A Practical Introduction to Federated Learning

Yaliang Li
Alibaba Group, Bellevue, WA USA
yaliang.li@alibaba-inc.com

Bolin Ding
Alibaba Group, Bellevue, WA USA
bolin.ding@alibaba-inc.com

Jingren Zhou
Alibaba Group, Bellevue, WA USA
jingren.zhou@alibaba-inc.com

ABSTRACT

As Internet users attach importance to their own privacy, and a number of laws and regulations go into effect in most countries, Internet products need to provide users with privacy protection. As one of the feasible solutions to provide such privacy protection, federated learning has rapidly gained popularity in both academia and industry in recent years. In this tutorial, we will start off with some real-world tasks to illustrate the topic of federated learning, and cover some basic concepts and important scenarios including cross-device and cross-silo settings. Along with it, we will give several demonstrations with popular federated learning frameworks. We will also show how to do the automatic hyperparameter tuning with federated learning to significantly save their efforts in practice. Then we dive into three parallel hot topics, Personalized Federated Learning, Federated Graph Learning, and Attack in Federated Learning. For each of them, we will motivate it with real-world applications, illustrate the state-of-the-art methods, and discuss their pros and cons using concrete examples. As the last part, we will point out some future research directions.

ACM Reference Format:
Yaliang Li, Bolin Ding, and Jingren Zhou. 2022. A Practical Introduction to Federated Learning. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3534678.3542631

1 BACKGROUND OF FEDERATED LEARNING

We first introduce fundamental concepts for federated learning [16, 25] (FL) and the necessity of FL in real-world tasks for exploiting isolated data without privacy leakage. Several important privacy-preserving technologies, including Homomorphic Encryption [17], Secure Multi-Party Computation [26] and Differential Privacy [4], will be briefly introduced in our tutorial, from the aspects of how these technologies protect the privacy of data (e.g., splitting the message into frames or adding noise into the exchanged message), and how to exploit these technologies in FL. Then we will show several standard and practical federated learning tasks as examples, e.g., aggregating records from different IoT devices for global prediction (i.e., horizontal FL [16]), and sharing the different features of overlapped app users among companies (i.e., vertical FL [7]). After that, based on the existing FL frameworks (such as TFF [2], FATE [25], and FederatedScope [24]), we will present two different ways to implement an FL procedure, including a sequential way and an event-driven way.

2 FEDERATED HYPERPARAMETER OPTIMIZATION

When Hyperparameter optimization (HPO) comes to FL, each attempt means several rounds of communication across participants, which can be very costly, especially for cross-device scenarios. Thus, it is necessary to let the community be aware of the uniqueness of HPO under the FL setting and promote the skill of federated HPO. At first, we will define the problem of federated HPO formally. Then we introduce the concept of low-fidelity HPO and highlight two ways to moderate fidelity in the FL setting. With this prerequisite, we demonstrate how to implement popular HPO algorithms such as Hyperband [11] to cooperate with an FL runner. The importance of this example is to show how the HPO component interacts with existing FL frameworks, including how to trigger an FL instance, specify the configuration of each attempt, and more importantly, how the quality of optimization changes along with the fidelity. Furthermore, we promote the weight-sharing view originated from neural architecture search [14] and applied to treating federated HPO very recently. We conclude this part by reviewing some recent works [9, 10, 28, 31], and introducing a new benchmark [21] for federated HPO.

3 PERSONALIZED FEDERATED LEARNING

Most of existing studies on Personalized FL control the extent a client learns from others and how to fuse the shared knowledge with local models. We will categorize the Personalized FL methods in the literature according to what is proposed to be different among clients, such as training configurations, submodules, training behaviors and the local models, and summarize their pros and cons from the aspects of effectiveness and efficiency. Then we will demonstrate how to plug a personalization module into a standard FL course, using several existing methods as examples, including pFedME [19], FedBN [13], FedEM [15], Ditto [12], etc. Further, we will also show how to monitor client-wise and global metrics to check the advantages of applying Personalized FL methods. Last but not least, we introduce a benchmark for Personalized FL [3], and discuss an extended task that further considers the heterogeneity of client-wise tasks [27].

