2020 IEEE 16th International Conference on Control & Automation (ICCA)
Oct 9-11, 2020 (Virtual) Sapporo, Hokkaido, Japan

A Survey on federated learning *

Li Li, *Member, IEEE*, Yuxi Fan, and Kuo-Yi Lin*, *Member, IEEE*

Abstract— **Federated learning (FL) is an emerging setting which implement machine learning in a distributed environment while protecting privacy. Research activities relating to FLhave grown at a fast rate recently in control. Exactly what activities have been carrying the research momentum forward is a question of interest to the research community. This study finds these research activities and optimization path of FL based on survey. Thus, this study aims to review related studies of FL to base on the baseline a universal definition gives a guiding for the future work. Besides, this study presents the prevailing FL applications and the evolution of federated learning. In the end, this study also identifies four research fronts to enrich the FL literature and help advance our understanding of the field. A comprehensive taxonomy of FL can also be developed through analyzing the results of this review.**

Index Terms—Federated learning, Literature survey, Citation analysis, Research front

I. INTRODUCTION

FL is a burgeoning machine learning scheme, aiming at tackling the problem of data island while preserving data privacy. It was put forward by Google in 2016 to predict user's text input within tens of thousands mobile devices while keeping data on devices [1]. The original process of FL is generally described as follows. Firstly, each device downloads a generic global model for local training. Secondly, the local model will be improved with local data belonging to different mobile devices and upload to cloud in an encryption mode. Thirdly, the averaged update of local models implemented in the cloud will be dispatched to device as a renewed global model. Finally, the above procedures repeat until the model achieves a certain desired performance or the final deadline arrives.

The emergence of this technology will solve the contradiction between data privacy and data sharing. Due to the property that data is not exposed to third central server, FL is appropriate for application when data are privacy-sensitive. These includes cases in industry applications [2][3] or mobile devices [4] that data are not available to be aggregated with legal concern.

* Research supported by National Key R&D Program of China, No. 2018YFE0105000, the National Natural Science Foundation of China under Grant No. 51475334, the Shanghai Municipal Commission of science and technology No. 19511132100 and the Fundamental Research Funds for the Central Universities of China under Grant No. 22120170077.

Li Li is with the Tongji University, Shanghai, 201804 China (corresponding author to provide phone: +86 18916087269; fax: 021-69589241; e-mail: lili@tongji.edu.cn).

Yuxi Fan is with the Tongji University, Shanghai, 201804 China (e-mail: 1830749@tongji.edu.cn).

Kuo-Yi Lin is with the Tongji University, Shanghai, 201804 China (e-mail: 19603@tongji.edu.cn).

Recently, a few scholars band together to publish a book to review advances and open problems in FL for hasty generalization of FL contribution [5]. Motivated by the promising prospects and increasing growth of FL research, this study aims to review related studies of FL to base on the baseline a definition in control.

This paper is organized as follows. Beside the introduction, we sketch the characteristics and basic overview of FL. In section III, we point out three challenges in FL along with relative improvement aiming at these shortages. Furthermore, we conclude indirect information leakage in FL and existing privacy-preserving method employed in FL. Section IV discusses realistic applications in both mobile devices and cross organizational architecture. Last but not the least, some frontier achievements are given, around these discussions we describe some promising direction of FL.

II. OVERVIEW OF FEDERATE LEARNING

A. Characteristics of FL

FL is highly related to distributed learning. FL of model update for Android clients presented in [1] is to some extent similar to distributed computation. To reveal difference between FL and distributed learning, we highlight following characteristics in FL.

Universality for cross-organizational scenarios. Essentially, FL proposed by Google is an encrypted distributed machine learning technology. Then the original concept of FL was extended to refer to all privacy-preserving decentralized collaborative machine learning techniques [6]. Therefore, it could be extended to bring cross-organizational enterprise into federal framework, thus intelligently construct joint model for multiple entities, multiple data sources, different feature dimensions.

Massively Non-IID distribution. In FL, available data in each node may be no more than the total number of nodes. While in distributed system, the main purpose is to increase degree of parallelism so as to alleviate computation or storage pressure in central server. In contrast with distributed system, which works primarily on IID data distribution, FL is concentrated on unbalanced and non-IID data because of the heterogeneity among device resources.

