INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 8, Issue 1 - V8I1-1463)
Available online at: https://www.ijariit.com

A survey on asynchronous distributed Federated Learning framework

Ashley Olebogeng Makgetho
ashleymakgetho@gmail.com
University of Science and Technology, Beijing, China

Huang Qiming
qmhuangcn@163.com
University of Science and Technology, Beijing, China

Ousman Manjang
manjangousman25@gmail.com
University of Science and Technology, Beijing, China

ABSTRACT

Through the hasty growth of data generated by intelligent IoT devices, Federated learning (FL) seems to be a promising technique which provides distributed Machine Learning (ML) amenities at the same time protecting data privacy. FL is the novel form of Artificial intelligence (AI) which builds on decentralized data set up and carry out training which brings learning to devices. It's mostly used in instances that involve security and privacy as the main concerns and empowers implementers to build secure learning environments. The federated averaging (FedAvg) is one the most used optimization algorithms that train models with a synchronized protocol. However, the algorithm is not realistic enough and communication efficiency issues tend to arise. The amount and distribution of collected data has a different training process because of varying sample sizes of devices. This paper carries out an in-depth review of FL and its asynchronous learning previous researches. Lastly, authors propose a privacy-preserving asynchronous FL framework for distributed healthcare care data that improves the model accuracy to health information. Although the framework is still being implemented it aims at guaranteeing improved communication amongst healthcare industry participants such as hospitals, clinics, laboratories, pharmaceuticals and many more facilities.

Keywords: *Artificial intelligence (AI), Federated averaging (FedAvg), Federated learning (FL), Machine learning (ML)*

1. INTRODUCTION

We live in an era where a lot of machine learning algorithms are used in various projects and collecting data is also an important aspect in order to have a good model performance. There are already existing applications such as google keyboard and google maps which are greatly used on daily basis [2] [3]. Through the high usage of internet of things (IOT) and social media there has been an increase of data generated in edge networks [8]. There have been regulatory restrictions such as GDPR, HIPAA, GLBA and CCPA which were formed to address issues of information privacy [4]. In 2018 the EU issued General Data Protection Regulation (GDPR) which means people's awareness of data privacy increases [7]. Systems entail various levels of information privacy dependent on use case and data leakage is one of the issues faced, because it can occur during storage, transmission or even at the sharing phase [5]. There has been existing works that try to tackle this issue include of K-anonymity and I-diversity however some works undertake that the malicious attackers have little knowledge which may not always be the case. The authors of McMahan et al. [1] suggested a distributed machine learning system that can assure data privacy. FL is a machine learning technique also known as collaborative learning which trains algorithm across decentralized networks that have devices with local data samples without sharing their actual datasets [6]. An example of FL approach is whereby various devices are interconnected and can communicate with an aggregator in order to perform neural network training. In FL, a global neural network is stored on a central server, and the training process is carried out using data stored locally among many nodes. This technique differs with traditional centralized machine learning approaches in which local datasets are uploaded to a server. It is gaining its popularity ever since it was introduced back in 2016 by Google. Federated learning tends to solve problems encountered in traditional machine learning models such as those related to data security and privacy. This technique resolves concerns with scalability, model accuracy, and training time, allowing it to be applied to the development of improved models that address complex Artificial Intelligence (AI) issues. Studies [11] [12] define FL as a system that parties collaborate to train models without having

to exchange raw data and the output is a ML model for each party involved in the set up. The encrypted models are used by ML during the process and at the same time the involved party's information is kept secure thus preserving privacy. Based on previous research, this technique may be classified into two types: synchronous and asynchronous federated learning [9] as depicted in figure 1. In synchronous FL a server chooses a part of nodes arbitrarily for iteration in every round When it comes to federated learning synchronization, McMahan et al. presented the FedAvg algorithm [10]. The technique trains models with only a few rounds of communication, while McMahan et al. contributed to the decrease of local update loading overhead by developing the concept of structured and sketched updates. However, in the Asynchronous FL set up a user node uploads and updates as long as the local iteration has been completed [9]. A global aggregation is performed often when local update is received and result calculations are returned to the node. The data required for training in AI model is huge therefore this has prompted distributed ML to arise.

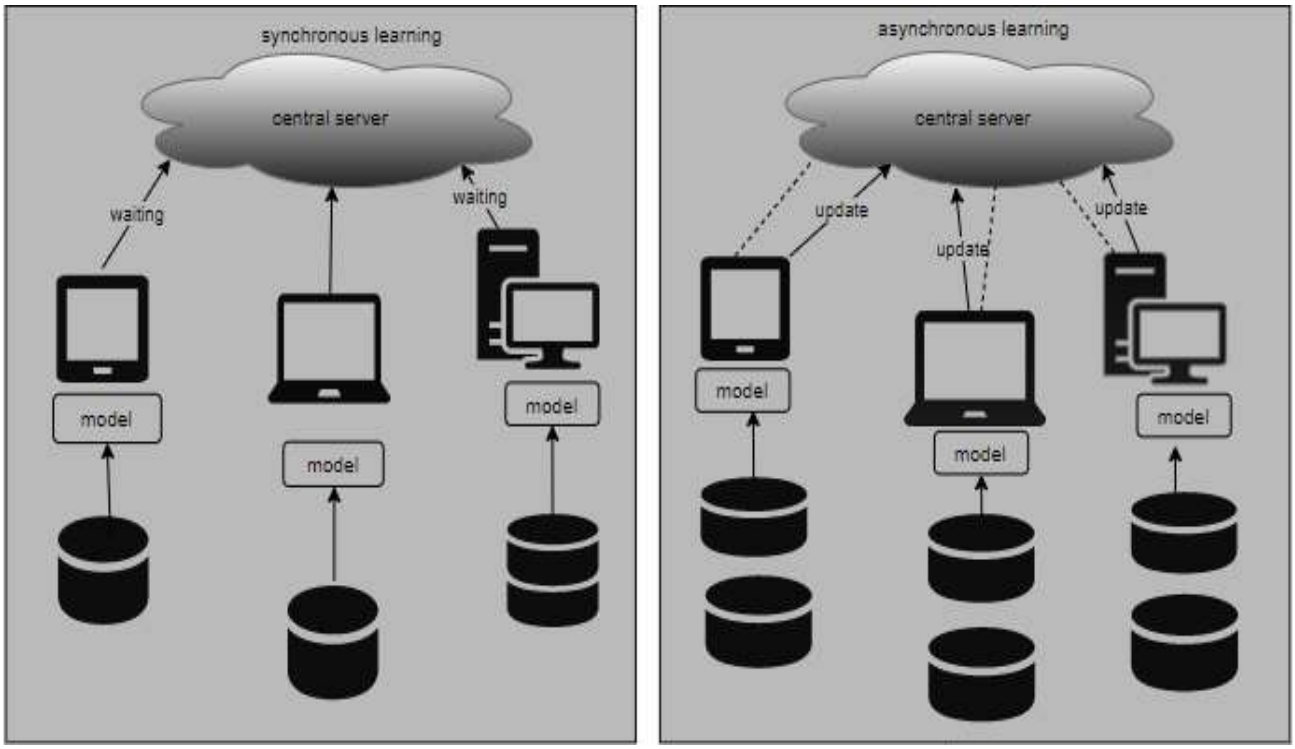

Figure -1: comparison of synchronous and asynchronous learning environments.

