A review of applications in federated learning

Li Li[1,2], Yuxi Fan[1], Mike Tse[3], Kuo-Yi Lin[1,2,*]

[1]College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China

[2]Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 201804, China

[3] Cardiff Business School, Cardiff University, Cardiff CF10 3AT, UK

Acknowledgements

This research was supported by National Key R&D Program of China, No. 2018YFE0105000, 2018YFB1305304, 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.

Correspondence addressee: Kuo-Yi Lin, Ph.D.
Associate Professor
College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China
Tel: +86 18217097362
Email: 19603@tongji.edu.cn

A review of applications in federated learning

Federated Learning (FL) is a collaboratively decentralized privacy-preserving technology to overcome challenges of data silos and data sensibility. Exactly what research is carrying the research momentum forward is a question of interest to research communities as well as industrial engineering. This study reviews FL and explores the main evolution path for issues exist in FL development process to advance the understanding of FL. This study aims to review prevailing application in industrial engineering to guide for the future landing application. This study also identifies six research fronts to address FL literature and help advance our understanding of FL for future optimization. This study contributes to conclude application in industrial engineering and computer science and summarize a review of applications in FL.

Keywords: Federated learning, Literature review, Citation analysis, Research front.

1.Introduction

With the development of storage capacity and processing power, the importance of data science in industrial engineering becomes more apparent. Recent years have seen the explosive development of artificial intelligence, machine learning, smart production and deep learning in industrial engineering (Li, Wang and Lin, 2020; Lin, 2018). However, there are two major challenges in this area as data science development. Firstly, data governance is the most significant aspect. Some data is privatized based on the legal concern. With the promulgation of General Data Protection Regulation (GDPR) (EU, 2018), users become the absolute owner of their own data. Any institutions or organizations do not have authority to employ user's own data unless they have agreement. Secondly, data silo is also a confronting problem that puts a limit on the development of modern industry since more training data would improve the training performance. For example, compared with the earliest AlphaGo, which used 160,000 sets of human chess data and could beat entry-level professional players. Alpha Zero (Holcomb et al., 2018) used 28.6 billion sets of human and machine-generated chess data, which could easily beat professional players. Besides, data annotation relies on experienced workers in some fields such as medical industry which may cause rareness of valid data. The scarcity of labeled data is also detrimental to industrial development. However, the emergence of FL happened to overcome these challenges in industry.

FL is a burgeoning machine learning scheme, aiming at tackling the problem of data island while preserving data privacy. It refers to multiple clients (such as mobile devices, institutions, organizations, etc.) coordinated with one or more central servers for decentralized machine learning settings. It was first put forward by Google in

2016 to predict user's text input within tens of thousands Android devices while keeping data on devices (McMahan et al., 2017). The original process of FL is generally described as Figure 1 shows. This kind of federated training approach called federated average (FedAvg), which is the baseline of FL in many other researches. Firstly, each device downloads a generic global model for the following local training. Secondly, the download global model will be improved by multiple local updates with local data, which belong to different mobile devices separately and then upload related gradient information 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 for dispersed devices. Due to the property that data are not exposed to third central server, FL is appropriate for application when data are privacy-sensitive. These includes cases in health care or mobile devices that data are not available to be aggregated with legal concern.

(Please insert Fig. 1 about here)

Recently, many scholars band together to publish papers to review advances and open problems in FL. The studies provide several further aspects to enhance FL contribution (Kairouz et al., 2019). Motivated by the promising prospects and increasing growth of FL research in industrial field, this study aims to review prevailing application of FL in industrial engineering to guide for the future landing application. This study concluded characteristic of FL and remained challenges to clarify various solutions that researchers have done to optimize FL. This study reviewed related studies of FL to base on the baseline a universal definition to identify fronts to address FL literature and help advance our understanding of FL for future optimization.

This paper is organized as follows. Beside the introduction, we sketch the overview of FL which include characteristics and mainstream open-source framework as well as categories in section 2. In section 3, we point out three challenges in FL along with relative improvement. Furthermore, we conclude indirect information leakage in FL and existing privacy-preserving method employed in FL. Section 4 discusses realistic applications in IOT devices and grounding application in industry engineering and healthcare. At the end, some frontier achievements are given, around these discussions we describe some promising direction of FL to give a guiding for future work.

2.Overview of Federate Learning

2.1 Characteristics of FL

FL is highly related to distributed learning. Traditional distributed system is made up of distributed computation, distributed storage. The first proposed FL of model update for Android clients is to some extent similar to distributed computation. Although FL put a great deal of emphasis on privacy protection, the latest researches of distributed machine learning also pay close attention to privacy-preserving distributed system. Distributed processing is to connect multiple computers in different locations via communication network under the control of center server, so that each computer undertakes different parts of the same task to complete it. Thus, the distributed processing is mainly aimed at accelerating processing stage, while FL focus on build a collaborative model without privacy leakage. To reveal difference between FL and distributed learning more specifically, we highlight following characteristics in FL.

Universality for cross-organizational scenarios. Essentially, FL proposed by Google is an encrypted distributed machine learning technology, that allows participants to build a joint training model but maintain underlying data locally. Then the original concept of FL was extended to refer to all privacy-preserving decentralized collaborative machine learning techniques (Yang, Q[b] et al., 2019). Therefore, FL is able to tackle not only horizontally partitioned data according to samples but also vertically partitioned data according to features in collaborative-learning setting. FL could be extended to bring cross-organizational enterprise into federal framework. For instance, bank that possess data of clients' purchasing power could cooperate with electronic business platform which possess data of product features, to recommend products. Thus, intelligently construct joint model for multiple entities, multiple data sources, different feature dimensions. This enable all to realize cross-platform and regional co-creation value on the premise of protecting data privacy.

Massively Non-Identically Independent Distribution (Non-IID). In FL, data is widespread in tens of thousands edge node or mobile devices. Available data in each node may no more than the total number of nodes. While in distributed system, the main purpose is to increase degree of parallelism to alleviate computation or storage pressure in central server. The number of nodes in distributed system couldn't reach the same order of magnitude as the FL. Nowadays, the world has entered an era of wearable devices which are used extensively for health monitoring (Edwards, 2019). Each device only generate several data and it cannot be compared with the total number of devices. Obviously, in this case, FL is more suitable for model improvement. In contrast with distributed system, which works primarily on balanced and 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, does not means complete decentralization, but there is no definitive center. Decentralization 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. The influence between nodes will generate a non-linear relationship through the network formed by client. 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 (Ho et al., 2013). This kind of centralized data processing method result in a double communication overhead. Because if some dataset scattered in different database are collected for training, these data should be copied and then transmitted to central server at first. And then central server will allocate data to each distributed client for distributed computation. It adds additional severe tests to system on computing power, storage and bandwidth. For cases in FL, each client is completely autonomous, data is not allocated by center and the training process is not governed by server. Therefore, FL is an integrated technology to combine machine learning models and data fusion through decentralized collaboration.

Equality of status for each node. In this cooperation framework, all parties enjoy equal status and certain dominion to achieve common prosperity. In terms of equality, whoever possesses the great mass of data has the dominant position in traditional distributed collaborative training. Thus, the development of collaborative learning in industrial field could be adversely affected by the preference on organizations with bulk of data or images with types of label. For joint training in deep learning network, those institutions with big data could manipulate the prediction model thus small and medium-sized organizations do not have impetus in joint training. However, in FL, position of these clients with small data sets would be promoted due to equality in all parties.

To sum up, FL is a decentralized technology that enable scattered clients or organizations to train a collaborative model autonomously, while keeping data localized. This method can support corporate organizations to share collaborative models without sharing any raw data.

2.2 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 (Google, 2019). 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 (Bonawitz et al., 2017). This TTF consists of FL API and Federated Core (FC) API. In detail, FL API offers a set of high-order interfaces make users can apply the included machine learning method to process federated training. FC API, the basic layer for federation learning, serving for distributed computation. Furthermore, it has been successfully applied in next word prediction or Emoji prediction in a mobile keyboard (Ramaswamy et al., 2019). In real application, it has achieved implementation over ten million of devices while hope to be highly scalable to deal with computation over billion of devices.

The other one is Federated AI Technology Enabler (FATE) created by Webank team (Webank[a], 2019). As the first open source industrial-level framework, it primarily serves for cross-organizational architecture. It provides enough privacy for client based on homomorphic encryption and secure multiparty computing. Besides, various machine learning algorithms such as logistic regression and deep learning, as well as transfer learning are able to be built on this federation system. In addition to these out of the box algorithms, most traditional methods can be adapted to this federal frame. At present, the Webank team has promoted the implementation of a series of FATE in credit risk control, object detection and anti-money laundering (Webank[b], 2019). These two frameworks are popular for FL in real application and further development on algorithm improvement.

2.3 Categorization of FL

Based on paper presented by Yang, Q[b] et al., (2019), FL largely falls into three groups, respectively, horizontal FL, vertical FL and federated transfer learning. Since data stored in different nodes or institutions mainly exist in a feature matrix form. Commonly, data consists many instances, and the horizontal axis of the sheet is regarded as client, while the vertical axis represents the characteristics of clients. Then we can divide FL based on data partition mode.

Horizontal FL. In the case of horizontal FL, 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. At present, the existing FL algorithms primarily aimed at application in smart devices or devices in the internet of things (IOT). FL in these scenarios usually could be classified into horizontal FL. Because data may significantly differ in sample space but have similar feature space simultaneously. As is mentioned above, the federated model solution for Android mobile phone update raised by Google (McMahan et al., 2017) 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., (2019) introduced hierarchical heterogeneous horizontal FL frame. The shortage of lack of label can be solved because heterogeneous domain adaptation would be adapted multiple times by using each participant as the target domain each time. This would do benefit to lack of data annotation in Electroencephalography (EEG) classification. In real application such as medical care, a large amount of work is inseparable from data collection. When it comes to cross-regional cooperation, it is almost impossible for each hospital to build a data pool for sharing. Thus, FL could construct a federal network for cross-regional hospitals with similar medical information to improve joint model as Figure 2 shows.

(Please insert Fig. 2 about here)

Vertical FL. Vertical FL is suitable for cases in which data is partitioned in the vertical direction according to feature dimension. All the parties hold homogeneous data which means they have partial overlap on sample ID whereas differ in feature space. For example, there was a medical institution, and they intend to identify illnesses such as diabetes mellitus in a predictive way. According to research, people who suffer from high blood pressure and obesity may be prone to developing type2 diabetes (Lee et al., 2020). Therefore, it could be analyzed in view of some rough dimensions, such as patients' age and weight as well as medical history. If there is a young man without obesity or high blood pressure, but intake more calories and lack of physical activity. He is also prone to get diabetes, but it couldn't be predicted and personalized due to lack of information. With development of FL, it can work with some companies which holds smartphone application data sets such as step counter or dietary structure. Further. They can cooperate with each other without demand for raw data transmission as Figure 3 shows. Generally, scholars deal with this problem through taking out the same entities with various characteristics to get a joint training. By contrast with horizontal FL, it is a more challenging work due to entity resolution（Gascón et al., 2017). Not quite as simple as situation in horizontal FL, aggregating all the dataset 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., (2018). Hardy et al., (2017) designed an end-to-end scheme on linear classifier and applied additive homomorphic encryption to defense honest-but-curious adversary for vertical FL. It is reported that current applications for parties with common sample space including traffic violation assessment and small enterprise credit risk assessment are based on FATE created by Webank team. In addition, Cheng et al., (2019) designed a secure framework called SecureBoost in the setting of vertically partitioned data set. However, the abovementioned 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.

