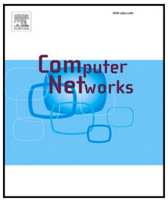
Computer Networks 212 (2022) 109048

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet



Review article

Fusion of Federated Learning and Industrial Internet of Things: A survey

Parimala Boobalan [a], Swarna Priya Ramu [a], Quoc-Viet Pham [b,1], Kapal Dev [c], Sharnil Pandya [d], Praveen Kumar Reddy Maddikunta [a], Thippa Reddy Gadekallu [a,*], Thien Huynh-The [e]

[a] *School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India*
[b] *Korean Southeast Center for the 4th Industrial Revolution Leader Education, Pusan National University, Busan, Republic of Korea*
[c] *Department of Institute of Intelligent Systems, University of Johannesburg, South Africa*
[d] *Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune, Maharashtra, India*
[e] *ICT Convergence Research Center, Kumoh National Institute of Technology, Gyeongsangbuk-do 39177, Republic of Korea*

ARTICLE INFO

Keywords:
Data storage
IIoT
Federated Learning
Data privacy
Data sharing
Resource management

ABSTRACT

Industrial Internet of Things (IIoT) lays a new paradigm for the concept of Industry 4.0 and paves an insight for new industrial era. Nowadays smart machines and smart factories use machine learning/deep learning based models for incurring intelligence. However, storing and communicating the data to the cloud and end device leads to issues in preserving privacy. In order to address this issue, Federated Learning (FL) technology is implemented in IIoT by the researchers nowadays to provide safe, accurate, robust and unbiased models. Integrating FL in IIoT ensures that no local sensitive data is exchanged, as the distribution of learning models over the edge devices has become more common with FL. Therefore, only the encrypted notifications and parameters are communicated to the central server. In this paper, we provide a thorough overview on integrating FL with IIoT in terms of privacy, resource and data management. The survey starts by articulating IIoT characteristics and fundamentals of distributed machine learning and FL. The motivation behind integrating IIoT and FL for achieving data privacy preservation and on-device learning are summarized. Then we discuss the potential of using machine learning (ML), deep learning (DL) and blockchain techniques for FL in secure IIoT. Further we analyze and summarize several ways to handle the heterogeneous and huge data. Comprehensive background on data and resource management are then presented, followed by applications of IIoT with FL in automotive, robotics, agriculture, energy, and healthcare industries. Finally, we shed light on challenges, some possible solutions and potential directions for future research.

1. Introduction

With an unprecedented increase in the number of Internet of things (IoT) devices and emerging applications, a large amount of traffic is created every day. Such an increase poses a great burden on the Internet network and also demands significant investments for the infrastructure upgrade. However, thanks to the development of big data analytics and artificial intelligence (AI) techniques such as deep learning and machine learning, the data collected can be effectively exploited for many purposes. From the communications perspective, the last few years has witnessed the emergence of AI applications in various fields. For example, ML is employed to investigate an efficient antenna selection in multi-antenna wireless systems [1], DL is used to handle the computation offloading problem in IoT systems with edge computing [2], and deep reinforcement learning (DRL) is used to optimize resource allocation problems at the network edge such as traffic classification, edge caching, network security, and data offloading [3]. However, the conventional AI models typically require central processing of data collected from all the users in the network, that is, users should upload their own data to a central server for training the learning model. However, a critical concern with the central learning is data privacy, i.e., some users want to keep track of their local data and do not want to transmit their local data to the central server. Training the learning model centrally requires a central cloud with immensely powerful computing capabilities and storage capacities. Meanwhile, recent advances in computing hardware and the proliferation of smart devices in our daily lives have shown that each

[*] Corresponding author.
E-mail addresses: parimala.m@vit.ac.in (P. Boobalan), swarnapriya.rm@vit.ac.in (S.P. Ramu), vietpq@pusan.ac.kr (Q.-V. Pham), kapal.dev@ieee.org (K. Dev), sharnil.pandya@sitpune.edu.in (S. Pandya), praveenkumarreddy@vit.ac.in (P.K.R. Maddikunta), thippareddy.g@vit.ac.in (T.R. Gadekallu), thienht@kumoh.ac.kr (T. Huynh-The).

[1] The work of Quoc-Viet Pham was supported by a National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) under Grant NRF-2019R1C1C1006143 and in part by BK21 Four, Korean Southeast Center for the 4th Industrial Revolution Leader Education.

https://doi.org/10.1016/j.comnet.2022.109048
Received 26 February 2022; Received in revised form 1 May 2022; Accepted 11 May 2022
Available online 21 May 2022

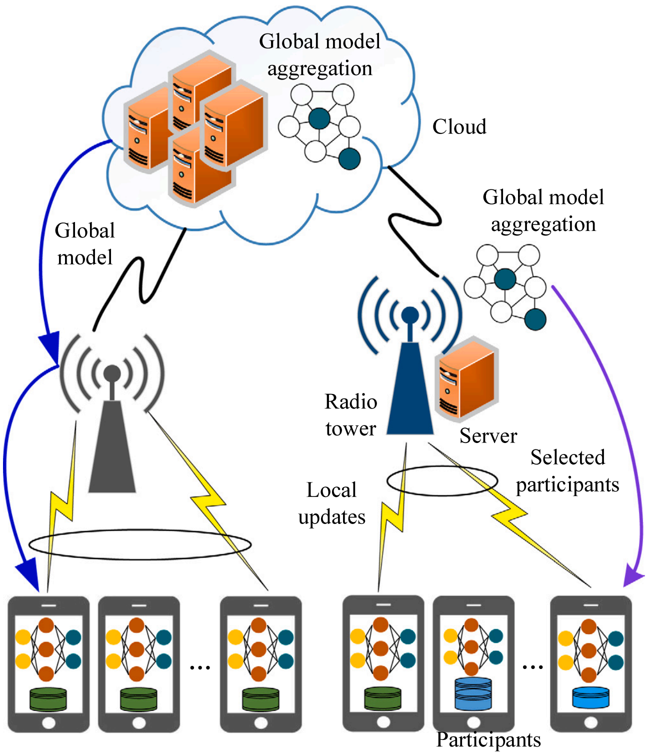


Fig. 1. An overview of FL in IoT systems.
Source: Adapted from [11].

IoT device can be equipped with reasonable computing and storage levels, which is closely comparable to a desktop computer ten years ago [4]. Therefore, the standard ML model is not readily applicable to large-scale IoT networks and cannot exploit the availability of distributed computing. This calls for a new learning model that leaves the training data distributed across individual IoT devices instead of being centralized.

Motivated by this issue, Google invented the concept of FL for on-device learning and data privacy preservation [5]. Using the FL approach, each IoT device can train its model based on locally collected data. Local data from IoT devices does not need to be transmitted to the centralized cloud. The centralized cloud only needs to collect the updated local training model from individual users. Thanks to its characteristics, FL has been adopted in many applications, for example, FL for improving Google keyboard suggestions [6], FL for healthcare [7], FL for smart city sensing [8], and FL for medicine [9]. To pictorially illustrate the concept of FL, an overview of FL in IoT systems is shown in Fig. 1. In general, each IoT device has its own dataset and the aggregation server can be either located at the network edge or a virtual cloud in the remote cloud computing system [10]. Each FL model has its own pros and cons, depending on various factors. For example, FL with the server at the network edge is suitable for applications requiring low latency, location awareness, and network contextual information [4] while cloud-based FL is suitable for applications with massive IoT devices over multiple regions and requirements of powerful computing/storage capabilities.

Although FL has found many benefits in enabling privacy-preserving learning solutions for IoT systems, it also poses significant challenges such as resource management, robustness, security, and incentive mechanisms. These challenges would be more significant and difficult to handle in *industrial IoT* (IIoT), in which much higher levels of security, safety, and reliability are demanded for IIoT services (e.g., attack detection, crowdsensing, and data offloading) and IIoT applications (e.g., smart factory, smart manufacturing, smart transportation, and smart healthcare) [12]. Despite the fast development of FL and the rapid emergence of many IIoT applications, there is a lack of a comprehensive survey that provides an overview of FL, IIoT, and applications of FL to IIoT. Motivated by this fact, this work focuses on providing the fundamentals of FL, IIoT, and reviewing state-of-the-art research works on FL for IIoT applications.

1.1. State-of-the-arts and contributions

Although there have been a number of surveys focusing on FL and IIoT, these two topics are usually studied separately. With regards to IIoT, the most well-known survey is presented in [13] with an introduction to IIoT, enabling technologies, major applications, and research outlooks. Similar surveys of IIoT on enabling technologies and applications can be found in [14], in which the term “Industry 4.0” is used instead. The roles of big data analytics in IIoT and a thorough classification, including data sources, big data tools, big data techniques, requirements, applications, and analytics types, are presented in [15]. More recently, a survey on IIoT security and the application of edge/fog computing is presented in [16]. With regards to FL, several surveys have been conducted over the last few years such as [17–19]. The two surveys in [17] present an introduction to FL along with general applications and challenges. Lim et al. [18] firstly provided fundamental features and unique characteristics of FL, and then reviewed three major aspects of FL at mobile edge networks, which are communication costs, resource allocation optimization, and privacy and security. Mothukuri et al. [19] tried to address four major questions, including the source of vulnerabilities, the kinds of security threats/attacks, comparison with security and privacy issues in conventional and distributed ML methods, and the defensive techniques. A review of computational offloading methods for UAV-enabled edge computing environment was done by Hudaet al. [20]. The detailed study of security and privacy aspects related to healthcare integrated IoT systems was done by Rasool et al. [21]. Zeb et al. [22] have conducted a review of a 5G-enabled digital twin for industries. We can also find interesting topics about FL in recent magazine articles. For example, the integration of FL into 6G wireless systems is discussed by Liu et al. [23], FL for intelligent fog/edge cloud radio access networks is introduced in Zhao et al. [24], a framework of reliable FL is proposed in [25], FL-based intelligent IoT over controllable communication channels via reconfigurable intelligent surfaces was studied by Yang et al. [26], and incentive resource allocation mechanisms using game-theoretic approaches are investigated by Khan et al. [27]. A recent survey on the integration of FL and IoT is provided [28]. This survey also provides a classification of FL-IoT studies, which are cloud server design, miner classification, edge collaboration, federated optimization schemes, incentive design, security and privacy, and global update strategies. A detailed review of the implementation challenges of FL in IoT resource-constrained environments was discussed by Imteaj et al. [29]. The privacy, security, and attack detection mechanisms related to FL were discussed by [30,31]. Integration of blockchain in IIoT environments and open issues were analyzed by Huo et al. [32]. The communication and networking-related challenges related to FL were researched by Wahab et al. [33]. A detailed review of potential resolution systems, their framework, and various approaches for IIoT environments were done by Ren et al. [34]. The survey of honeypots and honeynets in IoT, IIoT, Cyber–physical environments was conducted by Fanco et al. [35]. The review of the energy-efficient mechanisms and methodologies for IIoT systems was conducted by Mao et al. [36]. The review of various differential privacy applications, challenges, and future directions for IIoT systems was conducted by Jiang and his team Jiang et al. [37]. A review of FL-enabled digital twin was done by Ramu et al. [38]. A study of fog-enabled Industry 4.0 systems was done by Brik et al. [39]. A comparison analysis of existing surveys related to FL and IIoT is summarized in Table 1.

Table 1
Comparisons with existing review works related to FL and IIoT.

Ref. No.	Contributions	Limitations
[29]	The proposed study discusses the implementation challenges of FL in IoT resource-constrained environments.	The study did not discuss any ideas related to the fusion of FL and IIoT systems.
[30]	The paper discusses the privacy, security, and attack detection mechanisms related to FL, and its future directions.	The presented review work did not discuss FL-integrated IIoT technologies, applications, and future research directions.
[32]	The proposed review work has discussed the integration of blockchain in IIoT systems and its future directions.	The presented review work did not discuss the concept of FL-integrated IIoT, its key enabling technologies, applications, and future research directions.
[33]	This review work discusses the FL concepts, key enabling technologies, applications, and future directions related to communication and networking challenges.	The presented analysis did not discuss FL-integrated IIoT technologies, applications, and future research directions.
[34]	The paper surveyed the potential resolution systems, their frameworks, and various approaches for IIoT environments.	The surveyed paper did not discuss any ideas related to integrating FL into IIoT systems.
[35]	This review work discussed the review of honeypots and honeynets in IoT, IIoT, Cyber–physical systems.	The presented survey did not discuss the concept of FL-integrated IIoT, its key enabling technologies, applications, and future research directions.
[36]	Energy-efficient mechanisms and methodologies for IIoT systems are discussed in this review work.	The authors have not discussed the concept of FL-integrated IIoT, its key enabling technologies, applications, and future research directions.
[40]	This review work discussed various incentive mechanisms in detail for FL systems.	This review work did not discuss any ideas related to integrating FL into IIoT systems.
[37]	The proposed review discusses the differential privacy applications, challenges, and future directions for IIoT systems.	The conducted rigorous analysis did not discuss any methodologies, applications, and technologies related to integrating FL into IIoT systems.
This paper	In the presented review work, we have discussed a comprehensive review of several challenges associated with current studies on FL-IIoT for security, privacy, data and resource management, FL-IIoT applications, research challenges and further highlight open research issues that should be efficiently addressed to envision FL-IIoT solutions.	

Different from existing works [12], in this work, we set to provide a survey on the use of FL for IIoT applications. To the best of our knowledge, this work is the very first attempt on the integration of FL and IIoT in terms of security, privacy, data and resource management. The primary contributions offered by our work can be summarized as follows.

- We present the fundamentals of FL, the characteristics of IIoT, and the motivations for integrating FL with IIoT.
- We review the state-of-the-arts on FL-IIoT, including FL for enhancing security and privacy in IIoT, data management and resource management in IIoT, and promising application areas such as autonomous industry and healthcare industry.
- We discuss a number of challenges associated with current studies on FL-IIoT, and further highlight open research issues that should be efficiently addressed to fully realize FL-IIoT solutions.

1.2. Paper organization

This paper is organized as follows. In Section 2, we present the fundamentals of IIoT and FL and the motivations for FL-IIoT integration. Section 3 reviews the use of FL to provide secure solutions in IIoT. Next, Section 4 presents data management and resource allocation in IIoT and discusses the use of FL. Applications of FL for important sectors in IIoT are reviewed in Section 5. Challenges and future directions in the use of FL for IIoT are discussed in Section 6, and finally we conclude the paper in Section 7.

2. Industrial IoT and federated learning

This section provides a brief overview of basic concepts of IIoT, FL and motivation behind integration of FL with IIoT.

2.1. Fundamentals of federated learning

FL is a new branch of AI where the ML models are trained on the data located at the decentralized devices such as mobile phones and other smart devices, as shown in Fig. 1. The devices participating in the FL need not exchange their data samples, that ensures the privacy and security. The underlying principle in FL is that the ML models are trained on local data, then the parameters of the ML models are exchanged between the local nodes (devices) at regular intervals that enables the creation of a global ML model. The training data remains on the individual devices as only the model updates from the local devices are sent across to the central cloud storage. The global ML model resides at a server. Once all the devices send their models to the server, a combined model is created by averaging the parameters of the individual models. In this way, the individual devices learn collaboratively from a shared model. Some of the advantages of FL over traditional ML models are that FL ensures data privacy, reduced latency, reduced power consumption. Not only the FL ensures privacy of sensitive information of users, but also FL can deliver personalized ML models to the users with enhanced user experience [17,41].