1.1 Contributions

The purpose of this study entails a description of FL and its already existing applications with the aim of contributing towards the better understanding of this. machine learning technique. Furthermore, our study will take a deep look into Asynchronous learning which is one of the categories of FL and look at how it helps in the process of information privacy. Even though there have been tremendous studies aimed at FL our study will also contribute towards a better understanding of the topic more especially since Asynchronous federated learning has gained popularity. Although there are many researches directed towards distributed FL there hasn't been sufficient progress directed towards the tackling the issue of accuracy therefore our work aims at;

- providing a complete overview of FL looking into some of its technologies to show case its usage in various scenarios in order to help understand ways of applying the technique in different industries with the aim of privacy-preserving data.
- offering a wide-ranging overview of Asynchronous federated learning and identifying challenges encountered in the already existing works.
- Proposing an asynchronous FL framework that improves the model accuracy

1.2 Paper structure

The following is a breakdown of our paper's structure: section 1 introduces the and authors highlight how information privacy continues to shift in Machine learning towards federated learning over the years. Section 2 examines federated learning classes and application areas. Additionally, section 3 provides an insight into existing asynchronous FL frameworks that other authors have come up with. Section 4 introduces the features of our proposed asynchronous FL framework on distributed healthcare data. Lastly section 5 is the paper's summary.

2. FEDERATED LEARNING

The setting of the FL system may include evaluation metric for testing accuracy and model performance constraints. Given an assumption that we have ***i*** Parties $\mathbf{f_1 \ldots f_i}$ with data $\mathbf{d_1 \ldots d_i}$ and a ML model $\mathbf{M_t}$ which is trained through merging parties information as $\mathbf{d_t = d_1 \cup \ldots \cup d_i}$ [13]. The data owners who are the parties in this instance pass through the training process of the model represented as $\mathbf{M_f}$ and other parties $\mathbf{f_h}$ hide their data $\mathbf{d_h}$ from being viewed. Therefore, we can say FL gives them privileges to control their data access. Singh et al. suggest that FL system has to be as accurate when compared to the traditional method whereby if the accuracy of the model is specified as $\mathbf{M_{f\ and}}$ as $\mathbf{V_f}$ **it** has to be equivalent to the performance of $\mathbf{M_t}$, $\mathbf{V_t}$. The equation below suggests that δ represents accuracy loss that is non-negative number [13].

$$| V_f - V_t | < \delta$$

2.1 Data partition FL

This machine learning scheme is suitable in building models of data to be shared across innumerable domains and the fact that it can keep data private during the training process means that there will be leverage of smart features of ML by numerous domains [14]. It removes domain limits from the user's information and makes smart ML help available through collaborations and boosting the benefits of information across domains with similar and dissimilar interests. The statistical data resulting from user's data is useful for several applications of domains. In a given instance whereby a user is online, their activity is captured to make statistical data and such use cases are often used by applications in domains. Based on data partition, we may classify FL as Horizontal FL (HFL), Vertical FL (VFL), and Federated Transfer Learning (FTL).

2.1.1 Horizontal federated learning:

HFL can be useful in an instance where information has feature space that is identical though sample space may be separate. Assuming that we are in a healthcare setup, where organizations such pharmacies, hospitals etc. are involved, based on their regional operations there's the probability of their users varying and slight ordinary user set. Nevertheless, the feature space becomes the same since the nature of the healthcare is similar too. [16] describes a collaborative deep learning strategy in which numerous users learn an accurate neural-network model without sharing input datasets. Google in 2017 anticipated the usage of HFL framework towards Android phone mode updates in which phones locally update parameters and then upload them in the android system cloud [17]. [18] proposed a safe aggregation system that uses secure multiparty computation (SMC) to carry out an evaluation of model parameter updates that is secured by user's devices to ensure the privacy of aggregated user updates. On the other hand, instead of using SMC researchers opted to implement homomorphic encryption ensuring privacy preservation [18]. In 2019 Kim et al. came up with a system for mobile devices to update the local learning model in a blockchain based network setting. As figure 2 depicts, in stage 1 clients from a domain carry out training the global model. In stage 2, the trained global model is downloaded and each client utilizes it and training repeats again.

2.1.2 Vertical federated learning

VFL is an FL method which shares information amongst domains that are not related and is used in training global model [14]. Users that often use this approach tend to prefer having a third-party establishment that is responsible for providing encryption logic towards the shared data even though it's not obligatory as researchers in [20] implemented VFL without one. A scenario where VFL can be implemented is where a Company A and Company B that's responsible for digital marketing is involved. Company A would request to know utmost procured items from online domains in order to develop their model. In this approach various features are combined for privacy-preservation manner in order to shape the model considering data from both parties collaboratively. Hence why the party involved is considered to have identical identity and status. For the effective data partition Statistical analysis, gradient descent, safe linear regression and data mining are effective [21-26]. As figure 3 depicts in stage 1 clients involve ought to contribute towards the training of the global model through sharing of encrypted local model updates. In stage 2 the trained global model is then downloaded and used by each client nod. The training rounds will then continue after the second stage has been completed.

2.1.3 Federated transfer learning

In a scenario where sample and feature space differ FTL is implemented. Considering the same scenario given above about a company A and Company B we can assume the possibility of having both of them in 2 different countries. In this instance geographical location is an issue at hand and their group of users have small intersection and feature space has small overlap because the type of business varies therefore, we can implement FTL. Research works such as [27-31] demonstrate the usage of FTL with much recent works like [32][33][34] that combine both machine learning algorithms and FL. In figure 4 the first stage includes clients training a global model on the cloud using cryptographic methods, then follows the second stage that's all about pre-trained cloud model and knowledge transfer learning being applied so as to get modified models.