Decentralized technology. Decentralization, in a strictly technical sense, is only to dilute the awareness of the central node. There is no center to determine each client, and each client goes to influence central model. Parameter server, a typical distributed and centralized technology, mainly make use of central server which is dominating to dispatch data distribution and computation resource to obtain an efficient

collaborative model [7]. For cases in FL, each client is completely autonomous, data is not allocated by center and the training process is not governed by server.

To sum up, FL is a decentralized technology that enable scattered clients to train a collaborative model autonomously, while keeping data localized.

B. Open Source Framework

There have been two mainstream open-source frameworks for FL up to now and they are starting to take shape. One is TensorFlow Federated (TFF) framework at the service of machine learning or other computation demand for decentralized data [8]. It is the first self-contained framework designed at production level mainly for mobile devices. Specially, TFF integrates FedAvg for model update and Secure Aggregation for privacy concern [9]. Furthermore, it has been successfully applied in next word prediction or Emoji prediction in a mobile keyboard [10]. In practical application, it has realized implementation of tens millions of devices, and hopes to be highly scalable to deal with computation over billions of devices.

The other one is Federated AI Technology Enabler (FATE) created by Webank team [11]. As the first open source industrial-level framework, it primarily serves for cross-organizational architecture. It provides enough privacy based on homomorphic encryption and secure multiparty computing. At present, the Webank team has promoted the implementation of a series of FATE in credit risk control, object detection and anti-money laundering [12]. These frameworks are still in an early form due to limited bandwidth and incapable computation power on mobile devices.

C. Categorization of FL

Based on paper presented by Yang et al. [6], FL largely falls into three groups, respectively, horizontal FL, vertical FL and federated transfer learning.

Horizontal federated learning. In the case of horizontal federated learning, there is a certain amount of overlap between the feature of data spread across various nodes, while the data are quite different in sample space. As is mentioned above, the federated model solution for Android mobile phone update raised by Google [1]is typically a kind of horizontal FL since the data has the same feature dimension. In addition, to meet the challenge of limited labeled entities, Gao et al. [13] introduced hierarchical heterogeneous horizontal FL frame. When it comes to cross-regional cooperation, it is almost impossible for each client to build a data pool for sharing. Thus, federate learning could construct a federal network for cross-regional architecture with similar information to improve joint model.

Vertical federated learning. It is suitable for cases in which data is partitioned in the vertical direction according to feature dimension. Generally, scholars deal with this problem through taking out the same entities with various characteristics to get a joint training [14]. Aggregating all the data set in a common server to learn from the global model doesn't work on vertical FL since the correspondence between different owners is still an urgently need to be addressed. There comes a modified token-based entity resolution algorithm to preprocess vertical partitioned data, powered by Nock et al. [15]. Hardy et al. designed an end-to-end scheme on linear classifier for vertical federated learning [16]. However, the above-mentioned methods could only be applied in simple machine learning models such as logistic regression. Therefore, vertical FL still has much more room for improvement to be applied in more complicated machine learning approaches.

Federated transfer learning. In most cases, data shares neither sample space nor feature space. Thus, the main problem in this setting is lack of data labels with poor data quality. Transfer learning [17] enables to move the knowledge of one domain (i.e., the source domain) to another domain (i.e., the target domain) to achieve better learning results, which is appropriate for this situation. In this way, Liu et al. conceived federated transfer learning (FTL) to generalize FL to have broader application when it comes to common parties with small intersection [18]. For a real application, Chen et al. constructed a FedHealth model that gather data owned by different organizations and offer personalized service for healthcare through federated transfer learning [19]. However, FTL is an effective way to protect both data security and user privacy while breaking the barriers of data islands.

III. EVOLUTION OF FL

Recent works focus on algorithm optimization to improve efficiency and accuracy and participants' privacy to enhance data protection.