2.2. Characteristics of IIoT

IIoT is collection of people, sensors, machines, and computers that enable intelligent industrial operations with the help of ML and analytics [42]. IIoT enables intelligent devices and networked sensors to collect the data from the manufacturing plants and make use of predictions through AI and automate the decision making. IIoT offers a revival of industries that have been lately struggling due to several factors including shortage of skills from personnel. IIoT offers several benefits because of their minimal or nil dependency on human intervention. The machine learning algorithms and big data analytics can be used to analyze the data from IIoT devices, thereby enabling IIoT systems to make intelligent decisions by automatically learning from the data generated through the sensors. IIoT leads to smart machines that can achieve highest levels of accuracy that was not possible earlier because of human intervention. Thus, IIoT has huge potential to transform the way the products are manufactured and delivered, make factories smart and efficient, protect front line workers by providing better security in difficult conditions (where the machines will be working in tough conditions), reduce human errors, increase efficiency, thus, in turn, save huge amount of money [32,43]

Many sectors and applications such as manufacturing, transportation system, oil and gas, healthcare, agriculture, smart cities, energy are benefited by IIoT. One of the main benefits of IIoT is its ability to increase efficiency in operations through prediction. For instance, the sensors in a machine can automatically pinpoint specifically where the trouble is when it goes down and place a service request automatically. The sensors can predict the likelihood of a breakdown of machines before it happens by automatically sensing the sounds of a machine, temperature, vibrations, and other factors by using data analytics and ML. Hence, predictive maintenance reduces the idle time of a machine and also reduces the damages caused by faulty machines by fixing the issues before they are escalated. Another important advantage of IIoT is location tracking of equipment and tools. In many industries, the workers spend significant amount of time searching for tools, finished goods in an inventory. IIoT sensors save a lot of time by providing location services, thereby making it impossible to lose equipment, goods [44]. Until recently, manufacturing companies used to spend huge capital on purchasing and maintenance of equipment. But with the advent and advancements in IIoT, the industries can lease the equipment from the vendors. The owners can remotely monitor these equipments through the sensors, thereby delivering upgrades, repairs, and maintenance remotely. This will save both time and capital for industries. To summarize, the main benefits of IIoT are (1) reduced waste, (2) reduced maintenance costs, (3) energy savings, (4) workforce productivity gains, (5) improved service, and (6) revolutionary products and services.

2.3. Motivations of integrating federated learning with IIoT

The exponential growth in IIoT applications is being hindered by several issues such as security, privacy, and communication cost. Integration of FL with IIoT has the potential to solve many issues above. The main motivations behind integrating FL with IIoT applications are summarized as follows:

- **Security and Data Privacy Preservation**: In order to get valuable insights and patterns from the data generated by IoT devices in IIoT applications, ML and DL algorithms are used. To understand the complex patterns from the data generated, these algorithms are trained regularly by large datasets collected from different industries/locations. But moving these datasets to a central location for training the algorithms exposes the sensitive data to potential hackers, and intruders as the data from IIoT applications is quite sensitive and may reveal sensitive information about business models of industries [45]. The datasets need not be transferred to a central location for training the algorithms when FL is used. Hence it secures the sensitive data from potential attackers. This also preserves the privacy of sensitive data.
- **Reduced Communication Cost**: The sensors in IIoT applications generate large volumes of data. Transferring huge volumes of data to the remote cloud incurs significant communication cost. With FL, the entire data need not be transferred to the cloud. Instead, only the summarized results after application of the ML/DL models is transferred to the cloud. As only few instances of data are transferred to the central cloud, using FL will reduce the communication cost while transferring the data from the local devices to the cloud.
- **Improved Performance of the Network**: IIoT applications requires huge network infrastructure to handle the big data generated from the IIoT devices [46]. This may affect the performance of the network. In an FL setup, a ML/DL model is implemented on the data generated from IIoT applications that is stored in an edge device and only the summary of the results are transferred to the central location. Hence, the traffic in the network is reduced considerably. With reduced network traffic, the overall performance of the network can be significantly improved.
- **Scalability**: IIoT devices generate significant amount of data continuously. Conventional ML algorithms often fail to scale to the big data generated from IIoT devices. Integration of FL with IIoT enables DL algorithms to scale their learning as they need not be trained on large volumes of data generated locally. The central algorithm is trained only on summarized results from the individual edge devices. This facilitates a significant scalability of the central learning algorithms and train from the data generated at several edge/local devices [47].

3. Federated learning for security in IIoT

FL protects user data by training the model without transferring the data from client to the main server. The scope of security mechanism can be further enhanced while sharing the model updates and integrating the data generated from similar devices from various industries [48]. In this section, we discuss about providing different security mechanisms by integrating FL with ML and DL models in IIoT framework. FL with blockchain in IIoT is also presented in this section.

3.1. Federated machine learning for secure IIoT

The amount of data generated from a single industry will not be suitable for applying ML techniques. Generally, many industries would have similar IoT surveillance systems and so the data generated can be federated from all the devices that could lead to increase in the accuracy of ML algorithms. However, the key concern is the data security provided during leveraging of data. Federated ML (FML) has gained a lot of attention because the way it handles the privacy by decentralizing the data generated at the end user device and aggregation of ML models at a centralized server. The two types of attacks in IIoT namely eavesdroppers and hackers were discussed in [49].

1. *Eavesdroppers*: Attack happens when the data is transferred through communication channels. In this case, it may happen between IoT node and IoT sink, then from IoT sink to centralized server communication channels.
2. *Hackers*: Attack that happens in the centralized server. A stalker resides in the server and is capable of obtaining actual locations of a user.

As discussed in [49], initially the data is generated from IoT devices from every smart industry sent to the IoT sink. The sink is a repository for collecting the data from different IoT nodes in the industry through wired and wireless communications and it also encrypts the data which is sent to the centralized server. Then the server collects the information from multiple IoT sinks and federates the data. Finally, the smart industry decrypts the knowledge into an understandable format. In this case, both eavesdroppers and hackers attacks are not possible, as encrypted data is present in both the cases.

The main intention of FL is to collectively learn the global model without sacrificing the privacy of data. But sharing model updates to the server during training process may lead to an issue of leaking sensitive data like user personal information. In order to preserve privacy in ML models, number of methods are employed in FL.

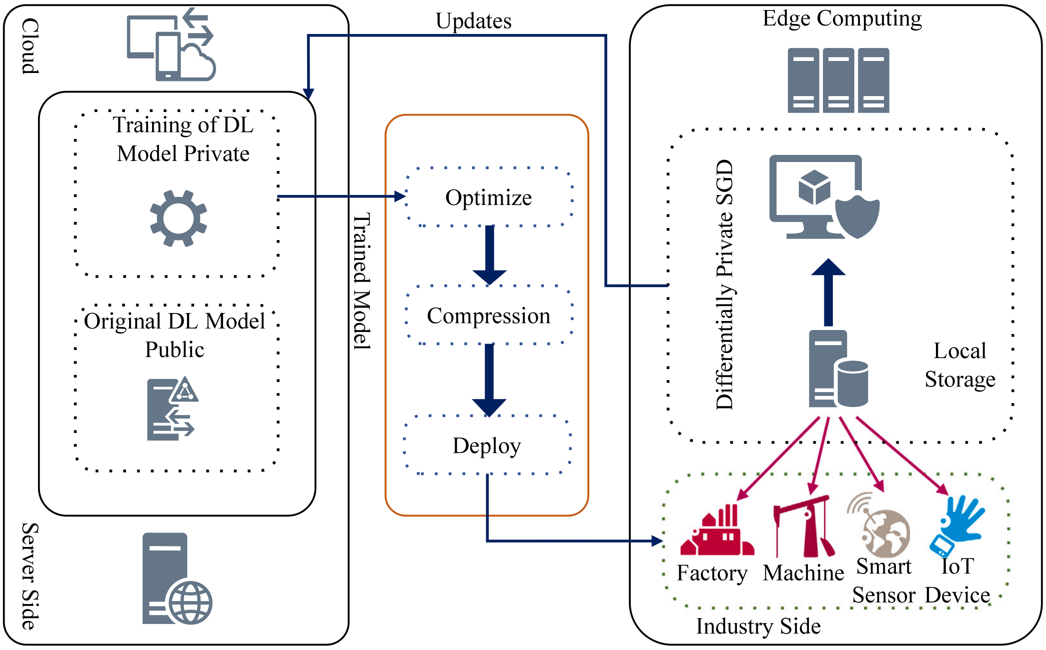


Fig. 2. Federated DL in IIoT.

- *Differential Privacy (DP)*: DP defines the amount of information about the data that can be available for third party analysis [50]. Information under DP can be categorized as general information which has the information about the entire population and the other type is private information that has the personal private data. Some of the properties of DP framework to provide sensitive personal information and privacy protection includes quantification of privacy loss, composition, group privacy and closure under post-processing.
- *Homomorphic Encryption:* Computation and analysis is done based on encryption technique so that the attacker finds very hard to find the original information [51].
- *Secure Multiparty Computation:* It is a model where multiple parties collaboratively compute without leaking any information to the third parties [52].

3.2. Federated deep learning in IIoT

Recently, integration of DL models with IoT and edge devices have become more popular which provide real time analytics with limited resources [53]. So, federated DL (FDL) enables Industry 4.0 companies to integrate DL in IoT devices and provides secure framework using FL as shown in Fig. 2. The main goal of FDL is to provide IIoT with advanced capabilities using optimized DL which would transform the Industry 4.0 factories into smart factories. Some of the parameters that are required to build the FDL model in IIoT are discussed below.

FDL Model: A FDL model can be implemented on both client and server side. In client side, the model is trained with locally generated data from the end device. On the server side, model present in cloud is continuously updated by integrating differentially gradients from each private network. Each local DL network takes its turn to continuously upload and download the currently updated gradients to the cloud model. So, distributed selective stochastic gradient descent approach can be applied in cloud model to frequently update local private model. The first decentralized model named “Model Chain” uses blockchain technology [54] to enable privacy preserving during data transfer in IIoT model. In addition to that, asynchronous stochastic gradient descent can also be used when a single model can be trained in parallel among all the devices in IIoT environment which in turn is aggregated and processed.

FDL Communication and Networking: The main advantage of using FDL is to run DL models in IIoT devices and also to involve DL models in decision making process. This kind of decentralizing DL process, enhance robustness, operational efficiency and reliability of IIoT devices. FDL provides two types of communication, namely, intra communication channel and inter-communication channel. Former transmits the data among all the tiers in the framework. FDL communicates between the IIoT and cloud tier where the optimized model in the cloud is deployed to the end device. However, the security and privacy must be maintained in the FDL during communication. In inter communication channel, the components in each layer communicates with each other in three different ways such as cloud, edge and end device. The main aim of FDL is to minimize intra-communication and maximize inter-communication which would significantly reduce the cost of communication in IIoT.

FDL privacy and security: Lim et al. [18] discussed about various security threats while FDL overcomes these threats by sharing the DL models from cloud to end devices. The security and privacy issues can be solved easily using data encryption [55]. However, IIoT-based systems often suffer security and privacy issues during data processing and analytics. The general principle behind FDL consists of training local DL models on local data samples and exchanging parameters like the weights and biases of a deep neural network. These updates are shared at some regular intervals between these local nodes to generate a global model shared by all local nodes. Anyway, FDL builds DL models that do not expose information about the data to the cloud. Security issue on the server side includes sharing of DL models on the cloud that leads to confidentiality and security risk. Security issue on the client side is done by encrypting the data during the training process before sending it to the cloud server. Some mechanisms namely DP and homomorphic encryption technique controls the amount of data to be shared on the cloud.

FDL Optimization: As the end devices have limited memory and computational requirements, DL models must be optimized so that they can be deployed on IIoT or end devices efficiently. DL models executed in all the IIoT nodes should be optimized so that memory and computational requirements is minimized which would subsequently improve

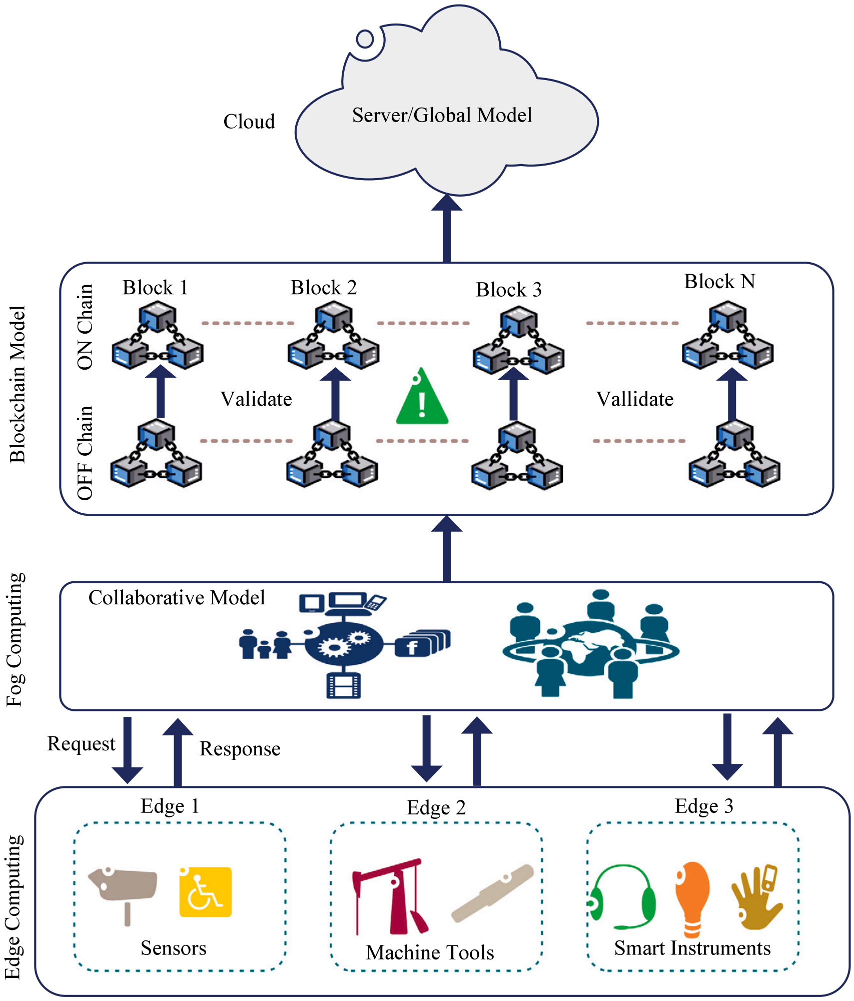


Fig. 3. Blockchain in IIoT using FL.

the performance of the system. In terms hardware optimization, GPU offers low power computation that decreases computation time. FPGA and Google's Tensor Processing Unit [56] are other DL devices that boosts the processing of DL network. In terms of memory optimization, algorithms like shared memory allocation algorithm can be used for DL models. Dynamic scheduling [57] is one of the key process used for optimizing the performance on the cloud server.

Recently, researchers in [58–61] proposed DL models in FL for IIoT networks in various application such as automobile and mobile network, traffic network and in image processing applications.

3.3. *Blockchain in IIoT using FL*

Generally, FL uses a central server to aggregate all the updates of ML and DL models and communicates the aggregated updates to the global model. Security breach can happen when the updates are transferred to the server or when the global model is sent to the clients [62]. So to overcome this security issue, the blockchain concept can be applied to store the model updates in immutable form [63–68]. Yin et al. in [69] presented a secured model that uses blockchain as a decentralized framework to replace the traditional central servers. Also, it stores the historical information as tamper resistant data. Lu et al. [70] investigated some security issues in node, block and framework which are designed using blockchain in FL.

The workflow of FL with blockchain is shown in Fig. 3. IIoT devices are designed in such a way that they train the model with local data. After training, they send the local updates to the edge node where all the local updates are consolidated. Edge nodes in turn send aggregated local updates to the fog node where the updates from all the edge node are aggregated. Then, the updates are converted into a block, which is verified based on the smart contract. The validated block is joined with blockchain which is tamper-proof and secured. Finally, all the local updates are received by the centralized server to train the global model. Then the updates of the global model is sent to all the edge node to update on all the end devices connected to the edge node [71].

The framework must be scalable for real world applications that can adapt to different kind of scenarios. In recent works [72–74], researchers have applied blockchain in image processing and vehicular networks using FL in IIoT framework. The works in [75–77] designed a FL based blockchain framework for health industry, railway network and defense organization. Generally, privacy preserving mechanism in FL using IIoT will be applied on end to edge device communication and from edge to collaborative model. Even in worst case also only the model updates can be tampered between end devices to collaborative engine. Recent studies of FL in ML, DL and blockchain in IIoT are summarized in Table 2.