4 FEDERATED GRAPH LEARNING

Although FL has been applied to various types of data, the importance and uniqueness of federated graph learning (FGL) make it deserve a dedicated part of this tutorial. At first, we enumerate several real-world FGL applications, including recommender systems [22], healthcare [29], anti-money laundering [18], etc., to attract our audiences. These applications show that the demand

KDD '22, August 14–18, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9385-0/22/08.
https://doi.org/10.1145/3534678.3542631

for federally handling graph data is very prevalent and intense. Then, we will show how to transform a vanilla graph dataset into its federal counterpart by splitting the original graph into subgraphs. We will compare two popular splitting strategies, random split and community-based split [1]. Next, we demonstrate how to implement a graph neural network (GNN) based on PyG [5], and how to integrate the developed GNN model into an FL framework and conduct FGL. Finally, we compare the performance of learned GNN models with the whole graph, client-wise subgraphs, and FGL. This example, at the same time, implies the potential advantages of completing each client-wise subgraph. Hence, we further present a recent work on this topic—FedSage+ [29], describing both its formulas and how to implement it with FederatedScope [24]. As FGL algorithms such as FedSage+ and GCFL+ [23] often call for exchanges of heterogeneous data across the FL participants and exhibit more complicated behaviors, their implementations give the audiences more insights into the event-driven design of FederatedScope. Finally, we will introduce a recent FL package [20] that is devoted to FGL.

5 ATTACK IN FEDERATED LEARNING

Applying privacy attacks directly can intuitively demonstrate the privacy-preserving strength of FL, which makes it an important part of this tutorial. Starting with the example of attacking FedAvg [16], one of the most popular FL algorithms, we will introduce the background of privacy attacks in FL: (1) the attack settings including the passive attack and the active attack, (2) the attack types, including membership inference attack, property inference attack, class representative attack, and training data and label inference attack, (3) the state-of-the-art privacy attack methods, including DMU-GAN [8]; DLG [32], iDLG [30], GRADINV [6]. Next, with the implemented attack methods in FederatedScope, we will demonstrate the privacy leakage in sharing the model parameter updates directly in FedAvg. Furthermore, by showing the confrontation between the privacy attacks and the defense strategies, we will illustrate how privacy attacks can guide the choice of defense strategies. Finally, we will also illustrate how to conveniently develop the user-customized attacker with FederatedScope.

6 CONCLUSIONS

We will conclude this tutorial, and point out a list of open problems and future research directions based on our experience in the industry and the trends in the academy.

REFERENCES

[1] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* (2008), P10008.

[2] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konečnỳ, Stefano Mazzocchi, Brendan McMahan, et al. 2019. Towards federated learning at scale: System design. *Proceedings of Machine Learning and Systems* 1 (2019), 374–388.

[3] Daoyuan Chen, Dawei Gao, Weirui Kuang, Yaliang Li, and Bolin Ding. 2022. pFL-Bench: A Comprehensive Benchmark for Personalized Federated Learning. *arXiv preprint arxiv.2206.03655* (2022).

[4] Cynthia Dwork, Aaron Roth, et al. 2014. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* 9, 3-4 (2014), 211–407.

[5] Matthias Fey and Jan E. Lenssen. 2019. Fast Graph Representation Learning with PyTorch Geometric. In *ICLR Workshop*.

[6] Jonas Geiping, Hartmut Bauermeister, Hannah Dröge, and Michael Moeller. 2020. Inverting gradients-how easy is it to break privacy in federated learning? *NeurIPS* 33 (2020), 16937–16947.

[7] Stephen Hardy, Wilko Henecka, Hamish Ivey-Law, Richard Nock, Giorgio Patrini, Guillaume Smith, and Brian Thorne. 2017. Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. *arXiv preprint arXiv:1711.10677* (2017).

[8] Briland Hitaj, Giuseppe Ateniese, and Fernando Perez-Cruz. 2017. Deep models under the GAN: information leakage from collaborative deep learning. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. 603–618.

[9] Mikhail Khodak, Renbo Tu, Tian Li, Liam Li, Nina Balcan, Virginia Smith, and Ameet Talwalkar. 2021. Federated Hyperparameter Tuning: Challenges, Baselines, and Connections to Weight-Sharing. *NeurIPS* 34 (2021).

[10] Antti Koskela and Antti Honkela. 2020. Learning Rate Adaptation for Differentially Private Learning. In *AISTATS*. 2465–2475.

[11] Liam Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. 2018. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research* 18-185 (2018), 1–52.

[12] Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. 2021. Ditto: Fair and robust federated learning through personalization. In *ICML*. 6357–6368.

[13] Xiaoxiao Li, Meirui Jiang, Xiaofei Zhang, Michael Kamp, and Qi Dou. 2021. FedBN: Federated learning on non-iid features via local batch normalization. *arXiv preprint arXiv:2102.07623* (2021).

