Confluence of Random Walks, Interacting Particle Systems, and Distributed Machine Learning: Federated Learning through Crawling over Networks

Seyyedali Hosseinalipour
Department of Electrical Engineering, University at Buffalo–SUNY
alipour@buffalo.edu

Invited Talk

FL emerged as a significant advancement in distributed machine learning (ML), initially popularized by Google for its application in ML training for smartphone keyboard prediction (Hard et al. 2018). The key innovation of FL is in shifting the ML training process to the location where the data originates, typically on edge devices. This approach aims to aggregate the ML models trained distributedly across the edge devices, where the aggregation schemes are generally categorized into *centralized* and *decentralized* (Martínez Beltrán et al. 2023). Pivotal aspects common to both of these architectures are (i) the principle of maintaining the data's locality and privacy, where the data in its raw format is never transferred across the network, and (ii) the existence of non-independent and identically distributed (non-iid) data across the nodes.

In FL with centralized aggregation, each participating node trains a local model on its own data and then sends the model updates, typically in the form of gradients or parameter updates, to a central server. The central server aggregates these updates, often by averaging, to update the global model. This updated global model is then sent back to the nodes for further training. This process iterates several times until a desired convergence criteria is met. In FL with decentralized aggregation, the traditional central server is eliminated or minimized, and the nodes collaborate in a peer-to-peer (P2P) – also known as device-to-device (D2D) – manner to train a global model. In this case, each node in the network trains a local model on its own data and then directly exchanges model information with neighboring nodes. These nodes aggregate the received information, often through techniques like distributed consensus (Liu, Chen, and Zhang 2022), to update their local models.

In this work, we aim to *unveil a new class of intermediate FL architectures between centralized and decentralized schemes called "FedCrawl."* FedCrawl takes advantage of benefits of D2D communications similar to decentralized schemes; however, it uses them in a nuanced way. FedCrawl is inspired by web crawlers, which effectively explore the websites to find updated/new content posted on the internet. The cornerstone of FedCrawl is its innovative conceptualization of neural networks (NNs) or other used ML models as autonomous entities with the capability to *move* or *jump* across nodes in the network through P2P connections.

We introduce five research aspects to study the nuanced intricacies governing random walker behavior in these environments. The first research aspect addresses the interplay between network topology and data distribution, emphasizing the importance of considering both factors for designing efficient random walks in FedCrawl. The second research aspect explores the applicability of node importance metrics in optimizing random walker paths for FedCrawl. We propose a dynamic perception-aware design, discussed in the third research aspect, where transition matrices adapt to the evolving state of random walkers, balancing exploration and exploitation. The fourth research aspect introduces innovative features like skipping, memory look-back, and caching/trailing to enhance random walker performance. Lastly, the fifth research aspect delves into the dynamics of multiple random walkers in networked environments, introducing the concept of multi-pole random walkers. Complementing these five research aspects, we present five conjectures, each introducing novel perspectives and methodologies in the domain of decentralized learning. These conjectures encompass areas such as temperature-based characterization of random walkers and network nodes, dynamic transition matrices, non-Markovian processes, and an evolutionary framework for random walker patterns.

References

Hard, A.; Rao, K.; Mathews, R.; Ramaswamy, S.; Beaufays, F.; Augenstein, S.; Eichner, H.; Kiddon, C.; and Ramage, D. 2018. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

Liu, W.; Chen, L.; and Zhang, W. 2022. Decentralized federated learning: Balancing communication and computing costs. *IEEE Transactions on Signal and Information Processing over Networks*, 8: 131–143.

Martínez Beltrán, E. T.; Pérez, M. Q.; Sánchez, P. M. S.; Bernal, S. L.; Bovet, G.; Pérez, M. G.; Pérez, G. M.; and Celdrán, A. H. 2023. Decentralized Federated Learning: Fundamentals, State of the Art, Frameworks, Trends, and Challenges. *IEEE Communications Surveys & Tutorials*, 25(4): 2983–3013.