(Please insert Fig. 3 about here)

Federated transfer learning. Unlike the scenarios in horizontal FL and vertical FL, 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 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 (Pan et al., 2010). In this way, Liu, Chen and Yang (2018) conceived federated transfer learning (FTL) to generalize FL to have broader application when it comes to common parties with small intersection. This is the first complete stack for FL based on transfer learning, including training, evaluation and cross validation. Besides this, the neural networks with additive homomorphic encryption technology in this frame could not only prevent privacy leakage but also provide comparable accuracy with traditional non-privacy-preserving method. However, communication efficiency remains an issue. Accordingly, Sharma et al., (2019) work hard on improvement for FTL. They made use of secret sharing technology instead of HE to further reduce overhead without decreasing the accurate rate. Furthermore, it could be extended to hinder malicious server. While in the previous work they assume that they the model is semi-honest. For a real application, Chen et al., (2019) constructed a FedHealth model that gather data owned by different organizations via FL and offer personalized service for healthcare through transfer learning. As shown in Figure 4, some disease diagnosis and treatment information in one hospital could be transferred to another hospital to help other disease diagnosis by FTL. The research in FTL is not yet mature, thereby there is still plenty of room for growth to make it more flexible with different data structure. Data islands and privacy protection issues are prominent problems encountered in the current large-scale industrialization of machine learning.

However, federated transfer learning is an effective way to protect both data security and user privacy while breaking the barriers of data islands.

(Please insert Fig. 4 about here)

3. Evolution of FL

The primitive framework of FL is FedAvg. Though it could deal with some lightweight Non-IID data. It is still faced with challenges of high communication overhead and structural heterogeneity. Recent works focus on algorithm optimization to improve efficiency and accuracy and participants' privacy to enhance data protection. In this section, this study discusses about the evolution and optimization in the following. We mainly explore development path in algorithm optimization level as well as security level.

3.1 Optimization

(Please insert Fig. 5 about here)

Since the term of FL was first proposed in 2016, drawing people's extensive attention, study about it has progressed. Although a lot of work had been done, there are still some challenges fail to be overcame for practical application. In terms of optimization for grounding application, high communication cost, statistical and structural heterogeneity are major issues faced by researchers currently (Li et al., 2020). In this section, we summarize the optimization path of FL according to development process and method categories to overcome these challenges. As Figure 5 shows, the algorithm optimization all based on the paper presented by McMahan et al., (2017). The first branch denotes the studies to deal with high communication cost.

The second one represents the evolution of overcome the challenge of statistical heterogeneity, while the third denotes structural heterogeneity. In the same branch, different symbols represent different ways to tackle the issue. 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.

(1) High communication cost.

By far, the key bottleneck of FL has been the difficulty of decreasing communication overhead when proceeding federal training (Yang et al., 2019). The most important feature of modern data is timeliness since the life cycle of this data is short and data iterative update speed is fast. To tackle with large masses of data and make FL flexible with explosive increasing data, reducing communication overhead should be given top priority. Meanwhile, effective efforts have been made in work including reducing communication rounds and improving model upload speed further reduce update time.

Reducing communication rounds. Due to unmatched download and upload speed, communication between server and clients is willing to be as little as possible to reduce upload times. The research of McMahan et al., (2017) is considered as the pioneering work on FL to make communication more efficient by increasing calculated quantity on each client between each communication round. They also pointed out that increase parallelism which means motivate more clients to join training on each round is an effective way. Inspired by Google, Nishio and Yonetani (2019) built FedCs framework to integrate the available clients to the utmost extent in each training round to make it efficiently in practice. Maximum mean discrepancy was inserted to FL algorithm to enforce local model to acquire more knowledge from other in training devices thus speed up convergence (Yao et al, 2018). Yurochkin et al., (2019) designed Bayesian Nonparametric FL framework, which is state of the art since it can aggregate local models into a federated model without extra parameters thus avoid unwanted communication rounds. The experiment shows that they can obtain satisfactory accuracy rating with only one communication round.

Decrease model update time. Even if the communication rounds are optimized, how to accelerate model update is a remained problem. Initially, McMahan et.al. proposed two strategies to reduce model update time (Konečný, 2017). One is structured update, which means transmit only part of the update **model** by means of low-rank model or in a random mask way. Likewise, an end-to-end neural network is a kind of structured update mode which maps update information into a lower-dimension space thus relieve pressure of communication (Li and Han, 2019). The other is sketched update, which refer to make use of compressed update model. Zhu and Jin (2019) optimized sparse evolutionary training (SET) thus convey only piece of parameters to server, which resemble the sketched update. Since in each round, each client manipulates fixed epochs, Jiang and Ying (2020) designed an adaptive method for local training. The local training epochs is decided by server according to training time and training loss, thus it will reduce local training time when loss is getting small. The above-mentioned algorithms all based on stochastic gradient descent (SGD), but this method could be inefficient if the function is anisotropic. Therefore, Liu et al., (2020) utilized momentum gradient descent to consider previous gradient information in each local training epoch to accelerate convergence speed. These algorithms are not fully suitable for all federal setting. Therefore, a more flexible communication-efficient method needs to be explored for high efficiency demand in medical industry.

(2) Statistical heterogeneity.

Traditional machine learning approach, implicitly or explicitly, assumes the data distribution is identically independent. This scenario is suitable for collecting all data and then training in a distributed way. However, data are collected from various devices or institutions thereby do not follow Identically Independent Distribution (IID). Skew characteristic and clinical validation may vary among different equipment version (Godinho et al., 2016). And data record form in across multiple horizontals could be totally different. Besides, there's may be a huge variety of data size in different nodes result in an unbalanced distribution. To tackle this problem, the general resolution is to focus on global model, or modify local training mode, or adding some extra procedure on data pre-processing stage.

Focus on global model. The first proposed FedAvg algorithm resolve this issue by averaging local upgrade on each device directly. In addition, Mohri et al., (2019) noticed previous work ignore the importance of fairness which may lead to bias centralized model. They improved global model to cope with any target distribution comprised by a mixture of different clients. As for aggregation stage, convergence behavior is another stressed issue. The existence of heterogeneity may lead to misconvergence of global model. Further Wang et al., (2019) discussed convergence bound of FL based on gradient-descent in Non-IID data background, and further bring forward an improved adaptive method to reduce loss function within constraints of resource budget. Moreover, Li, X. et al (2019). gave four kinds of convergence theorems with different parameters setting or premises for FedAvg in Non-IID situations. These studies fill a part of the theoretical gap in the research of convergence speed of a FL algorithm. Besides, it provides the effect of parameter adjustment on the convergence speed for the guidance.

Add extra data preprocessing procedure. For data pre-processing, Huang[a] et al., (2019) introduced clustering thought with FL and constructed a community-based FL method. By separating independent data into different clusters, then processing federated training on each community, the non-IID problem is thus can be resolved. However, one drawback is that it’s not suitable for massively data training due to high parameter conversion overhead. In hierarchical heterogeneous horizontal framework, it projects each embedding submanifold into a common embedding space to overcome data heterogeneity (Gao et al., 2019).

Modify local training mode. 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 (Smith et al., 2017). Zhao et al., (2018) did the similar work, they considered a solution to deal with non-iid data by sharing a small set of data among each local model. Huang[b] et al., (2019) also gained a good deal of enlightenment from the previous data sharing ideology to overcome Non-IID problem. They put cross-entropy loss into transmission process and assign different local update times for each client in each round.

(3) Structural heterogeneity.

In terms of 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, thereby cause unbalanced training time.

On the other hand, clients differ in network environment. The unreliable and unstable network may lead to devices' drop out. Up to now, methods to deal with structural heterogeneity mainly focus on resource allocation for heterogeneous devices and fault tolerance for devices prone to be offline.

Fault tolerance. The federated multi-task learning was constructed in the wake of Google's research on decentralized data training (Smith et al., 2017). To address the issue of stragglers (who is drop out or still training with an outdated global parameters), they considered influence with low participation in training process to resist device drop out. Enable FL system to be robust to dropped participants, scholars also designed secure aggregation protocol (Hao[a] et al., 2019) which is tolerant with arbitrary dropouts as long as surviving users are enough to join federate update. Li[b] et al., (2019) take stragglers into account and allow these devices to implement different locally update computation times. Wu et al., (2019) also fully considered device straggling phenomenon in heterogeneous network. They made use of a cache structure to store those unreliable user update thus alleviates their trustless impact on global model.

Resource allocation. For the sake of resource constraint, most of foregoing work devote to allocate resources properly to heterogeneous devices. For instance, Kang et al., (2019) took overhead in heterogeneous clients into consideration to motivate more high-quality devices to participate training process. And Tran et al., (2019) studied training accuracy and convergence time with influence of heterogeneous power constraints. Meanwhile, Chai et al., (2019) considered the impact of resource (e.g. CPU, memory, and network resources) heterogeneity on training time of FL. To address this issue, Li, T. et al., (2020) designed a fairness metrics to measure loss in devices and a q-Fair optimization goal to impel fair resource allocation in FL. In a nutshell, stragglers and heterogeneity run through FL framework. Therefore, in the future, optimization should continue to contribute to fault-tolerance and properly resource allocation to address this issue.

3.2 Security Analysis

(Please insert Fig. 6 about here)

In this section, we elaborate the evolution of privacy attack and enhancement in FL. As shown in Figure 6, the first branch indicates indirect privacy leakage in FL. And the other two branches show improvement trace for privacy enhancement for FL. One is privacy-preserving method on client side, and the other one is on the server side. These two branches intersect at a node which derive another branch to denote hybrid approach to enhance privacy. The thickness of the line also shows reference frequency of these papers. The thicker the line, the higher reference frequency of the paper. The details are as follows.

3.2.1 Privacy Risk

Though patients' private data never come out of the local storage during federated training process 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 (Bos et al., 2014). Since original data under the risk of being cracked by back deduction. Some investigators have considered to retrieve data in FL framework. The general attack types are mainly divided into three categories as bellow:

Data poisoning attack. Aiming at embedding some tainted data such as malicious samples or disguised data to destroy data integrity or give rise to the bias of training results. There are two main types of 'data poisoning' attack modes including model skew and feedback weaponization. Traditional machine learning approaches are vulnerable to data poisoning since adversarial could directly manipulate the triggers to misguide the global model. Nevertheless, these traditional data poisonings methods are less effective or may need many malicious participants when it comes to FL since malicious attackers have no direct access to raw data (Bagdasaryan et al., 2018). On the basis research of Bagdasaryan et al., (2018), Yang, Q[b]. et al., (2019) studied a novel and effective distributed backdoor attack. They divided an attack trigger into many slices and embedded each slice into different attackers instead of embedding a complete trigger into only one attacker. This new-fashioned mode throws a wrench in the old argument that FL is possible to avoid data poisoning. It also gives a new evaluation form for security analysis in FL.

Model poisoning（Also known as Adversarial attack). Model poisoning refer to make machine learning model to generate a wrong result by designing a specific input. Furthermore, it can be subdivided into Non-targeted adversarial attack and Targeted adversarial attack. The former one is a common type which lead to an incorrect consequence, and the other one is relatively difficult that aiming at injecting a specific type for input. In FL, secure aggregation is implemented, and aggregator is not familiar with the local update modes thus are not able to detect anomalies or verify correctness of local updates. According to this drawback, the backdoor can be inserted into federated environment by malicious participant through model-replacement methodology thus misunderstand the joint model. This novel attack method can be successfully employed in federated training tasks including image classification and word prediction (Bagdasaryan et al., 2018). Similarly, Bhagoji et al., (2019) attacked global model through few malicious adversaries to wrongly classified targeted model. This kind of attack obviously belong to targeted adversarial attack. In this case, they ensure convergence of integrated model and accuracy of most tasks. In addition, the results show Byzantine-resilient aggregation technology is weak to offense this type of attack in the federated setting. Then Zhang et al., (2019) give first attempt to generate model poisoning attack based on Generative Adversarial Nets (GAN). In this work, malicious participant pretended to be a benign agent. Then they assign a GAN architecture to generate training data as well as distributed a wrong label to induce benign client to be damaged. The existing methodologies aiming at defending poisoning attack are quite invalid in federated settings. 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 (also known as membership inference attacks) and reconstruction attacks. The first mentioned of two indicates to infer whether a client is contained in the data set. The latter advocates recover some features about an individual participant. With utilization of vulnerability of SGD, Nasr et al., (2019) designed a white-box membership inference attack method direct at neural network. Then it was successfully applied to federated setting to infer information via a curious server or any of a participant. The previous work focuses on malicious server assumption and unable to recover information on specific client because of invisibility of client update. In cases of this kind, Wang, Z. et al., (2019) built a general attack frame called mGAN-AI which could reconstruct private information for target client.