2.2 Data availability

On the basis of data availability and nodes we can classify FL as Cross-silo FL and Cross-device FL. Both of these classes are explained below separately.

2.2.1 Cross-silo FL

Clients tend to be at a smaller scale, with numbers ranging from 2-200 devices indexed and ready for training. Horizontal and vertical learning are ways data for training can be classified. Henceforth, communication and computation will be the most important aspects of machine learning. In comparison to cross-device FL, cross-silo FL is more versatile and is more likely to be employed in companies for ML model training [14].in vertical and transfer learning, encryption is ideal for restriction purposes. A framework by [36] built on FATE [35] indicates the implementation of cross-silo combined with HE and also [37] proposed gradient quantization that relies on batch encrypt algorithm with the aim of reducing communication and computation issues.

2.2.2 Cross-device FL

cross-device FL involves a large number of users in a comparable domain and with similar interests. Due to large number of users, transaction logs become difficult to keep, especially when many of them connect over unreliable networks, where training rounds are selected at random. As a result, this strategy would be extremely useful for IoT applications [38].

2.3 Application areas

FL continues to transform into various industries such as healthcare, finances and many others as time goes on. There are cases where FL can't be used to directly aggregate data for training the model. There are some instances where the data of the device is heterogenous and often traditional ML models can't work in such scenarios. Hence FL's applications in numerous industries can beneficial since it has found its way into edge computing, networking, robotics, cybersecurity, wireless communication and many others. Below are a few application areas of FL:

2.3.1 Google keyboard

FL has improved Gboard's query recommendations based on [39]. FL was used to power the functionality of the Google virtual keyboard. Users must meet a number of constraints in order to carry out FL processes, including environmental requirements, equipment standards, and language restrictions. The server also establishes a number of constraints, such as the target number of participants, the minimum user number required to run the round, the frequency with which training is conducted, the time limit for receiving user updates, and the percentage of users who must report back to commit a round. Training examples show that performance is high in the evening and that loss tends to occur. According to annotations, there is typically a little variation between the expected and actual query click-through rate during live deployment. Researchers in the work of [40] develop an advanced neural network model that outperforms centralized data model training. Furthermore, Ramaswamy et al. show the ability of a recurrent neural network to predict emojis from previous words on Google keyboard via federated learning.

2.3.2 Healthcare informatics

Medical organizations have a vast volume of data in their electronic health records (EHR), and it is frequently necessary to train a powerful medical model, however this is a difficult process because to the sensitive nature of the data [42]. The healthcare industry has transitioned from spread locations to a central database for analytical reasons, which has resulted in a slew of challenges, including rigorous rules, sensitive data that must be managed carefully during transfers, and so on. To solve these challenges, D. Liu et al. [43] suggested FADL, a FL-based technique in which the first layer of the neural network model is trained using data from all parties involved and the other levels model the data. Each cluster learns and shares a precise machine learning model, which enhances performance and efficiency. Through an intrusion detection system based on FL, Schneble et al. [45] have also contributed to the medical cyber physical systems business. Private information from the patient's devices, such as heart rate and others, is trained locally for global model advancements that can be utilized by other or the same patients to detect harmful activity. To create a high-performance model, similar patients with similar characteristics are clustered, and each cluster creates its own local and global model. Another federated transfer learning-based strategy has been proposed. Silva et al. [47] established another framework based on FL for biological data analysis, and this research work studies brain alterations from diverse disorders, including neurological ailments. With the deployment of FL, the authors of [48] and [49] established a medical imaging prediction framework for brain cancer segmentation, and their solutions facilitate multiple institution cooperation by sharing their locally computed models. Furthermore, differential privacy was used as a strategy for preventing information breaches in [49].

2.3.3 Wireless communication

Although previous methods have proven inappropriate for sophisticated ML communication, FL has a significant role in wireless communication [70][71][72]. Zhao et al [73] conducted research on the use of FL in the fog radio access network (F-RAN) with the goal of lowering data offloading and model training costs in edge computing. Other research, such as Khan et al [74], have been proposed to address issues in the deployment of FL at the edge by incentivizing interaction between participants and servers. Wang et al. [75] developed a deep reinforcement learning over a Federated architecture that increases computation in the edge network to provide better quality services such as quick content delivery. FL has also been included into other recently suggested efforts by vehicular ad hoc networks, such as [76] [77].

2.3.4 Data security mechanism

FL plays a significant part in data privacy and was developed with the goal of training machine learning models to keep personal data safe and secure. Several works based on FL with the goal of detecting assaults have been proposed. Another FL-based technique based on blockchain technology was proposed to handle the issue of data security and preventative maintenance of cloud-based intrusion detection systems [78]. Attack mitigation for smart city infrastructure can be done with blockchain FL [79]. Other studies to explore are the use of FL for privacy preservation in automotive cyber physical systems [80] and the use of FL in securing 5G heterogeneous networks with end-nodes and edge-nodes with attack detection capabilities [81]. FL has been applied in IoT systems with the goal of reducing IoT device vulnerabilities. Edge computing was designed to support and offload activities faced by IoT edge nodes because there are large computation loads on devices. Existing intrusion detection systems are ineffective when it comes to detecting corrupted IoT devices. Despite this, another study [82] offered an autonomous self-learning distributed system for detecting devices that had been compromised.

3. RELATED WORKS

In this section we shall look at the review of numerous articles from journals, conference proceedings, and online documents on several privacy preserving approaches presented based on asynchronous distributed federated learning that several authors have come up with.

FL aggregators are edge servers that work with independent nodes as local learners to train the same machine learning model utilizing locally accessible data. The aggregator receives and aggregates weight vectors from local learners in each round of training. The next stage is to use corroborating data to test the model's performance. In McMahan et al. FL.'s framework technique, a subset of nodes can be chosen at random to perform global aggregation in each round [1]. Further research has been conducted based on the idea that IoT devices are identical and can communicate in a static network. As a result, in IoT setups, synchronous aggregation is considered to produce inadequate weight updates [50] [51]. Before the aggregated weight vector can

be aggregated, the aggregator at the edge server must wait for weight vectors from all nodes; otherwise, the slowest node becomes the bottleneck, causing the training process to be delayed. On the actual world, IoT devices are diverse and have heterogeneous processing resources, and communication in wireless networks is dynamic and guaranteed. Wang et al [51] looked into synchronous distributed machine learning and devised an approach for determining a flexible global aggregation frequency based on available resources. Similarly, in [53], the authors presented a distributed strategy for improving performance by sharing local training data. Other applications, on the other hand, are concerned about data privacy, and IoT devices are unlikely to share local data with outsiders. Some synchronous learning assumptions state that all nodes must communicate weight vectors for each training round, which can cause network congestion Additionally, waiting for the slowest nodes in schemes may cause a considerable delay. Unlike previous studies that assume synchronous updates, the asynchronous node selection technique makes use of available resources for computation on each node, allowing the learning process to be completed quickly.