A. Optimization

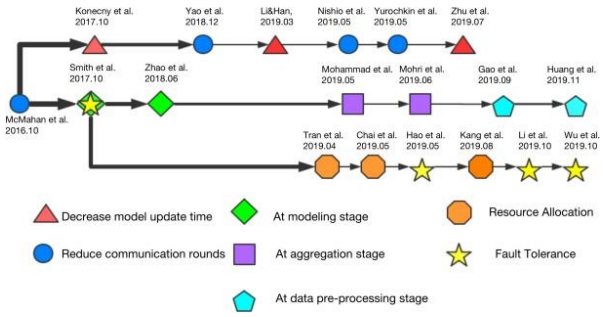

Figure 1. Optimization path to overcome three challenges in FL.

In terms of optimization, high communication cost, statistical and structural heterogeneity are major issues faced currently. In this section, we summarize the optimization path of FL to overcome these challenges as Fig 1 shows. The thickness of the line shows reference frequency of these papers in Google Scholar by other papers. The thicker the line, the higher reference frequency of the paper. The details of this optimization path are as follows.

High communication cost. By far, the key bottleneck of FL has been the difficulty of decreasing communication overhead. Meanwhile, effective efforts have been made in work including reducing communication rounds and improving model upload speed further reduce update time. The research of McMahan et al. Is the pioneering work to be more efficient by increasing calculated quantity on each client

between each communication round [1]. Nishio & Yonetani built FedCs framework to integrate the available clients to the utmost extent in each training round [20]. Maximum mean discrepancy was introduced to enforce local model to acquire more knowledge from other devices thus speed up convergence [21]. The Bayesian Nonparametric FL framework can aggregate local models without extra parameters thus avoid unwanted communication rounds [22]. Even if the communication rounds are optimized, unmatched download and upload speed is a remained problem. McMahan et.al. proposed two strategies to reduce model update time [23]. One is structured update, which means transmit only part of the update model. Likewise, a kind of structured update mode was proposed that maps update information into a lower-dimension space to relieve communication pressure [24]. The other is sketched update to make use of compressed update model. Zhu & Jin optimized sparse evolutionary training thus convey only piece of parameters to server [25]. These algorithms are not fully suitable for all federal setting; therefore, a more flexible communication-efficient method need to be explored.

Statistical heterogeneity. In reality, data are collected from various equipment thereby do not follow identically independent distribution. To tackle this problem, general researches focus on data pre-processing stage, modeling stage and aggregation stage. The first proposed FedAvg algorithm resolve this issue by averaging local upgrade on each device directly. In addition, Mohri et al. improved global model to cope with any target distribution comprised by a mixture of different clients [26]. In terms of aggregation stage, the existence of heterogeneity may lead to mis-convergence of global model. Further Wang et al. discussed convergence bound of FL in Non-IID data background [27]. For data pre-processing, Huang & Liu introduced clustering thought and constructed a community-based FL method [28]. In [13], it projects each embedding submainfold into a common embedding space to overcome data heterogeneity. Another idea is to optimize modeling way to achieve personalization for individual devices such as Mocha, which introduced multi-task learning to make utilization of shared representation [29]. Zhao et al. considered a solution to deal with non-iid data by sharing a small set of data among each local model [30].

Structural heterogeneity. It mainly refers to two aspects. On the one hand, the competence of computing and storage vary from nodes to nodes since different devices use various kinds of chip, and thereby cause unbalanced training time. On the other hand, unreliable and unstable network may lead to devices' drop out. In [29], Smith et al. considered influence with low participation in training process to resist device drop out. Scholars also designed secure aggregation protocol [9] [31] which is tolerant with arbitrary dropouts as long as surviving users are enough to join federate update. Li et al. took stragglers into account and allow these devices to implement variable locally update computation times [32]. Wu et al. made use of a cache structure to store those unreliable user update thus alleviates their impact on global model [33]. For the sake of resource constraint, Kang et al. took overhead in heterogeneous client into consideration to motivate more high-quality devices [34]. And Tran et al. studied training accuracy and convergence time with heterogeneous power constraints [35]. Meanwhile, Chai et al. considered the impact of resource heterogeneity on the training time of FL [36]. In the future, optimization should continue to contribute to fault-tolerance and properly resource allocation to address this issue.