[14] Hanxiao Liu, Karen Simonyan, and Yiming Yang. 2019. DARTS: Differentiable Architecture Search. In *ICLR*.

[15] Othmane Marfoq, Giovanni Neglia, Aurélien Bellet, Laetitia Kameni, and Richard Vidal. 2021. Federated multi-task learning under a mixture of distributions. *NeurIPS* 34 (2021).

[16] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. 1273–1282.

[17] Pascal Paillier. 1999. Public-key cryptosystems based on composite degree residuosity classes. In *International conference on the theory and applications of cryptographic techniques*. 223–238.

[18] Toyotaro Suzumura, Yi Zhou, Natahalie Baracaldo, Guangnan Ye, Keith Houck, Ryo Kawahara, Ali Anwar, Lucia Larise Stavarache, Yuji Watanabe, Pablo Loyola, et al. 2019. Towards federated graph learning for collaborative financial crimes detection. *arXiv preprint arXiv:1909.12946* (2019).

[19] Canh T Dinh, Nguyen Tran, and Josh Nguyen. 2020. Personalized federated learning with moreau envelopes. *NeurIPS* 33 (2020), 21394–21405.

[20] Zhen Wang, Weirui Kuang, Yuexiang Xie, Liuyi Yao, Yaliang Li, Bolin Ding, and Jingren Zhou. 2022. FederatedScope-GNN: Towards a Unified, Comprehensive and Efficient Package for Federated Graph Learning. In *Proc. of the SIGKDD International Conference on Knowledge Discovery and Data Mining*.

[21] Zhen Wang, Weirui Kuang, Ce Zhang, Bolin Ding, and Yaliang Li. 2022. FedHPO-B: A Benchmark Suite for Federated Hyperparameter Optimization. *arXiv preprint arxiv.2206.03966* (2022).

[22] Chuhan Wu, Fangzhao Wu, Yang Cao, Yongfeng Huang, and Xing Xie. 2021. Fedgnn: Federated graph neural network for privacy-preserving recommendation. *arXiv preprint arXiv:2102.04925* (2021).

[23] Han Xie, Jing Ma, Li Xiong, and Carl Yang. 2021. Federated graph classification over non-iid graphs. *NeurIPS* 34 (2021).

[24] Yuexiang Xie, Zhen Wang, Daoyuan Chen, Dawei Gao, Liuyi Yao, Weirui Kuang, Yaliang Li, Bolin Ding, and Jingren Zhou. 2022. FederatedScope: A Flexible Federated Learning Platform for Heterogeneity. *arXiv preprint arxiv.2204.05011* (2022).

[25] Qiang Yang, Yang Liu, Yong Cheng, Yan Kang, Tianjian Chen, and Han Yu. 2019. Federated learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 13, 3 (2019), 1–207.

[26] Andrew C Yao. 1982. Protocols for secure computations. In *annual symposium on foundations of computer science*. 160–164.

[27] Liuyi Yao, Dawei Gao, Zhen Wang, Yuexiang Xie, Weirui Kuang, Daoyuan Chen, Haohui Wang, Chenhe Dong, Bolin Ding, and Yaliang Li. 2022. A Benchmark for Federated Hetero-Task Learning. *arXiv preprint arxiv.2206.03436* (2022).

[28] Huanle Zhang, Mi Zhang, Xin Liu, Prasant Mohapatra, and Michael DeLucia. 2021. Automatic Tuning of Federated Learning Hyper-Parameters from System Perspective. *arXiv preprint arXiv:2110.03061* (2021).

[29] Ke Zhang, Carl Yang, Xiaoxiao Li, Lichao Sun, and Siu Ming Yiu. 2021. Subgraph federated learning with missing neighbor generation. *NeurIPS* 34 (2021).

[30] Bo Zhao, Konda Reddy Mopuri, and Hakan Bilen. 2020. iDLG: Improved deep leakage from gradients. *arXiv preprint arXiv:2001.02610* (2020).

[31] Yi Zhou, Parikshit Ram, Theodoros Salonidis, Nathalie Baracaldo, Horst Samulowitz, and Heiko Ludwig. 2021. FLoRA: Single-shot Hyper-parameter Optimization for Federated Learning. *arXiv preprint arXiv:2112.08524* (2021).

[32] Ligeng Zhu, Zhijian Liu, and Song Han. 2019. Deep leakage from gradients. *NeurIPS* 32 (2019).