4. Data management and resource management

The data generated by virtual objects as well as physical objects are heterogeneous and huge in volume such as logs in their SCADA systems [81], data from various business-oriented applications [82, 83],radio frequency data (RFID) [42], data from wearable devices [84], data from sensors [85], real time web data [86], media data and

Table 2
Overview of recent studies on FL with secure IIoT.

Ref	Framework	Algorithm	Application	Techniques	Contribution Inference	Limitations challenges
[49]	FMLin IIoT	FTM	Smart factories that belong to same vertical industry	FTM with homomorphic encryption which leverages the linear transformation	∘ Every factory shares its ciphertext data ∘ Mines the same knowledge from different factories ∘ Defend the attacks from distributed and centralized hackers	∘ Challenge to maintain balance between ML and data security ∘ FTM needs the centralized server to operate the federated mining
[50]		Primalchain	NLP and speech recognition	DP	∘ Uses FML to generate global representation in IIoT environment ∘ Blends DP, FML and blockchain	∘ Large scale ML latency is more ∘ Efficiency can be improved
[78]		Tensor Ridge Regression	Image processing	Parallel factor decomposition	∘ Projections of input tensor to more than one directions ∘ Multilinear mapping from tensorial input space to continuous output space	∘ High dimensional datasets leads to overtraining, high computational complexity and large memory requirement
[79]		STA	Computer vision and pattern recognition	Tensor Alignment Technique	Unsupervised tensor feature extraction Enhances the robustness in the alignment ∘ Works good for high dimensional space	Sparse projection learning methods are not investigates for tensor recognition Not efficient for high dimensional space
[80]		CPSS	Wireless IoT	∘ Privacy preserving high order singular value ∘ Decomposition based criterion	∘ Model allows users to utilize the storage and computing resources of cloud and fog without leaking any sensitive information	As it uses encrypted data, cost is very high
[58]	FDL in IIoT	DeepPAR DeepDPA	Smart factories	∘ Proxy re-encryption ∘ Group dynamic key management	∘ Protects each participants input privacy while preserving dynamic update ∘ Guarantees secrecy in group participants	Still security mechanism at edge device can be improved
[59]		Double Deep Q-network	Mobile networks	Decentralized cooperate edge caching Markov decision process	∘ Federated reinforcement learning deep framework enables fast training and decouple the learning process from cloud ∘ Reduces loss and average delay	∘ Decentralized loss can be maximized
[60]		FedGRU	Traffic network	∘ parameter aggregation ∘ Federated averaging algorithm ∘ Joint announcement protocol	∘ Random subsampling for participating organization is used to reduce the communication overhead ∘ Captures the spatiotemporal correlation of traffic flow data	Spatial temporal dependency can be better captured by applying graph convolutional network
[61]		ASTW-FedAVG	Image Processing	∘ Asynchronous strategy ∘ Weighted aggregation	∘ To reduce the communication between the client and server ∘ Asynchronous strategy is used to aggregate and update the parameters ∘ Previously trained local models are aggregated to enhance the performance	∘ Only server are involved in evolving the local model ∘ Communication cost can be reduced
[69]	Federated blockchain	Federated collaboration framework	Wearable sensor data	Private and public data center blockchain	∘ Framework for collaborating multiparty computation is designed ∘ Blockchain is used to overcome security issues	Large scale multiparty secure collaboration of IoT data is challenging
[72]		BlockFL	Wireless Communication Network	∘ On device ML ∘ Smart Consensus mechanism	∘ Verified local updates are generated as block and added to blockchain ∘ Each device generate global model from the updated new block	Proof-of-work can be hacked by the malicious miners
[73]		FL-Block	Image processing	∘ Decentralized privacy ∘ Poisoning attack	∘ Allow local updates to exchange with global blockchain model ∘ As the central authority is replaced by blockchain, privacy is achieved more	Tradeoff between privacy protection and efficiency should be optimized
[74]		BFL	Vehicular network	∘ Renewal reward approach ∘ VML	∘ Provide efficient communication of autonomous vehicles ∘ It forms representative groups from the overall vehicles	Choice of block updates will limit the increase in overall delay
[75]		CrowdSFL	Health Industry	Crowd sourcing encryption algorithm	∘ Data privacy is controlled by smart contracts ∘ The updates submitted by the local device cannot be leaked as it enables the decentralized blockchain platform	∘ Reward distribution mechanism is not defined ∘ Model can be optimized
[76]		SVM and Deep network	Railway network	SVM based on mixed kernel function and multi class model	∘ Distributed asynchronous collaborative ML is designed ∘ Smart contract involves all participant and generates new block	Accuracy and efficiency of the prediction can still be maximized
[77]		IoBT	Social Security through defense organization	∘ Multilayered model ∘ Defense led combat	∘ Distributed computing defense framework for sustainable society ∘ Optimize the data trained in local nodes	Participation of local node in training process can be rewarded based on their contribution

location data [42]. So, the traditional data management techniques fail to handle such heterogeneous and huge data. FL can be utilized to overcome these challenges as the technique is specifically designed to maintain data privacy by avoiding global sharing of data. This section investigates the contributions of researchers in the field of FL for data management and resource management.

4.1. Data storage in IIoT

Storing huge data and processing them to extract knowledge and reacting at the right time is the biggest challenge now [87]. The technologies like fog computing, cloud computing, and edge computing have given the researchers an opportunity to solve the issues with respect to data storage [88]. The work in [89] proposed a framework for processing the IIoT data. Almost all the data life cycle phases are integrated, starting from pre-processing to the retrieval of the requested data stored or archived using fog and cloud computing. The framework was designed to have entities, namely edge server, proxy server, and cloud server for taking care of data storage and processing. The time-sensitive data generated by the IIoT data sources is received by the edge servers where the data is processed at a fundamental level. If the data in the edge server is needed for future usage, then those are transmitted to the proxy server where the quality of the data is improvised by pre-processing, transforming into structured data so that they can be stored in the cloud server as historical data for further predictive analysis. Intelligent and smart applications in IIoT usually demand high computation and low delay in tasks. This can be achieved by using FL with edge servers where peer-to-peer edge offloading model is utilized and hence reliability of data is improved [90].

Borylo et al. [91] proposed and presented a scheme named dynamic resource provision by combining the cloud and edge computing schemes for the purpose of energy-awareness. In this scheme, the major limitation was that the amount of energy consumed was high. In order to overcome this limitation, Kaur et al. [92] designed an architecture consisting of three different layers, namely cloud layer, edge layer and software-defined network (SDN) layer. The SDN layer plays a major role in handling network congestion and hence minimizes energy consumption [93].

In [94], Singh et al. proposed a data storage method using bloom filter (BF) and fuzzy logic, namely Fuzzy Folded BF (FFBF). Two different BFs are compressed into a single BF by using fuzzy operations. The major advantage of this compression was that in a single BF, a large volume of data elements were stored. The computational cost was reduced due to the usage of the double hashing technique.

An integrated framework, namely IoT based Industrial Data Management System (IDMS), for an IIoT environment was proposed by Saqlain et al. [95] for handling and managing the huge industrial real-time data acquired from a smart manufacturing unit. For the purpose of handling emergency events, traditional communication protocols were used. The middle-ware, which was designed in the IDMS, provided a Service Oriented Architecture (SOA). The raw data was transformed into a structured format by the framework so that it can be archived in the cloud server where useful information and knowledge can be extracted.

In [96], Anton et al. presented a model that can smartly collect, aggregate, and analyze the data collected from heterogeneous sources. The model consists of three levels. The base level is responsible for collecting singular packets and the collected packets are analyzed here. The second level is responsible for aggregating all the packets that flow in the network and the information is extracted based on the application. The third level aggregates all the raw data and meta-information related to each application into a connected graph which gives an overall picture of the flow of packets in the network. An IIoT based intelligent control and management system was designed by Du et al. in [97] for the purpose of the motorcycle endurance test. The architecture designed in this system is made of four layers, namely the executive layer for the sake of data acquisition, the cognitive layer which takes care of the business logic, the network layer which acts as the communication layer and finally, the control layer. The design enables easy management, up-gradation, extension and compatibility.

Shu et al. [98] proposed a novel architecture termed as cloud integrated cyber physical system architecture (CCPSA) which comprises of three domains: one on the network and their communications, other dealing with the control and computational aspects and the final domain is a data-centric CPS. This work also focused on providing solutions for challenges like resource management, scheduling aspects of resources. Liu et al. [99] proposed a methodology for acquiring and collecting data in an efficient manner so that no data is missed due to the limited energy in the smart terminals using Ethereum blockchain and DRL. The DRL increases the amount of data collected at a time by the mobile terminals and the blockchain enables higher security.

Suppose there is a requirement for offloading data from multiple tasks at the same time, the task has to be completed with high power and reduced delay. To achieve this, Manzoor et al. [100] proposed a FL based framework-mobility and demand-aware proactive content offloading for improving the offloading ratio and also download link rate by 9.8% and 1.18% respectively. This approach is more efficient than the cloud based approach and shows an average increase of 6.7% in performance.

To summarize, most of the available data acquisition and data storage techniques as well as frameworks focus on specific issues, but acquiring and storing such huge heterogeneous industrial data is a challenging task. Also the data collected from various factories distributed at different floors and locations can be of different formats and volume [101]. Few challenges that can be listed out are data dynamicity, data visualization, data archiving. To meet these challenges a well-versed data management framework has to be devised that can efficiently handle the huge data.

4.2. Federated learning for data management in IIoT

Recent advancements have made it easier for the industrial giants to gather huge volumes of data from devices located in different geographical locations. The industrial applications need to process the huge data for making any decision by transferring them to a centralized server or location, which requires high bandwidth, but they fail because of a huge delay. To overcome this, few heterogeneous architectural based solutions were provided like usage of edge computing, fog computing, etc. But these solutions still have challenges like security, privacy, etc. The major issue with these kinds of architecture-based solutions is that most of the decisions are made locally, neglecting the global data, model and knowledge. This section discusses the various FL approaches coined by various researchers for handling the influence of local knowledge over the decisions.

Wang et al. [103] proposed a framework named "In-Edge AI" which imparted intelligence to the mobile edge network using DRL and FL. This framework is designed to reduce the volume of data that needs to be transferred to the central server and to provide a better data management scheme to adapt the heterogeneity. The framework was designed for cloud based robotic systems which gathered sensor data from heterogeneous resources. Yin et al. [69] proposed a framework based on FL for providing a secured collaboration of data across different geographical locations.

Zhu et al. [104] proposed a methodology to optimize their neural network model with the usage of evolutionary algorithms. This work used multi-objective functions for reducing the computational cost. FL is used for the purpose of solving the scalability issues in the DL models and also to increase the learning efficiency and performance of the global model using the data distributed across the clients.

A scheme by name Efficient and Privacy Enhanced FL (PEFL) is proposed in [106]. This scheme is best suited for industrial sectors dealing with sensitive data like healthcare, auto-pilot, auto-driving, industrial robots, etc. where decision making cannot be compromised due to hesitation in sharing data. The major advantage of this scheme is that the performance is not reduced even if there is a collision due to multiple entries. A cross domain sharing based scheduling scheme [108] was proposed for providing edge services with higher intelligence. The scheme integrated the block-chain and edge intelligence mechanisms to provide an intelligent IIoT framework. The proposed scheme is suitable for IIoT networks beyond 5G. An optimization technique using particle swarm optimization (PSO) and FL is proposed in [109] focusing on the hyper-parameter tuning for the DL models that are available locally in the smart city applications.

To summarize, there are various building blocks that can be used for customizing the traditional ML and FL algorithms to suit the IIoT applications. But there is no specific architecture that seems to be centered around specific industrial sector. The algorithms and methodologies discussed in the section can work for any type of industry once modified a bit keeping in mind various aspects such as efficiency, device set up, end users, end servers, structure and autonomy.

4.3. Federated learning for resource management in IIoT

The recent industrial revolution uses a wide variety of technologies like collaborative robotics, cloud and edge computing, cognitive computing, CPS, ML, etc.[118]. Integration of Industrial applications with IoT and ML leads to various advantages like optimized production costs, increased productivity, reduction of error, improvising the automation process, higher quality, and so on. But, the traditional ML aspects could not be used in real-time Industry 4.0 applications as data cannot be shared for training in a centralized ML server [18] as there is a risk of data leakage [119]. Hence most of the IIoT uses FL which is an efficient resource management technique for learning, adding new data and updating the aggregation server with the model updates. This section discusses about the various contribution of researchers in the field of FL for optimizing and managing resources over real-time processing in Industry 4.0 applications.

In [110], Chen et al. analyzed the performance of FL with the help of packet error rate over wireless networks. A closed-form expression was derived for calculating the convergence rate of FL. This was utilized for computing the optimal transmission power for allocating a resource block and selecting a user. In [111], Abad et al. proposed hierarchical

Table 3
Summary of recent research contributions on FL in IIoT for data storage-data management-resource management.

Ref.	Methodology	Research Focus	Techniques Used	Contributions	Proof of Concept
[89,91,92,102]	Data Storage	∘ Secure data storage ∘ Dynamic data gathering ∘ Reducing congestion ∘ Reducing energy consumption	∘ Middleware based solution ∘ SDN ∘ Interplay of fog and cloud computing ∘ latencyAware policy	∘ Retrieval feature tree is designed ∘ Reduction in storage space ∘ Carbon footprint of data centre is reduced ∘ Multi-objective evolutionary algorithm is proposed	∘ Gas density in factory ∘ Customer services and maintenance
[94,95]		∘ Effective data storage ∘ Acquiring huge data ∘ Provide intelligent management	∘ SOA based solution ∘ Distributed database server ∘ Data aggregation model	∘ FFBF service architecture is proposed ∘ An IDMS framework is proposed	∘ Manufacturing sector ∘ Healthcare ∘ Wearable devices
[96]		∘ Collection of data from various sources ∘ Aggregation of data ∘ Analysis of data	∘ Event Centric Model	∘ A data aggregation model is proposed ∘ Correlation model is proposed	∘ Maintenance Sector ∘ Interconnected Production Lines
[98,99]		∘ To handle device failure ∘ To provide interface for exchange of information and data	∘ Architecture based solution ∘ SDN ∘ CPS ∘ DRL	∘ Blockchain based DRL is proposed ∘ CCPSA is proposed ∘ Equipment utilization rate is increased	∘ Complex industrial applications ∘ Automated guided vehicles ∘ Industrial robots
[103]	Data Management	∘ To reduce the amount of data uploaded ∘ To provide intelligence to mobile edge nodes	∘ DRL for optimization ∘ FL for providing intelligence	∘ An "In-Edge AI" framework for improving deployment ∘ Deep Q-Learning for improving computation in edges	Xender's traces
[23]		∘ To provide knowledge fusion mechanism ∘ To provide a model for transfer learning	FL along with imitation learning	∘ Federated IL (FIL) is proposed ∘ A transfer learning approach using FIL is proposed	Self-driving cars
[69]		∘ To provide higher security while multiple parties try to collaborate	DL and FL along with blockchain	∘ A data collaboration framework is proposed using FDL ∘ A blockchain based secure mechanism for computation is proposed	∘ A private blockchain based on Libra protocol ∘ Raw data collected from wearable device
[70]		∘ To share data across multiple parties without any leakage	FL and blockchain	∘ Differential private FL is proposed	∘ Reuters data set ∘ 20 newsgroups data
[104]		∘ To optimize the structure of the model ∘ To minimize the computational cost ∘ To minimize the global model test errors	FL and multi-objective evolutionary algorithm	∘ Federated averaging algorithm is proposed ∘ A modified version of SET algorithm is proposed	Independent IID and non-IID data set
[105]		To provide better privacy in shared models	FL and MPC	∘ Two-phase MPC enabled FL is proposed ∘ Committee election algorithm is proposed	∘ Manufacturing unit ∘ Fault detection in electrical machines
[106,107]		To prevent exploitation of private data in sensitive applications	FL along with Augmented Learning	∘ Privacy enhanced FL is proposed for Industrial AI ∘ Aggregation decryption is utilized ∘ Model aggregation algorithm is proposed	∘ MNIST data set ∘ Healthcare and autopilot
[108]		∘ To provide intelligence to the edge network ∘ To prevent malicious attacks	Ubiquitous communication, blockchain along with DRL	∘ Credit Differentiated transaction approval scheme is proposed ∘ DRL for optimized resource scheduling	Five BS each having computing server and caching server
[109]		To optimize hyper-parameters in ML model	PSO along with FL	∘ PSO based parameter setting framework is proposed ∘ The number of hidden layers and neurons are optimized	Smart city traffic with City Pulse EU FP7 data set
[110,111]	Resource Management	∘ To optimize transmission power ∘ To reuse wireless resources	Hierarchical FL	∘ Closed form expression for computing convergence rate ∘ Optimal transmission power ∘ Minimized loss function	∘ Single BS scenario ∘ Heterogeneous cellular network consisting of MBS and SBS
[112]		To enable task offloading in edges	DL based FL	∘ Minimize cost due to delay, computation and energy consumption	Heterogeneous networks in edge computing
[113]		To optimize resource in cognitive IoT	Dispersed FL	Integer linear programming to minimize overall cost	Smart industry
[114]		∘ To evaluate ML model ∘ To provide structured workflows	IFL	∘ Model for adapting to diverse operating conditions ∘ Proposed IFL	∘ Synchronous framework ∘ Centralized FL
[115]		To mitigate unreliable communication	Blockchain based FL	Learning accuracy is balanced	DTWN
[116]		To provide dynamic resource allocation	DFRL	MAQL for dynamic slicing is proposed	Multi-industrial IoT
[117]		To perform image classification in exploration domain	Distributed FL	∘ GFC is deployed ∘ Communication cost is reduced ∘ Weighted zero forcing transmit precoding	UAVs

FL for optimizing the usage and allocation of resource in wireless applications. This work constructed a heterogeneous cellular network which consisted of a macro base station (MBS) and few small cell base stations (SBS).