B. Security Analysis

Though data never come out of the local storage which may alleviate privacy concerns. Nevertheless, the system is not sufficiently secure because the transmission of gradients and partial parameters may lead to indirect privacy leakage [37]. Some investigators have considered to retrieve data in FL framework. The general attack types are mainly divided into two categories as bellow:

Model poisoning. Model poisoning refer to make model to generate a wrong result by designing a specific input. In FL, aggregator is not familiar with the local update modes thus are not able to detect anomalies. According to this drawback, the backdoor can be inserted into federated environment through model-replacement methodology thus misunderstand the joint model [38]. Similarly, Bhagoji et al. attacked global model through few malicious adversaries so as to wrongly classified targeted model [39]. Then Zhang et al. gave first attempt to generate model poisoning attack based on Generative Adversarial Nets [40]. In the future work, to mitigate this type of attack for FL, anomaly detection in server side and concealment of classification results is a promising direction.

Inferring attack. The value of this type of attack mainly used to detect privacy records or restore training data through a white box or a black box. It can be broken down into tracing attacks and reconstruction attacks. Nasr et al. designed a white-box membership inference attack method to infer information via a curious server or any of a participant [41]. Wang et al. built a general attack frame called mGAN-AI which could reconstruct private information for target client [42]. To hinder this kind of attack, more stronger protection method should be explored, and data could be encrypted before upload to cloud.

C. privacy-preserving technology in FL

Faced with these vulnerabilities mentioned above, the existing privacy-preserving methods to enhance privacy guarantees mainly focus on information encryption for client or secure aggregation at server side as well as security protection for FL framework [43]. So, we will discuss novel privacy-preserving technologies based on this classification as bellows.

Privacy-preserving at client side. Differential Privacy (DP) acts as a means of privacy preservation for client. It will reduce chances for records to be identified while maximize query accuracy by introducing noise to blur raw data. Geyer et al. Leveraged DP on FL to conceal whether a client participant in the training process [44]. In addition, homomorphic encryption (HE) is also applied in FL frequently to hinder information leakage. In essence, it refers to an encryption mechanism that

parameters are encoded before adding or multiplying operation and performs equivalent result compare to uncode function. Liu et al. employed additively HE to modify neural network model and minimize the impact on training accuracy [18]. Training on these cryptographic models may raise additional communication overhead since more data such as private key should be conveyed.

Secure aggregation. Generally, Secure multi-party computation (SMC) is employed, which mainly concentrate on how to safely calculate a function for various client without a reliable third party. Bonawitz et al. proposed the first secure aggregation protocol with utilization of SMC [9]. In this agreement, only after enough devices update their model, can server receive the aggregated model. Occasionally, global model returned by clouds may not reliable or complete. Thereafter Xu et al. devised VerifyNet which can verify correctness of returned model from cloud [46]. They implemented variation of secret sharing combined with key agreement protocol to enhance confidentiality of gradients.

Protection method for FL framework. To enhance privacy for the framework, many hybrid approaches have been proposed. Hybrid-One scheme combine DP with MPC without compromising accuracy rate, which protect communication messages rely on MPC thus introduce less noise than traditional local DP [47]. But this method often results in unaffordable communication cost and long convergence time as HE can be. Then the efficient HybridAlpha [48] emerged at the right moment, which combined functional encryption with SMC protocol to achieve the highly-performance model without privacy sacrifice.

IV. Application

Even facing with the above limitations and severe challenges of recent research, early participants have seen significant opportunities of FL and have launched a series of related explorations and attempts. Study of FL can be divided into two main areas: the first one is the extension of Google's work on decentralized model for distributed mobile devices such as prediction on keyboard [49], and service placement for mobile edge computing [50]. Besides, FL is popular in IOT devices to overcome the challenge of functionality and connectivity such as energy demand prediction for electric vehicle network [51] and optimization for smart home architecture [52]. The data collected by various terminal equipment like application in smartphone and IoT instruments like wearable devices could be summarized to analyze after edge computing and further produce valuable applications in medical field.