There have also been some efforts by [54] [55] to improve asynchronous FL. Nishio et al presented an asynchronous node selection technique for global aggregation in [55], with the goal of resolving network communication problems. Whereas Lian et al [54] created an algorithm that speeds up convergence in ML task without the need for a central parameter sever. Lian et al. [57] also designed an algorithm in 2018 that is robust in heterogenous environment that is asynchronous decentralized stochastic gradient decent (AD-PSGD). There have been other schemes of stochastic optimization methods that try to solve computational costs. Unfortunately, existing ones are often synchronous and don't make efficient use of computational resources when there's load imbalance. Parallel asynchronous particle swarm optimization algorithm (PSO) was proposed in [59] to enhance computational efficiency. However, meta heurists like PSO one for instance don't necessarily guarantee ideal solution. Since end-devices have limited communication bandwidth, an upgraded FL approach was proposed [58] that proposed an asynchronous learning strategy on clients and temporally weighted aggregation of local models on the server side to decrease client-server communication. Different layers of deep neural networks (DNNs) are categorized into shallow and deep layers in their work, with the deep layers' parameters being changed less frequently than the shallow ones. On the server side, a temporary weighted aggregation strategy was added to make advantage of previously trained models in order to improve model accuracy and convergence. However, the approach is not feasible enough because clients need to advance their local models in order to have better learning performance and reduced communication costs.

Dijk et al. [60] suggested a paradigm that incorporates asynchronous federated learning as well as differential privacy. According to the suggested asynchronous federated learning, waiting times are eliminated, and overall network transmission is reduced. Their method tolerates biased input, which is a typical practice, and asynchronous FL with differential privacy does not appear to impair their model's accuracy more while increasing sample numbers. The framework records a decrease in communication between clients and servers, as well as a decrease in aggregated DP noise. Nevertheless, their framework does not absolutely guarantee security and privacy aggregation. Chen et al. [52] offered an existing technique based on asynchronous FL in order to address the issues described. They created an asynchronous FL model by constructing a 0-1 knapsack optimization problem that may represent heterogeneous computing resources in an unstable network in their research. One of the scheme's disadvantages is that numerous factors must be addressed, such as IoT device mobility and battery capacity. During movement, the wireless connection may be lost, causing instability. As a result, prospective service migration buffers for holding weight vectors must be considered in this research. Furthermore, their work must expand its parameters in terms of security and privacy issues in order to mitigate external adversaries by the application of proper distance measurements that verify received updates. A technique for FL of deep networks that is based on iterative model averaging was proposed [61]. In their work FedAvg trains high quality models through the usage of relatively partial rounds of communication and the model is to be applied in practical settings. Nonetheless the assumptions made are not accurate in heterogenous devices. Due to varying edge device sampling rates, the volume and distribution of recorded data varies depending on the training process. Different latency and system configurations, such as processing speed, exist on edge devices. Furthermore, difficulties at edge devices may result in a lack of contribution from other edge devices to the FL model. Therefore, it is not the best although it's the leading optimization method. Nevertheless, Chen et al [62] proposed an approach that tackles issues in [61]. They introduced an asynchronous online federated learning (ASO-Fed) system in which edge devices execute online learning with continuous local data flow and a central server collects model parameters from clients in their work. The work tackled the issue of computational loads in heterogenous edge devices including those that may lose connection. Nevertheless, the model is not enough due to communication bottleneck issues.

Lu et al. [63] suggested a strategy for resource sharing in vehicle networks dubbed differentially private asynchronous federated learning scheme. To make their FL system more secure and robust, they combined it with local differential privacy to protect the privacy of the updated local models. They also used a random distributed update approach to eliminate security vulnerabilities associated with centralized curators, as well as updates verification and weighted aggregation to promote convergence. However, in a real-world scenario, their scheme is insecure. A privacy-preserving asynchronous FL mechanism for edge network computing (PAFLM) has been presented [64], which allows several edge nodes to obtain an improved efficient FL without needing to reveal sensitive information. In contrast to typical distributed learning, their approach anticipated a compression technique between nodes and the parameter server during training without affecting accuracy. The nodes work asynchronously, without needing to wait for other nodes or synchronize the learning process. However, their model is insufficient, therefore they can propose a new attenuation function for their task. Authors in [58] and [64] express similar unhappiness with their work because, to their knowledge, their analyses do not take into account the influence of data imbalance on optimization. A research of data imbalance and asynchronous aggregation algorithm on the FL system was proposed by Diwangkara et al. [65]. They looked at how the asynchronous aggregation approach affected convergence time and test results. They also created an asynchronous aggregation technique by adapting a stale synchronous parallel algorithm. Asynchronous aggregation in FL enhances optimization performance, but it can also degrade it, and it has little effect on modest differences in server-wise update frequency and a relatively balanced data distribution. Their work also ought to be tested with more hyperparameter for effectiveness. To address

issues in synchronous communication systems, Hao et al. [9] suggested a semi-asynchronous FL framework in which the data expansion method is employed to successfully minimize stragglers in both synchronous and asynchronous communication models. Their design enables nodes in the loop to obtain updated data without having to wait for other nodes. Their work, however, does not entirely address the issue of coordinated FL communication. It is inefficient in the area of reducing communication overhead, which prevents a large quantity of communication in an asynchronous operation.