Huang et al. [112] proposed a novel approach named Deep-Q for the purpose of enabling resource allocation and task offloading for networks in edge computing. This work focused on offloading multiple tasks at the same time to the edge servers. The proposed technique helped in minimizing the cost due to delay, computation and energy consumption. In [113], Khan et al. proposed a novel dispersed FL to be employed in smart industries. This work focused on optimizing the resource in cognitive IoT based smart industrial applications. This work proposed a model which used distributed and hierarchical FL for offering optimized resource communication and robustness. Hiessl et al. [114] tried to adapt the available traditional FL for industrial environments and named it as Industrial FL (IFL). This work provided a collection of structured workflows for the requirements in the architecture of IFL. To get adapted to the diverse operating conditions of the industrial environment, the FL clients are not allowed to exchange the ML model parameters with the FL participants in the IFL architecture.

In [115], Lu et al. introduced the usage of blockchain based FL in the digital twin wireless networks (DTWN). The methodology helps in improving the reliability, security and data privacy by which the learning accuracy is balanced using the bandwidth allocation process. Messaoud et al. [116] focused on developing a FL based methodology to support the resource allocation strategy in multi-industrial IoT. This work proposed a novel deep federated reinforcement learning (DFRL) scheme for providing dynamic resource allocation and network management for the IIoT networks. The scheme provided IIoT slice based resource allocation with the help of a novel proposal Deep Federated Q Learning (DFQL). This work also proposed approaches namely Multi Agent deep Q-learning(MAQL) that dynamically slices

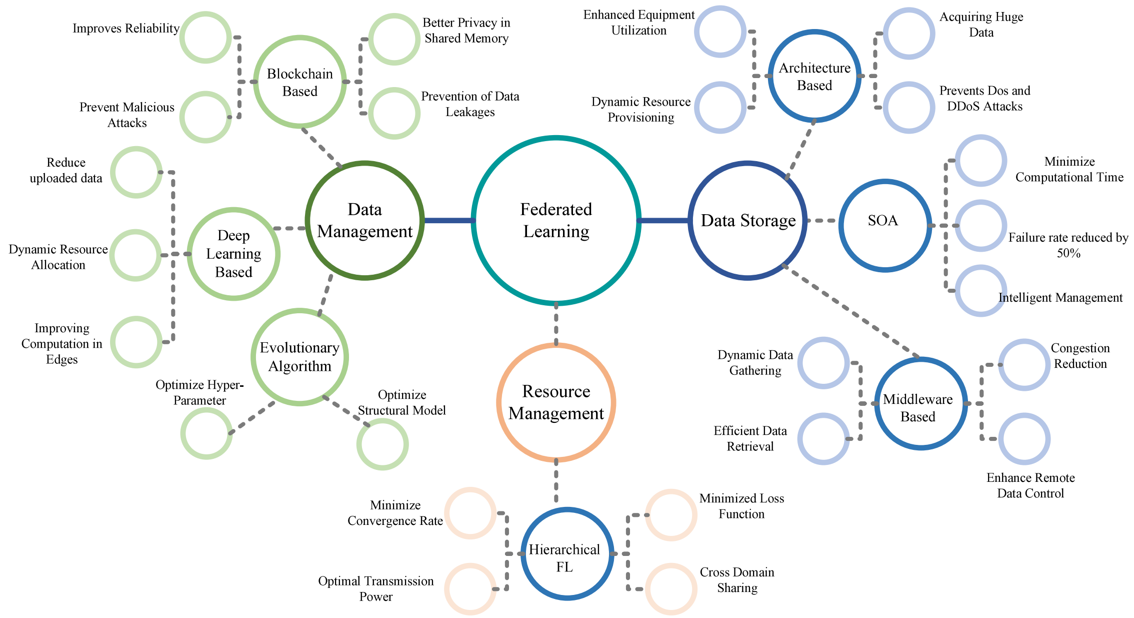


Fig. 4. Illustrative Usage of FL in IIoT for data/resource management.

the transmission power and spreading factor for maximizing the QoS requirements namely throughput and delay.

Unmanned aerial vehicles (UAVs) play a promising role in IIoT applications, but the characteristic of UAVs is that they have a limited payload and lesser flight time [120]. In [117], Zhang et al. investigated the UAVs for performing image classification tasks and proposed FL based multi-UAV systems. The image classification task is completed using the distributed FL. Multiple UAVs in the application area are co-ordinated with the use of ground fusion centre (GFC). Results show that the accuracy of image classification by UAV is high and communication cost is reduced relatively. Several resource allocation schemes have been recently carried out in UAV-enabled FL networks. For example, the work in [121] proposed to make use of the UAV to assist the FL process in the areas where the terrestrial computing infrastructure (e.g., macro base stations and edge computing servers) is not available. In particular, a UAV is deployed to wirelessly power energy-limited IoT/IIoT devices on the ground (i.e., aerial energy source) and perform model aggregation in the air (i.e., aerial learning server). The solution to this work was obtained by an efficient iterative algorithm, which was shown to be superior to several benchmarks that only optimize a subset of optimization variables. Moreover, this work was recently extended in [122] to consider a more practical energy harvesting model where the harvested power is modeled as a non-linear function. However, the aerial learning server (e.g., UAVs) and IIoT devices typically have limited battery capacities, thus demanding more sustainable solutions in which the devices can harvest external energy to perform the FL process and learning tasks. A few studies have considered this issue in UAV-enabled IoT/IIoT networks. For example, the work in [123] considered that a UAV is dispatched as the aerial learning server, and IoT devices are powered by renewable energy. Taking the system dynamics and the long-term energy limitation of the UAV into consideration, a DRL-based algorithm was developed with superior performance over the two baselines and the scheme using proximal policy optimization.

To summarize, resource optimization deals with optimizing the communication as well as computational resources required for enabling the FL systems to locally compute, communicate and update the global parameters. When these are optimized, the learning cost of FL model is significantly reduced. Various research contributions discussed in the section give solutions for reducing the computation cost and improving the resource allocation process. It is expected that IIoT would transform the way human beings live, work or play. The future of IIoT varies from basic automation to factory automation, high level of connectivity to wearable devices and smart applications [124]. In the near future, there will be a variety of applications which would be a part of our daily activities. There will be huge volume of data acquired by networks, upload and download of data via communication channel would be high. To handle these data, various solutions are proposed by recent researchers with respect to data storage, data management and resource management and is summarized in the Table 3. The usage of FL for data/resource management and data storage in IIoT is illustrated in Fig. 4. The data generated by these smart applications would be heterogeneous data and these data can be handled best using the convergent over-the-air (COTA) FL proposed in [125], where the updates of the model are transmitted in the form of analog signals over the multiple access channel. The centralized server would process the analog signals received by superposition principle. This methodology reduces the communication as well as overhead cost.

5. Applications

IIoT plays a significant role in our day to day modern life in various applications almost in every vertical such as smart automotive, smart industries, smart agriculture, smart energy, and smart healthcare. [126,127].

Algorithm 1 summarizes the detailed step-wise procedure of integrating FL with the IIoT applications as discussed in the following subsections: (i) selection of client nodes, (ii) model distribution, (iii) Uploading locally trained models, (iv) Federated aggregation process, (v) Model testing and update (see [128,129]). As described in Algorithm 1, FL models are being trained by edge IIoT devices which trains data generated by edge clients and the data fetched from industry partners. These models are regularly updated and sent to global cloud for aggregation and generating a new global ML model which is reliable and efficient.

5.1. Federated learning enabled IoT in smart automotive systems

Recently, in automobile industries, vehicles that are connected to IoT provide constant and simultaneous access to information to drivers and occupants while moving. However, the number of vehicles connected to each other over IoT keeps on increasing [131] which thrives

Algorithm 1: Algorithm for integrating FL with IIoT applications

Begin: The detailed step-wise process of integrated FL with IIoT applications is described below.

Step 1 - Selection of client nodes: In FL integrated IIoT applications, various edge clients are connected to other smart devices. The centralized cloud server collects and stores data generated from connected edge devices. The selection of edge client is done by the centralized cloud server for training the global ML model.

Step 2 - Model distribution This initial model is then distributed to edge clients for the purpose of updating and training the ML model based on clients' local raw data.

Step 3 - Uploading locally trained models These initial models are no longer initial as they are trained by the clients using their local raw data. The updated model is then uploaded to a centralized cloud server for federated averaging.

Step 4 - Federated Aggregation Process Having all the updated models by local edge clients, the central task executioner averages all these models to develop a new version of the global ML model which is better in terms of efficiency and accuracy than the initial global ML model. For averaging, FedAvg [128] is the best algorithm to solve this averaging problem yielding a well trained global model.

Step 5 - Model Testing and update This aggregated global model is then tested by the central task executioner with the data that belongs to other edge clients that have not participated in this training. Analyzing the testing results, some parameters are tuned to repeat the training process if needed, else the model is updated accordingly and distributed to all the edge clients.

Repeat Step 2 to Step 5 The above process is repeated until the global ML model attains the required accuracy [129]. Based on this, Bonawitz et al. [130] recommend that all local updates are aggregated securely and globally by using their weighted average. In this way, efficiency in FL-IIoT applications can be achieved with low cost.

End

for smart car emerging technologies with new requirements such as secure, robust and roadside infrastructures etc, which transforms the original concept of vehicular ad-hoc networks (VANETs) [132–134] into a new concept called the FL enabled Internet of Vehicles (IoV) [135]. Manufacturing of such smart vehicles in automotive industry must be capable of employing IIoT along with ML techniques paving the way for smart industry experience enabling efficient and sustainable production [136]. Computers, people, and machines coming together to execute industrial operations by capitalizing advanced data driven algorithms for groundbreaking business results is termed as IIoT. IIoT provides a way to transform business operations and processes by using advanced analytics to query large data sets. These gains results in optimized efficiency of productions, and thereby increasing profits. This connectivity allows for data collection, exchange, and analysis, potentially facilitating improvements in productivity and efficiency [10,137]. As high number of devices and machines are connected, classical cloud-based learning approaches are impractical and inadequate. But with FML enabled IIoT, privacy of the data is taken into consideration which further enhances the performance of IIoT devices and also accelerates the learning processes of the data which technique you are talking or building the base.

The IIoT devices are also becoming the main targets of malicious attacks leading to disastrous consequences. These attacks misleads the global models to predict undesirable outputs [138]. Various counter methods have been proposed such as ensemble diversity [139], Purifying Variational AutoEncoder (PuVAE) [140] and other training methods such as adversarial training, but none of them proved to be satisfactory as they focus only on a particular type of attacks. Alternative to these methods is to train the IIoT data locally in isolation to other devices but with this approach, resources required will be very high. On the other hand, with limited computation ability and local memory and insufficient data samples can result in further degrading the model. So to solve the issue of availability of IIoT device data in isolation and maintaining privacy of this data, a more advanced ML approach is developed which is termed as *Federated ML* [17].

FL enabled smart automobile industry has various automation systems, IIoT devices for storing corresponding data, control systems and various operating machines which are connected to each other through the internet. FML models are deployed to edge devices which updates the model based on demands and also check for any anomalies in production [141].

In smart automobile factory, automated machines and control systems drives the raw material used for production by communicating with each other cooperatively. Smart manufacturing enables factory managers to automatically collect and analyze data to make more informed decisions and optimize production in reduced costs. All the decisions depends upon harnessing of data and based on FL models, data will decide what to do and when to do a particular task. A graphical representation of FL and IoT-enabled smart automotive systems is depicted in Fig. 5.

5.2. Federated learning enabled robotics and IoT in smart industries

The concept of the smart industry comes from the rapid conversion of manufacturing tasks to intelligent manufacturing processes using technologies such as Artificial Intelligence of Internet of Things (AIoT). Advanced machine learning and deep learning methodologies along with IoT play an essential role in handling large amounts of data generated from industrial machines as part of routine production processes such as process modeling, manufacturing, maintenance, control, etc. The throughput of advanced AI algorithms has strong dependability on the volume and training of data. Due to privacy and security issues, it is not advisable to share a large amount of sensitive information over the industrial networks for AI methodologies. Integrating FL with IoT and AI allows smooth communication between industrial systems without leaking sensitive information or any kind of heavy data exchange over industrial networks [12,142]. In recent times, the integration of FL with robotics and automation, industrial edge-based IoT, and Industry 4.0 has envisioned the concept of smart industries. Robotics play a crucial role in performing various production tasks of industrial automation systems. However, industrial robotic systems suffer from the issues such as privacy, real-time data processing, etc. The FL can assist in resolving these issues by distributing intelligence among robotic systems instead of depending on the centralized server for processing. FL can employ AI methods in the local robots without any kind of latency issues using methodologies such as differential privacy. The FL can also assist in accumulating sensitive information from various robots to build a powerful AI methodology to improve the performance and efficiency of robots. Recently, it has been observed that technologies such as edge and fog computing are applied in Industrial IoT networks for pervasive and ubiquitous computing-based applications [143]. The primary objective of integrating edge/fog computing with IIoT applications is to mitigate the delay and smoothen the decision-making process. To achieve efficient communication and networking among IoT nodes, FL has been applied to enhance the distributed learning capacity and capabilities of IIoT edge-enabled devices.

