Towards Fair Federated Learning

Zirui Zhou
zirui.zhou@huawei.com
Huawei Technologies Canada
Burnaby, BC, Canada

Lingyang Chu
chul9@mcmaster.ca
MacMaster University
Hamilton, Ontario, Canada

Changxin Liu
chxliu@uvic.ca
University of Victoria
Victoria, BC, Canada

Lanjun Wang
lanjun.wang@huawei.com
Huawei Technologies Canada
Burnaby, BC, Canada

Jian Pei
jpei@cs.sfu.ca
Simon Fraser University
Burnaby, BC, Canada

Yong Zhang
yong.zhang3@huawei.com
Huawei Technologies Canada
Burnaby, BC, Canada

ABSTRACT

Federated learning has become increasingly popular as it facilitates collaborative training of machine learning models among multiple clients while preserving their data privacy. In practice, one major challenge for federated learning is to achieve fairness in collaboration among the participating clients, because different clients' contributions to a model are usually far from equal due to various reasons. Besides, as machine learning models are deployed in more and more important applications, how to achieve model fairness, that is, to ensure that a trained model has no discrimination against sensitive attributes, has become another critical desiderata for federated learning. In this tutorial, we discuss formulations and methods such that collaborative fairness, model fairness, and privacy can be fully respected in federated learning. We review the existing efforts and the latest progress, and discuss a series of potential directions.

CCS CONCEPTS

• **General and reference → Surveys and overviews**; • **Computing methodologies → Artificial intelligence**; • **Social and professional topics → Privacy policies; Soft intellectual property**.

KEYWORDS

federated learning, collaborative fairness, model fairness, data privacy, distributed learning, data leakage

ACM Reference Format:
Zirui Zhou, Lingyang Chu, Changxin Liu, Lanjun Wang, Jian Pei, and Yong Zhang. 2021. Towards Fair Federated Learning. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021, Virtual Event, Singapore*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3447548.3470814

Federated learning has been widely recognized as a fundamental technology towards next-generation artificial intelligence. It provides a distributed learning paradigm that facilitates collaborations among multiple data owners while protecting their data privacy. Lately, there has been a surge of research on federated learning. Federated learning has been successfully deployed in various application domains, such as internet-of-things, healthcare, and manufacturing. See [2, 3, 5] for some recent survey on federated learning.

One important aspect that every federated learning system has to accommodate is fairness in collaboration, since the contributions from different participants to the learning process can be substantially different [4]. In practice, many factors may affect the contribution from a participant, such as data volume, data quality, and rounds of communication with the coordinator. For example, participants with larger data volume and wider spectrum of data types, such as banks, governments, and tech giants, are deemed to have a much higher impact on the learnt model than participants with limited data resources. Moreover, some participants may act as free-riders and some may even be attackers that contribute negatively to the learning process. Therefore, developing a reward system that can fairly reflect the contribution of each participant is highly desirable and is a key to the prosperity of the ecosystem of federated learning. Currently, most federated learning systems are short of collaborative fairness, as they typically allow every participant to have access to the same global model that is trained with data from all participants. To solve this problem, one has to propose appropriate metrics that can measure the contribution of each participant and design novel incentive schemes so that participants with more contributions receive higher rewards. In addition, all the measurements and incentives need to be deployed in a way that data privacy of every participant is fully respected.

Another increasingly important aspect in federated learning is to ensure that the learnt model is fair, i.e., the model has no discrimination against certain population groups [1]. Indeed, as accurate machine learning models are deployed in more and more applications, enhancing and ensuring fairness in machine learning models becomes critical for AI for social good. Federated learning, as the foundation of the next-generation AI, has to be equipped with techniques to tackle this grand and interdisciplinary challenge. For general machine learning tasks, there are many recently proposed methods that ensure fairness of a learnt model. However, it is far from trivial to extend these methods to federated learning systems. In particular, these methods are mainly designed with an assumption of a unified available training data set, and thus cannot easily address the needs of federated learning, such as distributed collaboration, communication costs, and privacy preservation for

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08.
https://doi.org/10.1145/3447548.3470814

participating parties. As such, novel ideas and techniques have to be introduced so that federated learning can generate fair models.

In this tutorial, we aim to present a comprehensive survey on existing approaches that can lead to collaborative fairness as well as model fairness in federated learning systems. The tutorial consists of the following major modules.