On the other hand, as a disruptive method to preserve data privacy, FL has great prospect in financial, industry and healthcare. Since data in these areas are not available directly due to some constraints of laws and regulations. FL could be applied in cross-organizational architecture, such as credit card fraud prediction [53] and defect detection in visual inspection task [54] as well as environmental quality prediction [55]. Furthermore, it can be applied to handle information dispersed in various institutions such as hospitals or clinics to solve the problem about detached islands while keeping data on local database [56]. FL is also widely used in the area of biomedical imaging analysis. It has been put forward by Silva et al.[57] to extract features from magnetic resonance images (MRI) come from different medical centers. Studies[58] also demonstrated that FL can be applied in the domain of Natural language processing (NLP) to analyze valid information from health records.

V. Research Fronts

FL is in great potential with sustainable development. Current main trends are committed to security compliance, attack defense and algorithm efficiency, communication cost. In this section, we focus on some remarkable cutting-edge results to solve remained problems for better FL implementation, and briefly introduce some promising direction to lead future improvement in this area.

Asynchronous training mode. Generous asynchronous training mode in FL refer to asynchronous local update or asynchronous aggregation. Chen et al. designed an asynchronous approach for client model update. Layers in deep neural network are divided into deep layers and shallow layers with different update frequency [59]. At the server side, asynchronous online framework presented by Chen et al. updates central model in an asynchronous way by introducing dynamic learning step size [60]. Wu et al. proposed a semi-asynchronous protocol which allow straggling clients don't always go together with central server [33]. Gaining a good deal of enlightenment from this semi-asynchronous method, a combination of asynchronous mode and synchronous scheme is a promising direction.

Incentive mechanism. Incentive mechanism could be established to motivate some lazy or selfish clients. The cloud server would allocate the reward to each participant according to their contribution. And the client would maximize their utility to obtain more revenue. In this way, a benign cyclic effect would be formulated to obtain a satisfied model. The frameworks such as Stackelberg-based game theory enjoy wide popularity in motivation mechanism design. Sarikaya & Ercetin explored inventive mechanism in Stackelberg perspective to inspire workers to allocate more CPU for local training [61]. Khan et al. discussed Stackelberg-based incentive mechanism to set local iteration times adaptively to be effective as much as possible [62]. For future work, more frameworks like matching theory and auction theory can be introduced to cope with trade-off between number of participants and update latency.

Verification for returned model. In real application, client may wittingly or unwittingly transmit an erroneous model compel global model to deviate from normal trace. To detect this anomalous model update, Li et al. considered an autoencoder enable model parameters to be replaced by low-dimension vector as well as discover irregular weights update [63]. Muñoz-González&Lupu discussed adaptive FL to grub abnormal updates via a Hidden Markov Model to evaluate model quality [64]. Considering loss of accuracy in federated setting, it is better to design much more Byzantine fault-tolerant system based on fault detection to eliminate or reduce threats.

FL with block-chain technology. Block chain is essentially a distributed ledger, derived from Bitcoin [65], which is characterized by decentralization, immutability, traceability, collective maintenance, openness and transparency. Injecting block-chain technology into FL, Majeed&Hong envisioned a robust Fl-chain that could verify local model update [66]. To improve performance for FL, A dynamic weighting methods considered learning accuracy and participation frequency as training weight to motivate high-quality client to get involved in the training [67]. Besides, Block-Fl proposed by Kim et.al award client holding number of samples to reduce convergence time [68]. To sum up, incorporate block chain with FL could overcome the limitation of bandwidth in FL. Further, it could not only exchange updates while verify correctness to enhance security but also employ some activate mechanism to improve FL service. But introducing blockchain may cause more latency when exchange learning model. It would be better to design a blockchain-based FL to with low latency.

VI. Conclusion

This study contributes to a general review of FL. Amidst masses of literatures, we have concluded characteristic of FL and remained problems. Further, we give the main path of optimization trace to clarify various solutions that researchers have done to optimize FL mainly including privacy concerns and algorithm efficiency. Besides, we also sum up some applications in federated settings and some develop area with great potential. As a burgeoning technology, FL attracts increasing attention these days. This work would do benefit to researchers to overcome the remained challenges to make FL more mature.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", In Proceedings of the 20th I ternational Conference on Artificial Intelligence and Statistics , 2017, pp. 1273 – 1282.