5.3. Federated learning enabled IoT in smart agriculture

In general, smart agriculture discusses the concept of deploying low-cost agricultural sensors, actuators, and systems to automate the production processes of agriculture. The IoT-enabled edge devices capture a huge amount of data via open access channels such as the

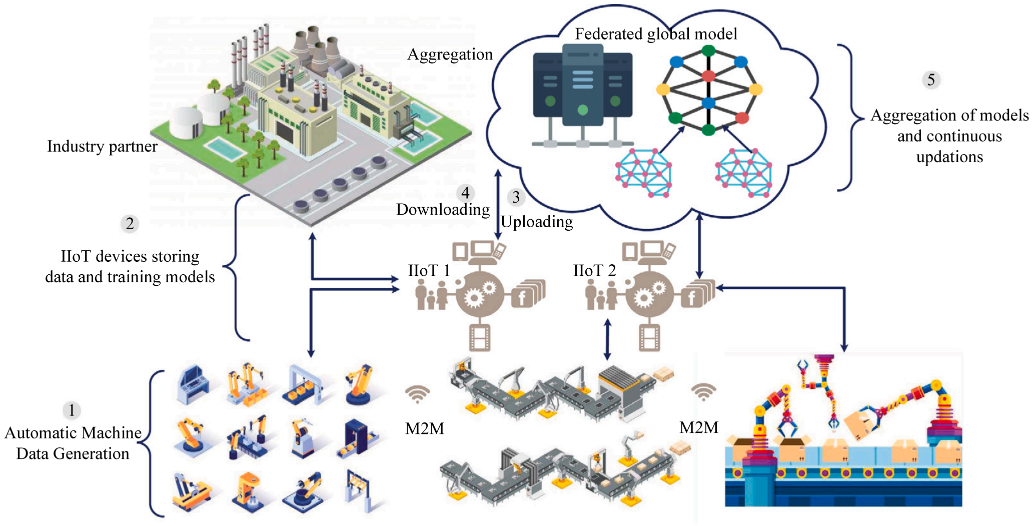

Federated global model
Aggregation
Industry partner
5
Aggregation of models and continuous updations
2
IIoT devices storing data and training models
4
Downloading
3
Uploading
IIoT 1
IIoT 2
1
Automatic Machine Data Generation
M2M
M2M

Fig. 5. A graphical representation of FL-enabled IoT in smart automotive systems.

Internet. Even though the employed IoT-enabled sensing devices enhance the capability of traditional agriculture, the captured raw data by IoT-enabled edge devices also get exposed to security and privacy issues due to its open-access nature [144]. The FL systems can also assist in improving the process of accurate selection and classification of crops. The security and privacy of sensitive spatial, temporal, and spectral information of crops are a serious concern for farmers. Generally, farmers are not open to sharing such sensitive information among agricultural communities and also hesitant to participate in technology-related discussions. Furthermore, for the surveillance of agriculture processes, and better decision-making processes, FL-enabled IoT systems can play a vital role by ensuring the security and privacy of various spatial and temporal remote sensing information during the agricultural monitoring process. The FL has the potential to train IoT-enabled edge servers locally for a particular farm owner, train them using the raw data collected by the agriculture IoT devices, to build a powerful AI algorithm. The build prediction model can be shared with multiple data owners along with ensuring the privacy of the locally trained agricultural data.

5.4. *Federated learning enabled IoT in smart energy*

In recent times, there has been the deployment of smart meters to ensure accurate measurement of the electricity consumption data. The smart metered data is regularly recorded at a specific time interval and location as per the need and requirements of the government authorities. Furthermore, the captured fine-grained electricity data can play a vital role for electricity companies to understand the customer's lifestyle, behavior, consumption, and needs. From the electricity company's perspective, the captured smart metered data can assist the companies in forecasting customer consumption and designing related tariffs [145]. Some fellow researchers have some smart metering approaches to analyze consumption behavior using various AI models. However, the privacy of fine-grained electricity data is a major concern for electricity companies. The competitive electricity companies are also not willing to share their consumer's data with their rivals. The European Union(EU) has introduced new regulations for the privacy of consumers' electricity data [146].

To address the issue of privacy, FL systems can be a blessing for electricity companies who are worried about the leaking of their consumer's fine-grained electricity data. FL integrated IoT systems can collect the electricity data from multi decentralized clients(smart meters), train them locally to preserve privacy, and summarize the sensitive electricity information of local smart metered nodes to build the global AI model for electricity vendors.

5.5. *Smart health care industry*

Healthcare industry is another area that we expect to benefit from the rising ML technique of FL. As the population is growing at alarming rate, there can be seen an alarming pressure on hospital staff and quality of service being delivered to patients by staff and medical practitioners. Thus, new efficient and reliable solutions are required from science and technology [147,148]. In the medical field, smart Industry with IIoT and ML techniques has played vital role providing wide range of applications in creating customized implants, tools and medical equipment. Smart healthcare industry improves the healthcare infrastructures and biomedical systems keeping in view the safety of patients, their vitals and real time monitoring. Such automated systems manufactured by smart medical industry make smart decisions through real-time communication and cooperation with humans, machines, sensors, and so forth. [149]. Medical Industry 4.0 manufactures smart devices which provides quality care to patients while addressing the issue of shortage of medical staff and beds especially in populated hospitals. IIoT has been widely identified as a potential solution to alleviate the pressures on healthcare systems [150] by assisting manufacturing and supply chain of medical devices. Remote-healthcare systems can monitor patients at home without any need of them visiting infirmary or hospitals. In essence, it can improve access to medical resources, while reducing the pressure on the medical system, and it can allow people to better control their health. Health systems have improved a lot from the earlier times, thus increasing average life expectancy of the individuals. As a result, there is increased population of elderly people living with chronic diseases, heart diseases, diabetes etc, thus placing huge demands of healthcare services. In order to address this issue, FL enabled IIoT systems are developed which can learn patients medical data such as day-to-day health, disease symptoms, medical reports and vitals and this data is shared with medical practitioners to subscribe them appropriate medication at their residence itself. There are a lot of IIoT based healthcare systems that focus on particular diseases and in-general monitoring of patients health.

6. Challenges and solutions

The IIoT based systems are generally heterogeneous and the structure is also too complex in nature. Though these systems are advantageous, there are so many technical challenges in implementing and deploying these types of applications in real time. When these challenges are met by providing solutions, these systems will surely rule the world in near future. Table 4 discusses the summary of challenges and future directions of FL-enabled IIoT systems.

Table 4
Summary of challenges and future directions for FL-enabled IIoT systems.

Challenges	Description	Future Directions
Effective Data Management	In FL-enabled IoT systems huge data is stored in heterogeneous devices, gateways, edge servers, fog servers, etc [151].	Efficient data management model has to be designed with efficient storage mechanism [152].
Heterogeneity and Interoperability	Integrating multiple heterogeneous vendor systems such as machines robots, etc is a major challenge.	Optimized resource sharing techniques can be developed [153].
Server-side attack	Preserving privacy and security of data is a major issue while communicating the model updates.	A fully homomorphic encryption algorithm can be designed to overcome the privacy issues while communicating the updates to the global model [154].
Optimization	When DL models are used, optimization is a major challenge due to limited memory, and computational power.	MPC in FL using secure aggregation protocol can be used.
Inference attack	The hacker can infer sensitive information by the results of authorized query.	Specialized policy for unauthorized users can be designed [155].
Client-side attack	The concern of security and privacy issues on the client side remains a challenging task in FL-enabled IIoT systems [156–158]	Integration of FL and Blockchain in IIoT applications can improve the security on client side [130].
Data privacy and leakage	Data privacy is a major concern in FL-enabled IIoT systems.	The design of effective and efficient trust and data privacy framework can solve the data leakage and privacy issues [50,70].
Communication overhead	Training in FL is a recursive process that involves intensive sharing of parameters in FL-enabled IIoT systems.	The use of gradient-descent FL methodology can be deployed [159].
Data sharing in horizontal FL	Due to availability of limited data samples,it is a challenging task to do data sharing in FL-enabled IIoT systems.	The Federated Tensor Mining to federate multisource data can be utilized [49].
Anomaly Detection	The failure of edge devices can impact the overall operation of IIoT production processes.	FL-enabled deep anomaly detection model can be developed [160].
Data Integrity	Due to data heterogeneity, it is a difficult task to maintain data integrity in FL-enabled IIoT systems.	The design of the FL-enabled Blockchain architecture can reduce the data integrity issues [161].
Data and Model attack	There exists a high possibility of attacks on changing the data labels used for training and in the model updates communicated at different levels of FL-enabled IIoT systems.	The design of federated defensive system can reduce the possibility of attacks [162].
Automation of Industrial processes	To automate and streamline industrial production processes, it is essential to reduce human efforts which is a big hurdle for achieving 100% automation in smart industries.	Integration of Cognitive Internet of Things in the routine industrial production processes can minimize the human efforts and intervention, and leakage of data [113,163].

6.1. *Effective data management*

The IIoT systems are highly heterogeneous in nature, huge in volume and also at times unstructured. The data gathered are usually streams of data at high velocity. Currently these huge streams of data are stored in heterogeneous devices, gateways, edge servers, fog servers or cloud servers. The huge data gathered is raw and needs to be processed before taking any decisions in real-time [151]. After processing, the processed data needs to be transmitted to the required destination in an efficient manner. The data once archived needs to be retrieved when requirement arises and hence the data has to be available at the right time for further analysis. While the data is archived, then arises the challenge for safe storage. Though IIoT is the buzz term in the recent days, most of the industries are still in a dilemma to implement these systems in real-time due to the above mentioned challenges.

Solution: To cope up with these challenges, an efficient data management model [152] has to be designed and utilized. The data model designed has to enable better storage mechanisms while archiving as well as sharing the sensitive data. It should also provide an efficient management scheme for handling the huge volume of data like processing, analyzing and retrieving the data at a higher speed.

6.2. *Heterogeneity and interoperability*

IIoT systems are typically comprised of heterogeneous multiple vendor systems [164] such as machines, robots, IoT devices, wearable sensors, actuators, networks which are either wired or wireless, 5G based cellular networks, broadband networks, storage servers which may be cloud servers or edge servers. Integrating these multiple technologies to build a whole system is a challenge.

Solution: The researchers need to concentrate on developing resource sharing techniques [153], synchronizing techniques and data sharing techniques to improve the interoperability [165] and collaboration.

6.3. *Server side attack*

Preserving privacy of data and model is major issue in decentralized framework. Security mechanism must be applied on Edge to end device communication and from edge to collaborative model. Even in worst case, the model updates can be tampered between end devices to collaborative engine. In FL, the updates from end devices are sent to the cloud for training the global model. However there is a high chance of stealing the updates from cloud by the malicious user which results in increase of inconsistency and noise in the model.

Solution: Fully Homomorphic Encryption (FHE) is best to overcome privacy threat challenge in the cloud. All operations can be encrypted using FHE except the activation functions. Even for activation function, higher degree polynomials or modified Chebyshev polynomial can be applied and then can be encrypted using FHE. Recently Phong et al. [154] designed a model that balances the trade-off between data privacy and accuracy. It also ensures no information is leaked from the cloud server.

6.4. Optimization

As the end devices have limited memory and computational power, DL models must be optimized so that they can be trained and deployed on IIoT or end devices efficiently. Optimizing the DL and ML models is an essential part to execute the models at the edge devices which remains a challenging task for the researchers even now. The end device have only limited power and resource to train the local model as they are shared by the other devices. Practically, it is difficult to provide a static connection for all end devices. So, we have to reduce the computational complexity and communication overhead in FL. Therefore, there should be a tradeoff between power consumption and resource allocation to train the model.

Solution: The main solution for the above challenge is to use MPC in FL by using secure aggregation protocol. The best possible approach would be to use the FHE based MPC that can be implemented with limited rounds. Generally FHE provides the confidentiality and privacy of the updates but threshold FHE can reduce the communication overhead.

6.5. Inference attack

Inference attack is an attack where the hacker can infer sensitive information by the results of authorized query. Information can be inferred directly from the model or industry from the frequent queries. Some of the privacy information or identification of the user can also be inferred from the model updates. Inference attacks can be inferring class representatives, membership, properties, training inputs and labels. There are two types of inference attacks namely identification attack where the users' personal information is leaked and matching attack which compares the two model updates [166].

Solution: The best way is to utilize differential policy by adding noise to the information such that the unauthorized user cannot distinguish the original information from noise. Therefore there should be a balance between the privacy of data and accuracy of the model. This can be done by choosing the parameter to balance the above two parameters. Orekondy [155] stated that domain specific data augmentation can provide effective results with minimal impact.

6.6. Client side attack

In general, privacy security mechanisms are computationally expensive techniques that results in delayed response. Security issue on server and client side remains a challenging task even when the data communication is restricted. Some mechanisms namely DP and homomorphic encryption technique controls the amount of data to be shared on the cloud. Model poisoning is one of the dangerous attack which introduces backdoor functionality [156] from the client side and poison the client model. Some of the popular backdoor attack models are label flipping attack [157] and Sybil attacks [158]. Based on the occurrence of their scenario, the complexity of the attack differs. One of the client device is compromised and the local model is trained using the backdoor data with their new techniques and submits the resulting model. Finally this poisoned client model is aggregated with other client model and replace the global model with the backdoor model.

Solution: To prevent backdoor attack combined anomaly detection algorithm, participant level DP and Byzantine-tolerant gradient descent can be used. [130] discussed about the secure aggregation as the updates from each end device are not visible to the aggregator. There is no one single solution to overcome the model poisoning. So it is difficult to design a best solution for various poisoning techniques. FedAvg one of the DP algorithm trains deep models on user partitioned data and provides required level of privacy for each participant.

6.7. Data leakage

Data leakage is one of the major barriers to data sharing in a distributed channel. Researchers are more focused on data leakage as sensitive data breaches can lead to increased risk factors like financial loss, online vandalism and future security costs. Providing secure, smart data transmission [167,168] in a distributed environment is a major concern.

Solution: Integrating FL and blockchain into IIoT can adequately improve data stabilization and data quality. Lu et al. [70] proposed a blockchain-based FL model that enables data privacy in IIoT applications. The research aims to enhance trust management while transferring sensitive data from source to destination. The researchers pay special focus by establishing a blockchain environment to resist data leaks and inhibit full access control of its owner. During this process, FL models are incorporated to conceptualize data models. However, smart mechanisms for enhancing the data utility and expanding the amount of resources for effective data sharing in a distributed setup should be incorporated.

6.8. Data privacy

Data privacy is one of the key issue when FL is used to train static streams and data streams. In general, privacy can be classified as global and local privacy. The model updates generated in each iteration on all the devices are not available to untrusted third parties other than the centralized server, whereas in local privacy, the updates are private to the server.

Solution: In recent work [50], a PriModChain framework is proposed for the implementation of trust and data privacy in IIoT. The proposed model implemented data privacy with a view to enhance the security and safety of the IIoT. The Smart Contract helps in providing transparency within the IIoT framework whenever a communication link is formulated between the distributed bodies and the central authorities. An interplanetary file system is used to achieve optimum latency, immutability, efficient data archiving and secure information sharing between the P2P network. Furthermore, the framework does not achieve optimum latency and thus curtails its use in the IIoT system. To improve resource optimization and data privacy for smart industries, the work in [113] unveils a dispersed federated approach for resource optimization and reliability. During this process, the proposed model aims to minimize FL costs, while optimization is achieved by a decomposing and relaxation-based algorithm. Later [113] indicated the convergence of sub-problems and remedied them using a stochastic optimization technique.

6.9. Communication overhead

Training in FL is a recursive process that involves intensive sharing of parameters among clients and servers. Consequently, to adopt FL to resource-constrained IIoT, minimizing communication overhead is a key issue requiring in-depth investigation. However, some works have been undertaken to mitigate the use of resources in FL by lowering the load distribution factor between source and destination.

Solution: The work in [159] presented gradient-descent FL that involves local updates and global convergence measures. In order to accomplish the desired exchange between local update and global accumulation, a control algorithm is introduced to significantly reduce the loss function in the context of budgetary resource concerns. Meanwhile, researchers in [169] exemplify ineffective client training with fairly low resources. The allocation of resources, data training and data authentication must be considered in order to improve the communication cost in FL.

6.10. *Data sharing in horizontal federated learning*

Industrial awareness is commonly observed using DL and data mining techniques. However some data is not easy to acquire one factory's information, as there are still few samples. If different alliance factories are able to gather their information together a lot more information could be extracted. In addition, security of information is a key issue for these factories. Traditional matrix-based techniques can ensure information security within a factory, but will not facilitate information sharing between factories, and therefore, due to lack of correlation, mining efficiency is weak.

Solution: A novel Federated Tensor Mining (FTM) approach is introduced in [49] to federate multisource data for tensor-based mining while ensuring security. The key contribution of FTM is that every factory only needs to share its ciphertext data and these ciphertexts are adequate for tensor-based knowledge mining due to its homomorphic attribution. Real-data-driven simulations demonstrate that FTM not only mines the same knowledge compared with the plaintext mining, but also is enabled to defend the attacks from distributed eavesdroppers and centralized hackers.

