Daniel Detore
Dr. Zewde
EE 441: Introduction to Wireless Systems
May 14 2025

Wireless Federated Learning: What It Is and How It's Changing AI

In recent years, artificial intelligence (AI) has become more and more popular, with AI systems claiming to improve many parts of the average person's life. The problem with making AI as widespread as its creators would like it to be is the sheer amount of data required, along with how much transmitting power it takes to send the data between client and server devices. One possible solution is wireless federated learning.

The most common method of training AI at the moment is centralized learning (CL). In this model, the steps are as such. First, all edge devices send all of their raw data to a central server. Then the server trains a model on all of the data. Finally, the server sends the entire model back to all of the edge devices [2]. With this method, however, we must transmit raw data, which can easily be eavesdropped. Even if we could ensure the data is safe, the transmission is slow and resource-intense. This problem is what federated learning (FL) aims to fix.

In FL, each edge device trains and shares its own model, and sends only the model (not the raw data) to a central server. Broadcasts of models between the edge devices and central server are much less expensive than that of data. Models are also obfuscated from their training data, so this method is much safer than CL. FL also has the benefit of distributed optimization, since more devices working concurrently does more work with less resources. Finally, the nature of FL allows devices which collect heterogeneous data to contribute to one global model [2].

One "round" of FL is executed as such: First, a subset of existing clients is selected, each of which downloads the current model. Then each selected client trains an updated model using their local data. The local models are sent from the selected clients to the server. Finally, the server aggregates these models (typically by averaging) to construct an updated global model [1].

FL was introduced in 2016 by Google research scientists McMahan and Ramage for "smarter models, lower latency, and less power consumption, all while ensuring privacy" [8]. Google used it internally for some time, but its practical use largely remained unstudied. Eventually, the idea of wireless federated learning (WFL) was presented. Much research was done in 2021 when the potential of 6G wireless technology made WFL "a promising decentralized solution for protecting the data privacy and meeting the low-latency demand" [3]. Studies on the topic have continued through today as wireless technology improves, along with the rising popularity of the issue of data privacy in AI training.

One of the fundamentals of implementing WFL is whether the signal should be digital or analog. According to [5], this choice depends on the layout of the network. Digital transmission is best for setups with sufficient radio resources and CSI uncertainties. However, analog transmission is best for setups with massive numbers of participating devices. Analog transmission also has the benefit of allowing the edge devices' models to be aggregated in the air. This setup saves computation power for the server, which makes WFL even less resource-intensive than CL.

To continue the comparison, WFL is independent of internet connection, which makes it useful for highly mobile devices like smartphones and cars. Since WFL is run on many devices simultaneously, the global model can adapt much more quickly and completely to major changes than a CL setup [4].

WFL is suited to networks with smartphones and cars, due to its ability to persist when devices lose connection to the server. Google's first implementation of this setup was in their phones' keyboards, which may be used to type sensitive personal information. It could also be ground-breaking for healthcare, since any raw data in healthcare is legally and ethically protected against sharing. For example, models could be trained to detect cancer without needing to send patient MRIs outside of the hospital. WFL is also built to handle heterogeneous data, which is especially useful for smart homes and power and water grids [3]. These are all diverse systems of many devices which collect many different types of data, and require many different types of output. Additionally, these devices usually have poor processing power and wireless connection, which most implementations of WFL can handle much better than CL.

Since WFL is a relatively new technology, there are many challenges for implementing it at scale. Ensuring the accurate broadcast of models is difficult, especially in bandwidth-constrained networks. Failure to accurately broadcast models could cause compounding inconsistencies between each node's models and negatively impact the global model [4]. The non-IID data of something like a smart home is difficult to process and combine into a useful global model. The devices involved might have processors too poor to actually create models or even collect data [7]. Perhaps most detrimental is that model broadcasts can still be peeked by eavesdroppers, and optimized setups are especially vulnerable. (Still, inherently much less information can be drawn out of the model than raw data.) Much research has gone into solving these problems as WFL has become more and more attractive to AI developers.

The future of WFL is uncertain, but some interesting trends and possible paths have arisen. A fully decentralized setup has been proposed, where model aggregation is done exclusively in the air between edge devices [4]. This will require robust privacy, which will likely include adversarial methods to prevent spoofing and encryption to prevent eavesdropping [4]. To implement WFL systems at scale, better methods must be found to combine more data from more heterogeneous sources into the global model. This requires collaborative work across wireless systems, neural network creation, and even creating the edge devices. These methods could be made possible by the implementation of 6G technology, which would allow better wireless communication, and thus more accurate model transmission.

WFL can be a massively useful tool across many fields like healthcare, smart homes, utilities, smartphones, and self-driving vehicles. Its privacy measures are inherently much more secure than centralized learning, but it is underdeveloped, so it is still unreliable and exploitable. Further research into WFL could completely change the state of AI and how we interact with it.

Bibliography

[1] Z. Yang, M. Chen, K.-K. Wong, P. H. Vincent, and S. Cui, “Federated Learning for 6G: Applications, Challenges, and Opportunities,” *arXiv.org*, 2021, Accessed: May 14, 2025. [Online]. Available: https://arxiv.org/abs/2101.01338

[2] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, “Federated Learning: Strategies for Improving Communication Efficiency,” *arXiv:1610.05492 [cs]*, Oct. 2017, Available: https://arxiv.org/abs/1610.05492

[3] Bouzinis, Pavlos S, Diamantoulakis, Panagiotis D, and G. K. Karagiannidis, “Wireless Federated Learning (WFL) for 6G Networks -- Part I: Research Challenges and Future Trends,” *arXiv.org*, 2021, Accessed: May 14, 2025. [Online]. Available: https://arxiv.org/abs/2105.00842

[4] M. B. Driss, E. Sabir, H. Elbiaze, and W. Saad, “Federated Learning for 6G: Paradigms, Taxonomy, Recent Advances and Insights,” *arXiv.org*, 2023, Accessed: Dec. 09, 2024. [Online]. Available: https://arxiv.org/abs/2312.04688

[5] J. Yao, W. Xu, Z. Yang, X. You, M. Bennis, and P. H. Vincent, “Wireless Federated Learning over Resource-Constrained Networks: Digital versus Analog Transmissions,” *arXiv.org*, 2024, Accessed: May 14, 2025. [Online]. Available: https://arxiv.org/abs/2405.17759

[6] S. Shaon, T. Nguyen, L. Mohjazi, A. Kaushik, and D. C. Nguyen, “Wireless Federated Learning over UAV-enabled Integrated Sensing and Communication,” *arXiv.org*, 2024, Accessed: May 14, 2025. [Online]. Available: https://arxiv.org/abs/2411.08918

[7] M. Asad, Safa Otoum, and Bassem Ouni, “Federated Learning in Vehicular Networks: A Review of Emerging Trends and Future Directions,” *2024 IEEE 10th World Forum on Internet of Things (WF-IoT)*, pp. 1–6, Nov. 2024, doi: https://doi.org/10.1109/wf-iot62078.2024.10811327.

[8] B. McMahan and D. Ramage, “Federated Learning: Collaborative Machine Learning without Centralized Training,” *research.google*, Apr. 06, 2017. https://research.google/blog/federated-learning-collaborative-machine-learning-without-centralized-training-data/