To hinder this kind of attack, more stronger protection method should be explored, and data could be encrypted before upload to cloud.

3.2.2 privacy-preserving technology in FL

The indirect privacy disclosure poses immense challenges on development of FL. Potential threats are usually from insider adversaries and outsider adversaries. Insider adversaries including honest-but-curious aggregator, colluding parties and malicious participants steal privacy during training process. The honest-but-curious aggregator means that the server will keep to the privacy agreement but have a try to explore more information about clients. Colluding parties or malicious participants are unreliable to transmit incorrect updates as well as learn additional information from other benign clients. Outsider adversaries refer to those who can peep intermediate outputs or users that have access to final model. Faced with these vulnerabilities, 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 (Ma et al., 2019). This study discusses novel privacy-preserving technologies based on this classification as bellows.

Privacy-preserving at client side. Differential privacy often acts as a means of enhancing privacy preservation for client. When querying data from database, it will reduce chances for records to be identified while maximize query accuracy as much as possible by introducing noise to blur raw data. For instance, since FedAvg is prone to be violated by differential attack, Geyer et al., (2019) leveraged differential privacy on FL to conceal whether a client participant in the training process. Likewise, to improve FedAvg, McMahan et al., (2018) also applied DP to this process by adding Gaussian noise to the global model. In federated online training for ranker using feedback from users, Kharitonov (2019) introduced ε-local differential privacy. Opposite to common algorithms, it is stricter since they protect user-lever privacy instead of imposing privacy-preserving technology after data aggregation.

In addition, homomorphic encryption is also a privacy policy applied in FL frequently to hinder information leakage during parameter exchange process among clients. Homomorphic encryption 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., (2018) employed additively homomorphic encryption to modify neural network model and minimize the impact on training accuracy. Ilias and Georgios (2019) also added homomorphic encryption to a more robust FL framework, which make it possible to compute aggregation on encrypted client. Training on these cryptographic models may raise additional communication overhead since more data such as private key should be conveyed.

Locality-sensitive hashing (LSH) is also a prevalent way to keep confidentiality (Gionis et al., 1999). All features would be mapped into an encryption form via p-stable hash function. The main advantage of this encryption mode is that similarity between two samples will be retained after hash representation. However, two different samples virtually impossible to hold similar hash values. Raw data wouldn't be exposed because many samples may have same outputs. Besides, LSH would not cause overmuch communication overhead like homomorphic encryption and reduce accuracy like differential privacy. Lee et al., (2018) make use of LSH to detect similar patients in federated settings. Recently, Li et al., (2020) build a practical gradient boosting decision trees rely on LSH. In the pre-processing stage, LSH would help find similar samples dispersed in different clients, and they will use the sum gradients of similar instances instead of only use the gradient of one instance when processing gradient updating.

Secure aggregation. 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., (2017) proposed the first secure aggregation protocol with utilization of secure multiparty computation. In this agreement, model update information of each device is unrevealed to central server. Only after enough devices update their model, can server receive the aggregated model. Owing to the quadratic communication cost, the above-mentioned protocol is not applicable for larger scale situations. By this way, Hao[a] et al., (2019) envisioned a more efficient privacy-preserving scheme for FL, which integrate differential privacy and lightweight homomorphic encryption technology. This protocol, mainly for stochastic gradient descent approach, is robust to curious-but-honest server and collusion between the cloud and server. Occasionally, global model returned by clouds may not reliable or complete. Because unreliable cloud server may be malicious to return a totally wronged model or may be lazy to convey a compressed but inaccurate model due to computational pressure. Thereafter Xu et al., (2020) devised VerifyNet, the first protocol that can verify correctness of returned model from cloud. For privacy guarantee, they implemented variation of secret sharing combined with key agreement protocol to enhance confidentiality of gradients. The up-to-date approach proposed by Chen et al., (2020) also concentrated on secure aggregation scheme. They add an extra public parameter dispatch to each client to force them training in a same way, thus detect malicious client easily when making an aggregation stage.

Protection method for FL framework. Although aforementioned algorithms could avoid adversary to invade central server or clients, the encrypted parameters may still cause information leakage by means of novel attack methods as 3.2.1 mentioned. To enhance privacy for the framework, many hybrid approaches have been proposed. However, the introduced noise of differential privacy may lead to decreased accuracy. To reduce noise, the Hybrid-One scheme combine the use of DP with MPC without compromising accuracy rate, which protect communication messages rely on MPC thus introduce less noise than traditional local DP (Truex et al.,2019). But this method often result in unaffordable communication cost and long convergence time as homomorphic encryption can be. Then the efficient HybridAlpha emerged at the right moment, which combined functional encryption with SMC protocol to achieve the highly-performance model without privacy sacrifice (Xu et al.,2019). Additionally, sketched algorithms are inherently suitable for FL since data identities are not stored, and extra mechanisms are needed to trace back original data. Inspired by this, Liu, Li, Smith and Sekar (2019) established relationship between FL and sketching algorithm to strength confidentiality.

4.Application

(Please insert Table. 1 about here)

FL takes hold as a prevailing scheme with the construction of collaborative model without legal concern. Even facing with the forementioned limitations and severe challenges, early participants have seen significant opportunities of FL and have launched a series of related explorations and attempts to apply FL in real life. In this section, we discuss several applications related to industry engineering or computer science.

4.1 Application for mobile devices

FL has been paid much attention to by the researchers since the concept was first put forward by Google to predict users' input from Gboard on Android devices. Further improvement for prediction on keyboard has been made through Chen, Mathews, Ouyang and Beaufays (2019), Leroy et al., (2019), Hard et al., (2019) and Yang, T. et al., (2018). Besides, emoji prediction is also a research hotspot (Ramaswamy et al.,2019). In addition, bring FL model to smart devices to predict human trajectory (Feng et al., 2020) or human behavior (Sozinov et al., 2018) is also a potential application.

Nowadays, although there is a rapid growth in storage capacity and computing power of mobile devices. It's difficult to satisfy the growing quality demand from mobile subscribers due to communication bandwidth limitation. Thus, most of comprehensive provider prefer to offer a service environment at the edge of the cellular network close to the customer instead of integrate cloud computing and cloud storage in core network so as to reduce network congestion. This technology is dubbed mobile edge computing (MEC), but it also faces increased risk of information leakage. One possible solution is the combination of FL and MEC, Wang, X. et al., (2019) investigate an 'In-Edge AI' framework which combine FL based on deep reinforcement learning with MEC system and further optimize resource allocation problem. Further, Qian et al., (2019) devoted to utilizing FL on MEC. They developed a privacy-aware service placement scheme to provide high-quality service by caching desired service on the edge server close to the users.

In this case, mobile devices not only refer to common smart phones but also refer to devices in IOT settings. Smart home is one of the important applicable fields of IoT. To better learn users' preference, devices in smart home architecture would upload some related data to cloud server which may lead to data breach. Therefore, Aïvodji et al., (2019) present a sufficient secure federated architecture to build joint models. Similarly, Yu et al., (2020) build a federated multi-task learning framework for smart home IOT to automatically learn users' behavior patterns, which could effectively detect physical hazards. Furthermore, Liu, B. et al., (2020) proposed a data fusion approach based on FL for robots imitation learning in robot networking. This method could be leveraged on self-driving cars to generate guide models and foresee various emergencies.

4.2Application in Industrial Engineering

Driven by the achievement of FL in data privacy protection, it is logical for industrial engineering to follow it with applications of FL. Since data in these areas are not available directly due to some constraints of laws and regulations. However, only when FL is leveraged to these areas, can we make use of these disperse dataset to acquire infinite benefits.

To the best of our knowledge, following with the rise of and maturation of FL, it could have widely popularization and application prospects in data-sensitive fields for industrial engineering. Take environment protection as a case in point, Hu et al., (2018) designed a novel environmental monitoring frame based on federated region learning FRL) for the sake of inconvenient interchangeable monitor data. Thus, monitoring data dispersed from various sensors could be utilized for superior performance of collaborative model. FL is also applied to visual inspection task (Han et al.,2019). It could not only help us solve the problem of lacking defective samples to detect defects in production tasks but also offered privacy guarantees for manufacturers. In image fields, vision -and-language is also a flashpoint, Liu, Wu, Ge, Fan and Zhou (2020) bring FL to acquire diversiform representation from federated tasks for better grounding applications. Apart from image detection and representation, FL is suitable for malicious attacks detection in communication system composed by Unmanned Aerial Vehicles (UAVs) (Mowla et al., 2020). Since the features of UAVs such as unbalanced data distribution and unreliable communication conditions are quite matching with challenges in FL. With the popularization of electric vehicles, Saputra et al., (2019) designed a federated energy demand prediction method for various charging stations to prevent energy congestion in transmission process. Moreover, Yang, Zhang, Ye, Li and C.-Z. Xu (2019) leveraged FL to transactions owned by different banks in order to detect credit card fraud efficiently, which is also a significant contribution to financial field. For text mining, Wang, Y. et al., (2020) exploit an industrial grade federated framework based on Latent Dirichlet Allocation. It has passed the assessment on real data for spam filtering and sentiment analysis.

To summarize, FL enable data owner to broaden the scope of data applications and improve model performance through iteration among different entities. In the future, FL technology would also support more industries to become more intelligent. The incorporation with FL in AI will build a federal ecosystem without data privacy concern.

4.3 Application in HealthCare

As a disruptive method to preserve data privacy, FL has great prospect in health care. Each medical institute might have a lot of patient data, but that may be far from enough to train their own prediction models (Szegedi, Kiss and Horváth,2019). Combination of FL and disease prediction is one of the good solutions to break down the barriers of analysis throughout different hospitals.

Electronic health records (EMR) contain lots of meaningful clinical concepts, Kim, et al., (2017) gave an attempt to use tensor factorization models for phenotyping analysis to obtain information concealed in health records without sharing patient-level data. It could be regarded as the first attempt for FL application in medical industry. Pfohl et al., (2019) explored differentially private learning for EMR in federated setting. And they further demonstrated the performance is comparable with training in a centralized setting. Huang[a] et al., (2019) make use of EMRs scattered across hospitals to predict mortality rate for heart disease patients. During training process, there is not any form of data or parameters transmission among hospitals' databases. Besides this, data consolidated from multiple remote clients into a central server is encoded in advance and the decoder will be abandoned at the end of training. In addition, Brisimi et al., (2018) also use EMRs to evaluate whether a patient with heart disease will be hospitalized based on a FL algorithm called cluster Primal Dual Splitting (cPDS). This prediction work can be accomplished either on health monitoring devices or hospitals holding these medical data without information leakage. With utilization of health records, Lee et al., (2018) proposed a federated patient hashing framework to detect similar patients scattered in different hospitals without sharing patient-level information. This patient matching method could help doctors to summarize general character and direct them to treat patient with more experience. In addition, Huang[b] et al., (2019) leveraged Loss-based adaptive boosting

Federated Averaging algorithm on drug usage extracted from MIMIC-III database to predict patient mortality rate. This research concerned computation complexity and communication cost as well as accuracy for each client therefore outperform baselines.