6.11. *Anomaly detection*

Since the failure of edge devices (i.e. anomalies) has a serious impact on the production process in IIoT, the relevant and reliable detection of anomalies is becoming extremely important. In addition, the data collected by the edge device contains tremendous private information which has to be secured.

Solution: To address this issue, Liu et al. [160] developed a communication-efficient FL-based deep anomaly detection model for sensing IIoT data. During this process, a FL model was first introduced to allow decentralized edge devices to train an anomaly detection model, which could enhance its classification accuracy. In the second step, an Attention Mechanism-based Convolutional Neural Network-Long Short Term Memory (AMCNN-LSTM) approach is formulated for accurate detection of anomalies. The AMCNN-LSTM method uses attention mechanism-based CNN units to identify relevant fine-grained features, to prevent memory loss and gradient dispersion risks. In the third step, a gradient compression technique to improve communication process is proposed. Liu et al. [170] proposed a CNN-LSTM FL-based framework for detecting anomalies in IIoT applications. To improve the data transmission efficacy of the framework, a gradient compression technique based on top-k selection has been used to minimize network overhead.

6.12. *Data integrity*

Blockchain technology is used in IIoT-based FL to provide data integrity. Blockchain-based FL systems face some of the challenges in the design phase to address the barriers of data heterogeneity in IIoT during earlier fault detection. However, existing security mechanisms do not focus on designing an effective framework and blockchain's reliability issue.

Solution: In [161], a new FL-based blockchain architecture is proposed to detect failures in IIoT applications. During this process, each client server continuously build a Merkle tree where each leaf node signifies the client data and saves the root tree to the blockchain. An on-chain system is designed to measure each client's impact based on the size of client data used in local model training.

6.13. *Data and model attack*

Even though data sharing is minimal in FL, there can be several attacks in the framework at different levels. Attacks can happen at any stage starting from the basic node level to the server level. They are as follows:

Data based attack: Changing the data or labels used in the training data falls under this category of the attack. Data attack can be clean label or dirty label attack. When unauthorized person cannot change the label of data it is known as clean label attack. In contrast in dirty label attack, the intruders attempt to include some number of samples which would mislead the ML process.

Model based attack: Changing the model updates communicated would cause the decrease in model efficiency. Hacking the model trained in local node results in leaking user identification data globally. Aggregated global model present in the centralized server is also prone to attack.

Solution: In order to provide solutions to these kinds of problems, Song et al. [162] developed federated defensive system for cloud-based IIoT architectures. The revised FL model and the implemented adversarial learning loss function, can

metabolize DNNs to handle existing adversarial threats. The obtained results were compared with two popular benchmarks and the results reveal that the improved method can not only increase overall protection against various existing adversarial attacks, but can also accurately detect the improper behavior of DNN induced by new attacks.

6.14. *Automation of industrial processes*

To automate and streamline industrial production processes, it is essential to reduce human efforts which is a big hurdle for achieving 100% automation in smart industries. In addition, the adaptability of the concept of centralized machine learning in smart industries foresees tough integration challenges due to data privacy-related restrictions, and leakage of sensitive industrial information [171].

Solution: Even though it is difficult to achieve 100% automation, Lim et al. [163] and Khan et al. [113] proposed to integrate the Cognitive Internet of Things in the routine industrial production processes to minimize human efforts and intervention. The Cognitive IoT employs and integrates cognitive computing methodologies with industrial automation systems to resolve the issue of leakage of information and assures trust and security to industrial automation systems.

7. Conclusion

The new paradigm of integrating FL on IIoT data was surveyed and several approaches and techniques associated with it are introduced. In specific, we highlighted the characteristics and benefits of IIoT in terms of FL and distributive learning. The motivation behind the integration of IIoT with FL is also discussed. Subsequently, several privacy preserving ML/DL and blockchain models employed in FL and IIoT are presented. Then, survey on algorithms to handle heterogeneous data is summarized. Contribution of researchers in the field of FL with data and resource management followed by several challenges and solutions are discussed. Reviewing automobile and healthcare applications on IIoT with FL would provide the researchers to understand the principal components of IIoT smart devices. We also identified various challenges, solutions and future research direction in path of FL for IIoT. This paper would enhance the readers to consolidate and combine the information with respect to IIoT, FL and privacy preserving mechanism, data management and storage mechanisms under a common new paradigm of Industry 4.0.

In future, researchers should focus on developing algorithms and methodologies to impose privacy constraints across edge device and data level. The privacy preserving model for FL should be designed to learn multiple geographically distributed training data sets without exploiting the sensitive information of the user. The existing architecture violates the usage of short packets for low latency communication. Therefore new generation of wireless systems such as 6G cellular networks can be integrated with FL enabled IIoT networks to provide more importance for privacy preservation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Joung, Machine learning-based antenna selection in wireless communications, IEEE Commun. Lett. 20 (11) (2016) 2241–2244.

[2] H. Li, K. Ota, M. Dong, Learning IoT in edge: Deep learning for the Internet of things with edge computing, IEEE Netw. 32 (1) (2018) 96–101.

[3] N.C. Luong, D.T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, D.I. Kim, Applications of deep reinforcement learning in communications and networking: A survey, IEEE Commun. Surv. Tutor. 21 (4) (2019) 3133–3174.

[4] Q.-V. Pham, F. Fang, V.N. Ha, M.J. Piran, M. Le, L.B. Le, W.-J. Hwang, Z. Ding, A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art, IEEE Access 8 (2020) 116974–117017.

[5] J. Konečný, H.B. McMahan, F.X. Yu, P. Richtárik, A.T. Suresh, D. Bacon, Federated learning: Strategies for improving communication efficiency, 2016, arXiv preprint arXiv:1610.05492.

[6] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, F. Beaufays, Applied federated learning: Improving google keyboard query suggestions, 2018, arXiv preprint arXiv:1812.02903.

[7] J. Xu, B.S. Glicksberg, C. Su, P. Walker, J. Bian, F. Wang, Federated learning for healthcare informatics, J. Healthc. Inform. Res. 5 (1) (2021) 1–19.

[8] J.C. Jiang, B. Kantarci, S. Oktug, T. Soyata, Federated learning in smart city sensing: Challenges and opportunities, Sensors 20 (21) (2020) 6230.

[9] M.J. Sheller, B. Edwards, G.A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R.R. Colen, et al., Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data, Sci. Rep. 10 (1) (2020) 1–12.

[10] S. Niknam, H.S. Dhillon, J.H. Reed, Federated learning for wireless communications: Motivation, opportunities, and challenges, IEEE Commun. Mag. 58 (6) (2020) 46–51.

[11] Q.-V. Pham, N.T. Nguyen, T. Huynh-The, L.B. Le, K. Lee, W.-J. Hwang, Intelligent radio signal processing: A survey, IEEE Access 9 (2021) 83818–83850.

[12] D.C. Nguyen, M. Ding, P.N. Pathirana, A. Seneviratne, J. Li, D. Niyato, H.V. Poor, Federated learning for industrial internet of things in future industries, IEEE Wirel. Commun. 28 (6) (2021) 192–199.

[13] Y. Lu, Industry 4.0: a survey on technologies, applications and open research issues, J. Ind. Inf. Integ. 6 (2017) 1–10.

[14] W.Z. Khan, M. Rehman, H.M. Zangoti, M.K. Afzal, N. Armi, K. Salah, Industrial internet of things: Recent advances, enabling technologies and open challenges, Comput. Electr. Eng. 81 (2020) 106522.

[15] M.H. ur Rehman, I. Yaqoob, K. Salah, M. Imran, P.P. Jayaraman, C. Perera, The role of big data analytics in industrial internet of things, Future Gener. Comput. Syst. 99 (2019) 247–259.

[16] K. Tange, M. De Donno, X. Fafoutis, N. Dragoni, A systematic survey of industrial internet of things security: Requirements and fog computing opportunities, IEEE Commun. Surv. Tutor. 22 (4) (2020) 2489–2520.

[17] Q. Yang, Y. Liu, T. Chen, Y. Tong, Federated machine learning: Concept and applications, ACM Trans. Intell. Syst. Technol. (TIST) 10 (2) (2019) 1–19.

[18] W.Y.B. Lim, N.C. Luong, D.T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, C. Miao, Federated learning in mobile edge networks: A comprehensive survey, IEEE Commun. Surv. Tutor. 22 (3) (2020) 2031–2063.

[19] V. Mothukuri, R.M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, G. Srivastava, A survey on security and privacy of federated learning, Future Gener. Comput. Syst. 115 (2020) 619–640.

[20] S.A. Huda, S. Moh, Survey on computation offloading in uav-enabled mobile edge computing, J. Netw. Comput. Appl. (2022) 103341.

[21] R.U. Rasool, H.F. Ahmad, W. Rafique, A. Qayyum, J. Qadir, Security and privacy of internet of medical things: A contemporary review in the age of surveillance, botnets, and adversarial ml, J. Netw. Comput. Appl. (2022) 103332.

[22] S. Zeb, A. Mahmood, S.A. Hassan, M.J. Piran, M. Gidlund, M. Guizani, Industrial digital twins at the nexus of nextg wireless networks and computational intelligence: A survey, J. Netw. Comput. Appl. (2022) 103309.

[23] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, D. Niyato, Federated learning for 6G communications: Challenges, methods, and future directions, China Commun. 17 (9) (2020) 105–118.

[24] Z. Zhao, C. Feng, H.H. Yang, X. Luo, Federated-learning-enabled intelligent fog radio access networks: Fundamental theory, key techniques, and future trends, IEEE Wirel. Commun. 27 (2) (2020) 22–28.

[25] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, M. Guizani, Reliable federated learning for mobile networks, IEEE Wirel. Commun. 27 (2) (2020) 72–80.

[26] K. Yang, Y. Shi, Y. Zhou, Z. Yang, L. Fu, W. Chen, Federated machine learning for intelligent IoT via reconfigurable intelligent surface, IEEE Netw. 34 (5) (2020) 16–22.

[27] L.U. Khan, S.R. Pandey, N.H. Tran, W. Saad, Z. Han, M.N. Nguyen, C.S. Hong, Federated learning for edge networks: Resource optimization and incentive mechanism, IEEE Commun. Mag. 58 (10) (2020) 88–93.

[28] L.U. Khan, W. Saad, Z. Han, E. Hossain, C.S. Hong, Federated learning for internet of things: Recent advances, taxonomy, and open challenges, IEEE Commun. Surv. Tutor. (2021).

[29] A. Imteaj, U. Thakker, S. Wang, J. Li, M.H. Amini, A survey on federated learning for resource-constrained iot devices, IEEE Internet Things J. 9 (1) (2021) 1–24.

[30] M. Alazab, S.P. RM, M. Parimala, P. Reddy, T.R. Gadekallu, Q.-V. Pham, Federated learning for cybersecurity: concepts, challenges and future directions, IEEE Trans. Ind. Inf. 18 (5) (2022) 3501–3509.

[31] K. Raja, K. Karthikeyan, B. Abilash, K. Dev, G. Raja, Deep learning based attack detection in iiot using two-level intrusion detection system, 2021.

[32] R. Huo, S. Zeng, Z. Wang, J. Shang, W. Chen, T. Huang, S. Wang, F.R. Yu, Y. Liu, A comprehensive survey on blockchain in industrial internet of things: Motivations, research progresses, and future challenges, IEEE Commun. Surv. Tutor. (2022).

[33] O.A. Wahab, A. Mourad, H. Otrok, T. Taleb, Federated machine learning: Survey, multi-level classification, desirable criteria and future directions in communication and networking systems, IEEE Commun. Surv. Tutor. 23 (2) (2021) 1342–1397.

[34] Y. Ren, R. Xie, F.R. Yu, T. Huang, Y. Liu, Potential identity resolution systems for the industrial internet of things: A survey, IEEE Commun. Surv. Tutor. 23 (1) (2020) 391–430.

[35] J. Franco, A. Aris, B. Canberk, A.S. Uluagac, A survey of honeypots and honeynets for internet of things, industrial internet of things, and cyber–physical systems, IEEE Commun. Surv. Tutor. 23 (4) (2021) 2351–2383.

[36] W. Mao, Z. Zhao, Z. Chang, G. Min, W. Gao, Energy efficient industrial internet of things: Overview and open issues, IEEE Trans. Ind. Inf. (2021).

[37] B. Jiang, J. Li, G. Yue, H. Song, Differential privacy for industrial internet of things: Opportunities, applications, and challenges, IEEE Internet Things J. 8 (13) (2021) 10430–10451.

[38] S.P. Ramu, P. Boopalan, Q.-V. Pham, P.K.R. Maddikunta, T.-H. The, M. Alazab, T.T. Nguyen, T.R. Gadekallu, Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions, Sustainable Cities Soc. (2022) 103663.

[39] B. Brik, M. Messaadia, M. Sahnoun, B. Bettayeb, M.A. Benatia, Fog-supported low latency monitoring of system disruptions in industry 4.0: A federated learning approach, ACM Trans. Cyber-Phys. Syst. (2022).

[40] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, S. Guo, A survey of incentive mechanism design for federated learning, IEEE Trans. Emerg. Top. Comput. (2021).

[41] R.S. Antunes, C.A. da Costa, A. Küderle, I.A. Yari, B. Eskofier, Federated learning for healthcare: Systematic review and architecture proposal, ACM Trans. Intell. Syst. Technol. (TIST) (2022).

[42] P. Mathur, Overview of IoT and IIoT, in: IoT Machine Learning Applications in Telecom, Energy, and Agriculture, Springer, 2020, pp. 19–43.

[43] S.A. Khowaja, K. Dev, N.M.F. Qureshi, P. Khuwaja, L. Foschini, Towards industrial private ai: A two-tier framework for data and model security, 2021, arXiv preprint arXiv:2107.12806.

[44] S. Liu, C. Guo, F. Al-Turjman, K. Muhammad, V.H.C. de Albuquerque, Reliability of response region: A novel mechanism in visual tracking by edge computing for IIoT environments, Mech. Syst. Signal Process. 138 (2020) 106537.

[45] V. Priya, I.S. Thaseen, T.R. Gadekallu, M.K. Aboudaif, E.A. Nasr, Robust attack detection approach for IIoT using ensemble classifier, Comput. Mater. Continua 66 (3) (2021) 2457–2470.

[46] S. Yarradoddi, T.R. Gadekallu, Federated learning role in big data, iot services and applications security, privacy and trust in iot, trust, Secur. Privacy Big Data (2022) 28.

[47] T.R. Gadekallu, Q.-V. Pham, T. Huynh-The, S. Bhattacharya, P.K.R. Maddikunta, M. Liyanage, Federated learning for big data: A survey on opportunities, applications, and future directions, 2021, arXiv preprint arXiv:2110.04160.

[48] Y. Guo, Z. Zhao, K. He, S. Lai, J. Xia, L. Fan, Efficient and flexible management for industrial internet of things: a federated learning approach, Comput. Netw. 192 (2021) 108122.

[49] L. Kong, X.-Y. Liu, H. Sheng, P. Zeng, G. Chen, Federated tensor mining for secure industrial internet of things, IEEE Trans. Ind. Inf. 16 (3) (2019) 2144–2153.

[50] P.C.M. Arachchige, P. Bertok, I. Khalil, D. Liu, S. Camtepe, M. Atiquzzaman, A trustworthy privacy preserving framework for machine learning in industrial IoT systems, IEEE Trans. Ind. Inf. 16 (9) (2020) 6092–6102.

[51] L. Kuang, L.T. Yang, J. Feng, M. Dong, Secure tensor decomposition using fully homomorphic encryption scheme, IEEE Trans. Cloud Comput. 6 (3) (2015) 868–878.

[52] G. Raja, Y. Manaswini, G.D. Vivekanandan, H. Sampath, K. Dev, A.K. Bashir, AI-powered blockchain-a decentralized secure multiparty computation protocol for IoV, in: IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS, IEEE, 2020, pp. 865–870.

[53] T. Huynh-The, C.-H. Hua, Q.-V. Pham, D.-S. Kim, MCNet: An efficient CNN architecture for robust automatic modulation classification, IEEE Commun. Lett. 24 (4) (2020) 811–815.