[2] M. Chen, O. Semiari, W. Saad, X. Liu, and C. Yin, "Federated Echo State Learning for Minimizing Breaks in Presence in Wireless Virtual Reality Networks," IEEE Transactions on Wireless Communications, 2020 ,vol. 19, no. 1, pp. 177 – 191.

[3] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Differentially Private Asynchronous Federated Learning for Mobile Edge Computing in Urban Informatics," IEEE Transactions on Industrial Informatics, 2020,vol. 16, no. 3, pp. 2134 – 2143.

[4] Mingqing Chen, Rajiv Mathews, Tom Ouyang, and Francoise Beaufays. Federated learning of out-of-vocabulary words. arXiv preprint .arXiv:1903.10635, 2019.

[5] P. Kairouz et al., " Advances and Open Problems in Federated Learning," arXiv:1912.04977 [cs, stat], Dec. 2019.

[6] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications", ACM Transactions on Intelligent Systems and Technology, 2019, 10(2):1-19.

[7] Q. Ho et al., "More Effective Distributed ML via a Stale Synchronous Parallel Parameter Server," Adv Neural Inf Process Syst, vol. 2013, pp. 1223 – 1231, 2013.

[8] Google. Tensorflow federated. https://www.tensorflow.org/federated, 2019.

[9] K. Bonawitz et al., "Practical Secure Aggregation for Privacy Preserving Machine Learning," Acm Sigsac Conference on Computer & Communications Security. ACM, 2017, pp.1175-1191.

[10] S. Ramaswamy, R. Mathews, K. Rao, and F. Beaufays, "Federated Learning for Emoji Prediction in a Mobile Keyboard, " arXiv:1906.04329 [cs], Jun. 2019.

[11] Webank.Federated AI Technology Enabler.(FATE) https://github.com/webankfintech/fate, 2019.

[12] Webank. https://cn.fedai.org/cases/.2019

[13] D. Gao, C. Ju, X. Wei, Y. Liu, T. Chen, and Q. Yang, "HHHFL: Hierarchical Heterogeneous Horizontal Federated Learning for Electroencephalography," arXiv:1909.05784 [cs, eess], Sep. 2019.

[14] A. Gascón et al., "Privacy-Preserving Distributed Linear Regression on High-Dimensional Data, " Proceedings on Privacy Enhancing Technologies, 2017.

[15] R. Nock et al., "Entity Resolution and Federated Learning get a Federated Resolution," arXiv:1803.04035 [cs], Mar. 2018.

[16] S. Hardy et al., "Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption," arXiv:1711.10677 [cs], Nov. 2017.

[17] S. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-Domain Sentiment Classification via Spectral Feature Alignment," presented at the Proceedings of the 19th International Conference on World Wide Web, WWW '10, 2010, pp. 751 – 760.

[18] Y. Liu, T. Chen, and Q. Yang, "Secure Federated Transfer Learning, " arXiv:1812.03337 [cs, stat], Dec. 2018.

[19] Chen, Y., Wang, J., Yu, C., Gao, W., & Qin, X. FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. arXiv:1907.09173 [cs], 2019.

[20] T. Nishio and R. Yonetani, "Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge," in ICC 2019 - IEEE International Conference on Communications (ICC), 2019, pp. 1 – 7.

[21] X. Yao, C. Huang, and L. Sun, "Two-Stream Federated Learning: Reduce the Communication Costs," in 2018 IEEE Visual Communications and Image Processing (VCIP), 2018, pp. 1 – 4.

[22] 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.

[23] M. Yurochkin, M. Agarwal, S. Ghosh, K. Greenewald, N. Hoang, and Y. Khazaeni, "Bayesian Nonparametric Federated Learning of Neural Networks," in International Conference on Machine Learning, 2019, pp. 7252 – 7261.

[24] H. Li and T. Han, "An End-to-End Encrypted Neural Network for Gradient Updates Transmission in Federated Learning," in 2019 Data Compression Conference (DCC), 2019, pp. 589 – 589.

[25] H. Zhu and Y. Jin, " Multi-Objective Evolutionary Federated Learning," IEEE Transactions on Neural Networks and Learning Systems, pp. 1 – 13, 2019.