Studies also demonstrated that FL can be applied in the domain of Natural language processing (NLP) to analyze valid information from health records. Liu, Dligach and Miller (2019) focus on need for unstructured data processing of clinical notes. It was the first attempt of NLP based on FL. They performed a two-stage federated training model contains pre-processing stage to predict a representation model for each patient and phenotyping training stage to study each kind of illness.

Recently, FL is also widely used in the area of biomedical imaging analysis. Federated principal components analysis (fPCA) has been put forward by Silva et al., (2019) to extract features from magnetic resonance images (MRI) come from different medical centers. Furthermore, Gao et al., (2019) proposed a hierarchical heterogeneous horizontal FL (HHHFL) framework for Electroencephalography (EEG) classification to overcome the challenge of limited labeled instances as well as the privacy constraint.

To the best of our knowledge, following with the rise of and maturation of FL, it could have very wide popularization and application prospects in data-sensitive fields in addition to the abovementioned fields. Table 1 shows application of FL has grown by leaps and bounds in 2019. Thus, it is optimistic that FL would have great potential in the future development. Currently, FL mainly contributes to horizontally collaborative training for landing applications, which means feature dimensions of each data are similar to each other. In the future, medical data in hospitals could be cooperated with other institutions such as insurance agent to obtain reasonable pricing. Therefore, vertically FL is a promising direction to be explored. Moreover, one problem is existing federal training mostly base on small set of organizations and is not able to extend to collaborative training for huge number of devices or institutions. Therefore, analysis of mobile devices data based on FL in an effective way should be progressed to generate more meaningful information.

5. Frontier achievements and Future work

FL is in great potential with sustainable development for landing application in industrial engineering and health care. Admittedly, many scholars have done arduous efforts to tackle challenges mentioned in section 3. To satisfy the situation with rapid development of IOT and increasing privacy concerns, it put forward rigorous demands for federated system design. Several research frontiers remain to be explored with FL. Current main trends are committed to security compliance establishment, attack defense and efficiency promotion as well as heterogeneities processing. In this section, we focus on some remarkable cutting-edge results to solve remained problems for better FL implementation in practical manufacturing application. Additionally, we also briefly introduce some promising direction to lead future improvement in this area.

***Asynchronous training mode*.** A basic choice on global model training mode is whether to take asynchronous or synchronous method. Recently, the synchronous training has already become the major form for FL due to superior performance of SGD in the central server settings compared to asynchronous way (Chen, Ning and Rangwala, 2019; Mohammad and Sorour, 2019) Prior optimization of FL mainly focusses on evolution of FedAvg in a synchronous fashion. However, this method relies on strong assumption of which is not realistic in practice. The heterogeneous resource in terms of different computation ability and various network settings and unbalanced data distribution would result in different training time and unknown communication cost. Based on previous work on asynchronous gradient descent, Sprague et al., (2019) compared asynchronous aggregation scheme with FedAvg and obtained basically satisfactory results. Generous asynchronous training mode in FL refer to asynchronous local update or asynchronous aggregation. At the client side, Chen, Sun and Jin (2019) 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. At the server side, asynchronous aggregation could be implemented. For instance, asynchronous online FL framework presented by Chen et al., (2019) updates central model in an asynchronous way by introducing feature learning and dynamic learning step size. Considering trade-off between advantages of synchronous update and asynchronous training, Wu et al., (2019) proposed a semi-asynchronous protocol which allow straggling clients don't always go together with central server. The main idea is that make stragglers join training properly to speed up training process with utilization of their slowly update model. Gaining a good deal of enlightenment from this semi-asynchronous method, a combination of asynchronous mode and synchronous scheme is a promising direction. In this way, can we reduce unwanted overhead and give little fault-tolerance to stragglers.

Gradient aggregation. Usually, in gradient aggregation stage, the gradient of global model is the sum of weighted gradient produced by each client. And the weight of each client is decided by the sample ratio. However, there is no evidence demonstrate that this weighted averaging gradient acquired from local clients is equivalent to real global gradient information due to biased estimation in local clients. Xiao et al., (2020) detect that the mutual information is increased which implies correlation between clients, while the distance of parameters is getting greater with increased iteration. This study shows gradient averaging is possible not a good manner for gradient aggregation. To eliminate gradient bias in local training stage, Yao et al., (2019) keep trace of dispatched global parameters in each local training epoch. Since local gradient update is a function of global parameters, then gradients can be aggregated in an unbiased way. To better learn aggregation mode in FL, Ji et al., (2019) introduce a recurrent neural network aggregator to automatically get an optimized way for gradient aggregation. In addition, Wang et al., (2019) designed a layer-wise aggregation mode to serially generate layer parameters in neural network for global model. Considering Non-iid distribution on clients, gradient aggregation in a simply averaging way isn't a good choice. It would be better if researchers can bring in some adaptive weight for each client or some machine learning method to learn how to aggregate these gradients in an effective way.

Incentive mechanism. For performance improvement, apart from optimization of resource allocation or novel architecture design, to establish an incentive mechanism to encourage more parties join into the training is also an effective way. The original FedAvg would select clients randomly. It seems that all clients are equally likely to go into the training. In fact, some lazy clients with high quality or some selfish clients afraid of power consumption may not attend the whole training process with a certain probability. Incentive mechanism could be established to motivate such 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 and Ercetin (2019) explore inventive mechanism in Stackelberg perspective to inspire workers to allocate more CPU for local training. Khan et al., (2019) discussed Stackelberg-based incentive mechanism to set local iteration times adaptively to be effective as much as possible. The crowdsourcing framework adopted two-stage stackelberg model to acquire utility maximization among clients and server (Pandey et al., 2019). 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. Most privacy-preserving method in FL rely on a strong assumption that clients are semi-honest which obey training rules but keeping curious about private data acquisition. However, realistic application gets the other kind. Client may wittingly or unwittingly transmit an erroneous model compel global model to deviate from normal trace. For instance, in wearable medical system, adversaries may generate plausible but not accurate data to attack the entire model (Cai and Venkatasubramanian, 2018). This kind of Byzantine problem is always encountered in FL. Thus, Byzantine fault-tolerant system should be developed which means even if certain clients don’t follow training protocol or be malicious to attack global model, the collaborative training can still work well.

To detect this anomalous model update, Li et al., (2019) considered an autoencoder enable model parameters to be replaced by low-dimension vector as well as discover irregular weights update. Muñoz-González and Lupu (2019) discussed adaptive FL to grub abnormal updates via a Hidden Markov Model to evaluate model quality. Traditional Byzantine fault-tolerant system is supported by some defense mechanism rather than malicious client detection. 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. As a novel technology, block-chain is developing fast abroad. In short, Block-chain is essentially a distributed ledger, derived from Bitcoin (Nakamoto. S, 2008), which is characterized by decentralization, immutability, traceability, collective maintenance, openness and transparency. Several blockchain-assisted schemes for industrial data sharing have been proposed, including quality surveillance of 3D-printed articles (Kennedy et al., 2017), consumption monitoring and privacy-preserving energy trading for smart grids (Aitzhan and Svetinovic, 2018) and emergency medical service for pre-hospital care (Hasavari and Song 2019). Existing studies based on block chain mainly focus on innovate medical information sharing system but training collaboratively to maximize data utilization has not been implemented. Recent research has proofed that block chain has potential to significantly transform some issues in FL. Blockchain and FL are auxiliary to each other. As an inherently secure distributed system, blockchain naturally suitable for developed with FL. Since blockchain framework is tolerant with malicious node and work normally as long as malicious nodes do not exceed 51% of the total.

Injecting block-chain technology into FL, Majeed and Hong (2019) envisioned a robust Fl-chain that could verify local model update. Although the security of entire architecture can be guaranteed with block chain technology, this security has nothing to do with privacy protection. There is no privacy concern in allusion to individual node. If there is a malicious clinic or hospital join in the collaborative training, it may spare no effort to snoop other participants' privacy information. Hence Ilias and Georgios (2019) utilized blockchain smart convention to coordinate all clients and additionally used homomorphic encryption to provide extra privacy guarantee. The blockchain-based privacy-preserving FL framework designed by Awan et al., (2019) also added a variation of the Paillier cryptosystem as an excess measure to forestall privacy leakage. Furthermore, take advantage of block chain, the contribution of each party to optimize global model could be traced, which make it possible for an incentive mechanism. Aforementioned FL frames based on block chain didn't give specific rewarding mechanism for clients to join training. To improve performance for FL, a dynamic weighting method had been proposed (Kim and Hong, 2019). It considered learning accuracy and participation frequency as training weight to motivate high-quality client to get involved in the training. Besides, Block-Fl proposed by Kim, H. et al., (2019) award client holding number of samples to reduce convergence time. To sum up, incorporate block chain with FL is auspicious since it is a decentralized technology thus doesn't need central server to predict global model anymore. Therefore, it 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.

Federated training for unsupervised machine learning. According to the analysis of research on FL, existing FL frameworks construct based on supervised learning method. For instance, FL have been effectively leveraged in neural network (Wang, S. et al., 2019; Hao[b] et al., 2019; Bonawitz et al., 2019) and SVM (Liu et al., 2019), as well as linear classifier (Hardy et al., 2017).

Actually, in most of cases where labeled data either do not exist, or with little existence, unsupervised learning methods are supposed to be applied. Thus, unsupervised learning is reasonable to be used to infer potential information in these messy data. For example, it has been widely used for image registration (Dalca, Balakrishnan, Guttag and Sabuncu, 2019; de Vos et al., 2019) and image classification (Ahn et al., 2019). Although researchers have made great progress on federated transfer learning to handle with dispersive data with few labels, the landing applications remain a bottleneck for unsupervised learning in federated framework. To tackle the challenge of limited number of labels, collaborative training has been employed in unsupervised area. Such method like Collaborative and Adversarial Network (CAN), a novel unsupervised domain adaptation approach, shows effectiveness and high performance (Zhang et al., 2018). Therefore, as a kind of collaborative training approach, FL has great potential on unsupervised learning area. Recently, van Berlo et al., (2020) introduced Federated Unsupervised Representation Learning which is a breakthrough in unsupervised FL. Through unsupervised representation learning during pre-training stage, the requirement of labeled data significantly reduced. This study also shows competitive performance compared with supervised learning and transfer learning. Therefore, it motivates future work towards the extension of federated framework on unsupervised learning.

6. Conclusion

This study contributes to conclude application in industrial engineering and computer science and summarize review of FL but not limited to applications. To our best knowledge, this work is the first time to summarize the development prospects of FL on industrial field. Amidst masses of literature, we have concluded characteristic of FL and remained challenges. 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 benefits to researchers to overcome the remained challenges of FL.

