução que equilibra o desempenho do modelo global e o consumo energético. A abordagem condicional, embora reduza ligeiramente a acurácia, oferece uma diminuição significativa no uso de energia, o que é crucial em dispositivos com restrições energéticas. Portanto, esta estratégia é especialmente relevante para redes sem fio e dispositivos IoT onde a eficiência energética deve ser considerada para contribuir para a sustentabilidade e eficácia operacional desses sistemas.

Referências

- Amannejad, Y. (2020). Building and Evaluating Federated Models for Edge Computing. *2020 16th International Conference on Network and Service Management (CNSM)*, pages 1–5.
- Asad, M., Moustafa, A., Ito, T., and Aslam, M. (2020). Evaluating the Communication Efficiency in Federated Learning Algorithms. *CoRR*, abs/2004.02738.
- Beutel, D. J., Topal, T., Mathur, A., Qiu, X., Parcollet, T., and Lane, N. D. (2020). Flower: A Friendly Federated Learning Research Framework. *CoRR*.
- Chen, H., Huang, S., Zhang, D., Xiao, M., Skoglund, M., and Poor, H. V. (2022). Federated Learning Over Wireless IoT Networks With Optimized Communication and Resources. *IEEE Internet of Things Journal*, 9(17):16592–16605.
- Chen, M., Yang, Z., Saad, W., Yin, C., Poor, H. V., and Cui, S. (2021). A Joint Learning and Communications Framework for Federated Learning Over Wireless Networks. *IEEE Transactions on Wireless Communications*, 20(1):269–283.
- ETSI (2022). Multi-Access Edge Computing (MEC) Framework and Reference Architecture. *ETSI Standard GS MEC 003*.
- Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X., and He, B. (2021). A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection. *IEEE Transactions on Knowledge and Data Engineering*, PP:1–1.
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. (2016). Communication-Efficient Learning of Deep Networks from Decentralized Data. *arXiv*.
- Nakayama, K. and Jenó, G. (2022). *Federated Learning with Python: Design and implement a federated learning system and develop applications using existing frameworks*. Packt Publishing.
- Wong, K. D. (2012). *Fundamentals of Wireless Communication Engineering Technologies*. Wiley Telecom. Chapters 1–25.
- Yang, Z., Chen, M., Wong, K.-K., Poor, H. V., and Cui, S. (2022). Federated Learning for 6G: Applications, Challenges, and Opportunities. *Engineering*, 8:33–41.
- Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., and Chandra, V. (2022). Federated Learning with Non-IID Data. *arXiv*.
- Zhu, G., Wang, Y., and Huang, K. (2020). Broadband Analog Aggregation for Low-Latency Federated Edge Learning. *IEEE Transactions on Wireless Communications*, 19(1):491–506.