Despite the fact that asynchronous FL local models are uploaded immediately once the update is completed, they are still vulnerable to poison attacks. Furthermore, because of the randomness of uploading time in asynchronous FL, scheduling device uploading time is a difficult challenge. Authors in [66] presented an asynchronous FL algorithm and studied its convergence rate when distributed amongst various devices that have data constraints which are relative to training the same model in a device. However, it was not feasible enough and Cong et al. [67] improved their framework and proposed an asynchronous FL optimization method that improves flexibility and scalability. For strong convex and non-convex related issues, their work has globally optimal near-linear convergence. Nevertheless, their work did not consider the robustness of the system. The are other studies that have been proposed such as that by Feng et al. [68]. To assure the security and efficiency of FL, they developed a blockchain-based asynchronous federated learning framework in their study. They evaluated the participating rank and proportion of the local model learned in the blockchain-based asynchronous FL of the devices using a unique entropy weight approach. However, blockchain technology is not always effective in delivering a tamper-proof and secure FL system that can record device scores and models on a distributed ledger. Lu et al. [69] proposed another framework based on blockchain and asynchronous FL. Their system architecture is based on FL in order to reduce transmission burden and meet provider privacy concerns. They created a hybrid based blockchain solution that combines permissioned blockchain with local directed acyclic graph to increase the security and dependability of model parameters (DAG). Additionally, they implemented asynchronous FL framework through the adoption of deep reinforcement learning (DRL) for the sole purpose of selection so as to enhance efficiency. this approach shows impressive results which confirm its effectiveness when compared to already existing ones.

Table-1: Summary of literature review

Author	Research work	Strength	weakness
Dijk et al.	Asynchronous Federated Learning with Reduced Number of Rounds and with Differential Privacy from Less Aggregated Gaussian Noise	The scheme eradicates waiting times and decreases network communication which saves costs as well.	Does not give assurance in terms of security and privacy.
Chen et al.	Towards asynchronous federated learning for heterogeneous edge-powered internet of things	Improved training efficiency for heterogenous IoT devices that are often in unstable communication networks.	Plentiful issues ought to be addressed such as device mobility and battery capacity. Lack of security and privacy mechanisms that ease external adversaries.
McMahan et al.	Communication-efficient learning of deep networks from decentralized data	Aims at model training decoupling without essentially gaining direct access towards the training data.	The model is not realistically accurate in heterogenous environments because of differing sampling rates and unavailable devices for instance.
Chen et al.	Asynchronous Online Federated Learning for Edge Devices with Non-IID Data	Unlike synchronized FL frameworks the scheme enables a wait-free computation and communication.	The scheme's model is however not enough to tackle communication bottleneck issues.
Lu et al.	Differentially Private Asynchronous Federated Learning for Mobile Edge Computing in Urban Informatics	Secure resource sharing over vehicular networks through an incorporation of differential privacy into gradient descent training process.	The framework is limited and in real-life scenario it is not secure enough.
Lu et al.	Privacy-Preserving Asynchronous Federated Learning Mechanism for Edge Network Computing	Enhanced privacy for sensitive information sharing.	Insufficient model which needs a new attenuation function.
Chen et al.	Communication-Efficient Federated Deep Learning with Layer-wise Asynchronous Model Update and Temporally Weighted Aggregation	Reduced communication costs and improved performance.	The approach doesn't take account for data imbalance.
Diwangkara et al.	Study of Data Imbalance and Asynchronous Aggregation	Improved convergence time through asynchronous	Inadequate testing hyperparameters.

	Algorithm on Federated Learning System	aggregation algorithm.	
Hao et al.	Time Efficient Federated Learning with Semi-asynchronous Communication	Improved convergence rate by getting updated data immediately.	Its efficient in reducing communication overhead.
Sprague et al.	Asynchronous Federated Learning for Geospatial Applications	Their work laid a ground for deploying huge scale FL.	It is not feasible enough.
Cong et al.	Asynchronous Federated Optimization	Improved flexibility and scalability	The scheme doesn't consider the system robustness
Feng et al.	Blockchain-based Asynchronous Federated Learning for Internet of Things	Reduced resource consumption and increased precision.	Blockchain doesn't always provide tamper-proof and security.
Lu et al.	Blockchain Empowered Asynchronous Federated Learning for Secure Data Sharing in Internet of Vehicles	Enhanced model efficiency	The framework needs to implement more secure privacy techniques

4. FEATURES OF THE PROPOSED FRAMEWORK

FL is good machine learning paradigm however it still faces challenges of model accuracy therefore more factors should be taken into consideration. In case where by the participants are not available during training that affects the distribution process. There have been existing studies such as [83] that relied on multiparty cryptographic scheme that assumed participants are always available. On another study by [84] they assume that participants can become unresponsive and the available ones can carry out partial execution though threshold encryption however that is not secure enough. Therefore, to solve such a problem we propose an asynchronous federated learning framework that will improve accuracy and preserve privacy of distributed data. The system will rely on the healthcare data which is one of the industries that deal with sensitive information. This will greatly benefit the healthcare industry through promotion of big data sharing that doesn't include complicated deterrents. The figure below portrays the asynchronous learning framework where by the server aggregates after it has received updates from the involved participants and then perform the feature learning. In this framework the central server has to update the central model after getting updates from the either of the available participants although others may not be available. The participants in question keep their copies of the central model in the memory which means that there's a possibility of copies varying from one participant to another. The server will then start the next iteration and distribution of new central model to participants that are ready after feature learning on the parameters in order to extract cross feature depiction.

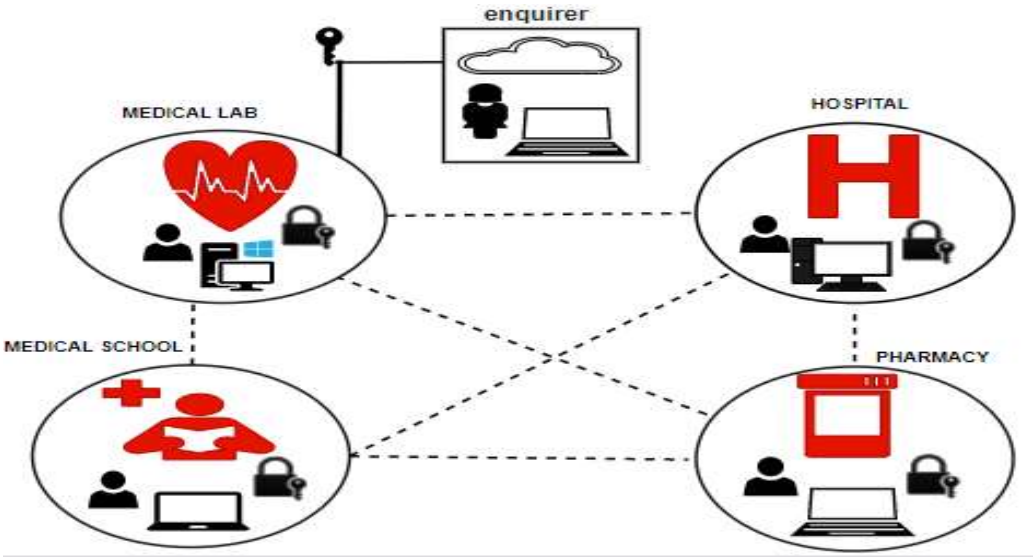

Figure-2: Architecture of the proposed framework

5. CONCLUSION

Federated learning technology was established to aid in extending ML benefits to domains that handle sensitive information. It's fair enough to state that it contributes a lot in privacy preservation however it has its limitations and challenges when it comes to model accuracy. In this paper we have carried out a comprehensive study on FL including its classification and some of its applications. Furthermore, the study extended out to a review of existing works that try to solve challenges in FL's accuracy through the use of asynchronous learning scheme. After evaluating numerous studies aligned with solving model accuracy issue,