REFERENCE

Ahn, E., Kumar, A., Feng, D., Fulham, M. and Kim, J. (2019). Unsupervised Deep Transfer Feature Learning for Medical Image Classification. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 1915–1918. https://doi.org/10.1109/ISBI.2019.8759275

Aitzhan, N. Z. and Svetinovic, D. (2018). Security and Privacy in Decentralized Energy Trading Through Multi-Signatures, Blockchain and Anonymous Messaging Streams. IEEE Transactions on Dependable and Secure Computing, 15(5), 840–852. https://doi.org/10.1109/TDSC.2016.2616861

Aïvodji, U. M., Gambs, S. and Martin, A. (2019). IOTFLA: A Secured and Privacy-Preserving Smart Home Architecture Implementing Federated Learning. 2019 IEEE Security and Privacy Workshops (SPW), 175–180. https://doi.org/10.1109/SPW.2019.00041

Awan, S., Li, F., Luo, B. and Liu, M. (2019). Poster: A Reliable and Accountable Privacy-Preserving Federated Learning Framework using the Blockchain. Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security - CCS ’19, 2561–2563. https://doi.org/10.1145/3319535.3363256

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D. and Shmatikov, V. (2018). How To Backdoor Federated Learning. ArXiv:1807.00459 [Cs]. Retrieved from http://arxiv.org/abs/1807.00459

Bhagoji, A. N., Chakraborty, S., Mittal, P. and Calo, S. (2019). Analyzing Federated Learning through an Adversarial Lens. International Conference on Machine Learning, 634–643. Retrieved from http://proceedings.mlr.press/v97/bhagoji19a.html

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., … Roselander, J. (2019). Towards Federated Learning at Scale: System Design. ArXiv:1902.01046 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1902.01046

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., … Seth, K. (2017). Practical Secure Aggregation for Privacy Preserving Machine Learning (No. 281). Retrieved from http://eprint.iacr.org/2017/281

Bos, J. W., Lauter, K. and Naehrig, M. (2014). Private predictive analysis on encrypted medical data. Journal of Biomedical Informatics, 50, 234–243. https://doi.org/10.1016/j.jbi.2014.04.003

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C. and Shi, W. (2018). Federated learning of predictive models from federated Electronic Health Records. International Journal of Medical Informatics, 112, 59–67. https://doi.org/10.1016/j.ijmedinf.2018.01.007

Cai, H. and Venkatasubramanian, K. K. (2018). Detecting data manipulation attacks on physiological sensor measurements in wearable medical systems. EURASIP Journal on Information Security, 2018(1), 13. https://doi.org/10.1186/s13635-018-0082-y

Chai, Z., Fayyaz, H., Fayyaz, Z., Anwar, A., Zhou, Y., Baracaldo, N., … Cheng, Y. (2019). Towards Taming the Resource and Data Heterogeneity in Federated Learning. 19–21. Retrieved from https://www.usenix.org/conference/opml19/presentation/chai

Chen, J., Pan, X., Monga, R., Bengio, S. and Jozefowicz, R. (2017). Revisiting Distributed Synchronous SGD. ArXiv:1604.00981 [Cs]. Retrieved from http://arxiv.org/abs/1604.00981

Chen, M., Mathews, R., Ouyang, T. and Beaufays, F. (2019). Federated Learning Of Out-Of-Vocabulary Words. ArXiv:1903.10635 [Cs]. Retrieved from http://arxiv.org/abs/1903.10635

Chen, Y., Luo, F., Li, T., Xiang, T., Liu, Z. and Li, J. (2020). A training-integrity privacy-preserving federated learning scheme with trusted execution environment. Information Sciences, 522, 69–79. https://doi.org/10.1016/j.ins.2020.02.037

Chen, Y., Ning, Y. and Rangwala, H. (2019). Asynchronous Online Federated Learning for Edge Devices. ArXiv:1911.02134 [Cs]. Retrieved from http://arxiv.org/abs/1911.02134

Chen, Y., Sun, X. and Jin, Y. (2019). Communication-Efficient Federated Deep Learning with

Asynchronous Model Update and Temporally Weighted Aggregation. ArXiv:1903.07424 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1903.07424

Chen, Y., Wang, J., Yu, C., Gao, W. and Qin, X. (2019). FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. ArXiv:1907.09173 [Cs]. Retrieved from http://arxiv.org/abs/1907.09173

Cheng, K., Fan, T., Jin, Y., Liu, Y., Chen, T. and Yang, Q. (2019). SecureBoost: A Lossless Federated Learning Framework. ArXiv:1901.08755 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1901.08755

Dalca, A. V., Balakrishnan, G., Guttag, J. and Sabuncu, M. R. (2019). Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. Medical Image Analysis, 57, 226–236. https://doi.org/10.1016/j.media.2019.07.006

de Vos, B. D., Berendsen, F. F., Viergever, M. A., Sokooti, H., Staring, M. and Išgum, I. (2019). A deep learning framework for unsupervised affine and deformable image registration. Medical Image Analysis, 52, 128–143. https://doi.org/10.1016/j.media.2018.11.010

Edwards, J. (2019). Medicine on the Move: Wearable devices supply health-care providers with the data and insights necessary to diagnose medical issues and create optimal treatment plans [Special Reports]. IEEE Signal Processing Magazine, 36(6), 8–11. https://doi.org/10.1109/MSP.2019.2930767

EU. Regulation of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive (general data protection regulation).(2018). https://eurlex.europa.eu/legal- content/EN/TXT. Accessed 26 December 2018.

Feng, J., Rong, C., Sun, F., Guo, D. and Li, Y. (2020). PMF: A Privacy-preserving Human Mobility Prediction Framework via Federated Learning. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4(1), 10:1–10:21. https://doi.org/10.1145/3381006

Gao, D., Ju, C., Wei, X., Liu, Y., Chen, T. and Yang, Q. (2019). HHHFL: Hierarchical Heterogeneous Horizontal Federated Learning for Electroencephalography. ArXiv:1909.05784 [Cs, Eess]. Retrieved from http://arxiv.org/abs/1909.05784

Gascón, A., Schoppmann, P., Balle, B., Raykova, M., Doerner, J., Zahur, S. and Evans, D. (2017). Privacy-Preserving Distributed Linear Regression on High-Dimensional Data. Proceedings on Privacy Enhancing Technologies, 2017(4), 345–364. https://doi.org/10.1515/popets-2017-0053

Geyer, R. C., Klein, T. and Nabi, M. (2018). Differentially Private Federated Learning: A Client Level Perspective. ArXiv:1712.07557 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1712.07557

Gionis, A.,Indyk, P., Motwani, R., et al., (1999). Similarity search in high dimensions via hashing. In Vldb, volume 99, 518–529.

Godinho, C., Domingos, J., Cunha, G., Santos, A. T., Fernandes, R. M., Abreu, D., … Ferreira, J. J. (2016). A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. Journal of Neuroengineering and Rehabilitation, 13, 24. https://doi.org/10.1186/s12984-016-0136-7

Google. Tensorflow federated. (2019) https://www.tensorflow.org/federated. Accessed 2019.

Han, X., Yu, H. and Gu, H. (2019). Visual Inspection with Federated Learning. In F. Karray, A. Campilho and A. Yu (Eds.), Image Analysis and Recognition (pp. 52–64). https://doi.org/10.1007/978-3-030-27272-2_5

Hao[a], M., Li, H., Luo, X., Xu, G., Yang, H. and Liu, S. (2019). Efficient and Privacy-enhanced Federated Learning for Industrial Artificial Intelligence. IEEE Transactions on Industrial Informatics, 1–1. https://doi.org/10.1109/TII.2019.2945367

Hao[b], M., Li, H., Xu, G., Liu, S. and Yang, H. (2019). Towards Efficient and Privacy-Preserving Federated Deep Learning. ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 1–6. https://doi.org/10.1109/ICC.2019.8761267

Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., … Ramage, D. (2019). Federated Learning for Mobile Keyboard Prediction. ArXiv:1811.03604 [Cs]. Retrieved from http://arxiv.org/abs/1811.03604

Hardy, S., Henecka, W., Ivey-Law, H., Nock, R., Patrini, G., Smith, G. and Thorne, B. (2017). Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. ArXiv:1711.10677 [Cs]. Retrieved from

http://arxiv.org/abs/1711.10677

Hasavari, S. and Song, Y. T. (2019). A Secure and Scalable Data Source for Emergency Medical Care using Blockchain Technology. 2019 IEEE 17th International Conference on Software Engineering Research, Management and Applications (SERA), 71–75. https://doi.org/10.1109/SERA.2019.8886792

Ho, Q., Cipar, J., Cui, H., Lee, S., Kim, J. K., Gibbons, P. B., … Xing, E. P. (2013). More Effective Distributed ML via a Stale Synchronous Parallel Parameter Server. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 26 (pp. 1223–1231). Retrieved from http://papers.nips.cc/paper/4894-more-effective-distributed-ml-via-a-stale-synchronous-parallel-parameter-server.pdf

Holcomb, S. D., Porter, W. K., Ault, S. V., Mao, G. and Wang, J. (2018). Overview on DeepMind and Its AlphaGo Zero AI. Proceedings of the 2018 International Conference on Big Data and Education - ICBDE '18, 67–71. https://doi.org/10.1145/3206157.3206174

Hu, B., Gao, Y., Liu, L. and Ma, H. (2018). Federated Region-Learning: An Edge Computing Based Framework for Urban Environment Sensing. 2018 IEEE Global Communications Conference (GLOBECOM), 1–7. https://doi.org/10.1109/GLOCOM.2018.8647649

Huang[a], L., Shea, A. L., Qian, H., Masurkar, A., Deng, H. and Liu, D. (2019). Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. Journal of Biomedical Informatics, 99, 103291. https://doi.org/10.1016/j.jbi.2019.103291

Huang[b], L., Yin, Y., Fu, Z., Zhang, S., Deng, H. and Liu, D. (2019). LoAdaBoost:Loss-Based AdaBoost Federated Machine Learning on medical Data. ArXiv:1811.12629 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1811.12629

Ilias, C. and Georgios, S. (2019). Machine Learning for All: A More Robust Federated Learning Framework. 544–551. Retrieved from http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0007571705440551

Ji, J., Chen, X., Wang, Q., Yu, L. and Li, P. (2019). Learning to Learn Gradient Aggregation by Gradient Descent. Proceedings of the Twenty-Eighth International Joint Conference on

Artificial Intelligence, 2614–2620. https://doi.org/10.24963/ijcai.2019/363

Jiang, P. and Ying, L. (2020). An Optimal Stopping Approach for Iterative Training in Federated Learning. 2020 54th Annual Conference on Information Sciences and Systems (CISS), 1–6. https://doi.org/10.1109/CISS48834.2020.1570616094

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., … Zhao, S. (2019). Advances and Open Problems in Federated Learning. ArXiv:1912.04977 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1912.04977

Kang, J., Xiong, Z., Niyato, D., Yu, H., Liang, Y.-C. and Kim, D. I. (2019). Incentive Design for Efficient Federated Learning in Mobile Networks: A Contract Theory Approach. 2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS), 1–5. https://doi.org/10.1109/VTS-APWCS.2019.8851649

Kennedy, Z. C., Stephenson, D. E., Christ, J. F., Pope, T. R., Arey, B. W., Barrett, C. A. and Warner, M. G. (2017). Enhanced anti-counterfeiting measures for additive manufacturing: Coupling lanthanide nanomaterial chemical signatures with blockchain technology. Journal of Materials Chemistry C, 5(37), 9570–9578. https://doi.org/10.1039/C7TC03348F

Khan, L. U., Tran, N. H., Pandey, S. R., Saad, W., Han, Z., Nguyen, M. N. H. and Hong, C. S. (2019). Federated Learning for Edge Networks: Resource Optimization and Incentive Mechanism. ArXiv:1911.05642 [Cs]. Retrieved from http://arxiv.org/abs/1911.05642

Kharitonov, E. (2019). Federated Online Learning to Rank with Evolution Strategies. WSDM 2019,249–257. https://doi.org/10.1145/3289600.3290968

Kim, H., Park, J., Bennis, M. and Kim, S.-L. (2019). Blockchained On-Device Federated Learning. ArXiv:1808.03949 [Cs, Math]. Retrieved from http://arxiv.org/abs/1808.03949

Kim, Y. J. and Hong, C. S. (2019). Blockchain-based Node-aware Dynamic Weighting Methods for Improving Federated Learning Performance. 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 1–4. https://doi.org/10.23919/APNOMS.2019.8893114