[26] M. Mohri, G. Sivek, and A. T. Suresh, "Agnostic Federated Learning, " in International Conference on Machine Learning, 2019, pp. 4615 – 4625.

[27] S. Wang et al., "Adaptive Federated Learning in Resource Constrained Edge Computing Systems," IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1205–1221, Jun. 2019,

[28] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, " Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records, " Journal of Biomedical Informatics, vol. 99, p. 103291, Nov. 2019.

[29] V. Smith, C.-K. Chiang, M. Sanjabi, and A. S. Talwalkar, "Federated Multi-Task Learning," in Advances in Neural Information Processing Systems 30, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 4424 – 4434.

[30] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated Learning with Non-IID Data," arXiv:1806.00582 [cs, stat], Jun. 2018.

[31] M. Hao, H. Li, G. Xu, S. Liu, and H. Yang, "Towards Efficient and Privacy-Preserving Federated Deep Learning," in ICC 2019 - IEEE International Conference on Communications (ICC), 2019, pp. 1 – 6.

[32] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, " Federated Optimization in Heterogeneous Networks, " arXiv:1812.06127 [cs, stat], Sep. 2019.

[33] W. Wu, L. He, W. Lin, RuiMao, and S. Jarvis, "SAFA: a Semi-Asynchronous Protocol for Fast Federated Learning with Low Overhead," arXiv:1910.01355 [cs], Oct. 2019.

[34] J. Kang, Z. Xiong, D. Niyato, H. Yu, Y.-C. Liang, and D. I. Kim, "Incentive Design for Efficient Federated Learning in Mobile Networks: A Contract Theory Approach," in 2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS), 2019, pp. 1 – 5.

[35] N. H. Tran, W. Bao, A. Zomaya, N. M. N.H, and C. S. Hong, " Federated Learning over Wireless Networks: Optimization Model Design and Analysis," in IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 2019, pp. 1387 – 1395.

[36] Z. Chai et al., "Towards Taming the Resource and Data Heterogeneity in Federated Learning," presented at the 2019 Conference on Operational Machine Learning (OpML 19), 2019, pp. 19 – 21.

[37] J. W. Bos, K. Lauter, and M. Naehrig, "Private predictive analysis on encrypted medical data," Journal of Biomedical Informatics, vol. 50, pp. 234 – 243, Aug. 2014.

[38] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How To Backdoor Federated Learning," arXiv:1807.00459 [cs], Jul. 2018.

[39] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, "Analyzing Federated Learning through an Adversarial Lens," p. 10.

[40] J. Zhang, J. Chen, D. Wu, B. Chen, and S. Yu, "Poisoning Attack in Federated Learning using Generative Adversarial Nets," in 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 2019, pp. 374 – 380.

[41] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-box Inference Attacks against Centralized and Federated Learning," in 2019 IEEE Symposium on Security and Privacy (SP), 2019, pp. 739 – 753.

[42] Z. Wang, M. Song, Z. Zhang, Y. Song, Q. Wang, and H. Qi, "Beyond Inferring Class Representatives: User-Level PrivacyLeakage From Federated Learning," in IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 2019, pp. 2512 – 2520.

[43] C. Ma et al., "On Safeguarding Privacy and Security in the Framework of Federated Learning," arXiv:1909.06512 [cs], Sep. 2019.

[44] R. Geyer, T. Klein, and M. Nabi, "Differentially Private Federated Learning: A Client Level Perspective," Dec. 2017.

[45] C. Ilias and S. Georgios, "Machine Learning for All: A More Robust Federated Learning Framework," presented at the 5th International Conference on Information Systems Security and Privacy, 2019, pp. 544 – 551.

[46] G. Xu, H. Li, S. Liu, K. Yang, and X. Lin, "VerifyNet: Secure and Verifiable Federated Learning," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 911 – 926, 2020.

[47] S. Truex et al., "A Hybrid Approach to Privacy-Preserving Federated Learning," in Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, New York, NY, USA, 2019, pp. 1 – 11.

[48] R. Xu, N. Baracaldo, Y. Zhou, A. Anwar, and H. Ludwig, "HybridAlpha: An Efficient Approach for Privacy-Preserving Federated Learning," in Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, 2019, pp. 13 – 23.