we propose a novel asynchronous federated learning framework for distributed healthcare data. To the best of our knowledge this is the first study that presented a review of existing asynchronous FL frameworks. As relatively asynchronous learning still has room for improvements and discussions.

6. REFERENCES

[1] J. Konecny, H. B. McMahan, F. X. Yu, P. Richt ´ arik, A. T. Suresh, and ´ D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016. [Online]. Available: https://arxiv.org/abs/1610.05492

[2] "Google AI Blog." https://ai.googleblog.com/2017/05/the-machine-intelligence-behind-gboard.html

[3] "Electronic frontier foundation." https://www.eff.org/deeplinks/2018/10/google-bug-more-about-cover-crime

[4] Petters J. Data Privacy Guide: Definitions, Explanations and Legislation, https://www.varonis.com/blog/data-privacy/; 2020 [accessed 17.06.21].

[5] Y. Lu, X. Huang, Y. Dai, S. Maharjan and Y. Zhang, "Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT," in IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4177-4186, June 2020, doi: 10.1109/TII.2019.2942190.

[6] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," A Practical Guide, 1st Ed., Cham: Springer International Publishing, 2017

[7] S. F. Lameh, W. Noble, Y. Amannejad and A. Afshar, "Analysis of Federated Learning as a Distributed Solution for Learning on Edge Devices," 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA), 2020, pp. 66-74, doi: 10.1109/IDSTA50958.2020.9264060.

[8] J. Hao, Y. Zhao and J. Zhang, "Time Efficient Federated Learning with Semi-asynchronous Communication," 2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS), 2020, pp. 156-163, doi: 10.1109/ICPADS51040.2020.00030.

[9] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial Intelligence and Statistics, pp. 1273–1282, 2017.

[10] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2):12, 2019.

[11] P.Kairouz, H. B .McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, et al. Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977, 2019.

[12] D. Jatain, V. Singh, N. Dahiya, A contemplative perspective on federated machine learning: Taxonomy, threats & vulnerability assessment and challenges, Journal of King Saud University - Computer and Information Sciences, 2021, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2021.05.016.

[13] V.Mothukuri, R.M. Parizi and S.Pouriyeh and Y.Huang and A. Dehghantanha, G.Srivastava, A survey on security and privacy of federated learning, Future Generation Computer Systems, vol. 115, pp 619-640, 2021, ISSN 0167-739X, https://doi.org/10.1016/j.future.2020.10.007

[14] S. Abdulrahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi and M. Guizani, "A Survey on Federated Learning: The Journey from Centralized to Distributed On-Site Learning and Beyond," in IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5476-5497, 1 April1, 2021, doi: 10.1109/JIOT.2020.3030072.

[15] Shokri, R., Shmatikov, V., 2015. Privacy-Preserving Deep Learning, in: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security - CCS '15. ACM Press, New York, New York, USA, pp. 1310–1321. https://doi.org/10.1145/2810103.2813687

[16] Mcmahan, H.B., Moore, E., Ramage, D., Com, B., 2012. Federated Learning of Deep Networks using Model Averaging Blaise AgüeraAg¨Agüera y Arcas. arXiv:1602.05629v1

[17] K. Bonawitz V. Ivanov B. Kreuter A. Marcedone H.B. McMahan S. Patel D. Ramage A. Segal K. Seth Practical Secure Aggregation for Privacy-Preserving Machine Learning, in 2017 ACM Press New York, New York, USA 1175 1191 10.1145/3133956.3133982

[18] Kim, H., Park, J., Bennis, M., Kim, S.-L., 2019. Blockchained On-Device Federated Learning. IEEE Commun. Lett. https://doi.org/10.1109/LCOMM.2019.2921755

[19] S. Yang, B. Ren, X. Zhou, L. Liu, Parallel distributed logistic regression for vertical federated learning without third-party coordinator, arXiv preprint arXiv:1911.09824

[20] W. Du, Y.S. Han, S. Chen, Privacy-preserving multivariate statistical analysis: Linear regression and classification, SIAM Proceedings Series. (2004), pp. 222-233, 10.1137/1.9781611972740.21

[21] Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, Yuan Gao, A survey on federated learning, Knowledge-Based Syst., 216 (2021), p. 106775, 10.1016/j.knosys.2021.106775

[22] Du, W., Atallah, M.J., 2001. Privacy-preserving cooperative statistical analysis. Proc. - Annu. Comput. Secur. Appl. Conf. ACSAC 2001-Janua, 102–110. https://doi.org/10.1109/ACSAC.2001.991526

[23] L. Wan, W.K. Ng, S. Han, V.C.S. Lee, Privacy-preservation for gradient descent methods, Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 775-783 (2007), 10.1145/1281192.1281275

[24] P. Schoppmann, B. Balle, J. Doerner, S. Zahur, D. Evans, Secure Linear Regression on Vertically Partitioned Datasets, IACR Cryptol. ePrint Arch. (2016), pp. 1-27

[25] J. Vaidya, C. Clifton, Privacy preserving association rule mining in vertically partitioned data. Proc. ACM SIGKDD Int, Conf. Knowl. Discov. Data Min., 639–644 (2002), 10.1145/775107.775142

[26] Y. Chen, J. Wang, C. Yu, W. Gao, X. Qin, Fedhealth: A federated transfer learning framework for wearable healthcare, CoRR abs/1907.09173. arXiv:1907.09173. URL http://arxiv.org/abs/1907.09173

[27] Y. Liu, T. Chen, Q. Yang, Secure federated transfer learning (2018). arXiv:1812.03337.