Kim, Y., Sun, J., Yu, H. and Jiang, X. (2017). Federated Tensor Factorization for Computational Phenotyping. Proceedings of the 23rd ACM SIGKDD International Conference on

Knowledge Discovery and Data Mining, 887–895. https://doi.org/10.1145/3097983.3098118

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T. and Bacon, D. (2017). Federated Learning: Strategies for Improving Communication Efficiency. ArXiv:1610.05492 [Cs]. Retrieved from http://arxiv.org/abs/1610.05492

Lee, J., Sun, J., Wang, F., Wang, S., Jun, C.-H. and Jiang, X. (2018). Privacy-Preserving Patient Similarity Learning in a Federated Environment: Development and Analysis. JMIR Medical Informatics, 6(2), e20. https://doi.org/10.2196/medinform.7744

Lee, S., Lacy, M. E., Jankowich, M., Correa, A. and Wu, W.-C. (2020). Association between obesity phenotypes of insulin resistance and risk of type 2 diabetes in African Americans: The Jackson Heart Study. Journal of Clinical & Translational Endocrinology, 19, 100210. https://doi.org/10.1016/j.jcte.2019.100210

Leroy, D., Coucke, A., Lavril, T., Gisselbrecht, T. and Dureau, J. (2019). Federated Learning for Keyword Spotting. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 6341–6345. https://doi.org/10.1109/ICASSP.2019.8683546

Li, H. and Han, T. (2019). An End-to-End Encrypted Neural Network for Gradient Updates Transmission in Federated Learning. 2019 Data Compression Conference (DCC), 589–589. https://doi.org/10.1109/DCC.2019.00101

Li, L., Wang, Y. and Lin, K. (2020). Preventive maintenance scheduling optimization based on opportunistic production-maintenance synchronization. Journal of Intelligent Manufacturing, 10.1007/s10845-020-01588-9.

Li, Q., Wen, Z. and He, B. (2020). Practical Federated Gradient Boosting Decision Trees. ICLR 2020-International Conference on Learning Representations.

Li, S., Cheng, Y., Liu, Y., Wang, W. and Chen, T. (2019). Abnormal Client Behavior Detection in Federated Learning. ArXiv:1910.09933 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1910.09933

Li, T., Sahu, A. K., Talwalkar, A. and Smith, V. (2020). Federated Learning: Challenges, Methods, and Future Directions. IEEE Signal Processing Magazine, 37(3), 50–60.

https://doi.org/10.1109/MSP.2020.2975749

Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A. and Smith, V. (2019). Federated Optimization in Heterogeneous Networks. ArXiv:1812.06127 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1812.06127

Li, T., Sanjabi, M., Beirami, A. and Smith, V. (2020). FAIR RESOURCE ALLOCATION IN FEDERATED LEARNING. ICLR 2020-International Conference on Learning Representations.

Li, X., Huang, K., Yang, W., Wang, S. and Zhang, Z. (2019, September 25). On the Convergence of FedAvg on Non-IID Data. Presented at the International Conference on Learning Representations. Retrieved from https://openreview.net/forum?id=HJxNAnVtDS

Lin, K. (2018). User Experience–Based Product Design for Smart Production to Empower Industry 4.0 in the Glass Recycling Circular Economy. Computers & Industrial Engineering, 125, 729–738.

Liu, B., Wang, L., Liu, M. and Xu, C.-Z. (2020). Federated Imitation Learning: A Novel Framework for Cloud Robotic Systems With Heterogeneous Sensor Data. IEEE Robotics and Automation Letters, 5(2), 3509–3516. https://doi.org/10.1109/LRA.2020.2976321

Liu, D., Dligach, D. and Miller, T. (2019). Two-stage Federated Phenotyping and Patient Representation Learning. Proceedings of the 18th BioNLP Workshop and Shared Task, 283–291. https://doi.org/10.18653/v1/W19-5030

Liu, F., Wu, X., Ge, S., Fan, W. and Zou, Y. (2020). Federated Learning for Vision-and-Language Grounding Problems.AAAI 2020-The Thirty-Fourth AAAI Conference on Artificial Intelligence,11572-11579.

Liu, W., Chen, L., Chen, Y. and Zhang, W. (2020). Accelerating Federated Learning via Momentum Gradient Descent. IEEE Transactions on Parallel and Distributed Systems, 31(8), 1754–1766. https://doi.org/10.1109/TPDS.2020.2975189

Liu, Y., Chen, T. and Yang, Q. (2018). Secure Federated Transfer Learning. ArXiv:1812.03337 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1812.03337

Liu, Z., Li, T., Smith, V. and Sekar, V. (2019). Enhancing the Privacy of Federated Learning with Sketching. ArXiv:1911.01812 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1911.01812

Ma, C., Li, J., Ding, M., Yang, H. H., Shu, F., Quek, T. Q. S. and Poor, H. V. (2019). On Safeguarding Privacy and Security in the Framework of Federated Learning. ArXiv:1909.06512 [Cs]. Retrieved from http://arxiv.org/abs/1909.06512

Majeed, U. and Hong, C. S. (2019). FLchain: Federated Learning via MEC-enabled Blockchain Network. 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 1–4. https://doi.org/10.23919/APNOMS.2019.8892848

McMahan, B., Moore, E., Ramage, D., Hampson, S. and Arcas, B. A. Y. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Artificial Intelligence and Statistics, 1273–1282. Retrieved from http://proceedings.mlr.press/v54/mcmahan17a.html

McMahan, H. B., Zhang, L., Ramage, D. and Talwar, K. (2018). LEARNING DIFFERENTIALLY PRIVATE RECURRENT LANGUAGE MODELS. ICLR 2018.

Mohammad, U. and Sorour, S. (2019). Adaptive Task Allocation for Asynchronous Federated Mobile Edge Learning. ArXiv:1905.01656 [Cs]. Retrieved from http://arxiv.org/abs/1905.01656

Mohri, M., Sivek, G. and Suresh, A. T. (2019). Agnostic Federated Learning. International Conference on Machine Learning, 4615–4625. Retrieved from http://proceedings.mlr.press/v97/mohri19a.html

Mowla, N. I., Tran, N. H., Doh, I. and Chae, K. (2020). Federated Learning-Based Cognitive Detection of Jamming Attack in Flying Ad-Hoc Network. IEEE Access, 8, 4338–4350. https://doi.org/10.1109/ACCESS.2019.2962873

Muñoz-González, L., Co, K. T. and Lupu, E. C. (2019). Byzantine-Robust Federated Machine Learning through Adaptive Model Averaging. ArXiv:1909.05125 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1909.05125

Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system(2008). https://bitcoin.org/bitcoin.pdf

Nasr, M., Shokri, R. and Houmansadr, A. (2019). Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-box Inference Attacks against Centralized and Federated Learning. 2019 IEEE Symposium on Security and Privacy (SP), 739–753. https://doi.org/10.1109/SP.2019.00065

Nishio, T. and Yonetani, R. (2019). Client Selection for Federated Learning with Heterogeneous

Resources in Mobile Edge. ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 1–7. https://doi.org/10.1109/ICC.2019.8761315

Nock, R., Hardy, S., Henecka, W., Ivey-Law, H., Patrini, G., Smith, G. and Thorne, B. (2018). Entity Resolution and Federated Learning get a Federated Resolution. ArXiv:1803.04035 [Cs]. Retrieved from http://arxiv.org/abs/1803.04035

Pan, S., Ni, X., Sun, J.-T., Yang, Q. and Chen, Z. (2010, January 1). Cross-Domain Sentiment Classification via Spectral Feature Alignment. 751–760. https://doi.org/10.1145/1772690.1772767

Pandey, S. R., Tran, N. H., Bennis, M., Tun, Y. K., Manzoor, A. and Hong, C. S. (2019). A Crowdsourcing Framework for On-Device Federated Learning. ArXiv:1911.01046 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1911.01046

Pfohl, S. R., Dai, A. M. and Heller, K. (2019). Federated and Differentially Private Learning for Electronic Health Records. ArXiv:1911.05861 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1911.05861

Qian, Y., Hu, L., Chen, J., Guan, X., Hassan, M. M. and Alelaiwi, A. (2019). Privacy-aware service placement for mobile edge computing via federated learning. Information Sciences, 505, 562–570. https://doi.org/10.1016/j.ins.2019.07.069

Ramaswamy, S., Mathews, R., Rao, K. and Beaufays, F. (2019). Federated Learning for Emoji Prediction in a Mobile Keyboard. ArXiv:1906.04329 [Cs]. Retrieved from http://arxiv.org/abs/1906.04329

Ramaswamy, S., Mathews, R., Rao, K. and Beaufays, F. (2019). Federated Learning for Emoji Prediction in a Mobile Keyboard. ArXiv:1906.04329 [Cs]. Retrieved from http://arxiv.org/abs/1906.04329

Saputra, Y. M., Hoang, D. T., Nguyen, D. N., Dutkiewicz, E., Mueck, M. D. and Srikanteswara, S. (2019). Energy Demand Prediction with Federated Learning for Electric Vehicle Networks. 2019 IEEE Global Communications Conference (GLOBECOM), 1–6. https://doi.org/10.1109/GLOBECOM38437.2019.9013587

Sarikaya, Y. and Ercetin, O. (2019). Motivating Workers in Federated Learning: A Stackelberg Game Perspective. ArXiv:1908.03092 [Cs]. Retrieved from

http://arxiv.org/abs/1908.03092

Sharma, S., Chaoping, X., Liu, Y. and Kang, Y. (2019). Secure and Efficient Federated Transfer Learning. ArXiv:1910.13271 [Cs]. Retrieved from http://arxiv.org/abs/1910.13271

Silva, S., Gutman, B. A., Romero, E., Thompson, P. M., Altmann, A. and Lorenzi, M. (2019). Federated Learning in Distributed Medical Databases: Meta-Analysis of Large-Scale Subcortical Brain Data. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 270–274. https://doi.org/10.1109/ISBI.2019.8759317

Smith, V., Chiang, C.-K., Sanjabi, M. and Talwalkar, A. S. (2017). Federated Multi-Task Learning. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett (Eds.), Advances in Neural Information Processing Systems 30 (pp. 4424–4434). Retrieved from http://papers.nips.cc/paper/7029-federated-multi-task-learning.pdf

Sozinov, K., Vlassov, V. and Girdzijauskas, S. (2018). Human Activity Recognition Using Federated Learning. 2018 IEEE Intl Conf on Parallel Distributed Processing with Applications, Ubiquitous Computing Communications, Big Data Cloud Computing, Social Computing Networking, Sustainable Computing Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom), 1103–1111. https://doi.org/10.1109/BDCloud.2018.00164

Sprague, M. R., Jalalirad, A., Scavuzzo, M., Capota, C., Neun, M., Do, L. and Kopp, M. (2019). Asynchronous Federated Learning for Geospatial Applications. In A. Monreale, C. Alzate, M. Kamp, Y. Krishnamurthy, D. Paurat, M. Sayed-Mouchaweh, … R. P. Ribeiro (Eds.), ECML PKDD 2018 Workshops (pp. 21–28). https://doi.org/10.1007/978-3-030-14880-5_2

Szegedi, G., Kiss, P. and Horváth, T. (2019). Evolutionary federated learning on EEG-data. ITAT 2019-Information technologies -- Applications and Theory, 71-78.