[49] D. Leroy, A. Coucke, T. Lavril, T. Gisselbrecht, and J. Dureau, " Federated Learning for Keyword Spotting," in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 6341 – 6345.

[50] Y. Qian, L. Hu, J. Chen, X. Guan, M. M. Hassan, and A. Alelaiwi, "Privacy-aware service placement for mobile edge computing via federated learning," Information Sciences, 2019, vol. 505, pp. 562 – 570.

[51] Y. M. Saputra, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, and S. Srikanteswara, " Energy Demand Prediction with Federated Learning for Electric Vehicle Networks, " arXiv:1909.00907 [cs, eess], Sep. 2019.

[52] U. M. Aïvodji, S. Gambs, and A. Martin, "IOTFLA : A Secured and Privacy-Preserving Smart Home Architecture Implementing Federated Learning," in 2019 IEEE Security and Privacy Workshops (SPW), 2019, pp. 175 – 180.

[53] W. Yang, Y. Zhang, K. Ye, L. Li, and C.-Z. Xu, "FFD: A Federated Learning Based Method for Credit Card Fraud Detection," in Big Data – BigData 2019, Cham, 2019, pp. 18 – 32.

[54] X. Han, H. Yu, and H. Gu, "Visual Inspection with Federated Learning," in Image Analysis and Recognition,, 2019, pp .52–64.

[55] B. Hu, Y. Gao, L. Liu, and H. Ma, "Federated Region-Learning: An Edge Computing Based Framework for Urban Environment Sensing, " in 2018 IEEE Global Communications Conference (GLOBECOM), 2018, pp. 1 – 7.

[56] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. Ch. Paschalidis, and W. Shi, " Federated learning of predictive models from federated Electronic Health Records, " International Journal of Medical Informatics, vol. 112, pp. 59 – 67, Apr. 2018.

[57] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, "Federated Learning in Distributed Medical Databases: Meta-Analysis of Large-Scale Subcortical Brain Data," in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Apr. 2019, pp. 270–274.

[58] D. Liu, D. Dligach, and T. Miller, "Two-stage Federated Phenotyping and Patient Representation Learning," in Proceedings of the 18th BioNLP Workshop and Shared Task, Italy, Aug. 2019, pp. 283–291.

[59] Y. Chen, X. Sun, and Y. Jin, "Communication-Efficient Federated Deep Learning with Asynchronous Model Update and Temporally Weighted Aggregation," arXiv:1903.07424 [cs, stat], Mar. 2019.

[60] Y. Chen, Y. Ning, and H. Rangwala, " Asynchronous Online Federated Learning for Edge Devices," arXiv:1911.02134 [cs], Nov. 2019.

[61] Y. Sarikaya and O. Ercetin, " Motivating Workers in Federated Learning: A Stackelberg Game Perspective," arXiv:1908.03092 [cs], Aug. 2019.

[62] L. U. Khan et al., "Federated Learning for Edge Networks: Resource Optimization and Incentive Mechanism, " arXiv:1911.05642 [cs], Nov. 2019.

[63] S. Li, Y. Cheng, Y. Liu, W. Wang, and T. Chen, "Abnormal Client Behavior Detection in Federated Learning," arXiv:1910.09933 [cs, stat], Oct. 2019.

[64] L. Muñoz-González, K. T. Co, and E. C. Lupu, "Byzantine-Robust Federated Machine Learning through Adaptive Model Averaging," arXiv:1909.05125 [cs, stat], Sep. 2019.

[65] Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system.https://bitcoin.org/bitcoin.pdf,2008.

[66] U. Majeed and C. S. Hong, " FL chain: Federated Learning via MEC-enabled Blockchain Network, " in 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 2019, pp. 1 – 4.

[67] S. Wu and J. Du, "Electronic medical record security sharing model based on blockchain, " in Proceedings of the 3rd International Conference on Cryptography, Security and Privacy - (ICCSP), Kuala Lumpur, Malaysia, 2019, pp. 13 – 17.

[68] Y. J. Kim and C. S. Hong, "Blockchain-based Node-aware Dynamic Weighting Methods for Improving Federated Learning Performance, " in 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 2019, pp. 1 – 4.