[54] N. Deepa, Q.-V. Pham, D.C. Nguyen, S. Bhattacharya, B. Prabadevi, T.R. Gadekallu, P.K.R. Maddikunta, F. Fang, P.N. Pathirana, A survey on blockchain for big data: Approaches, opportunities, and future directions, Future Gener. Comput. Syst. 131 (2022) 209–226.

[55] R.U. Khan, X. Zhang, M. Alazab, R. Kumar, An improved convolutional neural network model for intrusion detection in networks, in: 2019 Cybersecurity and Cyberforensics Conference, CCC, IEEE, 2019, pp. 74–77.

[56] Y.E. Wang, G.-Y. Wei, D. Brooks, Benchmarking TPU, GPU, and CPU platforms for deep learning, 2012, arXiv preprint arXiv:1907.10701.

[57] H.-D. Cho, P.D.P. Engineer, K. Chung, T. Kim, Benefits of the big, LITTLE Archit. (2012).

[58] X. Zhang, X. Chen, J.K. Liu, Y. Xiang, Deeppar and deepdpa: Privacy preserving and asynchronous deep learning for industrial IoT, IEEE Trans. Ind. Inf. 16 (3) (2019) 2081–2090.

[59] X. Wang, C. Wang, X. Li, V.C. Leung, T. Taleb, Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching, IEEE Internet Things J. 7 (10) (2020) 9441–9455.

[60] Y. Liu, J. James, J. Kang, D. Niyato, S. Zhang, Privacy-preserving traffic flow prediction: A federated learning approach, IEEE Internet Things J. 7 (8) (2020) 7751–7763.

[61] Y. Chen, X. Sun, Y. Jin, Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation, IEEE Trans. Neural Netw. Learn. Syst. 31 (10) (2019) 4229–4238.

[62] J. Song, W. Wang, T.R. Gadekallu, J. Cao, Y. Liu, Eppda: An efficient privacy-preserving data aggregation federated learning scheme, IEEE Trans. Netw. Sci. Eng. (2022).

[63] W. Wang, H. Xu, M. Alazab, T.R. Gadekallu, Z. Han, C. Su, Blockchain-based reliable and efficient certificateless signature for iiot devices, IEEE Trans. Ind. Inf. (2021).

[64] M.K. Hasan, M. Akhtaruzzaman, S.R. Kabir, T.R. Gadekallu, S. Islam, P. Magalingam, R. Hassan, M. Alazab, M.A. Alazab, Evolution of industry and blockchain era: Monitoring price hike and corruption using biot for smart government and industry 4.0, IEEE Trans. Ind. Inf. (2022).

[65] B. Deebak, F.H. Memon, K. Dev, S.A. Khowaja, W. Wang, N.M.F. Qureshi, Tab-sapp: A trust-aware blockchain-based seamless authentication for massive iot-enabled industrial applications, IEEE Trans. Ind. Inf. (2022).

[66] B. Deebak, F.H. Memon, S.A. Khowaja, K. Dev, W. Wang, N.M.F. Qureshi, C. Su, Lightweight blockchain based remote mutual authentication for ai-empowered iot sustainable computing systems, IEEE Internet Things J. (2022).

[67] B. Deebak, F.H. Memon, K. Dev, S.A. Khowaja, N.M.F. Qureshi, Ai-enabled privacy-preservation phrase with multi-keyword ranked searching for sustainable edge-cloud networks in the era of industrial iot, Ad Hoc Netw. 125 (2022) 102740.

[68] D.K. Soother, S.M. Ujjan, K. Dev, S.A. Khowaja, N.A. Bhatti, T. Hussain, Towards soft real-time fault diagnosis for edge devices in industrial iot using deep domain adaptation training strategy, J. Parallel Distrib. Comput. 160 (2022) 90–99.

[69] B. Yin, H. Yin, Y. Wu, Z. Jiang, FDC: A secure federated deep learning mechanism for data collaborations in the internet of things, IEEE Internet Things J. 7 (7) (2020) 6348–6359.

[70] Y. Lu, The blockchain: State-of-the-art and research challenges, J. Ind. Inf. Integ. 15 (2019) 80–90.

[71] T.R. Gadekallu, Q.-V. Pham, D.C. Nguyen, P.K.R. Maddikunta, N. Deepa, B. Prabadevi, P.N. Pathirana, J. Zhao, W.-J. Hwang, Blockchain for edge of things: applications, opportunities, and challenges, IEEE Internet Things J. 9 (2) (2021) 964–988.

[72] H. Kim, J. Park, M. Bennis, S.-L. Kim, Blockchained on-device federated learning, IEEE Commun. Lett. 24 (6) (2019) 1279–1283.

[73] Y. Qu, L. Gao, T.H. Luan, Y. Xiang, S. Yu, B. Li, G. Zheng, Decentralized privacy using blockchain-enabled federated learning in Fog computing, IEEE Internet Things J. 7 (6) (2020) 5171–5183.

[74] S.R. Pokhrel, J. Choi, Federated learning with blockchain for autonomous vehicles: Analysis and design challenges, IEEE Trans. Commun. 68 (8) (2020) 4734–4746.

[75] Z. Li, J. Liu, J. Hao, H. Wang, M. Xian, CrowdSFL: A secure crowd computing framework based on blockchain and federated learning, Electronics 9 (5) (2020) 773.

[76] G. Hua, L. Zhu, J. Wu, C. Shen, L. Zhou, Q. Lin, Blockchain-based federal learning for intelligent control in heavy haul railway, IEEE Access 8 (2020) 176830–176839.

[77] P.K. Sharma, J.H. Park, K. Cho, Blockchain and federated learning-based distributed computing defence framework for sustainable society, Sustainable Cities Soc. 59 (2020) 102220.

[78] W. Guo, I. Kotsia, I. Patras, Tensor learning for regression, IEEE Trans. Image Process. 21 (2) (2011) 816–827.

[79] Z. Lai, W.K. Wong, Y. Xu, C. Zhao, M. Sun, Sparse alignment for robust tensor learning, IEEE Trans. Neural Netw. Learn. Syst. 25 (10) (2014) 1779–1792.

[80] J. Feng, L.T. Yang, X. Liu, R. Zhang, Privacy-preserving tensor analysis and processing models for wireless internet of things, IEEE Wirel. Commun. 25 (6) (2018) 98–103.

[81] B. Babayigit, H. Sattuf, An IIoT and web-based low-cost SCADA system for industrial automation, in: 2019 11th International Conference on Electrical and Electronics Engineering, ELECO, IEEE, 2019, pp. 890–894.

[82] S. Leminen, M. Rajahonka, R. Wendelin, M. Westerlund, Industrial internet of things business models in the machine-to-machine context, Ind. Mark. Manag. 84 (2020) 298–311.

[83] S.B. Prathiba, G. Raja, S. Anbalagan, K. Dev, S. Gurumoorthy, A.P. Sankaran, Federated learning empowered computation offloading and resource management in 6G-V2X, IEEE Trans. Netw. Sci. Eng. (2021).

[84] S. Munirathinam, Industry 4.0: Industrial internet of things (IIoT), in: Advances in Computers, Vol. 117, Elsevier, 2020, pp. 129–164.

[85] T. Reddy, S.P. RM, M. Parimala, C.L. Chowdhary, S. Hakak, W.Z. Khan, et al., A deep neural networks based model for uninterrupted marine environment monitoring, Comput. Commun. 157 (2020) 64–75.

[86] M. Parimala, R. Swarna Priya, M. Praveen Kumar Reddy, C. Lal Chowdhary, R. Kumar Poluru, S. Khan, Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach, in: Software: Practice and Experience, 2020.

[87] S. Sarkar, S. Agrawal, T. Baker, P.K.R. Maddikunta, T.R. Gadekallu, Catalysis of neural activation functions: Adaptive feed-forward training for big data applications, Appl. Intell. (2022) 1–20.

[88] Q.-V. Pham, H.T. Nguyen, Z. Han, W.-J. Hwang, Coalitional games for computation offloading in NOMA-enabled multi-access edge computing, IEEE Trans. Veh. Technol. 69 (2) (2020) 1982–1993.

[89] J.-S. Fu, Y. Liu, H.-C. Chao, B.K. Bhargava, Z.-J. Zhang, Secure data storage and searching for industrial IoT by integrating fog computing and cloud computing, IEEE Trans. Ind. Inf. 14 (10) (2018) 4519–4528.

[90] L. Tang, B. Tang, L. Tang, F. Guo, J. Zhang, Reliable mobile edge service offloading based on p2p distributed networks, Symmetry 12 (5) (2020) 821.

[91] P. Borylo, A. Lason, J. Rzasa, A. Szymanski, A. Jajszczyk, Energy-aware fog and cloud interplay supported by wide area software defined networking, in: 2016 IEEE International Conference on Communications, ICC, IEEE, 2016, pp. 1–7.

[92] K. Kaur, S. Garg, G.S. Aujla, N. Kumar, J.J. Rodrigues, M. Guizani, Edge computing in the industrial internet of things environment: Software-defined-networks-based edge-cloud interplay, IEEE Commun. Mag. 56 (2) (2018) 44–51.

[93] R.H. Jhaveri, S.V. Ramani, G. Srivastava, T.R. Gadekallu, V. Aggarwal, Fault-resilience for bandwidth management in industrial software-defined networks, IEEE Trans. Netw. Sci. Eng. 8 (4) (2021) 3129–3139.

[94] A. Singh, S. Garg, K. Kaur, S. Batra, N. Kumar, K.-K.R. Choo, Fuzzy-folded bloom filter-as-a-service for big data storage in the cloud, IEEE Trans. Ind. Inf. 15 (4) (2018) 2338–2348.

[95] M. Saqlain, M. Piao, Y. Shim, J.Y. Lee, Framework of an IoT-based industrial data management for smart manufacturing, J. Sensor Actuator Netw. 8 (2) (2019) 25.

[96] S.D. Anton, D. Fraunholz, J. Zemitis, F. Pohl, H.D. Schotten, Highly scalable and flexible model for effective aggregation of context-based data in generic IIoT scenarios, 2019, arXiv preprint arXiv:1906.03064.

[97] S. Du, B. Liu, H. Ma, G. Wu, P. Wu, IIoT -based intelligent control and management system for motorcycle endurance test, IEEE Access 6 (2018) 30567–30576.

[98] Z. Shu, J. Wan, D. Zhang, D. Li, Cloud-integrated cyber–physical systems for complex industrial applications, Mob. Netw. Appl. 21 (5) (2016) 865–878.

[99] C.H. Liu, Q. Lin, S. Wen, Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning, IEEE Trans. Ind. Inf. 15 (6) (2018) 3516–3526.

[100] S. Manzoor, A.N. Mian, A. Zoha, M.A. Imran, Federated learning empowered mobility-aware proactive content offloading framework for fog radio access networks, Future Gener. Comput. Syst. (2022).

[101] P.D.U. Coronado, R. Lynn, W. Louhichi, M. Parto, E. Wescoat, T. Kurfess, Part data integration in the shop floor digital twin: Mobile and cloud technologies to enable a manufacturing execution system, J. Manuf. Syst. 48 (2018) 25–33.

[102] S.A. Khowaja, K. Dev, P. Khowaja, P. Bellavista, Toward energy-efficient distributed federated learning for 6G networks, IEEE Wirel. Commun. 28 (6) (2021) 34–40.

[103] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, M. Chen, In-edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning, IEEE Netw. 33 (5) (2019) 156–165.

[104] H. Zhu, Y. Jin, Multi-objective evolutionary federated learning, IEEE Trans. Neural Netw. Learn. Syst. 31 (4) (2019) 1310–1322.

[105] R. Kanagavelu, Z. Li, J. Samsudin, Y. Yang, F. Yang, R.S.M. Goh, M. Cheah, P. Wiwatphonthana, K. Akkarajitsakul, S. Wang, Two-phase multi-party computation enabled privacy-preserving federated learning, in: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGRID, IEEE, 2020, pp. 410–419.

[106] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, S. Liu, Efficient and privacy-enhanced federated learning for industrial artificial intelligence, IEEE Trans. Ind. Inf. 16 (10) (2019) 6532–6542.

[107] M.A. Jarwar, S.A. Khowaja, K. Dev, M. Adhikari, S. Hakak, Neat: A resilient deep representational learning for fault detection using acoustic signals in iiot environment, IEEE Internet Things J. (2021).

[108] K. Zhang, Y. Zhu, S. Maharjan, Y. Zhang, Edge intelligence and blockchain empowered 5G beyond for the industrial internet of things, IEEE Netw. 33 (5) (2019) 12–19.

[109] B. Qolomany, K. Ahmad, A. Al-Fuqaha, J. Qadir, Particle swarm optimized federated learning for industrial IoT and smart city services, 2020, arXiv preprint arXiv:2009.02560.

[110] M. Chen, Z. Yang, W. Saad, C. Yin, H.V. Poor, S. Cui, A joint learning and communications framework for federated learning over wireless networks, IEEE Trans. Wireless Commun. 20 (1) (2021) 269–283.

[111] M.S.H. Abad, E. Ozfatura, D. Gunduz, O. Ercetin, Hierarchical federated learning across heterogeneous cellular networks, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2020, pp. 8866–8870.

[112] L. Huang, X. Feng, C. Zhang, L. Qian, Y. Wu, Deep reinforcement learning-based joint task offloading and bandwidth allocation for multi-user mobile edge computing, Digit. Commun. Netw. 5 (1) (2019) 10–17.

[113] L.U. Khan, M. Alsenwi, I. Yaqoob, M. Imran, Z. Han, C.S. Hong, Resource optimized federated learning-enabled cognitive internet of things for smart industries, IEEE Access 8 (2020) 168854–168864.

[114] T. Hiessl, D. Schall, J. Kemnitz, S. Schulte, Industrial federated learning–requirements and system design, in: International Conference on Practical Applications of Agents and Multi-Agent Systems, Springer, 2020, pp. 42–53.

[115] Y. Lu, X. Huang, K. Zhang, S. Maharjan, Y. Zhang, Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks, IEEE Trans. Ind. Inf. 17 (7) (2021) 5098–5107.

[116] S. Messaoud, A. Bradai, O. Ben Ahmed, P. Quang, M. Atri, M.S. Hossain, Deep federated Q-learning-based network slicing for industrial IoT, IEEE Trans. Ind. Inf. 17 (8) (2021) 5572–5582.

[117] H. Zhang, L. Hanzo, Federated learning assisted multi-UAV networks, IEEE Trans. Veh. Technol. (2020).

[118] F. Tao, Q. Qi, L. Wang, A. Nee, Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: correlation and comparison, Engineering 5 (4) (2019) 653–661.

[119] F. Learning, A step closer towards confidential AI, 2020.

[120] Q.-V. Pham, T. Huynh-The, M. Alazab, J. Zhao, W.-J. Hwang, Sum-rate maximization for UAV-assisted visible light communications using NOMA: Swarm intelligence meets machine learning, IEEE Internet Things J. 7 (10) (2020) 10375–10387.

[121] Q.-V. Pham, M. Zeng, R. Ruby, T. Huynh-The, W.-J. Hwang, UAV communications for sustainable federated learning, IEEE Trans. Veh. Technol. 70 (4) (2021) 3944–3948.

[122] Q.-V. Pham, M. Le, T. Huynh-The, Z. Han, W.-J. Hwang, Energy-efficient federated learning over UAV-enabled wireless powered communications, IEEE Trans. Veh. Technol. (2022).

[123] Q.V. Do, Q.-V. Pham, W.-J. Hwang, Deep reinforcement learning for energy-efficient federated learning in UAV-enabled wireless powered networks, IEEE Commun. Lett. 26 (1) (2022) 99–103.

[124] H.K. Narsani, P. Raut, K. Dev, K. Singh, C.-P. Li, Interference limited network for factory automation with multiple packets transmissions, in: 2021 IEEE 18th Annual Consumer Communications & Networking Conference, CCNC, IEEE, 2021, pp. 1–6.

[125] T. Sery, N. Shlezinger, K. Cohen, Y.C. Eldar, Over-the-air federated learning from heterogeneous data, IEEE Trans. Signal Process. 69 (2021) 3796–3811.