Tran, N. H., Bao, W., Zomaya, A., N.H, N. M. and Hong, C. S. (2019). Federated Learning over Wireless Networks: Optimization Model Design and Analysis. IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 1387–1395. https://doi.org/10.1109/INFOCOM.2019.8737464

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. and Zhou, Y. (2019). A Hybrid Approach to Privacy-Preserving Federated Learning. Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, 1–11. https://doi.org/10.1145/3338501.3357370

van Berlo, B., Saeed, A. and Ozcelebi, T. (2020). Towards federated unsupervised representation learning. Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking, 31–36. https://doi.org/10.1145/3378679.3394530

Wang, H., Yurochkin, M., Sun, Y., Papailiopoulos, D. and Khazaeni, Y. (2019, September 25). Federated Learning with Matched Averaging. Presented at the International Conference on Learning Representations. Retrieved from https://openreview.net/forum?id=BkluqlSFDS

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T. and Chan, K. (2019). Adaptive Federated Learning in Resource Constrained Edge Computing Systems. IEEE Journal on Selected Areas in Communications, 37(6), 1205–1221. https://doi.org/10.1109/JSAC.2019.2904348

Wang, X., Han, Y., Wang, C., Zhao, Q., Chen, X. and Chen, M. (2019). In-Edge AI: Intelligentizing Mobile Edge Computing, Caching and Communication by Federated Learning. IEEE Network, 33(5), 156–165. https://doi.org/10.1109/MNET.2019.1800286

Wang, Y., Tong, Y. and Shi, D. (2020). Federated Latent Dirichlet Allocation: A Local Differential Privacy Based Framework. AAAI 2020,6283-6290.

Wang, Z., Song, M., Zhang, Z., Song, Y., Wang, Q. and Qi, H. (2019). Beyond Inferring Class Representatives: User-Level Privacy Leakage From Federated Learning. IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 2512–2520. https://doi.org/10.1109/INFOCOM.2019.8737416

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

Webank[b]. FedAI ecosystem. (2019), https://cn.fedai.org/cases/.2019, Accessed 2019.

Wu, W., He, L., Lin, W., RuiMao and Jarvis, S. (2019). SAFA: A Semi-Asynchronous Protocol for Fast Federated Learning with Low Overhead. ArXiv:1910.01355 [Cs]. Retrieved from

http://arxiv.org/abs/1910.01355

Xiao, P., Cheng, S., Stankovic, V. and Vukobratovic, D. (2020). Averaging Is Probably Not the Optimum Way of Aggregating Parameters in Federated Learning. Entropy, 22(3), 314. https://doi.org/10.3390/e22030314

Xu, G., Li, H., Liu, S., Yang, K. and Lin, X. (2020). VerifyNet: Secure and Verifiable Federated Learning. IEEE Transactions on Information Forensics and Security, 15, 911–926. https://doi.org/10.1109/TIFS.2019.2929409

Xu, R., Baracaldo, N., Zhou, Y., Anwar, A. and Ludwig, H. (2019). HybridAlpha: An Efficient Approach for Privacy-Preserving Federated Learning. Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, 13–23. https://doi.org/10.1145/3338501.3357371

Yang, K., Jiang, T., Shi, Y. and Ding, Z. (2019). Federated Learning Based on Over-the-Air Computation. ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 1–6. https://doi.org/10.1109/ICC.2019.8761429

Yang, Q.[a], Liu, Y., Chen, T. and Tong, Y. (2019). Federated Machine Learning: Concept and Applications. ACM Transactions on Intelligent Systems and Technology, 10(2), 1–19. https://doi.org/10.1145/3298981

Yang, Q.[b], Liu, Y., Cheng, Y., Kang, Y., Chen, T. and Yu, H. (2019). DBA: DISTRIBUTED. Synthesis Lectures on Artificial Intelligence and Machine Learning, 13(3), 1–207. https://doi.org/10.2200/S00960ED2V01Y201910AIM043

Yang, T., Andrew, G., Eichner, H., Sun, H., Li, W., Kong, N., Ramage, D. and Beaufays, F. (2018). Applied Federated Learning: Improving Google Keyboard Query Suggestions. ArXiv:1812.02903 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1812.02903

Yang, W., Zhang, Y., Ye, K., Li, L. and Xu, C.-Z. (2019). FFD: A Federated Learning Based Method for Credit Card Fraud Detection. In K. Chen, S. Seshadri and L.-J. Zhang (Eds.), Big Data – BigData 2019 (pp. 18–32). https://doi.org/10.1007/978-3-030-23551-2_2

Yao, X., Huang, C. and Sun, L. (2018). Two-Stream Federated Learning: Reduce the Communication Costs. 2018 IEEE Visual Communications and Image Processing (VCIP), 1–4. https://doi.org/10.1109/VCIP.2018.8698609

Yao, X., Huang, T., Zhang, R.-X., Li, R. and Sun, L. (2019). Federated Learning with Unbiased Gradient Aggregation and Controllable Meta Updating. ArXiv:1910.08234 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1910.08234

Yu, T., Li, T., Sun, Y., Nanda, S., Smith, V., Sekar, V. and Seshan, S. (2020). Learning Context-Aware Policies from Multiple Smart Homes via Federated Multi-Task Learning. 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), 104–115. https://doi.org/10.1109/IoTDI49375.2020.00017

Yurochkin, M., Agarwal, M., Ghosh, S., Greenewald, K., Hoang, N. and Khazaeni, Y. (2019). Bayesian Nonparametric Federated Learning of Neural Networks. International Conference on Machine Learning, 7252–7261. Retrieved from http://proceedings.mlr.press/v97/yurochkin19a.html

Zhang, J., Chen, J., Wu, D., Chen, B. and Yu, S. (2019). Poisoning Attack in Federated Learning using Generative Adversarial Nets. 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), 374–380. https://doi.org/10.1109/TrustCom/BigDataSE.2019.00057

Zhang, W., Ouyang, W., Li, W. and Xu, D. (2018). Collaborative and Adversarial Network for Unsupervised Domain Adaptation. 3801–3809. Retrieved from http://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_Collaborative_and_Adversarial_CVPR_2018_paper.html

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D. and Chandra, V. (2018). Federated Learning with Non-IID Data. ArXiv:1806.00582 [Cs, Stat]. Retrieved from http://arxiv.org/abs/1806.00582

Zhu, H. and Jin, Y. (2019). Multi-Objective Evolutionary Federated Learning. IEEE Transactions on Neural Networks and Learning Systems, 1–13. https://doi.org/10.1109/TNNLS.2019.2919699

Table 1. Application in various fields

Researchers	Application domain	Studies	Pros	Constraint
Applications in mobile devices				
Chen, Mathews, Ouyang and Beaufays. 2019	Smart phone keyboard	learn out- of-vocabulary words	expanding the vocabulary of a keyboard without exporting sensitive text	Strongly relies on a learned probabilistic mode
Leroy et al. 2019	Smart phone voice assistant	Learn embedded wake word detector	using an adaptive averaging strategy in place of standard weighted model averaging	Do not show robustness to back- ground noise
Hard et al.2019	Smart phone keyboard	next-word prediction in a virtual keyboard	train an RNN model from scratch in the server and federated environments and achieve recall improvements	Still have high communication cost
Yang et al. 2018	Smart phone keyboard	improve virtual keyboard search suggestion quality	being easily trainable given the convexity of the error function by logistic regression model	impractical to train models with a large number of parameters
Ramaswamy et al. 2019	Smart phone keyboard	predict emoji from text typed on a keyboard	achieve better performance than a server trained model	client cache contents are different, and metrics cannot be compared across experiments
Wang et al. (2019)	Mobile edge computing	optimizing MEC, caching and communication	discussed the potential of inte-grating the Deep Reinforcement Learning and FL framework with the mobile edge system	how to distribute the huge computation load on heterogenous scenarios are still unexplored.
Qian et al. 2019	Mobile edge computing	Privacy-aware service placement for mobile edge computing	propose a privacy-aware ser- vice placement (PSP) scheme to meet users' service demands	Not able to be used for several edge clouds
Feng et al.2020	Mobile devices motion sensors	Privacy-preserving Human Mobility Prediction	Using group optimization strategy, reduce the performance degradation	only consider the basic mobility model for the simplicity
Sozinov et al. 2018	Smart devices motion sensors	Human Activity Recognition	identifies and rejects erroneous clients	producing models with slightly worse, accuracy compared to centralized models
Aïvodji et al. 2019	Smart home IOT	Design a sufficient secure federated smart home setting	combines FL with secure data aggregation	rather complex architecture to implement
Yu et al. 2020	Smart home IOT	learn users' behavior patterns	effectively detect physical hazards	Not flexible with mapping mechanism for diverse deployment
Liu et al. 2020	Robot network	robots imitation learning	increases imitation learning efficiency of local robots in cloud robotic systems	Need to further work on convergence justification of the fusion process
Applications in industrial engineering				
Hu et al. 2018	environment protection	environmental monitoring frame based on federated region learning	considers the regional characteristics during the distribution of training samples to improve the inference accuracy	Need to be extended to multi-layer structures instead of two-layer structure
Han et al. 2019	Image detection	provide manufactures with the service in automated defect inspection	solve the problem of lacking defective samples to detect defects	need quick model deployment to serve various industries
Liu et al. 2020	Image representation	obtain various types of image representations from different tasks	Be validated on three kinds of FL settings	more beneficial for the smaller dataset than the larger one in horizontal FL
Mowla et al. 2020	Unmanned Aerial Vehicles	malicious attacks detection in communication	enhance the model with a client group prioritization technique leveraging the	Need to improve the reliability of the global updates in this architecture.

		system of UAVs	Dempster–Shafer theory	
Saputra et al. 2019	Electrical vechicles	federated energy demand prediction	applied the clustering-based energy demand learning method for to further improve the prediction accuracy	Need to be more stable and flexible.
Yang et al. 2019	Financial field	detect credit card fraud	achieves an average of test AUC 10% higher than traditional method.	Should take more reliable measurements into account to protect the privacy
Wang et al. 2020	text mining	spam filtering and sentiment analysis	Using Random Response with Priori (RRP), which provides theoretical guarantees on both data privacy and model accuracy.	noise from our perturbing mechanism will slightly influence the overall performance
Applications in Health Care				
Brisimi et al. 2018	Predict future hospitalizations for patients	Cluster Primal Dual Splitting algorithm	Yield classifiers using relatively few features	Need more iterations for convergence
Silva et al.2019	MRI Analysis	Provide federated analysis framework compatible with the standard ENIGMA pipelines	Deal with variability of high dimensional features efficiently	Only tested in limited dataset
Liu et al.2019	Extraction of clinical notes	Two-stage federated NLP method	Adding pre-processing step to improve accuracy	Not suitable for small questionable cases
Gao et al.2019	EEG Classification	Design a hierarchical heterogeneous horizontal FL framework	The first EEG classifier over heterogeneous EEG data	Only work on 3 different datasets
Li and Liu 2019	Predict mortality and hospital stay time	Introduce community-based FL and evaluate it on non-iid icu EMRs	converged to higher predictive accuracy in less communication rounds than the baseline FL model	Model parameters of community will lead to extra communication overhead
Pfohl et al.2019	Clinical prediction	Establish efficacy of FL over centralized and local learning	Perform FL in a differentially private manner	Underestimate of privacy cost
Huang[b] et al.2019	Mortality prediction over drug utilization data	Adaptive boosting method	Alleviate non-iid by introducing data-sharing technology	Training on iid data outperform non-iid data
Kim et al.2017	Analysis of computational phenotypes	Federated tensor factorization for privacy preserving computational phenotyping	Summarized information does not disclose the patient data	Only Accurate with small or skewly distributed data
Lee et al. 2018	Similar patient matching	Federated patient hashing framework	Avoid security attack from reverse engineering	Inevitable computational complexity