[28] S. J. Pan, Q. Yang, A survey on transfer learning, IEEE Transactions on knowledge and data engineering 22 (10) (2009) 1345–1359.

[29] H. Yang, H. He, W. Zhang, X. Cao, Fedsteg: A federated transfer learning framework for secure image steganalysis, IEEE Transactions on Network Science and Engineering (2020) 1–1.
[30] Y. Liu, Y. Kang, C. Xing, T. Chen, Q. Yang, A secure federated transfer learning framework, IEEE Intelligent Systems (2020) 1–1. 41 Jour
[31] C. Nadiger, A. Kumar, S. Abdelhak, Federated reinforcement learning for fast personalization. Proc. - IEEE 2nd Int, Conf. Artif. Intell. Knowl. Eng. AIKE, 2019 (2019), pp. 123-127, 10.1109/AIKE.2019.00031
[32] Boyi Liu, Lujia Wang, Ming Liu, Lifelong Federated Reinforcement Learning: A Learning Architecture for Navigation in Cloud Robotic Systems, IEEE Robot. Autom. Lett., 4 (4) (2019), pp. 4555-4562, 10.1109/LSP.2016.10.1109/LRA.2019.2931179
[33] Liu, Y., Kang, Y., Zhang, X., Li, L., Cheng, Y., Chen, T., Hong, M., Yang, Q., 2019. A Communication Efficient Collaborative Learning Framework for Distributed Features. arXiv:1912.11187.
[34] Fate framework from webank, web. URL https://fate.fedai.org
[35] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, Y. Liu, Batchcrypt: Efficient homomorphic encryption for cross-silo federated learning.
[36] D. Alistarh, D. Grubic, J. Li, R. Tomioka, M. Vojnovic, Qsgd: Communication-efficient sgd via gradient quantization and encoding, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems 30, Curran Associates, Inc., 2017, pp. 1709–1720.
[37] Yang, T., Andrew, G., Eichner, H., Sun, H., Li, W., Kong, N., Ramage, D., Beaufays, F., 2018. Applied Federated Learning: Improving Google Keyboard Query Suggestions. arXiv:1812.02903.
[38] Yang et al., "Applied federated learning: Improving google keyboard query suggestions," 2018. [Online]. Available: arXiv:1812.02903.
[39] A. Hard et al., "Federated learning for mobile keyboard prediction," 2018. [Online]. Available: arXiv:1811.03604.
[40] S. Ramaswamy, R. Mathews, K. Rao, and F. Beaufays, "Federated learning for emoji prediction in a mobile keyboard," 2019. [Online]. Available: arXiv:1906.04329.
[41] Q.Xia, W.Ye, Z.Tao, J.Wu, Q.Li, "A survey of federated learning for edge computing: Research problems and solutions", in Hig-Confidence Computing, vol 1, no. 1, 2021, 100008, ISSN 2667-2952, https://doi.org/10.1016/j.hcc.2021.100008.
[42] D. Liu, T. Miller, R. Sayeed, and K. Mandl, "FADL: Federated autonomous deep learning for distributed electronic health record," 2018. [Online]. Available: arXiv:1811.11400.
[43] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, "Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records," J. Biomed. Informat., vol. 99, Mar. 2019, Art. no. 103291.
[44] 6] W. Schneble, "Federated learning for intrusion detection systems in medical cyber-physical systems," Ph.D. dissertation, Dept. Comput. Sci., Univ. Washington, Bothell, WA, USA, 2018.
[45] Y. Chen, J. Wang, C. Yu, W. Gao, and X. Qin, "FedHealth: A federated transfer learning framework for wearable healthcare," 2019. [Online]. Available: arXiv:1907.09173.
[46] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, "Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data," in Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI), 2019, pp. 270–274.
[47] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multiinstitutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation," in Proc. Int. MICCAI Brainlesion Workshop, 2018, pp. 92–104.
[48] W. Li et al., "Privacy-preserving federated brain tumour segmentation," in Proc. Int. Workshop Mach. Learn. Med. Imag., 2019, pp. 133–141.
[49] Jianmin Chen, et al., Revisiting distributed synchronous SGD, in: International Conference on Learning Representations Workshop Track, 2016
[50] Shiqiang Wang, et al., When edge meets learning: adaptive control for resource constrained distributed machine learning, in: IEEE Conference on Computer Communications (INFOCOM'18), 2018, pp. 63–71.
[51] Z. Chen and W.Liao and K.Hua and C.Lu and W.Yu, Towards asynchronous federated learning for heterogeneous edge-powered internet of things, Digital Communications and Networks, 2021, ISSN 23528648, https://doi.org/10.1016/j.dcan.2021.04.001
[52] Yue Zhao, et al., Federated Learning with Non-IID Data, 2018 arXiv preprint arXiv: 1806.00582
[53] Xiangru Lian, et al., Asynchronous decentralized parallel stochastic gradient descent, in: Proceedings of the 35th International Conference on Machine Learning, vol. 80, 2018. Stockholm Sweden.
[54] Takayuki Nishio, Ryo Yonetani, Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge, 2018 arXiv preprint arXiv:1804.08333.
[55] M. Van Dijk, N. V Nguyen, and T. N. Nguyen, "Asynchronous Federated Learning with Reduced Number of Rounds and with Differential Privacy from Less Aggregated Gaussian Noise," pp. 1–52.
[56] X. Lian, W. Zhang, C. Zhang, and J. Liu, "Asynchronous decentralized parallel stochastic gradient descent," 35th Int. Conf. Mach. Learn. ICML 2018, vol. 7, pp. 4745–4767, 2018.
[57] Y. Chen, X. Sun, and Y. Jin, "Communication-Efficient Federated Deep Learning with Layerwise Asynchronous Model Update and Temporally Weighted Aggregation," IEEE Trans. Neural Networks Learn. Syst., vol. 31, no. 10, pp. 4229–4238, 2020, doi: 10.1109/TNNLS.2019.2953131.
[58] B. Il Koh, A. D. George, R. T. Haftka, and B. J. Fregly, "Parallel asynchronous particle swarm optimization," Int. J. Numer. Methods Eng., vol. 67, no. 4, pp. 578–595, 2006, doi: 10.1002/nme.1646.
[59] M. Van Dijk, N. V Nguyen, and T. N. Nguyen, "Asynchronous Federated Learning with Reduced Number of Rounds and with Differential Privacy from Less Aggregated Gaussian Noise," pp. 1–52.