[126] B. McMahan, D. Ramage, Federated learning: Collaborative machine learning without centralized training data, Google Res. Blog 3 (2017).

[127] M. Abu-Elkheir, M. Hayajneh, N.A. Ali, Data management for the internet of things: Design primitives and solution, Sensors 13 (11) (2013) 15582–15612.

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

[129] T.T. Anh, N.C. Luong, D. Niyato, D.I. Kim, L.-C. Wang, Efficient training management for mobile crowd-machine learning: A deep reinforcement learning approach, IEEE Wirel. Commun. Lett. 8 (5) (2019) 1345–1348.

[130] 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: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175–1191.

[131] R. James, et al., The internet of things: a study in hype, reality, disruption, and growth, raymond james US research, Technol. Commun. Ind. Rep. (2014).

[132] S. Zeadally, R. Hunt, Y.-S. Chen, A. Irwin, A. Hassan, Vehicular ad hoc networks (VANETS): status, results, and challenges, Telecommun. Syst. 50 (4) (2012) 217–241.

[133] J.T. Isaac, S. Zeadally, J.S. Camara, Security attacks and solutions for vehicular ad hoc networks, IET Commun. 4 (7) (2010) 894–903.

[134] J. Guerrero-Ibáñez, C. Flores-Cortés, S. Zeadally, Vehicular ad-hoc networks (VANETS): architecture, protocols and applications, in: Next-Generation Wireless Technologies, Springer, 2013, pp. 49–70.

[135] Z. Du, C. Wu, T. Yoshinaga, K.-L.A. Yau, Y. Ji, J. Li, Federated learning for vehicular internet of things: Recent advances and open issues, IEEE Open J. Comput. Soc. 1 (2020) 45–61.

[136] M. Zolanvari, M.A. Teixeira, L. Gupta, K.M. Khan, R. Jain, Machine learning-based network vulnerability analysis of industrial internet of things, IEEE Internet Things J. 6 (4) (2019) 6822–6834.

[137] K. Dev, S.A. Khowaja, P.K. Sharma, B.S. Chowdhry, S. Tanwar, G. Fortino, Ddi: A novel architecture for joint active user detection and iot device identification in grant-free noma systems for 6G and beyond networks, IEEE Internet Things J. (2021).

[138] G. Xu, H. Li, Y. Dai, K. Yang, X. Lin, Enabling efficient and geometric range query with access control over encrypted spatial data, IEEE Trans. Inf. Forensics Secur. 14 (4) (2018) 870–885.

[139] T. Pang, K. Xu, C. Du, N. Chen, J. Zhu, Improving adversarial robustness via promoting ensemble diversity, 2019, arXiv preprint arXiv:1901.08846.

[140] U. Hwang, J. Park, H. Jang, S. Yoon, N.I. Cho, Puvae: A variational autoencoder to purify adversarial examples, IEEE Access 7 (2019) 126582–126593.

[141] A. Husaković, E. Pfann, M. Huemer, Robust machine learning based acoustic classification of a material transport process, in: 2018 14th Symposium on Neural Networks and Applications, NEUREL, IEEE, 2018, pp. 1–4.

[142] S. Agrawal, S. Sarkar, O. Aouedi, G. Yenduri, K. Piamrat, S. Bhattacharya, P.K.R. Maddikunta, T.R. Gadekallu, Federated learning for intrusion detection system: concepts, challenges and future directions, 2021, arXiv preprint arXiv:2106.09527.

[143] Y. Qu, S.R. Pokhrel, S. Garg, L. Gao, Y. Xiang, A blockchained federated learning framework for cognitive computing in industry 4.0 networks, IEEE Trans. Ind. Inf. 17 (4) (2020) 2964–2973.

[144] P. Kumar, G.P. Gupta, R. Tripathi, Pefl: Deep privacy-encoding based federated learning framework for smart agriculture, IEEE Micro. (2021).

[145] M. Wazid, A.K. Das, N. Kumar, J.J. Rodrigues, Secure three-factor user authentication scheme for renewable-energy-based smart grid environment, IEEE Trans. Ind. Inf. 13 (6) (2017) 3144–3153.

[146] C. Briggs, Z. Fan, P. Andras, Federated learning for short-term residential energy demand forecasting, 2021, arXiv preprint arXiv:2105.13325.

[147] A. Challoner, G.H. Popescu, Intelligent sensing technology, smart healthcare services, and internet of medical things-based diagnosis, Am. J. Med. Res. 6 (1) (2019) 13–18.

[148] R. Gupta, S. Tanwar, S. Tyagi, N. Kumar, Tactile-internet-based telesurgery system for healthcare 4.0: An architecture, research challenges, and future directions, IEEE Netw. 33 (6) (2019) 22–29.

[149] S. Wang, J. Wan, D. Zhang, D. Li, C. Zhang, Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination, Comput. Netw. 101 (2016) 158–168.

[150] P. Gope, T. Hwang, BSN-care: A secure IoT-based modern healthcare system using body sensor network, IEEE Sens. J. 16 (5) (2015) 1368–1376.

[151] B. Diène, J.J. Rodrigues, O. Diallo, E.H.M. Ndoye, V.V. Korotaev, Data management techniques for internet of things, Mech. Syst. Signal Process. 138 (2020) 106564.

[152] X. Zheng, J. Lu, S. Sun, D. Kiritsis, Decentralized industrial IoT data management based on blockchain and IPFS, in: IFIP International Conference on Advances in Production Management Systems, Springer, 2020, pp. 222–229.

[153] P.L.G. Ramírez, M. Taha, J. Lloret, J. Tomás, An intelligent algorithm for resource sharing and self-management of wireless-IoT-gateway, IEEE Access 8 (2019) 3159–3170.

[154] Y. Aono, T. Hayashi, L. Wang, S. Moriai, et al., Privacy-preserving deep learning via additively homomorphic encryption, IEEE Trans. Inf. Forensics Secur. 13 (5) (2017) 1333–1345.

[155] S. Rahimian, T. Orekondy, M. Fritz, Sampling attacks: Amplification of membership inference attacks by repeated queries, 2020, arXiv preprint arXiv:2009.00395.

[156] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, V. Shmatikov, How to backdoor federated learning, in: International Conference on Artificial Intelligence and Statistics, PMLR, 2020, pp. 2938–2948.

[157] R. Tourani, S. Misra, T. Mick, G. Panwar, Security, privacy, and access control in information-centric networking: A survey, IEEE Commun. Surv. Tutor. 20 (1) (2017) 566–600.

[158] A.K. Mishra, A.K. Tripathy, D. Puthal, L.T. Yang, Analytical model for sybil attack phases in internet of things, IEEE Internet Things J. 6 (1) (2018) 379–387.

[159] S. Wang, T. Tuor, T. Salonidis, K.K. Leung, C. Makaya, T. He, K. Chan, Adaptive federated learning in resource constrained edge computing systems, IEEE J. Sel. Areas Commun. 37 (6) (2019) 1205–1221.

[160] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang, M.S. Hossain, Deep anomaly detection for time-series data in industrial IoT: A communication-efficient on-device federated learning approach, IEEE Internet Things J. 8 (8) (2021) 6348–6358.

[161] W. Zhang, Q. Lu, Q. Yu, Z. Li, Y. Liu, S.K. Lo, S. Chen, X. Xu, L. Zhu, Blockchain-based federated learning for device failure detection in industrial IoT, IEEE Internet Things J. (2020).

[162] Y. Song, T. Liu, T. Wei, X. Wang, Z. Tao, M. Chen, FDA3: Federated defense against adversarial attacks for cloud-based IIoT applications, IEEE Trans. Ind. Inf. (2020).

[163] W.Y.B. Lim, Z. Xiong, J. Kang, D. Niyato, C. Leung, C. Miao, X. Shen, When information freshness meets service latency in federated learning: A task-aware incentive scheme for smart industries, IEEE Trans. Ind. Inf. 18 (1) (2020) 457–466.

[164] N.E. Petroulakis, E. Lakka, E. Sakic, V. Kulkarni, K. Fysarakis, I. Somarakis, J. Serra, L. Sanabria-Russo, D. Pau, M. Falchetto, et al., Semiotics architectural framework: End-to-end security, connectivity and interoperability for industrial IoT, in: 2019 Global IoT Summit, GIoTS, IEEE, 2019, pp. 1–6.

[165] M. Platenius-Mohr, S. Malakuti, S. Grüner, J. Schmitt, T. Goldschmidt, File-and API-based interoperability of digital twins by model transformation: An IIoT case study using asset administration shell, Future Gener. Comput. Syst. 113 (2020) 94–105.

[166] L. Melis, C. Song, E. De Cristofaro, V. Shmatikov, Inference attacks against collaborative learning, vol. 13, 2018, arXiv preprint arXiv:1805.04049.

[167] L. Zong, F.H. Memon, X. Li, H. Wang, K. Dev, End-to-end transmission control for cross-regional industrial internet of things in industry 5.0, IEEE Trans. Ind. Inf. 18 (6) (2021) 4215–4223.

[168] F. Khan, A. ur Rehman, Y. Zhang, S. Mastorakis, H. Song, M.A. Jan, K. Dev, A secured and reliable continuous transmission scheme in cognitive harq-aided internet of things, IEEE Internet Things J. 8 (19) (2021) 14835–14844.

[169] T. Nishio, R. Yonetani, Client selection for federated learning with heterogeneous resources in mobile edge, in: ICC 2019-2019 IEEE International Conference on Communications, ICC, IEEE, 2019, pp. 1–7.

[170] Y. Liu, N. Kumar, Z. Xiong, W.Y.B. Lim, J. Kang, D. Niyato, Communication-efficient federated learning for anomaly detection in industrial internet of things, in: 2020 IEEE Global Communications Conference, GLOBECOM, IEEE, 2020, pp. 1–6.

[171] P.K.R. Maddikunta, Q.-V. Pham, B. Prabadevi, N. Deepa, K. Dev, T.R. Gadekallu, R. Ruby, M. Liyanage, Industry, 5.0: A survey on enabling technologies and potential applications, J. Ind. Inf. Integ. (2021) 100257.

Parimala M. is working as an Associate Professor in VIT University and completed her PhD in 2017. She has a experience of more than 13 years in field of teaching and research. She worked as a Co-Investigator for 3 years in a project on Spatio-temporal transmission model of Influenza A (H1N1) using clustering technique funded by the Indian Council of Medical Research (ICMR), Government of India under award number 32/1/2010-ECD-I. Research has been carried out in interdisciplinary settings of Data Mining, epidemiology and social network analysis. She is an Expertise in modeling disease dynamics and mining pattern from social network. She published research papers in reputed journals, written book chapters and presented papers in reputed conferences.

Swarna Priya R. M is currently working as Associate Professor in School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India. She received her B.E. in Computer Science and Engineering from Periyar University, Tamil Nadu, M.Tech. in Software Engineering from Anna University, Chennai, Tamil Nadu and Ph.D. in Computer Science and Engineering from Vellore Institute of Technology, Vellore, India. Her thesis was on Hyperspectral Satellite Imagery and its processing. She has an experience of 16 years in academic and 2 years in industry. She has more than 30 international/national publications. Dr. Priya is a member of Indian Society of Technical Education (ISTE), Computer Society of India (CSI), Indian Science Congress and the Institution of Engineers. Her current research interests include Machine Learning, Federated Learning, Internet of Things, Brain Computer Interface, Cognitive Computing and Advanced Image Processing.

Quoc-Viet Pham received the B.S. degree in electronics and telecommunications engineering from the Hanoi University of Science and Technology, Vietnam, in 2013, and the Ph.D. degree in telecommunications engineering from Inje University, Republic of Korea, in 2017. He has been a Research Professor with Pusan National University, Republic of Korea, since Jan. 2020. He is specialized in applying convex optimization, game theory, and machine learning to analyze and optimize edge computing and future wireless communications. He has been granted the Korea NRF Funding for outstanding young researchers for the term 2019–2024. He is an editor of Journal of Network and Computer Applications (Elsevier), Scientific Reports (Nature), and Frontiers in Communications and Networks, and a lead guest editor of the IEEE Internet of Things Journal. He was also the recipient of the Best Ph.D. Dissertation Award in engineering from Inje University in 2017, the Top Reviewer Award from the IEEE Transactions on Vehicular Technology in 2020, and the golden globe award 2021 from the Ministry of Science and Technology (Vietnam).

Kapal Dev is Senior Researcher at MTU, Ireland working on a Industrial project where is Investigating the application of DLT and Smart Contracts in the manufacturing sector. He was a Post-doc at Trinity College Dublin (TCD). He is Associate Editor (AE) in Springer Wireless Networks, Elsevier Physical Communication, IET Quantum Communication, IET Networks, Topic Editor in MDPI Network. He is Guest Editor (GE) in Q1 journals; IEEE TII, IEEE TNSE, IEEE TGCN, IEEE Standard Communication Magazine, Elsevier COMCOM and COMNET. He contributed(ing) as Lead chair in one of ACM MobiCom 2021, Globecom 2021, CCNC 2021 workshops, TPC member of several top conferences. He contributed as PI for Erasmus + International Credit Mobility (ICM), Capacity Building for Higher Education, and H2020 Co-Fund projects. His research interests include Blockchain, Wireless Networks and Artificial Intelligence.

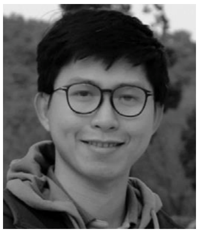
Sharnil Pandya (Senior Member, IEEE) is an Associate Professor at Computer Science and AIML dept, at Symbiosis Institute of Technology, and Research Faculty at Symbiosis Centre for Applied AI(SCAAI). His research interests include Smart Sensing, Ambient Healthcare Systems, Acoustics and Sound, and Computer Vision. He has also worked on two national-level funded projects: Defense Research and Development Organization(DRDO), India, and Department of Science & Technology(DST), worth 16 and 10 lacs. He has published more than 80 international/national publications. He has guided/guiding numerous national/international undergraduate, post-graduate and PhD students from computer engineering and other disciplines. In addition to academics and research, he is also a reviewer for reputed journals such as IEEE Transactions on Industrial Electronics, IEEE Internet of Things, IEEE Sensors journal, Advanced Engineering Informatics, Elsevier, Transactions on Emerging Telecommunications Technologies, and JBHI.

Praveen Kumar Reddy is currently working as Assistant Professor Senior in the School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu, India. He obtained his B.Tech. in CSE from Jawaharlal Nehru Technological University, India, M.Tech. in CSE from VIT University, Vellore, Tamil Nadu, India, and completed his Ph.D. in VIT, Vellore, Tamil Nadu, India. He was a Visiting Professor with the Guangdong University of Technology, China, in 2019. He worked with Alcatel-Lucent as a software developer in 2013. He produced more than 75 international/national publications.

Thippa Reddy Gadekallu is currently working as Associate Professor in School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu, India. He obtained his B.Tech. in CSE from Nagarjuna University, India, M.Tech. in CSE from Anna University, Chennai, Tamil Nadu, India and completed his Ph.D in VIT, Vellore, Tamil Nadu, India. He has more than 14 years of experience in teaching. He has published more than 80 international/national publications. Currently, his areas of research include Machine Learning, Internet of Things, Deep Neural Networks, Blockchain, Computer Vision.

Google Scholar: https://scholar.google.com/citations?user=nQFCxmkAAAAJ&hl=en&oi=ao

Researchgate:https://www.researchgate.net/profile/Thippa_Gadekallu



Thien Huynh-The received the B.S. degree in electronics and telecommunication engineering from the Ho Chi Minh City University of Technology and Education, Vietnam, in 2011, and the Ph.D. degree in computer science and engineering from Kyung Hee University (KHU), South Korea, in 2018. From March to August 2018, he was a Postdoctoral researcher with Kyung Hee University. He is currently a Postdoctoral Research Fellow with the ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea. His current research interests include radio signal processing, digital image processing, computer vision, machine learning, and deep learning. He was awarded with the Superior Thesis Prize from KHU.