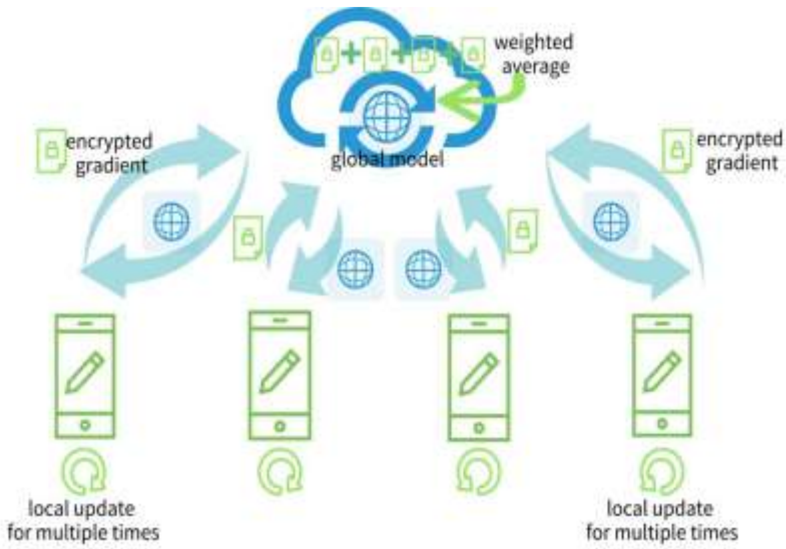

Fig. 1. Illustration of FL framework proposed by Google

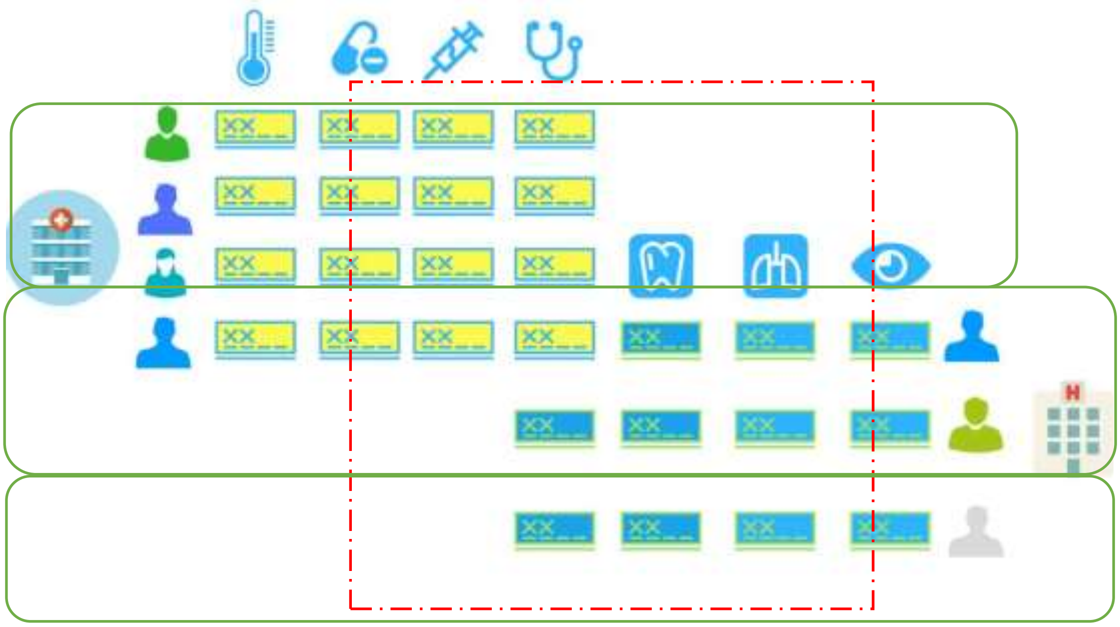
Fig. 2. An application sample of Horizontal FL

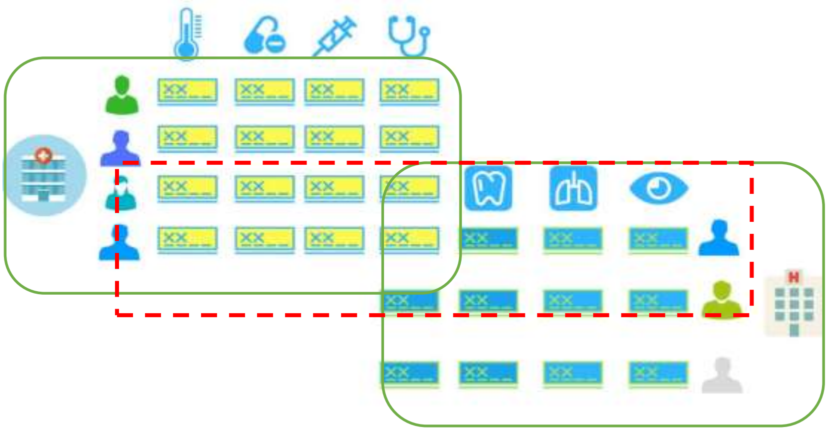

Fig. 3. An application sample of Vertical FL

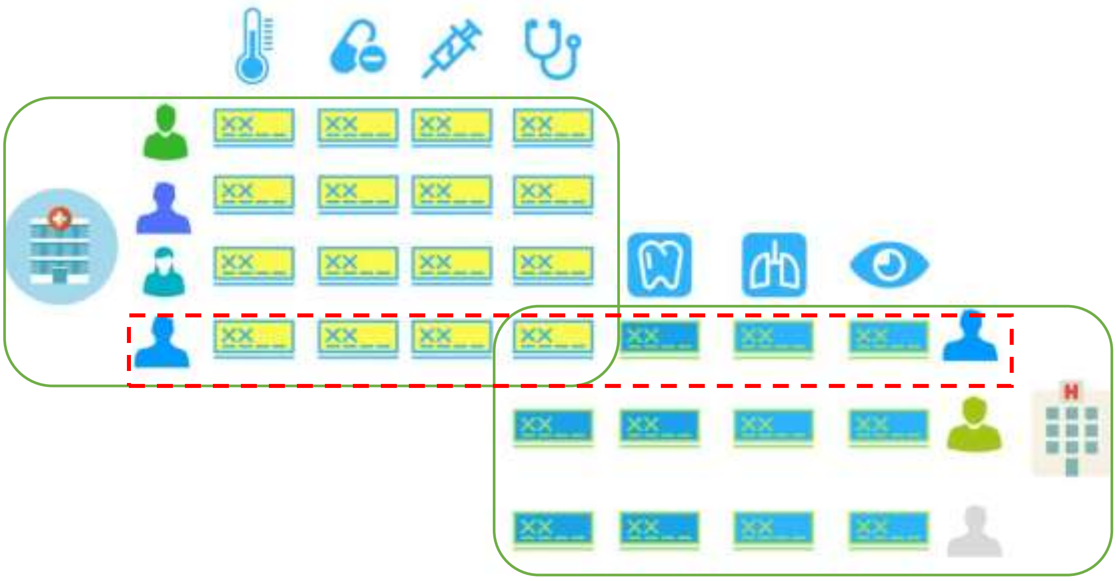

Fig. 4. An application sample of federated transfer learn

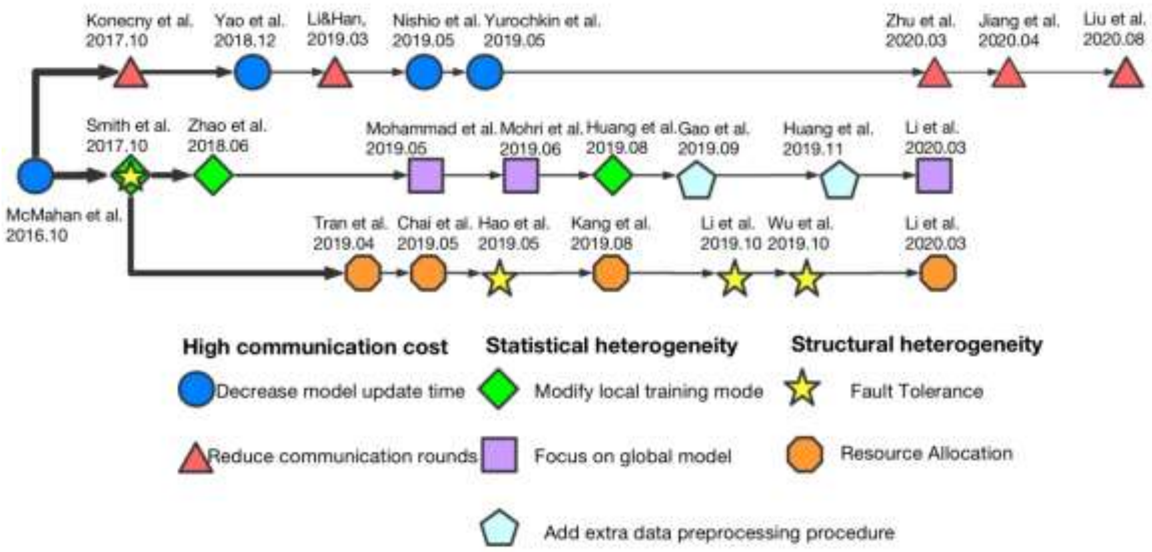

Fig. 5. Optimization path to overcome three challenges in FL

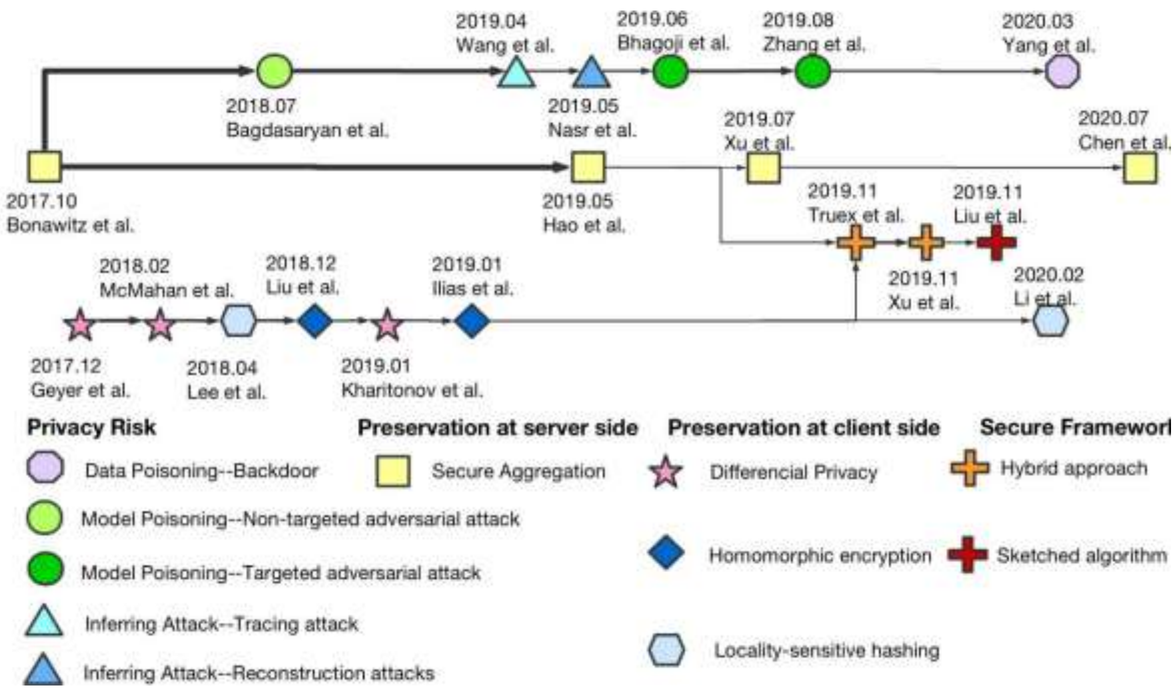

Fig. 6. The evolution of privacy attack and enhancement in FL