[60] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. 20th Int. Conf. Artif. Intell. Stat., vol. 54. Fort Lauderdale, FL, USA, Apr. 2017, pp. 1273–1282.
[61] Y. Chen, Y. Ning, M. Slawski and H. Rangwala, "Asynchronous Online Federated Learning for Edge Devices with Non-IID Data," 2020 IEEE International Conference on Big Data (Big Data), 2020, pp. 15-24, doi: 10.1109/BigData50022.2020.9378161.
[62] Y. Lu, X. Huang, Y. Dai, S. Maharjan and Y. Zhang, "Differentially Private Asynchronous Federated Learning for Mobile Edge Computing in Urban Informatics," in IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 2134-2143, March 2020, doi: 10.1109/TII.2019.2942179.
[63] X. Lu, Y. Liao, P. Lio and P. Hui, "Privacy-Preserving Asynchronous Federated Learning Mechanism for Edge Network Computing," in IEEE Access, vol. 8, pp. 48970-48981, 2020, doi: 10.1109/ACCESS.2020.2978082.
[64] S. S. Diwangkara and A. I. Kistijantoro, "Study of Data Imbalance and Asynchronous Aggregation Algorithm on Federated Learning System," 2020 International Conference on Information Technology Systems and Innovation (ICITSI), 2020, pp. 276-281, doi: 10.1109/ICITSI50517.2020.9264958.
[65] Sprague, Michael & Jalalirad, Amir & Scavuzzo, Marco & Capota, Catalin & Neun, Moritz & Do, Lyman & Kopp, Michael. (2019). Asynchronous Federated Learning for Geospatial Applications. 21-28. 10.1007/978-3-030-14880-5_2.
[66] Xie, Cong & Koyejo, Sanmi & Gupta, Indranil. (2019). Asynchronous Federated Optimization, Department of Computer Science, University of Illinois Urbana-Champaign.
[67] L. Feng, Y. Zhao, S. Guo, X. Qiu, W. Li and P. Yu, "Blockchain-based Asynchronous Federated Learning for Internet of Things," in IEEE Transactions on Computers, doi: 10.1109/TC.2021.3072033.
[68] Y. Lu, X. Huang, K. Zhang, S. Maharjan and Y. Zhang, "Blockchain Empowered Asynchronous Federated Learning for Secure Data Sharing in Internet of Vehicles," in IEEE Transactions on Vehicular Technology, vol. 69, no. 4, pp. 4298-4311, April 2020, doi: 10.1109/TVT.2020.2973651.
[69] Solmaz Niknam, Harpreet S. Dhillon, Jeffrey H. Reed, "Federated Learning for Wireless Communications: Motivation, Opportunities, and Challenges", IEEE Commun. Mag., 58 (6) (2020), pp. 46-51, 10.1109/MCOM.3510.1109/MCOM.001.1900461
[70] Stefano Savazzi, Monica Nicoli, Vittorio Rampa, "Federated Learning with Cooperating Devices: A Consensus Approach for Massive IoT Networks", IEEE Internet Things J., 7 (5) (2020), pp. 4641-4654, 10.1109/JIoT.648890710.1109/JIOT.2020.2964162
[71] Chuan Ma Jun Li Ming Ding Howard H. Yang Feng Shu Tony Q. S. Quek H. Vincent Poor on Safeguarding Privacy and Security in the Framework of Federated Learning IEEE Netw. 34 4 2020 242 248 10.1109/MNET.65 10.1109/MNET.001.1900506
[72] Zhao, Z., Feng, C., Yang, H.H., Luo, X., 2020. Federated-Learning-Enabled Intelligent Fog Radio Access Networks: Fundamental Theory, Key Techniques, and Future Trends. IEEE Wirel. Commun. 27, 22–28. https://doi.org/10.1109/MWC.001.1900370
[73] Latif U. Khan Shashi Raj Pandey Nguyen H. Tran Walid Saad Zhu Han Minh N. H. Nguyen Choong Seon Hong Federated Learning for Edge Networks: Resource Optimization and Incentive Mechanism IEEE Commun. Mag. 58 10 2020 88 93 10.1109/MCOM.35 10.1109/MCOM.001.1900649
[74] Wang, X., Han, Y., Wang, C., Zhao, Q., Chen, X., Chen, M., 2018. In-Edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning. arXiv:1809.07857.
[75] Close Elbir, A.M., Coleri, S., 2020. Federated Learning for Vehicular Networks. arXiv 2006.01412.
[76] Dongdong Ye, Rong Yu, Miao Pan, Zhu Han, "Federated Learning in Vehicular Edge Computing: A Selective Model Aggregation Approach", IEEE Access, 8 (2020), pp. 23920-23935, 10.1109/Access.628763910.1109/ACCESS.2020.2968399
[77] Xinhong Hei, Xinyue Yin, Yichuan Wang, Ju Ren, Lei Zhu, "A trusted feature aggregator federated learning for distributed malicious attack detection", Comput. Secur., 99 (2020), p. 102033, 10.1016/j.cose.2020.102033
[78] K. Demertzis Blockchained Federated Learning for Threat Defense 2021 arXiv:2102.12746
[79] Hichem Sedjelmaci, Fateh Guenab, Sidi Mohammed Senouci, Hassnaa Moustafa, Jiajia Liu, Shuai Han, "Cyber Security Based on Artificial Intelligence for Cyber-Physical Systems", IEEE Netw., 34 (3) (2020), pp. 6-7, 10.1109/MNET.6510.1109/MNET.2020.9105926
[80] Yunkai Wei Sipei Zhou Supeng Leng Sabita Maharjan Yan Zhang 35 2 2021 88 94
[81] T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan and A. Sadeghi, "DÏoT: A Federated Self-Learning Anomaly Detection System for IoT," 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 756-767, doi: 10.1109/ICDCS.2019.00080.
[82] Sav S, Pyrgelis A, Troncosco-Pastoriza R.J, Froelicher D, Bossuat J, Sousa J, Hubaux J-P. "POSEIDON: Privacy-Preserving Federated Neural Network Learning", Cornell University vol.2, pp.1-24, 30 September 2020. https://arxiv.org/pdf/2009.00349.pdf
[83] Froelicher D, Troncosco-Pastoriza R.J, Pyrgelis A, Sav S, Sousa J, Bossuat J-P, Hubaux J-P. "Scalable Privacy-Preserving Distributed Learning", Cornell University vol.1, pp.1-24, May 2020. https://arxiv.org/pdf/2005.09532v1.pdf