A Quick Review of Federated Learning and Fairness. We start with motivating examples of federated learning and introduce the definition and key features of federated learning. Then, we elaborate the differences of horizontal federated learning and vertical federated learning and discuss some real world applications. Finally, we briefly introduce the concepts of collaborative fairness and model fairness in federated learning.

Collaborative Fairness in Federated Learning. We start with some motivating examples of uneven contributions in both horizontal and vertical federated learning and emphasize the importance of building a collaboratively fair federated learning system. Then, we introduce some metrics that quantify the contributions of participants and measure collaborative fairness of a system. We also discuss some challenges of achieving collaborative fairness in a federated learning system. Finally, we review the existing approaches towards collaboratively fair federated learning systems and discuss their strengths and limitations.

Model Fairness in Federated Learning. We start with reviewing a number of fairness measurements and existing methods for training a fair model in general machine learning tasks. Then, we discuss the challenges of extending these fairness measurements and methods to the setting of federated learning and introduce some intrinsic difficulties of training a fair model in both horizontal and vertical federated learning. Finally, we discuss some possible directions that may lead to model fairness in federated learning systems.

Open Challenges. Both collaborative fairness and model fairness in federated learning are just at their infant stage. Also, the business settings of federated learning are evolving rapidly and constantly bringing new challenges to this ever-expanding field. We discuss a series of open challenges at the end of this tutorial.

This tutorial is self-contained. We do not assume any background knowledge and make the essential ideas highly accessible to practitioners and new graduate students. The full version of this tutorial covers over 100 references, mostly published recently. The references are mainly from the field of federated learning, data evaluation, fairness, and game theory.

SPEAKER BIBLIOGRAPHIES

Zirui Zhou is a Senior Principal Researcher at Huawei Technologies Canada. Before joining Huawei, he was an assistant professor in the Department of Mathematics at Hong Kong Baptist University. He received his Ph.D. in Systems Engineering and Engineering Management from Chinese University of Hong Kong. His research interest includes numerical optimization and its applications in machine learning. His research works on convex analysis, optimization theory, and provable non-convex methods have been published in top-tier journals and conferences.

Lingyang Chu is an Assistant Professor at the Department of Computing and Software of McMaster University. Before joining McMaster University, he is a principal researcher at Huawei Technologies Canada. His research interest is in the area of data mining and machine learning. He is especially interested in designing novel algorithms to interpret deep neural networks and to protect data privacy and model fairness in modern machine learning frameworks. His research works in large scale graph mining and interpretable AI have been published in top-tier venues.

Changxin Liu is a Ph.D. candidate at the University of Victoria, Canada. In March 2021, he joined Huawei Technologies Canada (Vancouver Research Center) as a research intern. His research interest focuses on distributed optimization and control for networked systems. He is an active reviewer for more than 10 journals and conferences and an Outstanding Reviewer for IEEE Trans. Cybernetics in 2018.

Lanjun Wang now is a Senior Principal Researcher and technical leader of Vancouver Research Centre, Huawei Canada. She received her Ph.D. degree in Electronic Engineering from Tsinghua University in 2011, and in the same year, she joined IBM Research-China as a Staff Researcher. Before that, she was a member of Academic Talent Program 2001 at Tsinghua University and received B.S degree in Math and Physics in 2005. Lanjun's research interests include data mining, analytics, and modeling. She has issued about 20 patents and led more than ten top-quality publications.

Jian Pei is currently a Professor in the School of Computing Science and an associate member of the Department of Statistics and Actuarial Science at Simon Fraser University, Canada. His general research areas include data science, big data, data mining, and database systems. His expertise is on developing effective and efficient data analysis techniques for novel data intensive applications. He is recognized as a Fellow of the Royal Society of Canada (RSC), the Canadian Academy of Engineering, ACM and IEEE.

Yong Zhang currently is a Distinguished Researcher at Huawei Technologies Canada and leading the big data and intelligence platform lab at Vancouver research center. Prior to that, he was a postdoctoral research fellow at Stanford University, USA. His research interests include large scale numerical optimization and machine learning. His research works have been published in top-tier journals and conferences.

REFERENCES

[1] Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2017. Fairness in machine learning. *NIPS Tutorial* (2017).
[2] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. 2019. Advances and open problems in federated learning. *arXiv preprint arXiv:1912.04977* (2019).
[3] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. 2020. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine* 37, 3 (2020), 50–60.
[4] Lingjuan Lyu, Xinyi Xu, Qian Wang, and Han Yu. 2020. Collaborative fairness in federated learning. In *Federated Learning*. Springer, 189–204.
[5] 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.